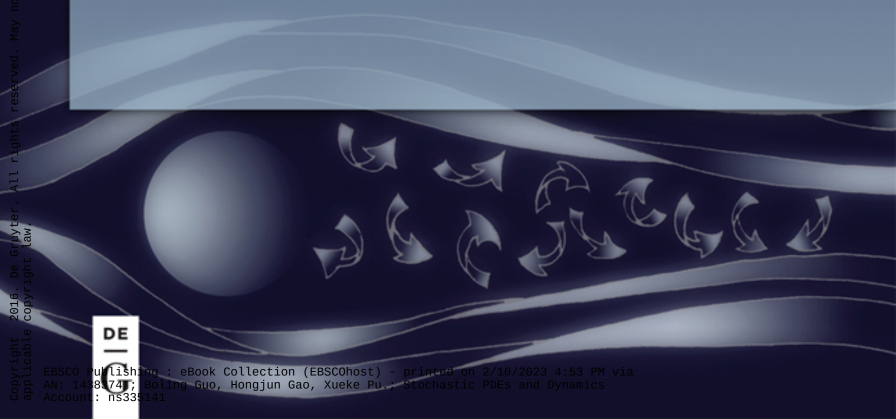
DE GRUYTER

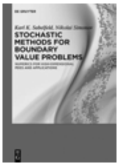
Boling Guo, Hongjun Gao, Xueke Pu

STOCHASTIC PDES AND DYNAMICS

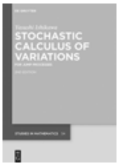


Boling Guo, Hongjun Gao, Xueke Pu
Stochastic PDEs and Dynamics

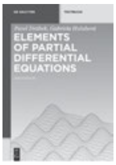
Also of interest



Stochastic Methods for Boundary Value Problems
Karl K. Sabelfeld, Nikolai A. Simonov, 2016
ISBN 978-3-11-047906-5, e-ISBN 978-3-11-047916-4

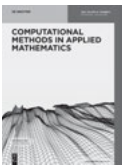


Stochastic Calculus of Variations. For Jump Processes
Yasushi Ishikawa, 2016
ISBN 978-3-11-037776-7, e-ISBN 978-3-11-039232-6



Elements of Partial Differential Equations
Pavel Drábek, Gabriela Holubová, 2014
ISBN 978-3-11-031665-0, e-ISBN 978-3-11-037404-9

Journal of Numerical Mathematics
Ronald H. W. Hoppe, Yuri Kuznetsov (Editor-in-Chief)
ISSN 1570–2820, e-ISSN 1569–3953



Computational Methods in Applied Mathematics
Carsten Carstensen (Editor-in-Chief)
ISSN 1609–4840, e-ISSN 1609–9389

Radon Series on Computational and Applied Mathematics
Ulrich Langer, Hansjörg Albrecher, Heinz W. Engl, Ronald H. W. Hoppe,
Karl Kunisch, Harald Niederreiter, Christian Schmeiser (Eds.)
ISSN 1865–3707

Boling Guo, Hongjun Gao, Xueke Pu

Stochastic PDEs and Dynamics

DE GRUYTER

Authors
Prof. Boling Guo
Laboratory of Computational Physics
Institute of Applied Physics & Computational Mathematics
Huayuan Road 6
100088 Beijing
China

Prof. Hongjun Gao
Nanjing Normal University
Institute of Mathematics, School of Mathematics Sciences
Wenyuan Road 1
210023 Nanjing
China

Dr. Xueke Pu
Chongqing University
College of Mathematics and Statistics
Shazheng Street 174
404100 Chongqing
China

ISBN 978-3-11-049510-2
e-ISBN (PDF) 978-3-11-049388-7
e-ISBN (EPUB) 978-3-11-049243-9
Set-ISBN 978-3-11-049389-4

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: Integra Software Services Pvt. Ltd.
Cover image: Chong Guo, guochong200805@163.com
Printing and binding: CPI books GmbH, Leck
♾ Printed on acid-free paper
Printed in Germany

www.degruyter.com

Preface

The past two decades has witnessed the emergence of stochastic nonlinear partial differential equations (SPDEs) and their dynamics in different fields such as physics, mechanics, finance and biology etc. For example, there are corresponding SPDEs descriptions for the atmosphere-oceanic circulation, for the nonlinear wave propagation in random media, for the pricing of risky assets, and the law of fluctuation of stock market prices. In early 1970s, mathematicians such as Bensoussan, Temam, Pardoux, just to name only a few, initiated the mathematical studies of SPDEs and the studies of corresponding random dynamical systems began slightly later afterwards. In the middle 90's, Crauel, Da Prato, Debussche, Flandoli, Schmalfuss, Zabczyk *et al* established the framework of random attractors, Hausdorff dimension estimates and Invariant measure theory of random dynamic systems with applications to stochastic nonlinear evolutionary equations. Recently, there are theoretical and numerical aspects of nonlinear SPDEs has been developed, resulting in many fruitful achievement and subsequently, many monographs were published.

The authors of this book had been working in the fields of nonlinear SPDEs and random dynamics as well as stochastic processes such as Lévy process and fractional Wiener process for more than a decade. Seminars were held and discussions had been going with scholars all over the world since then. Interesting and preliminary results were made on some mathematical problems in climate, ocean circulation and propagation of nonlinear waves in random media.

The aim of book is twofold. First, to give some preliminaries that are of importance to SPDEs. Second, to introduce latest recent results concerning several important SPDEs such as Ginzburg-Landau equation, Ostrovsky equation, geostrophic equations and primitive equations in climate. Materials are presented in a concise way, hoping to bring readers into such an interesting field of modern applied mathematics.

Chapter one introduces preliminaries in probability and stochastic processes, and Chapter 2 briefly presents the stochastic integral and Ito formula, which plays a vital role in stochastic partial differential equations. Chapter 3 discusses the Ornstein-Ulenbeck process and some linear SDEs. Chapter 4 establishes the basic framework of stochastic dynamic systems. In Chapter 5, latest results on several SPDEs emerging from various physics backgrounds are given.

Last but not the least, I would like to take the opportunity to express sincere gratitudes to Dr. Mufa Chen, Member of Chinese Academy of Sciences, and Dr. Jian Wang at Fuzhou University, from whom we benefited constantly in preparing this book.

Boling Guo
August 20, 2016

Contents

1 Preliminaries

This chapter contains some preliminaries in probability and stochastic processes, especially some basic properties of the Wiener process, Poisson process and Lévy process. Because many contents in this chapter can be found in the other literature, we here only give the conclusions and omit the proofs.

1.1 Preliminaries in probability

1.1.1 Probability space

There are many uncertainties and randomness in our natural world and social environments. A lot of observations and tests are asking for the research in random phenomena. A random experiment must contain certain properties, usually requiring that (i) the experiment can be repeated arbitrarily many times under the same conditions and (ii) the outcome for the experiment may be more than one and all the possible outcomes are known, but one can't accurately predict which outcome would appear in one trial.

The random experiment is usually called test for short and is expressed as E. Each possible result in E is called a basic event or a sample point, expressed as ω. The set of all sample points in E, denoted by Ω, is called the space for basic event, and the set of sample points is called event and is expressed in capital letters $A, B, C, \cdots$. The event A occurs if and only if one of the sample points in A occurs.

Take the "roll the dice" game as a simple example. The outcome can't be predicted in the experiment when the dice was rolled once, but certainly it is one of the outcomes "point one," $\cdots$, "point six." Hence, $\Omega = \{1, 2, 3, 4, 5, 6\}$ consists of six elements, representing the six possible outcomes in the "roll the dice" experiment. In this experiment, "roll prime number point" is an event and consists of three basic events 2, 3, 5, which we denote as $A = \{2, 3, 5\}$.

In practice, various manipulations such as intersection, union or complement to subsets are needed. It is nature to ask that whether the result is still an event after such manipulations. This leads to the concept of σ-algebra.

Definition 1.1.1. *Let Ω be a sample space, then the set $\mathscr{F} = \{A : A \subset \Omega\}$ is called a* σ-algebra *if it satisfies*

(1) $\Omega \in \mathscr{F}$;

(2) *if* $A \in \mathscr{F}$, *then* $A^c := \Omega \backslash A \in \mathscr{F}$;

(3) *if* $A_i \in \mathscr{F}$, *then* $\cup_{i=1}^{\infty} A_i \in \mathscr{F}$.

Then $(\Omega, \mathscr{F})$ is called a measurable space and each element in $\mathscr{F}$ is measurable. Two trivial examples are $\mathscr{F} = \{\emptyset, \Omega\}$ and $\mathscr{F}$ contains all the subsets of Ω. These two examples are too special to be studied in mathematics. The first one is so small that we cannot get enough information that we are interested in, while the latter one is so big that it is difficult to define a probability measure on it. Therefore, we need to consider other intermediate σ-algebra we are interested in. For a collection $\mathcal{C}$ of subsets of Ω, we denote $\sigma(\mathcal{C})$ the σ-algebra generated by $\mathcal{C}$, that is the smallest σ-algebra containing $\mathcal{C}$.

Definition 1.1.2. *Let $(\Omega, \mathscr{F})$ be a measurable space and P be a real valued function defined on the event field $\mathscr{F}$. If P satisfies*
(1) *for each $A_i \in \mathscr{F}$, it has $P(A_i) \geq 0$;*
(2) *$P(\Omega) = 1$;*
(3) *for $A_i \in \mathscr{F} (i = 1, 2, \cdots, \infty)$ with $i \neq j$, $A_iA_j \doteq A_i \cap A_j = \emptyset$, it has*

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i),$$

then P is called probability measure, and probability for short.

The above three properties are called Kolmogorov's axioms, named after Andrey Kolmogorov, one of the greatest Russian mathematicians. Such triple $(\Omega, \mathscr{F}, P)$ is called a measure space or a probability space in probability theory. In Kolmogrov's probability theory, $\mathscr{F}$ doesn't have to include all the possible subsets of Ω, but only includes the subsets we are interested in. In such a measure space, $\mathscr{F}$ is usually called an event field and the element in $\mathscr{F}$ is called an event or a measurable set. The event $A = \Omega$ is called certain event since the possibility for A to occur is $P(\Omega) = 1$ and the event $A = \emptyset$ is called impossible event accordingly since $P(\emptyset) = 0$ thanks to properties (2) and (3).

In the following, we regard Ω in $(\Omega, \mathscr{F}, P)$ as sample space, $\mathscr{F}$ as event field in Ω, and P as a determinate probability corresponding to $(\Omega, \mathscr{F})$. The properties in the definition are called non negativity, normalization, and complete additivity of probability, respectively.

We also note that for a fixed sample space Ω, many σ-algebra can be constructed (hence not unique), but not every σ-algebra is an event field. For example, let $\mathscr{F}_1 = \{\emptyset, \Omega\}$ and $\mathscr{F}_2 = \{\emptyset, A, A^c, \Omega\}$, where $A \subset \Omega$. By definition, $\mathscr{F}_1$ is certainly an event field, but $\mathscr{F}_2$ is not necessarily an event field since A is possibly not measurable under P.

After introducing the probability space, the relations and operations among events and the conditional probability can be considered. Two events A and B are called mutually exclusive, if both A, B can't occur in the same experiment (but it is possible that neither of them occurs). If any two events are exclusive, then these events are called mutually exclusive pairwise.

Theorem 1.1.1. *The probability for the sum of some of mutual exclusion events is equal to the sum of every event, i.e.,*

$$P(A_1 + A_2 + \cdots) = P(A_1) + P(A_2) + \cdots,$$

if $A_i \cap A_j = \emptyset$ *whenever* $i \neq j$.

The conditional probability measures the probability of an event under the assumption that another event has occurred. For example, we are interested in the probability of "prime number occurs" in the "roll the dice" game, under the assumption that we have known that the outcome in one trial is an odd number. We give the definition below.

Definition 1.1.3. *Let* $(\Omega, \mathscr{F}, P)$ *be a probability space,* $A, B \in \mathscr{F}$ *and* $P(B) \neq 0$. *Then the conditional probability of* A *given* B *or the probability of* A *under the condition* B *is defined by*

$$P(A|B) = P(AB)/P(B), \tag{1.1.1}$$

where $P(AB) = P(A \cap B) = P(A \text{ and } B \text{ both occur})$.

We also take the "roll the dice" as an example. Consider the event A="prime point occurs," B="odd point occurs," and C="even point occurs," that is

$$A = \{2, 3, 5\}, \quad B = \{1, 3, 5\}, \quad C = \{2, 4, 6\}.$$

It can be calculated that the (unconditional) probability of A is 1/2. Now, if the event B is assumed to have occurred, then we ask for the probability of A. By the definition, the conditional probability of A under B is $P(A|B) = P(AB)/P(B) = P(\{3, 5\})/P(B) = 2/3$. Similarly, if the event C is assumed to have occurred, or we know that C has occurred, then the conditional probability of A is $P(A|C) = 1/3$.

Another important concept in probability is independence. Consider two events A and B. Generally speaking, the conditional probability $P(A)$ of A is different from $P(A|B)$. If $P(A|B) > P(A)$, then the occurrence of B enlarges than the probability of A. Otherwise, if $P(A) = P(A|B)$, then the occurrence of B has no influence on A. In the latter case, the events A, B are said to be independent and

$$P(AB) = P(A)P(B). \tag{1.1.2}$$

Definition 1.1.4. *If* A, B *satisfy eq.* (1.1.2)*, then* A, B *are said to be independent.*

Definition 1.1.5. *Let* $A_1, A_2, \cdots$ *be at most countably many events. If for any finite events* $A_{i_1}, A_{i_2}, \cdots, A_{i_m}$*, there holds*

$$P(A_{i_1} A_{i_2} \cdots A_{i_m}) = P(A_{i_1}) P(A_{i_2}) \cdots P(A_{i_m}), \tag{1.1.3}$$

then the events $A_1, A_2, \cdots,$ *are said to be independent.*

It is noted that the events in a subset of independent events are also independent.

Theorem 1.1.2. *Let $A_1, A_2, \cdots, A_n$ be independent, then*

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2)\cdots P(A_n).$$

Next, we introduce the law of total probability and Bayes formula. Let $B_1, B_2, \cdots$ be at most countably many events, mutually exclusive, and at least one of them happens in an experiment. That is, $B_iB_j = \emptyset$ (impossible event), when $i \neq j$ and $B_1 + B_2 + \cdots = \Omega$ (certain event). Given any event A, noting Ω is a certain event, one gets $A = A\Omega = AB_1 + AB_2 + \cdots$, where $AB_1, AB_2, \cdots$ are mutually exclusive as $B_1, B_2, \cdots$ are mutually exclusive. Hence, by Theorem 1.1.1

$$P(A) = P(AB_1) + P(AB_2) + \cdots, \tag{1.1.4}$$

and by the definition of conditional probability, we have $P(AB_i) = P(B_i)P(A|B_i)$, which follows that

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots. \tag{1.1.5}$$

This formula is called the law of total probability. By eqs (1.1.4) and (1.1.5), the probability of $P(A)$ is decomposed into the sum of many parts. It can be understood that the events B_i is a possible cause leading to A. The probability of A, under the possible cause B_i, is the conditional probability $P(A|B_i)$. Intuitively, the probability of A, $P(A)$, must be between the smallest and largest $P(A|B_i)$ under this mechanism, and also because the probabilities of $P(B_i)$ are different in all kinds of causes, the probability $P(A)$ should be a weighted average of $P(A|B_i)$, with the weight being $P(B_i)$.

Under the assumption of the law of total probability, one has

$$\begin{aligned} P(B|A) =& P(AB_i)/P(A) \\ =& P(B_i)P(A|B_i)/\sum_j P(B_j)P(A|B_j). \end{aligned} \tag{1.1.6}$$

This formula is called Bayes formula. Formally, it is just a simple deduction of the conditional probability and the law of total probability. It is famous for its explanation in reality and philosophical significance. For $P(B_i)$, it is the probability of B_i under no further information. Now, if it has new information (we know that A has occurred), then the probability of B_i has a new estimate. If the event of A is viewed as a result, and $B_1, B_2, \cdots$ are the possible causes of A, then we can formally view the law of total probability as “from the reason to result,” while Bayes formula can be viewed as “from the result to reasons.” In fact, a comprehensive set of statistical inference methods has been developed by the idea, which is called “Bayes statistics.”

1.1.2 Random variable and probability distribution

Random variable, as the name indicates, is a variable whose value is determined randomly. Strictly speaking, given a probability space $(\Omega, \mathscr{F}, P)$, random variable X is defined as a measurable mapping from Ω to R^d. When $d \geq 2$, X is usually called a random vector, and d is the dimension. Random vectors can be divided into discrete and continuous types according to the value of random variables. The research of random variable is the content in probability, because in a random experiment, what is concerned are variables, which are usually random and are often associated with certain problems of interests.

Next, we consider the distribution of random variable.

Definition 1.1.6. *Let X be a random variable. Then the function*

$$F(x) = P(X \leq x), \quad -\infty < x < \infty, \tag{1.1.7}$$

is called the probability distribution function of X, where $P(X \leq x)$ denotes the probability of the event $\{\omega : X(\omega) \leq x\}$.

Here, it doesn't request the random variable to be discrete or continuous. It's obvious that the distribution function has the following properties: (1) $F(x)$ is a monotonically nondecreasing function, (2) $F(x) \to 0$ as $x \to -\infty$, and (3) $F(x) \to 1$ as $x \to \infty$.

First, let us consider a discrete random variable X taking possible values $a_1, a_2, \cdots$. Then $p_i = P(X = a_i),\ i = 1, 2, \cdots$ is called the probability function of X. An important example of the discrete distribution is the Poisson distribution. If X is a non-negative integer-valued random variable with its probability function $p_i = P(X = i) = e^{-\lambda}\lambda^i/i!$, then X is said to subject to Poisson distribution, denoted by $X \sim P(\lambda)$, where $\lambda > 0$ is a constant.

For the distribution of continuous random variable, it can't be described as the discrete ones. One method to describe continuous random variable is to use distribution function and probability density function.

Definition 1.1.7. *Let $F(x)$ be the distribution function of a continuous random variable X, then the derivative $f(x) = F'(x)$ of $F(x)$, if exists, is called the probability density function of X.*

The density function $f(x)$ has the following properties:
(1) $f(x) \geq 0$.
(2) $\int_{-\infty}^{\infty} f(x)dx = 1$.
(3) For any $a < b$, there holds

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x)dx.$$

An important example of continuous distribution is the normal distribution, whose probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$

The associated random variable is usually denoted by $X \sim N(\mu, \sigma^2)$.

The above conclusions can be generalized to random vectors. Consider a d-dimensional random vector $X = (X_1, \cdots, X_n)$, whose components $X_1, \cdots, X_n$ are one-dimensional random variable. For $A \subset R^n$, $X \in A$ denotes $\{\omega : X(\omega) \in A\}$.

Definition 1.1.8. *A nonnegative function $f(x_1, \cdots, x_n)$ on R^n is said to be the probability density function of X if*

$$P(X \in A) = \int_A f(x_1, \cdots, x_n)dx_1 \cdots dx_n, \tag{1.1.8}$$

for any $A \subset R^n$.

We remark that similar to the one-dimensional case, we can introduce the probability distribution function

$$F(x_1, x_2, \cdots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n),$$

for any random vector $X = (X_1, \cdots, X_n)$. For the random vector X, each competent X_i is one-dimensional random variable and has its own one-dimensional distribution function F_i, for $i = 1, \cdots, n$, which are called the “marginal distribution” of distribution F or of random vector X. It is easy to see that the marginal distribution is completely determined by the distribution F. For example, let $X = (X_1, X_2)$ with probability density function $f(x_1, x_2)$. Since $(X_1 \leq x_1) = (X_1 \leq x_1, X_2 < \infty)$, we have

$$F_1(x_1) = P(X_1 \leq x_1) = \int_{-\infty}^{x_1} dt_1 \int_{-\infty}^{\infty} f(t_1, t_2)dt_2,$$

and the probability density function of X_1 is given by

$$f_1(x_1) := \frac{dF_1(x_1)}{dx_1} = \int_{-\infty}^{\infty} f(x_1, x_2)dx_2.$$

Similarly, in the multi dimensional case, we have for $X = (X_1, \cdots, X_n)$ that

$$f(x_1) := \frac{dF_1(x_1)}{dx_1} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \cdots, x_n)dx_2 \cdots dx_n.$$

We recall the conditional probability, based on which independence is introduced. Now, we discuss the conditional probability distribution and the independence of random variables. We also take the continuous random variables for example. Given two random variables X_1 and X_2, we let $f(x_1, x_2)$ be the probability density function of the two-dimensional random vector $X = (X_1, X_2)$. Consider the conditional distribution of X_1 under the condition $a \le X_2 \le b$. Since

$$P(X_1 \le x_1 | a \le X_2 \le b) = P(X_1 \le x_1, a \le X_2 \le b)/P(a \le X_2 \le b),$$

by the marginal distribution function f_2 of X_2, it follows

$$P(X_1 \le x_1 | a \le X_2 \le b) = \int_{-\infty}^{x_1} dt_1 \int_a^b f(t_1, t_2) dt_2 / \int_a^b f_2(t_2) dt_2,$$

which is the conditional distribution function of X_1. The conditional density function can be obtained by derivative on x_1, i.e.,

$$f_1(X_1 | a \le X_2 \le b) = \int_a^b f(x_1, t_2) dt_2 / \int_a^b f_2(t_2) dt_2.$$

It is interesting to consider the limited case $a = b$. In this limit, we obtain

$$\begin{aligned} f_1(x_1|x_2) &= f_1(x_1 | X_2 = x_2) \\ &= \lim_{h \to 0} f_1(x_1 | x_2 \le X_2 \le x_2 + h) \\ &= \lim_{h \to 0} \int_{x_1}^{x_2+h} f(x_1, t_2) dt_2 / \lim_{h \to 0} \int_{x_1}^{x_2+h} f_2(t_2) dt_2 \\ &= f(x_1, x_2)/f_2(x_2). \end{aligned}$$

This is the conditional density function of X_1 under the condition $X_2 = x_2$, and we need $f_2(x_2) > 0$ such that the above equality makes sense. It can be rewritten as

$$f(x_1, x_2) = f_2(x_2) f_1(x_1|x_2), \tag{1.1.9}$$

corresponding to the conditional probability formula $P(AB) = P(A)P(B)$. In higher dimensional case, $X = (X_1, \cdots, X_n)$ with probability density function $f(x_1, \cdots, x_n)$, one has

$$f(x_1, \cdots, x_n) = g(x_1, \cdots, x_k) h(x_{k+1}, \cdots, x_n | x_1, \cdots, x_k),$$

where g is the probability density of $(X_1, \cdots, X_k)$, and h is the conditional probability density of $(X_{k+1}, \cdots, X_n)$ with the condition $X_1 = x_1, \cdots, X_k = x_k$. The formula can also be regarded as definition of the conditional probability density h. Integrating eq. (1.1.9) w.r.t. x_2, we have

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)dx_2 = \int_{-\infty}^{\infty} f_1(x_1|x_2)f_2(x_2)dx_2. \tag{1.1.10}$$

Next, we discuss the independence of random variables. By the above notations, if $f_1(x_1|x_2)$ depends only on x_1 and is independent of x_2, then the distribution of X_1 is completely unrelated with the value of X_2. That is the stochastic variables X_1 and X_2 are independent in probability.

Definition 1.1.9. *Let $f(x_1, \cdots, x_n)$ be the joint probability density function of the n-dimensional random variable $X = (X_1, \cdots, X_n)$, and the marginal density functions of X_i are $f_i(x_i), i = 1, \cdots, n$. If*

$$f(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n),$$

then the stochastic variables $X_1, \cdots, X_n$ are mutually independent or independent for short.

The concept of independence of variables can also be considered in the following view. If $X_1, \cdots, X_n$ are independent, then the probabilities of the variables are not affected by other variables, hence the events

$$A_1 = (a_1 \le X_1 \le b_1), \cdots, A_n = (a_n \le X_n \le b_n)$$

are independent.

1.1.3 Mathematical expectation and momentum

The probability distribution of random variable we introduced above is the most complete characterization of the probability properties of random variables. Next, we consider the mathematical expectation, momentum, and related topics. Let us first consider the mathematical expectation.

Definition 1.1.10. *If X is a discrete random variable, taking countable values $a_1, a_2, \cdots$ with probability distribution $P(X = a_i) = p_i, i = 1, 2, \cdots$, and $\sum_{i=1}^{\infty} |a_i|p_i < \infty$, then $EX = \sum_{i=1}^{\infty} a_i p_i$ is called the mathematical expectation of X.*

If X is a continuous random variable with probability density function $f(x)$ and $\int_{-\infty}^{\infty} |x|f(x)dx < \infty$, then $E(X) = \int_{-\infty}^{\infty} xf(x)dx$ is defined as the mathematical expectation of X.

Next, we consider the conditional mathematical expectation of random variables. Let X, Y be two random variables, we need to compute the expectation $E(Y|X = x)$ or simply $E(Y|x)$ of Y under the given condition $X = x$. Suppose that the joint density of

(X, Y) is given, then the conditional probability density function of Y is $f(y|x)$ under the given condition $X = x$. We then have by definition that

$$E(Y|x) = \int_{-\infty}^{\infty} yf(y|x)dy.$$

The conditional mathematical expectation reflects the mean change of Y with respect to x. Hence, $E(Y|X)$ is a random variable and changes with X. In statistics, the conditional expectation $E(Y|x)$ is regarded as a function of x and is usually termed as the "regression function" of Y to X.

From conditional mathematical expectation, we can get an important formula to the unconditional mathematical expectation. Recall the law of total probability $P(A) = \sum_i P(B_i)P(A|B_i)$. This can be understood as to find the unconditional conditional probability $P(A)$ from the conditional probability $P(A|B_i)$ of A. In this regard, $P(A)$ is the weighted average of conditional probability $P(A|B_i)$ with weight being the probability $P(B_i)$. By analogy, the unconditional expectation of Y should be equal to the weighted average of the conditional expectation $E(Y|x)$ of x with weight proportional to the probability density $f_1(x)$ of X, i.e.,

$$E(Y) = \int_{-\infty}^{\infty} E(Y|x)f_1(x)dx.$$

The proof is not difficult and omitted here. Recalling that right-hand side (RHS) member of this formula is just the mathematical expectation of the random variable $E(Y|X)$ with respect to X, hence we have

$$E(Y) = E(E(Y|X)).$$

Next, we consider the conditional expectation under σ-subalgebra of $\mathscr{F}$.

Definition 1.1.11. *Let* $(\Omega, \mathscr{F}, P)$ *be a probability space and* $\mathscr{G} \subset \mathscr{F}$. *If* $X : \Omega \to R^n$ *is an integrable random variable, then* $E(X|\mathscr{G})$ *is defined as a random variable satisfying*

(i) $E(X|\mathscr{G})$ *is* $\mathscr{G}$ *measurable;*

(ii) $\int_A XdP = \int_A E(X|\mathscr{G})dP\ \forall A \in \mathscr{G}$.

The conditional mathematical expectation has the following properties:

Proposition 1.1.1.

(i) *Let* X *be* $\mathscr{G}$ *measurable, then* $E(X|\mathscr{G}) = X$ *almost surely (a.s.)*

(ii) *Let* a, b *be constants, then* $E(aX + bY|\mathscr{G}) = aE(X|\mathscr{G}) + bE(Y|\mathscr{G})$ *a.s.*

(iii) *Let* X *is* $\mathscr{G}$ *measurable and* XY *be integrable, then* $E(XY|\mathscr{G}) = XE(Y|\mathscr{G})$ *a.s.*

(iv) *Let* X *be independent of* $\mathscr{G}$, *then* $E(X|\mathscr{G}) = E(X)$ *a.s.*

(v) *Let* $\mathscr{E} \subset \mathscr{G}$, *then*

$$E(X|\mathscr{E}) = E(E(X|\mathscr{G})|\mathscr{E}) = E(E(X|\mathscr{E})|\mathscr{G})\ a.s.$$

(vi) *Let $X \le Y$ a.s., then $E(X|\mathscr{G}) \le E(Y|\mathscr{G})$ a.s.*

The proof can be found, for example, in Ref. [89]. Also, the following Jensen's inequality is cited without proof. The more general Jensen's inequality can be found in standard real analysis textbooks.

Lemma 1.1.1. *Let $\Phi : R \to R$ be convex and $E(|\Phi(X)|) < \infty$, then*

$$\Phi(E(X|\mathscr{G})) \le E(\Phi(X)|\mathscr{G}).$$

In order to depict the dispersion of random variables, the concept of variance is introduced.

Definition 1.1.12. *Let X be a random variable with distribution F, then*

$$\sigma_X^2 = Var(X) = E(X - EX)^2$$

is called the variance of X or F, and σ_X the standard deviation of X or F.

It is easy to show $\mathrm{Var}(X) = E(X^2) - (EX)^2$. As a generalization, one can consider the kth moment of X around c, defined as $E[(X - c)^k]$, where c is a constant and k is a positive integer. In particular, it is called the origin moment when $c = 0$, and central moment when $c = EX$.

Now, let us introduce the concepts of covariance and correlation. Still take two-dimensional random vector (X, Y) for example. Set

$$EX = m_1, EY = m_2;\ \ \mathrm{Var}(X) = \sigma_1^2,\ \ \mathrm{Var}(Y) = \sigma_2^2.$$

Definition 1.1.13. $E[(X - m_1)(Y - m_2)]$ *is called the covariance of X, Y, denoted by $Cov(X, Y)$. $Cov(X, Y)/(\sigma_1\sigma_2)$ is the correlation of X, Y, denoted by $\rho(X, Y)$.*

From Schwarz inequality, it leads to $\mathrm{Cov}(X, Y) \le \sigma_1\sigma_2$ and hence $-1 \le \rho(X, Y) \le 1$. If X, Y are independent, then $\mathrm{Cov}(X, Y) = 0$, and hence $\rho(X, Y) = 0$. But the zero correlation $\rho(X, Y) = 0$ does not necessarily imply the independence of X, Y. The reason is that the correlation is just the "linearly dependent coefficient," which does not characterize the general (nonlinear) relationship between X and Y. But it is true when (X, Y) obeys the two-dimensional normal distribution. That is to say, when X, Y are two-dimensional normal random variables, $\rho(X, Y) = 0$ implies the independence of X and Y.

Finally, let us simply introduce the law of large numbers and central limit theorem. In probability, if $X_1, X_2, \cdots, X_n$ are some random variables, it is usually very difficult to compute the distribution of $X_1 + \cdots + X_n$. Then one asks naturally whether we can find some approximate computations. In fact, this idea is not only possible but also very convenient, since in many cases the limit distribution of the sum is just the normal distribution. For this reason, the normal distribution plays an important and a special role in probability. We usually call such a result "central limit theorem," which roughly states that the distribution of the sum converges to normal distribution. There is another limit theorem, "the law of large numbers," which roughly states that "frequency converges to probability" in statistics. Let us consider the n times independent and repeated experiments and observe whether the event A occurs in each experiment. Set

$$X_i = \begin{cases} 1, & A \text{ happens in the } i\text{th experiment}, \\ 0, & A \text{ does not happen in the } i\text{th experiment}, \end{cases} \qquad i = 1, 2, \cdots, n. \tag{1.1.11}$$

Then A occurred $\sum_i X_i$ times in all the n times experiments, i.e., the frequency is $p_n = \sum_i^n X_i / n \bar{X}_{(n)}$. If $P(A) = p$, then "frequency converges to probability" is to say, in some sense, p_n is close to p when n is large enough. This is the law of large numbers in general case, which happens when the experiments are observed for a large number of times.

Theorem 1.1.3. *Let $X_1, X_2, \cdots, X_n, \cdots$ be independent and identically distributed (iid) random variables. Suppose that they have the same mean value μ and variance σ^2, then for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} P(|\bar{X}_{(n)} - \mu| \geq \varepsilon) = 0. \tag{1.1.12}$$

In probability, this convergence is called "$\bar{X}_{(n)}$ converges to μ in probability," which is different from that in calculus. Here we also omit the proof, which can be found in a general probability textbook. It is proper here to introduce the useful Chebyshev inequality.

Chebyshev inequality: If $\mathrm{Var}(Y)$ exists, then

$$P(|Y - EY| \geq \varepsilon) \leq \mathrm{Var}(Y)/\varepsilon^2.$$

This is one of the pioneer law of large numbers proved by Bernoulli in 1713, which is usually called the Bernoulli law of large numbers. The above theorem requires the existence of variances of $X_1, X_2, \cdots$, but when these random variables obey the same distribution, such requirement is not necessary. This is the following Khintchine theorem, which takes the Bernoulli law of large numbers as a special example.

Theorem 1.1.4. *Let random variable $X_1, X_2, \cdots$ are mutually independent, obey the same distribution, and have mathematical expectation $EX_k = \mu$ for $k = 1, 2, \cdots$, then for any $\varepsilon > 0$, one has*

$$\lim_{n\to\infty} P(|\bar{X}_{(n)} - \mu| < \varepsilon) = 1.$$

Now we make the central limit theorem mentioned earlier more precise. Let $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt$ be the distribution function of the standard normal random variable $X \sim N(0,1)$.

Theorem 1.1.5. *Let $X_1, X_2, \cdots, X_n, \cdots$ be iid random variables, $EX_i = \mu$, $Var(X_i) = \sigma^2$ with $0 < \sigma^2 < \infty$. Then for any real number x, one has*

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{n}\sigma}(X_1 + \cdots + X_n - n\mu) \le x\right) = \Phi(x). \tag{1.1.13}$$

This theorem is usually called the Lindbergh-Lévy theorem, which states that when n is sufficiently large, the standardized variable of the sum $\sum_{k=1}^{n} X_k$ of iid random variables $X_1, X_2, \cdots, X_n$ with mean μ and variance σ^2 approximately obeys the standard normal distribution, i.e., approximately,

$$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \sim N(0,1).$$

Theorem 1.1.6. *Let $X_1, X_2, \cdots, X_n, \cdots$ be iid random variables, and X_i satisfies the distribution*

$$P(X_i = 1) = p,\ P(X_i = 0) = 1 - p, \quad 0 < p < 1.$$

Then for any real number x, one has

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{np(1-p)}}(X_1 + \cdots + X_n - np) \le x\right) = \Phi(x).$$

It is called the De Moivre–Laplace theorem, which is the pioneer central limit theorem.

At the end of this section, we collect some useful convergence concepts. They are often encountered in probability and stochastic analysis.

Definition 1.1.14. *Let $\{X_n; n \ge 1\}$ be a random variable sequence, if there is random variable X such that for any $\varepsilon > 0$, $\lim_{n\to\infty} P(|X_n - X| \ge \varepsilon) = 0$, then the random variable sequence $\{X_n; n \ge 1\}$ is said to converge to X in probability, denoted as $X_n \to^{P} X$.*

If $P(\omega : \lim_{n\to\infty}(X_n(\omega) - X(\omega)) = 0) = 1$, then $\{X_n; n \ge 1\}$ is said to converge to X a.s. or with probability 1, denoted as $X_n \to X$ a.s.

Let $X, X_n (n \ge 1)$ have finite pth moment for some $p \ge 1$. If $\lim_{n\to\infty} E(|X_n - X|^p) = 0$, then $\{X_n; n \ge 1\}$ is said to converge to X in L^p norm, denoted as $X_n \to^{L_p} X$. In particular, when $p = 2$, $\{X_n; n \ge 1\}$ is said to converge to X in mean square.

One can show that both convergence in mean square and convergence a.s. imply convergence in probability, but the inverse is not true. Convergence in mean square and convergence a.s. cannot imply each other. But convergence in probability implies convergence a.s. up to a subsequence by Riesz theorem.

1.2 Some preliminaries of stochastic process

In this section, we will simply introduce some basic concepts of stochastic process. The stochastic process, roughly speaking, is a class of random variables $X = \{X(t,\omega)\}_{t\geq 0}$ which is defined on the same probability space $(\Omega, \mathscr{F}, P)$ and depended on a parameter t. Throughout $X(t,\omega)$ is usually simplified as $X_t(\omega)$ or X_t.

Definition 1.2.1. *Let $(\Omega, \mathscr{F}, P)$ be a probability space and T be a given parameter set. If for any $t \in T$, there is a random variable $X(t,\omega)$, then $\{X(t,\omega)\}_{t\in T}$ is a stochastic process on $(\Omega, \mathscr{F}, P)$, denoted by $\{X(t)\}_{t\in T}$. The set of all possible states of $X(t)$ is called the state space or phase space, denoted as $\mathscr{S}$. T is called the parameter set, usually representing time.*

From another view, the stochastic process $\{X(t,\omega)\}_{t\in T}$ can be regarded as a function of two variables T and Ω. For fixed $t \in T$, $X(t,\omega)$ is a random variable on $(\Omega, \mathscr{F}, P)$, while for fixed $\omega \in \Omega$, $X(t,\omega)$ is a usual function defined on T, which is called a sample function or a sample path of $\{X(t,\omega)\}_{t\in T}$.

Here, we have to mention the Kolmogorov theorem, whose starting point and conclusion are as follows. To translate a real problem to a stochastic process model, first one has to establish a probability space $(\Omega, \mathscr{F}, P)$ and on which a group of random variables, such that any finite-dimensional joint distribution of these random variables has the same distribution to the real problem we observed. Generally speaking, it is not difficult to construct $(\Omega, \mathscr{F})$. For example, one can let $\Omega = \mathscr{S}^T$, and $\mathscr{F} = \Sigma^T$, where Σ is the minimal σ-algebra generated by the open subset of $\mathscr{S}$. But it is usually difficult to construct the probability measure P. Furthermore, we only know some finite-dimensional distribution in real problems, and we need some compatible conditions in order to guarantee that P is uniquely determined by these finite-dimensional distributions.

Assume there is a stochastic process $X = \{X_t\}_{t\in T}$ on $(\Omega, \mathscr{F}, P)$, taking values in the state space $(\mathscr{S}, \Sigma)$. For any integer n and $t_1, t_2, \cdots, t_n \in T$, the joint distribution of random variables $(X_{t_1}(\cdot), \cdots, X_{t_n}(\cdot))$ is a probability measure $p_{t_1,\cdots,t_n}(\cdot)$ on $(\mathscr{S}^n, \Sigma^n)$ and

$$p_{t_1,\cdots,t_n}(\cdot) := P(\omega : (X_{t_1}(\cdot), \cdots, X_{t_n}(\cdot)) \in B), \quad B \in \Sigma^n,$$

where Σ^n is the σ-algebra generated by the open subset of $\mathscr{S}^n$. Here $p_{t_1,\cdots,t_n}(\cdot)$ is called the marginal measure of X at times $t_1, t_2, \cdots, t_n$. It is obvious that for $n \geq 1$ and $t_1, t_2, \cdots, t_n \in T$, $t_i \neq t_j$, $i \neq j$, $p_{t_1,\cdots,t_n}(\cdot)$ has the following properties:

(K.1) $p_{t_1,\cdots,t_n,\cdots,t_{n+m}}(B \times \mathscr{S}^m) = p_{t_1,\cdots,t_n}(B)$ for $B \in \Sigma^n$;

(K.2) Let $\tau : \tau\{1, 2, \cdots, n\} = \{i_1, i_2, \cdots, i_n\}$ be a permutation on $\{1, 2, \cdots, n\}$, and set $\tau B = \{(x_{i_1}, x_{i_2}, \cdots, x_{i_n}); (x_1, x_2, , \cdots, x_n) \in B\}$, then

$$p_{t_1,\cdots,t_n}(B) = p_{t_{i_1},\cdots,t_{i_n}}(\tau B).$$

Conditions (K.1) and (K.2) are the Kolmogorov compatible conditions. Here the problem is that given a class of distributions satisfying the compatible conditions, can we find a probability measure P such that (K.1) and (K.2) hold. Kolmogorov theorem gives an affirmative answer.

Consider a complete separable metric space $\mathscr{S}$ and the σ-algebra Σ generated by its open subset. Such a space $(\mathscr{S}, \Sigma)$ is called a Polish space. The following set is called cylindrical set:

$$\Gamma_{t_1,\cdots,t_n} = \{\omega = (\omega_t, t \in T) \in \mathscr{S}^T; (\omega_{t_1}, \cdots, \omega_{t_n}) \in B\}, \qquad B \in \Sigma^n,$$

where $t_1, \cdots, t_n \in T$ are mutually different. Set

$$\mathscr{C} = \{\Gamma_{t_1,\cdots,t_n}; B \in \Sigma^n, \forall t_1, \cdots, t_n \in T \text{ mutually different}, n \geq 1\},$$

$\Omega = \mathscr{S}^T$ and $\mathscr{F} = \sigma(\mathscr{C})$.

Theorem 1.2.1 (Kolmogorov). *Let $(\mathscr{S}, \Sigma)$ be a Polish space and probability distributions*

$$\{p_{t_1,\cdots,t_n}(\cdot); n \geq 1, \forall t_1, \cdots, t_n \in T \textit{ mutually different}\}$$

satisfy compatible conditions, then there is a unique probability measure P in $(\Omega, \mathscr{F})$, such that for all $n \geq 1$, $t_1, \cdots, t_n \in T$ with $t_i \neq t_j$, $i \neq j$ and $B \in \Sigma^n$, there holds

$$P\{\omega = (\omega_t); ((\omega_{t_1}, \cdots, \omega_{t_n}) \in B)\} = p_{t_1,\cdots,t_n}(B).$$

Next, we introduce four special types of stochastic processes.

1. Process with independent increments: Let $t_1 < t_2 < \cdots < t_n$, $t_i \in T$, $1 \leq i \leq n$. If the increments $X_{t_1}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}}$ are mutually independent, then X is called a process with independent increments. If for any $0 \leq s < t$, the distribution of increment $X_t - X_s$ depends only on $t - s$, then X is said to have stationary increment.

2. Markov process: If for any $t_1 < t_2 < \cdots < t_n < t$ and x_i, $1 \leq i \leq n$, one has

$$P(X_t \in A | X_{t_1} = x_1, X_{t_2} = x_2, \cdots, X_{t_n} = x_n) = P(X_t \in A | X_{t_n} = x_n),$$

then $X = \{X_t\}_{t \in T}$ is called a Markov process. The formula implies that the current state doesn't depend on the past states. In other words, the Markov process is memoryless.

3. Gauss process: If for any positive integer n and $t_1, t_2, \cdots, t_n \in T$, $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$ is an n-dimensional normal variable, then the stochastic process $\{X_t\}_{t\in T}$ is called a normal process or a Gauss process. The Wiener process (c.f. Definition 1.4.1) $\{W_t\}_{t\in T}$ is a special case of normal process.

4. Stationary process: If for any constant τ and positive integer n, $t_1, t_2, \cdots, t_n \in T, t_1 + \tau, t_2 + \tau, \cdots, t_n + \tau \in T$, $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$ and $(X_{t_1+\tau}, X_{t_2+\tau}, \cdots, X_{t_n+\tau})$ have the same joint distribution, then stochastic process $\{X_t\}_{t\in T}$ is called a (strict) stationary process.

If, furthermore, EX_t = constant and for any $t, t + h \in T$, the covariance $E(X_t - EX_t)(X_{t+h} - EX_{t+h})$ exists and is independent of t, then the process $\{X_t\}_{t\in T}$ is called a stationary process.

Example 1.2.1 (white noise). *Let $X = \{X_n\}_{n\in\mathbb{Z}}$ be a real- or complex-valued random variable sequence, $EX_n = 0$ $E\{|X_n|^2\} = \sigma^2 < \infty$ and $EX_nX_m = \delta_{nm}\sigma^2$, where $\delta_{nm} = 1$ if $n = m$ and 0 otherwise. Then it is easy to see that X is a stationary process. In this example, $X = \{X_n\}_{n\in\mathbb{Z}}$ is usually called white noise.*

1.2.1 Markov process

For our applications in partial differential equations, we mainly consider the Markov process with continuous time, while sometimes we also use discrete time Markov process as simple examples. First, we consider the case of discrete state space $\mathscr{S}$. Such a Markov process with continuous time and discrete state space is also called a Markov chain.

Since the Markov process is memoryless, as mentioned earlier, the state at a future time t depends only on the present known state at time s ($s \leq t$), but not the past state. The transition probability $p(s, t, i, j) = P(X_t = j|X_s = i)$ denotes the probability of the process lying in j at time $t \geq s$, given the process lying in state i at time s. If the transition probability depends only on $t - s$, then the continuous-time Markov chain is called homogeneous and denoted by $p_{ij}(t - s) = p(s, t, i, j)$ for simplicity and $\mathbf{P}(t-s) = (p_{ij}(t-s))$ for $i, j \in \mathscr{S}$ and $t \geq s$ is called the transition probability matrix. It is obvious that $\forall s \leq \tau \leq t$,

$$p(s, t; i, j) = \sum_{k\in I} p(s, \tau; i, k)p(\tau, t; k, j). \tag{1.2.1}$$

This equality is the Chapman–Kolmogorov (C-K) equation of continuous-time Markov chain. We have already known that when the process starts from i at time s, the probability of the process to lie in $k \in \mathscr{S}$ at time $s \leq \tau \leq t$ is $p(s, \tau; i, k)$. The left-hand side of eq. (1.2.1) represents the probability of the process that lies in j at time t starting from i at time s. While the RHS of (1.2.1) is the summation of the probabilities of the process that lies in j at time t starting from i at time s with the probability of the process that lies in k at time τ ($s \leq \tau \leq t$) starting from i at time s being $p(s, \tau; i, j)$, which is already known. This equality is valid due to the law of total probability. The C-K equation also reflects the memoryless property of the Markov process.

The transition probability matrix $\mathbf{P}(s,t)$ has the following two important properties:

(1) $\mathbf{P}(s,t)$ is a nonnegative matrix, and

$$\mathbf{P}(s,t)\mathbf{1} = \mathbf{1} \quad (\mathbf{1} = (1,1,\cdots,1,\cdots,)^T).$$

(2) $\mathbf{P}(s,\tau)\mathbf{P}(\tau,t) = \mathbf{P}(s,t)\ \forall s < \tau < t.$

Thus, the time-homogeneous Markov chain corresponds to $\mathbf{P}(s,t+s) = \mathbf{P}(0,t)\ \forall s,t \in T$, abbreviated as $\mathbf{P}(t)$, and the C-K equation becomes

$$\mathbf{P}(s)\mathbf{P}(t) = \mathbf{P}(s+t) \quad \forall s,t \in T.$$

Therefore, $\{\mathbf{P}(s)\}_{s\in T}$ is a continuous-time semigroup.

In what follows, we always suppose the Markov process is time homogeneous, if not specified. In the discrete-time case, we set $T = \mathbb{Z}^+$ and we have $\mathbf{P}(n) = (\mathbf{P}(1))^n$ for any $n \in \mathbb{Z}^+$, i.e., $\mathbf{P}(n) = e^{n\ln\mathbf{P}}$, where $\mathbf{P}$ is the one-step transition probability matrix. Then a natural question is whether a similar formula holds in the continuous-time case $t \in R$? That is, is there a matrix $\mathbf{Q}$, independent of time t, such that $\mathbf{P}(t) = e^{t\mathbf{Q}}$. If such a matrix $\mathbf{Q}$ exists, then it satisfies

$$\mathbf{P}'(0) = \lim_{t\to 0}\frac{\mathbf{P}(t)-\mathbf{P}(0)}{t} = \lim_{t\to 0}\frac{e^{t\mathbf{Q}}-\mathbf{I}}{t} = \mathbf{Q},$$

where we set $\lim_{t\downarrow 0}\mathbf{P}(t) = \mathbf{I}$, the identity matrix. Such a matrix $\mathbf{P}$ is said to be canonical and suggests us to study the derivatives of $\mathbf{P}(t)$ at $t = 0$.

Theorem 1.2.2. *Let transition matrix be canonical, then*

$$\lim_{t\downarrow 0}\frac{p_{ii}(t)-1}{t} = -q_i(=: q_{ii}) \tag{1.2.2}$$

exists (possibly being ∞), and

$$\lim_{t\downarrow 0}\frac{p_{ij}(t)}{t} = q_{ij}(i \neq j) \tag{1.2.3}$$

exists and is finite.

The proof can be found in Ref. [214]. From Fatou lemma, we have

$$0 \le q_{ij} \le q_i \le +\infty \quad \sum_{j\neq i} q_{ij} \le q_i.$$

Generally, if a matrix $\mathbf{Q} = (q_{ij})$ satisfies

(Q.1) $q_{ii} = -q_i \leq 0$ (possibly being $-\infty$),

(Q.2) $0 \leq q_{ij} < +\infty$ $(i \neq j)$,

(Q.3) $\sum_{j\neq i} q_{ij} \leq q_i$.

then the matrix $\mathbf{Q}$ is a called a Q-matrix .

By the theorem, for the canonical Markov process, matrix $\mathbf{Q} = (q_{ij})$ is a Q-matrix. If furthermore $\sum_{j\neq i} q_{ij} = q_i < +\infty$, $\mathbf{Q}$ is conservative.

Definition 1.2.2. *For some Q-matrix* $\mathbf{Q}$*, if there is a Markov chain such that eqs* (1.2.2) *and* (1.2.3) *hold, then the Markov chain is a Q-process with matrix* $\mathbf{Q}$.

Using the $\mathbf{Q}$ matrix, we can deduce the differential equation satisfied by the transition probability of any time interval t, which can be used to solve the probability matrix. From C-K equation

$$p_{ij}(t+h) = \sum_{k\in\mathscr{S}} p_{ik}(h)p_{kj}(t), \tag{1.2.4}$$

we have

$$p_{ij}(t+h) - p_{ij}(t) = \sum_{k\neq i} p_{ik}(h)p_{kj}(t) - (1 - p_{ii}(h))p_{ij}(t).$$

Dividing h on both sides and then taking limit $h \to 0$, we get

$$\lim_{h\to 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \lim_{h\to 0} \sum_{k\neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) + q_{ii}p_{ij}(t).$$

If we can change the order of the limit and the summation, then we obtain the Kolmogorov backward equation

$$p'_{ij}(t) = \sum_{k\neq i} q_{ik}p_{kj}(t) + q_{ii}p_{ij}(t). \tag{1.2.5}$$

Theorem 1.2.3. *Suppose* $\sum_{k\in\mathscr{S}} q_{ik} = 0$*, then the Kolmogorov backward equation* (1.2.5) *holds for all* i, j *and* $t \geq 0$.

The reason why it is called the backward equation is that in the C-K equation (1.2.4)

$$\begin{aligned} p_{ij}(t+h) &= \sum_{k\in\mathscr{S}} P\{X_{t+h} = j | X_0 = i, X_h = k\} \\ &\quad \times P\{X_h = k | X_0 = i\} = \sum_{k\in\mathscr{S}} p_{ik}(h)p_{kj}(t), \end{aligned}$$

when we compute probability distribution at time $t + h$, we take condition back to time h. If we take condition probability only back to time t, we have

$$p_{ij}(t+h) = \sum_{k \in \mathscr{S}} p_{kj}(h) p_{ik}(t), \tag{1.2.6}$$

which results to the Kolmogorov forward equation

$$p'_{ij}(t) = \sum_{k \neq i} p_{ik}(t) q_{kj} + p_{ij}(t) q_{jj}. \tag{1.2.7}$$

Since limit and summation are not interchangeable in general, eq. (1.2.7) only holds in the case of finite state space.

Theorem 1.2.4. *Suppose the state space of Markov chain with canonical transition matrix* $\mathbf{P}$ *is finite, then the Kolmogorov forward equation* (1.2.7) *holds.*

Assume the state space of the Markov chain is countable and the transfer matrix $\mathbf{P}$ is canonical, then Q-matrix must be conservative and hence both the Kolmogorov forward and backward equations hold. They can be rewritten in the following simple form:

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}, \quad \mathbf{P}'(t) = \mathbf{Q}\mathbf{P}(t).$$

Given initial condition $\mathbf{P(0)} = \mathbf{I}$, there exists a unique solution $\mathbf{P}(t) = e^{t\mathbf{Q}}$. This indicates that for any finite-dimensional conservative Q-matrix $\mathbf{Q} = (q_{ij})$, there must be a unique transition matrix $\{\mathbf{P}(t)\}_{t \in R^+}$ such that $\mathbf{P}'(0+) := (p'_{ij}(0+)) = \mathbf{Q}$. It can be proved that the solutions obtained from the forward equation and the backward equation are identical.

In applications, we need also to consider the probability distribution $p_j(t) = P(X_t = j)$ of a process X at t. From

$$p_j(t+\tau) = \sum_{i \in \mathscr{S}} p_i(\tau) p_{ij}(t), \tag{1.2.8}$$

we have by taking $\tau = 0$

$$p_j(t) = \sum_{i \in \mathscr{S}} p_i p_{ij}(t).$$

Multiplying both sides of Kolmogorov forward equation (1.2.7) by p_i and taking summation with i yield

$$\sum_{i \in \mathscr{S}} p_i p'_{ij}(t) = \sum_{i,k \in \mathscr{S}} p_i p_{ik} q_{kj} = \sum_{k \in \mathscr{S}} p_k(t) q_{kj}.$$

The equation is usually called the Fokker–Planck equation, which holds in finite state space for the above reason. Hence, we have

Theorem 1.2.5. *Let $\mathscr{S}$ be a finite state space. The probability distribution $p_j(t)$ for $j \in \mathscr{S}$ of a homogeneous Markov process at t satisfies the following Fokker–Planck equation:*

$$p_j'(t) = -p_j(t)q_j + \sum_{k \neq j} p_k(t)q_{kj}. \tag{1.2.9}$$

For the probability distribution of a process, there is a special but important case, i.e., the invariant probability measure.

Definition 1.2.3. *Suppose $\pi = (\pi_1, \cdots, \pi_k)$ with $\pi_i \geq 0$ for all $i \in \{1, \cdots, k\}$ and $\pi \neq 0$. If $\sum_i \pi_i p_{ij} = \pi_j$ for all j, then π is an invariant measure of $\mathbf{P}$. If furthermore, $\sum_j \pi_j = 1$, then it is an invariant probability measure.*

Example 1.2.2. *Birth and death process.*

Let $\{X_t\}_{t\geq 0}$ be a Markov chain with state space $\mathscr{S} = \{0, 1, 2, \cdots\}$. If $\mathbf{P}(t) = p_{ij}(t)$ satisfies,

$$\begin{cases} p_{i,i+1}(h) = \lambda_i h + o(h), & \lambda_i \geq 0, i \geq 0, \\ p_{i,i-1}(h) = \mu_i h + o(h), & \mu_i \geq 0, i \geq 1, \\ p_{ii}(h) = 1 - (\lambda_i + \mu_i)h + o(h), & \mu_0 = 0, i \geq 0, \\ p_{ij}(h) = o(h), & |i - j| \geq 2, \end{cases} \tag{1.2.10}$$

where h is small enough, then X is called a birth and death process. Here, λ_i is the birth rate and μ_i is the death rate. If $\mu_i \equiv 0$, X is a pure birth process and X is a pure death process for $\lambda_i \equiv 0$. The $\mathbf{Q}$ matrix is easy to be written out (leave to the reader) and is obvious a conservative Q-matrix. The transition matrices $\mathbf{P}(t)$ and $\mathbf{Q}$ satisfy the forward and backward equations

$$p_{ij}'(t) = -p_{ij}(t)(\lambda_i + \mu_i) + p_{i,j-1}(t)\lambda_{j-1} + p_{i,j+1}(t)\mu_{j+1}, \tag{1.2.11}$$

$$p_{ij}'(t) = -(\lambda_i + \mu_i)p_{ij}(t) + \lambda_i p_{i+1,j}(t) + \mu_i p_{i-1,j}(t). \tag{1.2.12}$$

The probability distribution satisfies the Fokker–Planck equation

$$\begin{cases} p_0'(t) = -p_0(t)\lambda_0 + p_1(t)\mu_1, \\ p_j'(t) = -p_j(t)(\lambda_j + \mu_i) + p_{j-1}(t)\lambda_{j-1} + p_{j+1}(t)\mu_{j+1}. \end{cases} \tag{1.2.13}$$

Example 1.2.3. *A special example.*

Take $\mathscr{S} = \{0, 1\}$ in the above birth and death process, then we get

$$\mathbf{Q} = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}.$$

By Kolmogorov forward equation

$$p'_{00}(t) = \mu p_{01}(t) - \lambda p_{00}(t) = (-\lambda + \mu)p_{00}(t) + \mu.$$

Since the initial value is $p_{00}(0) = 1$, we have by solving this equation that

$$p_{00}(t) = \mu_0 + \lambda_0 e^{-(\lambda+\mu)t},$$

where $\lambda_0 = \frac{\lambda}{\lambda+\mu}$ and $\mu_0 = \frac{\mu}{\lambda+\mu}$. Similarly, we obtain

$$\begin{aligned} p_{00}(t) =& \mu_0 + \lambda_0 e^{-(\lambda+\mu)t}, \quad p_{01}(t) = \lambda_0(1 - e^{-(\lambda+\mu)t}), \\ p_{11}(t) =& \lambda_0 + \mu_0 e^{-(\lambda+\mu)t}, \quad p_{10}(t) = \mu_0(1 - e^{-(\lambda+\mu)t}). \end{aligned}$$

If the process has initial distribution

$$\pi_0 = p_0 = P(X_0 = 0) = \mu_0, \;\; \pi_1 = p_1 = P(X_0 = 1) = \lambda_0,$$

then the distribution at time t is

$$p_0(t) = p_0 p_{00}(t) + p_1 p_{10}(t) = \mu_0, \quad p_1(t) = p_0 p_{01}(t) + p_1 p_{11}(t) = \lambda_0.$$

Hence, $\pi = (\mu_0, \lambda_0)$ is an invariant measure of the system. Since $\pi_0 + \pi_1 = 1$, π is also an invariant probability measure.

Finally, let us consider the general Markov process and give some basic concepts. The space $(\mathscr{S}, \Sigma)$ is called a *Polish space*, if $\mathscr{S}$ is a complete separable metric space with distance $d(x, y)$ for any $x, y \in \mathscr{S}$, and Σ is the σ-algebra generated by all the open subsets of $\mathscr{S}$. Let $B(x, \delta)$ denote the open ball in $\mathscr{S}$ centered at x with radius δ. Consider a Markov process $X = \{X_t\}_{t\in R^+}$ on a probability space $(\Omega, \mathscr{F}, P) : (\Omega, \mathscr{F}, P) \to (\mathscr{S}, \Sigma)$. Under suitable regularity conditions, the transition probability is given by

$$p(s, t; x, A) = P(X_t \in A|X_s = x) \quad \forall s \le t, x \in \mathscr{S}, A \in \Sigma. \tag{1.2.14}$$

It is called a time-homogeneous Markov process, if $p(s, t; x, A)$ depends only on $t - s$. If there holds $p(s, t; x, A) = \int_A \rho(s, t; x, y)dy$ for some ρ, then $\rho(s, t; x, y)$ is called a transition probability density function. In particular, if the process is time homogeneous, then $\rho(s, t; x, y)$ is simplified to $\rho(t, x, y)$. The corresponding C-K equation is given by

$$p(s,t;x,A) = \int_{\mathscr{S}} p(\tau,t;y,A)p(s,\tau;x,dy) \quad \forall x \in \mathscr{S}, 0 \le s \le \tau \le t < \infty, A \in \Sigma.$$

The similar expressions can be found for the density function.

Notice that eq. (1.2.14) can be rewritten as

$$p(s,t;x,A) = E(1_A(X_t)|X_s = x) \quad \forall s \le t, x \in \mathscr{S}, A \in \Sigma. \tag{1.2.15}$$

If we extend the index function to a general bounded measurable function, we obtain

$$(P_{s,t}f)(x) = E(f(X_t)|X_s = x) \quad \forall s \le t, x \in \mathscr{S}, \tag{1.2.16}$$

where $f \in B_b(\mathscr{S})$ is a bounded measurable function from $(\mathscr{S}, \Sigma)$ to $(R^1, \mathscr{B}_1)$ with norm $\|f\| := \sup_{x \in \mathscr{S}} |f(x)|$. Under such a norm, $B_b(\mathscr{S})$ is a Banach space. Then, it is easy to show from conditional expectation that

$$(P_{s,t}f)(x) = \int_{\mathscr{S}} f(y)p(s,t;x,dy) \quad \forall s \le t, x \in \mathscr{S}. \tag{1.2.17}$$

If the Markov process is time homogeneous, then $P_{s,t} = P_{0,t-s} =: P_{t-s}$. Denote

$$B_0(\mathscr{S}) = \{f \in B_b(\mathscr{S}) : \ \|P_t f - f\| \to 0, \text{ as } t \to 0\}.$$

Then it is not difficult to obtain that $\{P_t\}_{t \ge 0}$ is a strong continuous contract positive operator semigroup from $B_0(\mathscr{S})$ to itself. Such a semigroup is called a Markov transition semigroup and $B_0(\mathscr{S})$ is the strong continuous center. If $\lim_{t \downarrow 0}(P_t f - f)/t = \mathscr{A} f$ exists in norm, then $\mathscr{A}$ is called the infinitesimal generator or generator for short. In such a way, P_t defines a (possibly unbounded) operator $\mathscr{A}$ with domain $\mathscr{D}(\mathscr{A}) := \{f : \lim_{t\downarrow 0}(P_t f - f)/t \text{ exists}\}$. Obviously, $\mathscr{D}(\mathscr{A}) \subset B_0(\mathscr{S})$.

Suppose $\{\mathscr{F}_t\}_{t \ge 0}$ is a nondecreasing σ-algebra flow on the probability space $(\Omega, \mathscr{F}, P)$. Consider random variable $\tau : \Omega \to [0, \infty]$. If for any $t \in [0, \infty]$, the event $\{\tau \le t\}$ is $\mathscr{F}_t$ measurable, then τ is called a stopping time. If

$$E(1_A(X_{t+\tau})|\mathscr{F}_\tau) = p(t, X_\tau, A) \quad \text{for a.e. } \omega \in \{\omega : \tau(\omega) < +\infty\},$$

for all $t > 0$ and $A \in \Sigma$, then the time-homogeneous Markov process $X = \{X_t\}_{t \ge 0}$ is said to be a strong Markov process. Intuitively speaking, in a Markov process, the current time t is a number of the parameter set $T = R^+$, while in a strong Markov process, the current time τ could be any stopping time. The strong Markov property is important in stochastic process. One may wonder under what conditions can a Markov process be a strong Markov one. This is true for a Markov process with discrete-time parameters, but for a strong Markov process, it fails. However, for a Markov process with the sample path being right continuous and its transition function satisfying the Feller property, it is transfered to a strong Markov process [214].

Definition 1.2.4. *The transition function $\{p(t; x, A)\}$ or the corresponding transition semigroup P_t is said to satisfy the Feller (resp. strong Feller) property if $P_t f$ is a bounded continuous function for all $t > 0$ and bounded continuous (resp. bounded measurable) function f. If the semigroup of a Markov process is a Feller semigroup (resp. strong Feller semigroup), then it is a Feller process (resp. strong Feller process).*

Let $\mathcal{M}_1(E)$ denote the sets of probability measures on $(\mathscr{S}, \Sigma)$. For $t \geq 0, \mu \in \mathcal{M}_1(\mathscr{S})$, we define

$$P_t^*\mu(A) = \int_{\mathscr{S}} P(t; x, A)\mu(dx), \quad t \geq 0, A \in \Sigma. \tag{1.2.18}$$

Definition 1.2.5. *A probability measure $\mu \in \mathcal{M}_1(\mathscr{S})$ is invariant with respect to $\{P_t\}$, if*

$$P_t^*\mu = \mu \quad \forall t \geq 0.$$

1.2.2 Preliminaries on ergodic theory

In thermodynamics, the limit $\lim_{T\to} \frac{1}{T}\int_0^T f(\theta_t\omega)$, if exists, is usually taken to be the value of the physical quantity of a system. Then in what sense does this limit exist. It is a basic mathematical problem in statistic mechanics. Another especially interesting problem is whether such time average tends to the space average $Ef(\omega) := \int f(\omega)dP(\omega)$.

For $\{f(\theta_k\omega); k = 0, 1, 2, \cdots\}$ being iid, this problem was solved by employing the strong law of large numbers. For general cases, Von Neumann and Birkhoff established the existence of the limit in the sense of mean square convergence and almost everywhere convergence.

Consider the invertible measurable transformation group $\theta_t : \Omega \to \Omega$, $t \in R$, on a probability space $(\Omega, \mathscr{F}, P)$. It is called measurable if $\theta_t^{-1}A := \{\omega : \theta_t\omega \in A\} \in \mathscr{F}$ for all $A \in \mathscr{F}$. If, furthermore, $P(\theta_t A) = P(A)$ for all $A \in \mathscr{F}$ and $t \in R$, then it is called a measure-preserving transformation and $S = (\Omega, \mathscr{F}, P, \theta_t)$ is called a dynamical system. The group $\{\theta_t\}_{t\in R}$ induces a linear transformation in space $\mathcal{H} = L^2(\Omega, \mathscr{F}, P)$:

$$U_t\xi(\omega) = \xi(\theta_t\omega), \quad \xi \in \mathcal{H}, \omega \in \Omega, t \in R.$$

Since θ_t is measure preserving, U_t is unitary, from which $\|U_t\| = 1$ and hence equicontinuity condition is satisfied. By taking advantage of the sequentially weak compactness of $\mathcal{H}$, one obtains the mean ergodic theorem of J. Von Neumann:

$$s - \lim_{T\to\infty} \frac{1}{T}\int_0^T U_t x = x_0 \in \mathcal{H} \tag{1.2.19}$$

holds for any $\xi \in \mathcal{H}$. The proof can be found, for example, in Ref. [260].

Next, we consider the Birkhoff's ergodic theorem. $A \in \mathscr{F}$ is called an invariant set of θ_t or the dynamical system S, if

$$P(\theta_t A \cap A) = P(A) = P(\theta_t A) \quad \forall t \in R,$$

or equivalently $U_t \chi_A = \chi_A$, P-a.s.

Definition 1.2.6. *If $P(A) = 0$ or $P(A) = 1$ for any invariant set, then the measure-preserving transformation θ_t or the dynamical system $S = (\Omega, \mathscr{F}, P, \theta_t)$ is called ergodic.*

Theorem 1.2.6. *Let $(\Omega, \mathscr{F}, P)$ be a probability space, on which there is a measure-preserving transformation $\Theta : \Omega \to \Omega$. Then for any $\xi \in \mathcal{H}$, there is $\xi^* \in \mathcal{H}$ such that*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \xi(\Theta^k(\omega)) = \xi^*(\omega), \quad \xi^*(\omega) = \xi^*(\Theta(\omega)), \qquad P\text{-a.s.}$$

and $E\xi = E\xi^*$, where $E\xi$ is the expectation of ξ.

This theorem ensures the convergence almost everywhere of time average. One can also show that $S = (\Omega, \mathscr{F}, P, \theta_t)$ is ergodic if and only if

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T P(\theta_t A \cap B)dt = P(A)P(B) \quad \forall A, B \in \mathscr{F}.$$

This is also usually used as definition of ergodicity.

Definition 1.2.7. *A stochastic process $X = \{X_t\}_{t\in I}$ is stochastically continuous if*

$$\lim_{s\to t, s\in I} P(|X(t) - X(s)| > \varepsilon) = 0 \quad \forall \varepsilon > 0, \forall t \in I.$$

A Markov semigroup $\{P_t\}_{t\geq 0}$ is continuous if $\lim_{t\to 0} P_t(x, B(x, \delta)) = 1$ for all $x \in \mathscr{S}$ and $\delta > 0$.

If P_t is stochastically continuous, then for every $x \in \mathscr{S}$ and $T > 0$, $R_T(x, \Gamma) = \frac{1}{T}\int_0^T P_t(x, \Gamma)dt$ defines a probability measure on Σ. For any $\nu \in \mathcal{M}_1(\mathscr{S})$, define $R_T^*\nu(\Gamma) = \int_{\mathscr{S}} R_T(x, \Gamma)\nu(dx)$ for any $\Gamma \in \Sigma$, then it is obvious for any $\varphi \in B_b(\mathscr{S})$,

$$\langle R_T^*\nu, \varphi\rangle = \frac{1}{T}\int_0^T \langle P_t^*\nu, \varphi\rangle dt.$$

In this sense, we denote $R_T^*\nu = \frac{1}{T}\int_0^T P_t^*\nu dt$. The following theorem is due to Krylov–Bogoliubov.

Theorem 1.2.7. *Let $\{P_t\}_{t\geq 0}$ be a Feller semigroup. If for some $\nu \in \mathcal{M}_1(\mathscr{S})$ and sequence $T_n \nearrow \infty$, $R^*_{T_n}\nu \to \mu$ weakly as $n \to \infty$, then μ is an invariant measure of $\{P_t\}_{t\geq 0}$.*

The proof can be found in Ref. [76].

Corollary 1.2.1. *If for some $\nu \in \mathcal{M}_1(\mathscr{S})$ and a sequence $T_n \nearrow \infty$, $\{R^*_{T_n}\nu\}$ is tight as $n \to \infty$, then $\{P_t\}_{t\geq 0}$ has an invariant measure.*

A class of measures $\{\mu_t\}_{t\in T}$ is tight, if for all $\varepsilon > 0$, there exists a compact subset $K_\varepsilon \in \mathscr{S}$, such that $\mu_t(\mathscr{S} \setminus K_\varepsilon) < \varepsilon$ for all $t \in T$. This corollary can be proved from the following Tychonoff theorem.

Proposition 1.2.1. *Let $(\mathscr{S}, \Sigma)$ be a separable metric space, on which there is a class of probability measures $\{\mu_n\}_{n\geq 1}$. If $\{\mu_n\}_{n\geq 1}$ is tight, then μ_n has weakly convergent subsequences.*

Now, let us discuss the relation between the invariant measure and ergodic theory. Let $\{P_t\}_{t\geq 0}$ be a Markov semigroup and μ be one of its invariant measures. Denote $\mathscr{F}^\mu = \Sigma^R$. For any finite set $I = \{t_1, \cdots, t_n\} \subset R$, $t_1 < t_2 < \cdots < t_n$, define a probability measure P^μ on $(\mathscr{S}^I, \Sigma^I)$ by

$$\begin{aligned} P^\mu_I(\Gamma) = &\int_{\mathscr{S}} P_{t_1}(x, dx_1) \cdots \int_{\mathscr{S}} P_{t_{n-1}-t_{n-2}}(x_{n-2}, dx_{n-1}) \\ &\times \int_{\mathscr{S}} P_{t_n - t_{n-1}}(x_{n-1}, dx_n)\chi_\Gamma(x_1, \cdots, x_n) \ \ \forall \Gamma \in \Sigma^I. \end{aligned}$$

For this finite distribution, there is a unique probability measure P^μ on (Ω, P^μ) such that

$$P^\mu(\{\omega : (\omega_{t_1}, \cdots, \omega_{t_n}) \in \Gamma\}) = P^\mu_I(\Gamma), \ \ \Gamma \in \Sigma^{\{t_1,\cdots,t_n\}},$$

thanks to Kolmogorov theorem. Define process

$$X_t(\omega) = \omega(t), \ \ \mathscr{F}^\mu_t = \sigma\{X_s : s \leq t\}, \quad \omega \in \Omega, t \in R.$$

This process is Markovian in that

$$\begin{aligned} P^\mu(X_{t+h} \in \Gamma | \mathscr{F}^\mu_t) &= P^\mu(X_{t+h} \in \Gamma | \sigma(X_t)) \\ &= P_h(X_t, \Gamma), \quad P^\mu\text{-a.s.}, \Gamma \in \Sigma^R. \end{aligned}$$

More generally,

$$\begin{aligned} P^\mu(X_{t+\cdot} \in \Gamma | \mathscr{F}^\mu_t) &= P^\mu(X_{t+\cdot} \in \Gamma | \sigma(X_t)) \\ &= P^{X_t}(\Gamma), \quad P^\mu\text{-a.s.}, \ \Gamma \in \Sigma^R. \end{aligned}$$

Introduce the invertible measurable transformation group $\theta_t : \Omega \to \Omega$ such that $(\theta_t\omega)(s) = \omega(t+s)$ for all $t, s \in R$. Since μ is invariant, the process $\{X_t\}_{t\in R}$ is stationary, i.e., $P^\mu(X \in \theta_t\Gamma) = P^\mu(X \in \Gamma)$ for all $t \in R$ and $\Gamma \in \Sigma^R$, where $\theta_t\Gamma = \{\omega : \theta_t^{-1}\omega \in \Gamma\}$ and θ_t is measure preserving. Hence, $S^\mu = (\Omega, \mathscr{F}, P^\mu, \theta_t)$ defines a dynamical system on the $\mathscr{S}$-valued function space $\Omega = \mathscr{S}^R$, called the canonical dynamical system associated with P_t and μ. As above, associated with θ_t a linear transformation U_t on $\mathcal{H}^\mu = L^2(\Omega, \mathscr{F}^\mu, P^\mu)$ can be induced by

$$U_t\xi(\omega) = \xi(\theta_t\omega), \quad \xi \in \mathcal{H}, \omega \in \Omega, t \in R.$$

It can be shown that if P_t is a stochastically continuous Markov semigroup and μ its invariant measure, then so it is with the process $\{X_t\}_{t\in R}$ in $(\Omega, \mathscr{F}^\mu, P^\mu)$.

Definition 1.2.8. *The invariant measure $\mu \in \mathcal{M}_1(E)$ of P_t is ergodic, if the corresponding S^μ is ergodic.*

Theorem 1.2.8. *Let $\{P_t\}_{t\geq 0}$ be a stochastically continuous Markovian semigroup and μ its invariant measure. Then the following conditions are equivalent:*

(i) *μ is ergodic;*

(ii) *if $\varphi \in L^2_{\mathbb{C}}(\mathscr{S}, \mu)$, or $\varphi \in L^2(\mathscr{S}, \mu)$) and*

$$P_t\varphi = \varphi, \quad \mu\text{-a.s. for all } t > 0,$$

then φ is constant μ-a.s.;

(iii) *if for a set $\Gamma \in \mathscr{F}^\mu$ and any $t > 0$, $P_t\chi_\Gamma = \chi_\Gamma$, μ-a.s., then either $\mu(\Gamma = 1)$ or $\mu(\Gamma) = 0$;*

(iv) *for any $\varphi \in L^2_{\mathbb{C}}(\mathscr{S}, \mu)$ (resp. $L^2(\mathscr{S}, \mu)$),*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T P_s\varphi ds = \langle\varphi, 1\rangle \ \text{ in } L^2_{\mathbb{C}}(\mathscr{S}, \mu)(\text{resp. } L^2(\mathscr{S}, \mu)).$$

The proof can be found in Ref. [76].

1.3 Martingale

Let $(E, \mathscr{E})$ be a measurable space and $(\Omega, \mathscr{F}, P)$ be a complete probability space, where $\mathscr{F}$ is the σ-algebra of Ω. If the class of σ-subalgebra $\{\mathscr{F}_t\}_{t\in I}$ of $\mathscr{F}$ satisfy $\mathscr{F}_s \subseteq \mathscr{F}_t$ when $s \leq t$, then $\{\mathscr{F}_t\}_{t\in I}$ is called a filtration. If, furthermore,

$$\mathscr{F}_t = \mathscr{F}_{t+} := \bigcap_{s>t} \mathscr{F}_s, \quad \forall t \in I,$$

then the filtration $\{\mathscr{F}_t\}_{t\in I}$ is called right continuous.

Definition 1.3.1. *If X_t is measurable w.r.t. $\mathscr{F}_t$ for all $t \in I$, then the E-valued stochastic process $X = \{X_t\}_{t \in I}$ is called adapted, usually denoted as $X_t \in \mathscr{F}_t$.*

Denote $\mathscr{F}_t^X = \sigma\{X(s) : 0 \le s \le t\}$ to be the σ-algebra generated by the process, then the process $\{X_t\}_{t \in R}$ is obviously adapted under such a filtration and hence $\mathscr{F}_t^X$ is usually called the natural filtration. Obviously, if X is $\{\mathscr{F}_t\}_{0 \le t \le \infty}$ adapted, then $E(X(s)|\mathscr{F}_s) = X(s)$ (a.s.). Intuitively, process is adapted means that $\mathscr{F}_t$ contains all the information to determine the behavior of X till t. Hence, if X is $\mathscr{F}_t$ adapted, then $\mathscr{F}_t^X \subseteq \mathscr{F}_t$ for any $t \ge 0$.

Definition 1.3.2. *Let $\{X_t\}_{t \in I}$ be an E-valued stochastic process defined on a probability space $(\Omega, \mathscr{F}, P)$. An E-valued stochastic process $\{Y_t\}_{t \in I}$ is called a modification or version of X if $P(X_t = Y_t) = 1$ for any $t \in I$. If furthermore Y has P-a.s. continuous sample path, then we say X has a continuous modification.*

Obviously, the modification of X and X itself have the same finite-dimensional distributions.

Definition 1.3.3. *Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}, P)$ be a filtrated probability space. For a random variable $\tau : \Omega \to [0, +\infty]$, if $\{\omega : \tau \le t\} \in \mathscr{F}_t$ for any $t \in I$, then we say τ is a stopping time (w.r.t. $\{\mathscr{F}_t\}$).*

Let $(B, |\cdot|_B)$ be a Banach space whose norm is usually simplified to be $|\cdot|$ if no confusions arise. An E-valued random variable X is said to be p-integrable if

$$E|X|_B^p := \int_\Omega |X(\omega)|_B^p P(d\omega) < \infty.$$

If $p = 1$, we say it is integrable and when $p = 2$ we say it is square integrable. The p-integrability of a stochastic process can be similarly defined.

Definition 1.3.4. *A B-valued stochastic process $\{X_t\}_{t \in I}$ is right continuous, if $X(t+) := \lim_{s \downarrow t, s \in I} X(s) = X(t)$ for all $t \in I$. A B-valued process X is continuous if $X(t) = \lim_{s \to t, s \in I} X(s)$ for all $t \in I$. A B-valued process X is càdlàg (continue à droite et limites à gauche, i.e., left limit and right continuous) if X is right continuous and the left limit exists, i.e., $X(t-) := \lim_{s \uparrow t, s \in I} X(s)$ exists for any $t \in I$.*

Definition 1.3.5. *A family of B-valued random variables $\{X_t\}_{t \in I}$ is called uniformly integrable if*

$$\lim_{r \to \infty} \sup_t \int_{\{|X_t| \ge r\}} |X_t|_B dP = 0.$$

Obviously, if $I \subset R$ is compact, then the stochastic continuous process X is uniformly stochastically continuous, i.e., for any $\varepsilon > 0$ there exists $\delta > 0$ such that $P(|X(t) - X(s)|_B > \varepsilon) < \varepsilon$ for all $t, s \in I$ and $|t - s| \le \delta$.

Let $\mathcal{P}_I$ be the predictable σ-field, i.e., the minimal σ-algebra generated by all the subsets of the form $((s, t] \cap I) \times A$ of $I \times \Omega$, where $s, t \in I, s < t$ and $A \in \mathscr{F}_s$. In particular, if $I = [0, \infty)$, we denote $\mathcal{P}_{[0,\infty)}$ by $\mathcal{P}$ for brevity.

Definition 1.3.6. *An $(E, \mathscr{E})$-valued stochastic process is called predictable if it is measurable as a map from $(I \times \Omega, \mathcal{P}_I)$ to $(E, \mathscr{E})$.*

The predictable process is obvious adapted, and for a continuous adapted process, we have the following result (refer to Ref. [202, P.27]).

Proposition 1.3.1. *Let $X = \{X_t\}_{t\in[0,T]}$ be a B-valued stochastically continuous adapted process on $[0, T]$, then X has a predictable modification on $[0, T]$.*

The following regularity result is useful in stochastic differential equations (see Refs [140, 254]).

Theorem 1.3.1. *(Kolmogorov–Loève–Chentsov). Let (E, ρ) be a complete metric space and $\{X_v\}_{v\in R^d}$ be a family of E-valued random variables. If there are $a, b, c > 0$ such that $E[(\rho(X_v, X_u))^a] \le c|u - v|^{d+b}$ for all $u, v \in R^d$, then there exists another family of random variables Y such that $X(u) = Y(u)$ a.s. for all u. Furthermore, Y is locally Hölder continuous a.s. with Hölder index $\alpha \in (0, b/a)$.*

Definition 1.3.7. *A B-valued integrable stochastic process $\{X_t\}_{t\in I}$ is called a martingale with respect to $\mathscr{F}_t$, if it's $\{\mathscr{F}_t\}$-adapted and $E(X_t|\mathscr{F}_s) = X_s$ for all $t, s \in I, t \ge s$.*

Usually a martingale is denoted by $M = M_t$ or $M = M(t), t \in I$. A martingale is called continuous if it has a continuous modification. Let X be an $\mathbb{R}^1$-valued martingale. It is called a supermartingale (resp. submartingale) if it is $\{\mathscr{F}_t\}$-adapted and $E(X_t|\mathscr{F}_s) \le$ (resp. $\ge$)X_s for all $t, s \in I, t \ge s$.

Definition 1.3.8. *A B-valued stochastic process $\{X_t\}_{t\ge0}$ is called a local martingale with respect to $\{\mathscr{F}_t\}$, if it is $\{\mathscr{F}_t\}$-adapted and there exists a sequence of stopping times $\tau_n \nearrow \infty$ a.s. such that each of the process $\{X^{\tau_n}\}_{t\ge0} := \{X(t \wedge \tau_n)\}_{t\ge0}$ is a martingale.*

Any martingale is clearly a local martingale. But it is not necessarily true for the reverse.

Definition 1.3.9. *An adapted stochastic process $\{X_t\}_{t\ge0}$ is said to have finite variations, if $\{X(t, \omega)\}_{t\ge0}$ has finite variation for almost all $\omega \in \Omega$. A process $\{X_t\}_{t\ge0}$ is said to have* *local finite variations, if there is a sequence of* $\{\mathscr{F}_t\}$ *stopping times* $\tau_n \nearrow \infty$ *such that* $\{X^{\tau_n}\}_{t\geq 0} := \{X(t \wedge \tau_n)\}_{t\geq 0}$ *has finite variations.*

Definition 1.3.10. *A* $\{\mathscr{F}_t\}$*-adapted B-valued stochastic process* $\{X_t\}_{t\geq 0}$ *is a semimartingale with respect to* $\{\mathscr{F}_t\}$*, if* $X(t)$ *has decomposition* $X(t) = M(t) + V(t)$ *for* $t \geq 0$*, where* M *is a local martingale and* V *is an adapted process with local finite variation.*

Proposition 1.3.2. *Let* X *be a B-valued integrable process. If for all* $t, s \in I$, $t > s$*, the random variable* $X_t - X_s$ *is independent of* $\mathscr{F}_s$*, then the process* $Y(t) := X_t - EX_t$, $t \in I$ *is a martingale.*

This proposition is a direct consequence of conditional expectation. That is $E(X|\mathscr{G}) = EX$, P-a.s., if the random variable X is independent of the σ-subalgebra $\mathscr{G}$ of $\mathscr{F}$ (see Proposition 1.1.1).

Theorem 1.3.2. *Let* $X = \{X_t\}_{t\in R^+}$ *be an adapted* càdlàg *process. If* $E(|X_\tau|) < \infty$ *and* $E(X_\tau) = 0$ *for any stopping time* τ*, then* X *is a uniformly integrable martingale.*

Proof. Let $0 \leq s < t < \infty$, $\Lambda \in \mathscr{F}_s$ and $u_\Lambda = u$ for $\omega \in \Lambda$ and $u_\Lambda = \infty$ otherwise. Then for any $u \geq s$, u_Λ is a stopping time. Noting that

$$\int_\Lambda X_u dP = \int X_{u_\Lambda} dP - \int_{\Omega\setminus\Lambda} X_\infty dP = -\int_{\Omega\setminus\Lambda} X_\infty dP,$$

where $E(X_{u_\Lambda}) = 0$ for $u \geq s$, one has for $\Lambda \in \mathscr{F}_s (s < t)$ that

$$E(X_t 1_\Lambda) = E(X_s 1_\Lambda) = -E(X_\infty 1_{\Omega\setminus\Lambda}).$$

Hence, $E(X_t|\mathscr{F}_s) = X_s$ and X is a martingale for $0 \leq t \leq \infty$. ∎

Proposition 1.3.3. *Let* $\{M(t)\}_{t\in[0,T]}$ *be a B-valued martingale and* g *be a monotonic increasing convex function, which maps* $[0,\infty)$ *to itself. If* $E(g(|M(t)|_B)) < \infty$ $(\forall t \in [0,T])$*, then* $\{g(|M(t)|_B)\}_{t\in[0,T]}$ *is a submartingale.*

The proof can be found in Ref. [75]. It follows that if $E|M(t)|_B^p < \infty$, then $|M(t)|_B^p \leq E(|M(t)|_B|\mathscr{F}_s)$ and hence the process $\{|M(t)|_B\}_{t\in[0,T]}$ is a submartingale. Using this proposition and the maximal inequality of submartingales (c.f.[190]), one has the following martingale inequality [75].

Theorem 1.3.3.
(i) *Let* $\{M_t\}_{t\in I}$ *be a B-valued martingale and* I *be a countable set, then for any* $\lambda > 0$*, it holds*

$$P\left(\sup_{t\in I} |M_t|_B \geq \lambda\right) \leq \lambda^{-p} \sup_{t\in I} E(|M_t|_B^p), \quad p \geq 1,$$

and

$$E\left(\sup_{t\in I}|M_t|_B^p\right) \le \left(\frac{p}{p-1}\right)^p \sup_{t\in I} E(|M_t|_B^p), \quad p > 1.$$

(ii) *If I is not countable, the above estimates still hold when the martingale $M(t)$ is continuous.*

The following optional sampling theorem plays an important role in the theory of submartingales, which we cited here for reference. The proof can be found in Kallenberg [140].

Theorem 1.3.4 (Doob's optional samplin.g).
(i) *Let $\{X_n\}_{n=1,2,\cdots,k}$ be a submartingale (resp. supermartingale, martingale) with respect to $\{\mathscr{F}_n\}$ and $\tau_1, \cdots, \tau_m$ be an increasing sequence of $\{\mathscr{F}_n\}$ stopping times taking values in the finite set $\{1, 2, \cdots, k\}$, then $\{X_{\tau_i}\}_{i=1,2,\cdots,m}$ is a submartingale (resp. supermartingale, martingale) w.r.t. $\{\mathscr{F}_{\tau_i}\}$.*
(ii) *Let $\{X_t\}_{t\in[0,T]}$ be a right-continuous submartingale (resp. supermartingale, martingale) w.r.t. $\{\mathscr{F}_t\}$ and $\tau_1, \cdots, \tau_m$ be an increasing sequence of stopping times taking values in $[0, T]$. Then $\{X_{\tau_i}\}_{i=1,2,\cdots,m}$ is a submartingale (resp. supermartingale, martingale) w.r.t. $\{\mathscr{F}_{\tau_i}\}$.*

Now, we introduce Doob's submartingale inequality.

Theorem 1.3.5 (Doob's inequality). *Let $\{X_t\}_{t\ge 0}$ be a right-continuous submartingale. Then*

$$rP(\sup_{t\in[0,T]} X_t \ge r) \le EX^+(T) \quad \forall r > 0, \forall T \ge 0,$$

where $X^+(t) := \max\{X(t), 0\}$ is the positive part of $X(t)$.

Proof. Select an increasing sequence $\{Q_k\}$ of finite subsets of $[0, T]$ containing T such that $Q = \bigcup_k Q_k$ is dense in $[0, T]$. Since for any $\varepsilon \in (0, r)$,

$$\left\{\sup_{t\in[0,T]} X_t \ge r\right\} \subset \bigcup_k \{\max_{t\in Q_k} X_t \ge r - \varepsilon\},$$

it follows that

$$P\left(\sup_{t\in[0,T]} X_t \ge r\right) \le \frac{1}{r-\varepsilon} EX_T^+.$$

The result follows by taking $\varepsilon \downarrow 0$. ∎

What follows concerns with regularities of submartingales of Doob. Similar results hold for supermartingales. The proof can be found in Ref. [219].

Theorem 1.3.6. *Any stochastically continuous submartingale $\{X_t\}_{t\in I}$ has a càdlàg modification.*

By the above two theorems, we can obtain the following conclusion.

Theorem 1.3.7. *Let $\{M(t)\}_{t\geq 0}$ be a stochastically continuous, square-integrable martingale, taking values in a Hilbert space $(U, \langle\cdot,\cdot\rangle_U)$. Then M has a càdlàg modification (still denote by M) such that*

$$P\left(\sup_{t\in[0,T]} |M(t)|_U \geq r\right) \leq \frac{E|M(T)|_U^2}{r^2} \quad \forall T \geq 0, \forall r > 0, \tag{1.3.1}$$

and

$$E \sup_{t\in[0,T]} |M(t)|_U^\alpha \leq \frac{2}{2-\alpha}(E|M(T)|_U^2)^{\alpha/2} \quad \forall T \geq 0, \forall \alpha \in (0,2). \tag{1.3.2}$$

The proof can be found in Ref. [204]. Next, we introduce Doob–Meyer theorem. Given filtration $\{\mathscr{F}_t\}$ and $T \in [0,\infty)$, we denote $\Sigma_{[0,T]}$ to be the family of all stopping times such that $P(\tau \leq T) = 1$. Let $X = \{X(t)\}_{t\geq 0}$ be a right-continuous submartingale with respect to filtration $\{\mathscr{F}_t\}$. It is said to belong to class (DL) if for any $T \in [0,\infty)$ the random variable $\{X(\tau)\}_{\tau\in\Sigma_{[0,T]}}$ is uniformly integrable.

Theorem 1.3.8 (Doob–Meyer). *For any càdlàg submartingale X in (DL), there is a unique decomposition $X(t) = N(t) + A(t), t \geq 0$, where N is a martingale and $A(t)$ is a predictable increasing process with $A(0) = 0$ a.s.*

The proof can be found in Kallenberg [140], Rogers and Williams [219] or Jakubowski [138]. Applying this theorem, it is easy to show that $|M|_B^2 = \{|M(t)|_B^2\}_{t\geq 0}$ is a submartingale in (DL) when $M = \{M(t)\}_{t\geq 0}$ is a B-valued right-continuous square-integrable martingale.

Denote $\mathcal{M}^2(B)$ to be the space of all B-valued square-integrable martingale processes $M = \{M(t)\}_{t\geq 0}$ with respect to $\{\mathscr{F}_t\}$ such that $\{|M(t)|_B\}_{t\geq 0}$ is càdlàg. When $B = R$, we denote $\mathcal{M}^2(R) = \mathcal{M}^2$. If M is stochastically continuous, then Doob's regularity results show that the submartingale $|M|_B^2$ has a càdlàg modification. If B is a Hilbert space, Theorem 1.3.7 also implies that M has a càdlàg modification. In fact, in Hilbert space, we can always suppose that the elements in $\mathcal{M}^2(B)$ are càdlàg. By Doob–Meyer decomposition, if $M \in \mathcal{M}^2(B)$, then there is a unique predictable increasing process $\{\langle\langle M,M\rangle\rangle_t\}_{t\geq 0}$ such that $\langle\langle M,M\rangle\rangle_0 = 0$ and $\{|M(t)|_B^2 - \langle\langle M,M\rangle\rangle_t\}_{t\geq 0}$ is a martingale.

The process $\{\langle\langle M, M\rangle\rangle_t\}_{t\geq 0}$ is usually called the angle bracket or predictable variation process of M. Let $M, N \in \mathcal{M}^2(B)$, then we define

$$\langle\langle M, N\rangle\rangle := \frac{1}{4}(\langle\langle M+N, M+N\rangle\rangle - \langle\langle M-N, M-N\rangle\rangle).$$

Since for $M, N \in \mathcal{M}^2$,

$$M(t)N(t) = \frac{1}{4}(|M(t)+N(t)|^2 - |M(t)-N(t)|^2),$$

the process $\{M(t)N(t) - \langle\langle M, N\rangle\rangle_t\}_{t\geq 0}$ is a martingale. More generally, if $(U, \langle\cdot,\cdot\rangle_U)$ is a Hilbert space and $M, N \in \mathcal{M}^2(U)$, then $\{\langle M(t), N(t)\rangle_U - \langle\langle M, N\rangle\rangle_t\}_{t\geq 0}$ is a martingale.

The more general Doob–Meyer decomposition theorem can be found in Ref. [211].

Theorem 1.3.9. *Any submartingale $X = \{X_t\}_{t\in I}$ has the decomposition $X(t) = X_0 + N(t) + A(t)$, where $A = \{A_t\}_{t\in I}$ is an adapted increasing process and $N = \{N_t\}_{t\in I}$ is a local martingale.*

Given a process $X = \{X_t\}_{t\geq 0}$ and a stopping time τ, we denote X^τ to be the process $\{X_{t\wedge\tau}\}_{t\geq 0}$. For a certain class of processes $\mathcal{X}$ (such as martingale, supermartingale and so on), $\mathcal{X}_{\text{loc}}$ denotes the corresponding local class, i.e., there is a stopping time sequence $\tau_n \uparrow \infty$ such that $X^{\tau_n} \in \mathcal{X}$ for all $n \in \mathbb{N}$. We also denote $\mathcal{M}_{\text{loc}}(B)$ and $\mathcal{M}^2_{\text{loc}}(B)$ to be the classes of local martingales and local square-integrable martingales, respectively. Let $\mathcal{BV}$ be the class of all real-valued adapted càdlàg processes, which has bounded variation in finite time intervals.

Definition 1.3.11. *The càdlàg-adapted real-valued process X is called a semimartingale, if it can be expressed in the form of $X = M + A$, where $M \in \mathcal{M}^2$ and $A \in \mathcal{BV}$. X is called a local semimartingale if $M \in \mathcal{M}^2_{loc}$ and $A \in \mathcal{BV}_{loc}$.*

Theorem 1.3.10. *For any $M \in \mathcal{M}^2$, there exists an adapted càdlàg increasing process $[M, M]$ such that*

(i)

$$[M, M]_t = \lim_{n\to\infty} \sum_j (M(t^n_{j+1} \wedge t) - M(t^n_j \wedge t))^2$$

in the sense of $L^1(\Omega, \mathscr{F}, P)$, for any partition $\Pi_n :\ 0 < t^n_0 < t^n_1 \cdots$ of $[0, \infty)$ such that $t^n_k \to \infty$ when $k \to \infty$ and $\lim_{n\to\infty} \sup_j (t^n_{j+1} - t^n_j) = 0$.

(ii) *$M^2 - [M, M]$ is a martingale.*

(iii) *If M has continuous trajectories, then $\langle\langle M, M\rangle\rangle = [M, M]$.*

The process $[M, M]$ here is called the quadratic variation process of M. For the proof of this theorem, the readers may refer to Métivier and Peszat and Zabczyk [183, 204].

Similarly, we can also define

$$[M, N] := \frac{1}{4}([M + N, M + N] - [M - N, M - N])$$

for $M, N \in \mathcal{M}^2$. The following Burkholder–Davis–Gundy (BDG) inequality can be found in Kallenberg and Peszat and Zabczyk [140, 204].

Theorem 1.3.11. *For any $p > 0$, there is a constant $C_p \in (0, \infty)$ such that for any real-valued continuous martingale M with $M_0 = 0$ and for every $T \geq 0$,*

$$C_p^{-1} E\langle\langle M, M\rangle\rangle_T^{p/2} \leq E \sup_{t\in[0,T]} |M_t|^p \leq C_p E\langle\langle M, M\rangle\rangle_T^{p/2}.$$

For discontinuous martingales, one has the similar BDG inequality.

Theorem 1.3.12. *For any $p \geq 1$, there is a constant $C_p \in (0, \infty)$ such that for any càdlàg real-valued square-integrable martingale M with $M_0 = 0$ and $T \geq 0$,*

$$C_p^{-1} E[M, M]_T^{p/2} \leq E \sup_{t\in[0,T]} |M_t|^p \leq C_p E[M, M]_T^{p/2}.$$

Next, let us discuss the construction of $\mathcal{M}^2(B)$. Fix $T > 0$ and denote $\mathcal{M}_T^2 = \mathcal{M}_T^2(B)$ for short. The following theorem is given in Ref. [75].

Theorem 1.3.13. *The space $(\mathcal{M}_T^2, |M|_{\mathcal{M}_T^2})$ is a Banach space under norm*

$$|M|_{\mathcal{M}_T^2} = (E \sup_{t\in[0,T]} |M(t)|_B^2)^{\frac{1}{2}}.$$

Example 1.3.1 (212, p. 33). *Let $B = \{B_t\}_{t\geq 0}$ be a Brownian motion in R^3 and $B_0 = x \in R^3 (x \neq 0)$. Let $u(y) = \frac{1}{|y|}, y \in R^3$, then u is the superharmonic function in R^3. By Itô formula (Chapter 2), $X_t = u(B_t)$ is a positive supermartingale. On the other hand, let $\tau_n = \inf\{t > 0 : |B_t| \leq \frac{1}{n}\}$. Since the function u is harmonic outside the sphere $B(0, \frac{1}{n})$, $u(B_{t\wedge\tau_n})$ is a martingale by Itô formula. Since $u(B_{t\wedge\tau_n})$ is bounded by n, it is uniformly integrable. Furthermore, since $B_0 = x \neq 0$, we have $\lim_{t\to\infty} E(u(B_t)) = 0$ by properties of Brownian motion. But it's obvious that $E(u(B_0)) = \frac{1}{|x|}$. So $u(B_t)$ cannot be a martingale.*

A nature question is under what conditions can a local martingale be a martingale. To answer this question, we define the maximal function $M_t^* = \sup_{s\leq t} |M_s|$ and $M^* = \sup_s |M_s|$.

Theorem 1.3.14. *Suppose $M = \{M_t\}_{t\geq 0}$ is a local martingale and $E(M_t^*) < \infty$ for any $t \geq 0$, then M is a martingale. If furthermore $E(M^*) < \infty$, then M is a uniformly integrable martingale.*

Proof. Let $(\tau_n)_{n\geq 1}$ be a stopping time sequence such that $\tau_n \uparrow \infty$ as $n \to \infty$. If $s \leq t$, then $E(M_{t\wedge\tau_n}|\mathscr{F}_s) = M_{s\wedge\tau_n}$, which is followed by $E(M_t|\mathscr{F}_s) = M_s$ thanks to dominated convergence theorem. If furthermore $E(M^*) < \infty$, then $\{M_t\}_{t\geq 0}$ is uniformly integrable since $|M_t| \leq M^*$. ∎

1.4 Wiener process and Brown motion

In 1828, the Scottish botanist Robert Brown observed that the pollen particles suspended in water make irregular motion, which was later explained as the random collision of liquid molecules. While mathematicians regard it as a stochastic process denoted by $B_t(\omega)$, representing the position of the pollen particle ω at the time t. A breakthrough in this direction has been made by Kolmogorov, who gave the precise mathematical description for the phenomenon in theory in 1918 and some trajectory properties of the Brownian motion. These pioneer works made the Brown motion be broadly and extensively studied during the first half of the 20th century and thus leads to the rapid development of Brownian motion. Nowadays, the Brownian motion has become an important branch in probability and it also plays an important role in applied mathematics and modern analysis.

The existence of Brownian motion was first established by Wiener in 1923, which was based on Daniell's method of constructing measures in infinite-dimensional spaces. Later the Brownian motion was constructed by Fourier series by assuming only the existence of iid Gaussian random variable sequences. See also the construction by Lévy and Ciesielski [63, 169].

To construct $\{B_t\}_{t\geq 0}$, we note by Kolmogorov theorem that it suffices to find the probability measures $\{\nu_{t_1,\cdots,t_{t_k}}\}$, satisfying the intuitive fact observed. Fix $x \in R^d$ and define

$$p(t,x,y) = (2\pi t)^{-n/2} \exp\left\{-\frac{|x-y|^2}{2t}\right\}, \quad y \in R^d, t > 0.$$

For $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k$, we define the following measure on R^{dk},

$$\begin{aligned}\nu_{t_1,\cdots,t_k}(F_1 \times \cdots \times F_k) = \int_{F_1\times\cdots\times F_k} & p(t_1,x,x_1)p(t_2-t_1,x_1,x_2) \\ & \cdots p(t_k-t_{k-1},x_{k-1},x_k)dx_1\cdots dx_k,\end{aligned} \tag{1.4.1}$$

where dy is the standard Lebesgue measure under the convention $p(0,x,y)dy = \delta_x(y)$. Using Kolmogorov consistency condition (K.1), we extend the definition to all finite sequences t_i. Since $\int_{R^d} p(t,x,y)dy = 1$ for all $t \geq 0$, Kolmogorov consistency condition (K.2) holds. By Kolmogorov theorem, there is a probability space $(\Omega, \mathscr{F}, P^x)$ and on which a stochastic process $\{B_t\}_{t\geq 0}$ with finite-dimensional distribution (1.4.1),

$$P^x(B_{t_1} \in F_k, \cdots, B_{t_k} \in F_k) = \int_{F_1\times\cdots\times F_k} p(t_1, x, x_1) \\ \cdots p(t_k - t_{k-1}, x_{k-1}, x_k)dx_1 \cdots dx_k. \tag{1.4.2}$$

Such process is called the Brown motion starting from x and $P^x(B_0 = x) = 1$. It's worth noting that such Brown motion is not unique, i.e., there are more than one $(\Omega, \mathscr{F}, P^x, B_t)$ such that eq. (1.4.2) holds. This does not affect our discussion. In fact, we can select any such Brown motion for our discussion. As we will see, the path of Brown motion is continuous a.s., hence we can regard almost all $\omega \in \Omega$ and the continuous function $t \mapsto B_t(\omega)$ from $[0,\infty)$ to R^d to be the same. In this viewpoint, Brownian motion is just the space $C([0,\infty), R^d)$ with a certain probability measure P^x determined by eqs (1.4.1) and (1.4.2). Such selected Brownian motion is called the "canonical Brownian motion." Besides intuition, such selection facilitates detailed analysis of the measures on $C([0,\infty), R^d)$. More about the measures in infinite-dimensional space can be referred to Ref. [260, Chapter 8].

The above defined Brownian motion has the following properties: (1) it is a Gaussian process, i.e., any finite-dimensional joint distribution is normal, (2) it has independent increments and (3) the trajectory is continuous (more specifically, there is a continuous modification). The following gives a more general definition of Wiener process. We assume we are given a probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, P)$.

Definition 1.4.1. *The R^d-valued stochastic process $B = \{B_t\}_{t\geq 0}$ is called a d-dimensional Wiener process or Brownian motion, if*
(i) *for $0 \leq s < t < \infty$, the increment $B_t - B_s$ is independent of $\mathscr{F}_s$,*
(ii) *for any $s, t > 0$, $B_{s+t} - B_s \sim N(0, Ct)$ is a Gaussian random variable with expectation 0 and variance matrix Ct.*

If $P(B_0 = x) = 1$, then B is called a Brownian motion starting from x. In particular, when $d = 1, C = \sigma^2 = 1$ and $x = 0$, B is called a canonical Brownian motion and denoted by $B_t \sim N(0,t)$.

Let us consider the one-dimensional case. Intuitively, $B_t(\omega)$ denotes the position of pollen ω at t. Suppose the liquid is isotropic, then it is natural to assume that the displacement $B_{s+t} - B_s$ from time s to $s+t$ is the sum of many independent small ones

$$B_{s+t} - B_s = (B_{s+\frac{t}{n}} - B_s) + \cdots + (B_{s+\frac{nt}{n}} - B_{s+\frac{(n-1)t}{n}}).$$

By central limit theorem, it's nature to assume that $B_{s+t} - B_s \sim N(0, \sigma^2 t)$, where σ is independent of s, t and x.

There is only one thing not mentioned in the definition, i.e., the continuity of the trajectories. In fact (i) and (ii) determine the distribution of B, but continuity cannot be completely determined by the distribution of B. By Kolmogorov–Loève–Chentsov (KLC) Theorem 1.3.1, we can prove the continuity and even the Hölder continuity of the trajectories.

Consider the d-dimensional Brownian motion. Let $t > s \geq 0$, then for all integers $m = 1, 2, \cdots,$

$$\begin{aligned} E(|B_t - B_s|^{2m}) &= \frac{1}{(2\pi r)^{d/2}} \int_{R^d} |x|^{2m} e^{-\frac{|x|^2}{2r}} dx \quad (r = t - s > 0) \\ &= \frac{1}{(2\pi)^{d/2}} r^m \int_{R^d} |y|^{2m} e^{-\frac{|y|^2}{2}} dy \quad (y = x/\sqrt{r}) \\ &= Cr^m = C|t-s|^m. \end{aligned}$$

Then applying the KLC Theorem ($a = 2m, d = 1, b = m-1$) shows that the Brown motion B is Hölder continuous (a.s.) with the Hölder index $0 < \alpha < \frac{b}{a} = \frac{1}{2} - \frac{1}{2m}$ for all m. This leads to the following result.

Theorem 1.4.1. *The sample path $t \mapsto B_t(\omega)$ of Brownian motion is α-Hölder continuous for all $\alpha \in (0, 1/2)$ in $[0, T]$, for almost all ω and any $T > 0$.*

This also implies any Wiener process has continuous sample path (a.s.). One may ask naturally whether the α here can be taken larger. The answer is disappointing, since for any $\alpha > \frac{1}{2}$, the sample path $t \mapsto B_t(\omega)$ is nowhere α-Hölder continuous for a.a. ω. As a direct corollary in real analysis, Brown motion is of unbounded variation on any finite interval for a.a. ω.

Even so, it still has the following convergence result in the mean square sense. Define a partition Π_n of the interval $[s, t]$:

$$\Pi_n : s = t_0^n < t_1^n < \cdots < t_{m_n-1}^n < t_{m_n}^n = t,$$

and let $\|\Pi_n\| := \max_{1 \leq k \leq m-1} |t_{k+1}^n - t_k^n|$.

Theorem 1.4.2. *Let $B = \{B_\tau\}_{\tau \in [s,t]}$ be a Brownian motion and $S_n := \sum_k |B_{t_{k+1}^n}(\omega) - B_{t_k^n}(\omega)|^2$, then $S_n \to t - s$ in the sense of $L^2(\Omega, \mathscr{F}, P)$ as $\|\Pi_n\| \to 0$ $(n \to \infty)$.*

Proof. Under the above notation,

$$S_n - (t - s) = \sum_k (B_{t_{k+1}^n} - B_{t_k^n})^2 - (t_{k+1}^n - t_k^n).$$

Since the terms in the summation are mutually independent with mean value 0, we have

$$E(S_n-(t-s))^2=\sum_k E[(B_{t^n_{k+1}}-B_{t^n_k})^2-(t^n_{k+1}-t^n_k)]^2$$
$$=\sum_k E[(Y^2-1)(t^n_{k+1}-t^n_k)]^2.$$

Here, Y obeys the standard normal distribution $N(0,1)$, and hence the above formula becomes

$$E[S_n-(t-s)]^2\le E(Y^2-1)^2\|\Pi_n\|\sum_k((t^n_{k+1}-t^n_k))$$
$$=E(Y^2-1)^2(t-s)\|\Pi_n\|\to 0,$$

completing the proof. ∎

In fact, for the canonical Brown motion $B=\{B_t\}_{t\in[s,t]}$, $\lim_{n\to\infty} S_n=t-s$ a.s. (see Protter [211, p. 18]).

The following proposition can be seen as an equivalent definition of the Brownian motion. Let us only consider the canonical Brownian motion.

Proposition 1.4.1. *Let $B_0=0$. $B=\{B_t\}_{t\ge 0}$ is a Brownian motion if and only if it is Gaussian and satisfies*

$$EB_t=0,\quad E(B_tB_s)=t\wedge s. \tag{1.4.3}$$

Proposition 1.4.2. *Let $\{B_t;t\in R^+\}$ be a Brownian motion and $B_0=0$. Then $\{B_{s+t}-B_s\}_{t\in R^+}$, $\{\frac{1}{\sqrt{\lambda}}B_{\lambda t}\}_{t\in R^+}$, $\{tB_{\frac{1}{t}}\}_{t\in R^+}$ and $\{B_{T-s}-B_T\}_{s\in[0,T]}$ still obey the Brown distribution.*

The proof is not difficult. For example

$$E(tB_{\frac{1}{t}}sB_{\frac{1}{s}})=ts\left(\frac{1}{t}\wedge\frac{1}{s}\right)=ts\frac{1}{t\vee s}=t\wedge s.$$

We can see from Theorem 1.4.2 that when t is very small, $B_{t+s}-B_s$ is approximately $\sqrt{|t|}$. There are more refined estimates for $B_{s+t}-B_s$, among which the following log log-type estimate should be cited here. The proof can be found in Ref. [214].

Proposition 1.4.3. *For Brownian motion $\{B(t)\}_{t\in R}$, there holds*

$$\overline{\lim}_{t\downarrow 0}\frac{B(s+t)-B(s)}{\sqrt{t\log\log(1/t)}}=1\ a.s., \tag{1.4.4}$$

$$\overline{\lim}_{t\uparrow\infty}\frac{B(s+t)-B(s)}{\sqrt{t\log\log t}}=1\ a.s. \tag{1.4.5}$$

Corollary 1.4.1. *For Brownian motion $\{B(t)\}_{t\geq 0}$, there holds*

$$\underline{\lim}_{t\downarrow 0}\frac{B(s+t)-B(s)}{\sqrt{t\log\log(1/t)}}=-1\ a.s.,$$
$$\underline{\lim}_{t\uparrow\infty}\frac{B(s+t)-B(s)}{\sqrt{t\log\log t}}=-1\ a.s.$$

Theorem 1.4.3 (The martingale of Brownian motion). *Let $\{B_t\}_{t\in R^+}$ be a standard $\mathscr{F}_t$-Brownian motion, then B_t, B_t^2-t and $\exp(\sigma B_t-(\sigma^2/2)t)$ are all $\mathscr{F}_t$-martingales.*

The proof can be found in Refs [192, 214]. On the other hand, we have the following martingale characterization of Brownian motion, whose proof will be omitted.

Theorem 1.4.4. *Let $X=\{X_t\}_{t\geq 0}$ be an adapted process, with continuous sample path and mean value 0 and for any $1\leq i,j\leq d, s,t\geq 0$, the covariance matrix is $EX_i(t)X_j(s)=a_{ij}(s\wedge t)$, where $A=(a_{ij})$ is a $d\times d$ positive-definite matrix. Then the following assertions are equivalent:*
(1) *X is a Brownian motion with covariance matrix A,*
(2) *X is a martingale, $\langle\langle X_i,X_j\rangle\rangle(t)=a_{ij}t\ \forall 1\leq i,j\leq d, t\geq 0$,*
(3) *for any $u\in R^d$, the process $\{e^{i(u,X(t))+\frac{t}{2}(u,au)}\}_{t\geq 0}$ is a martingale.*

Next, we consider the Markov property, semigroup and generator of Brownian motion. Since Brownian motion is a time-homogeneous process with independent increments, it's also time-homogeneous Markov process, whose transition function is denoted by $p(t,x,A)$. Since

$$\begin{aligned}
p(t,B_0,A) &:= E(1_A(B_t)|\mathscr{F}_0)\\
&= E(1_A(B_t-B_0+x)|\mathscr{F}_0)|_{x=B_0}\quad (B_0\in\mathscr{F}_0)\\
&= E(1_A(B_t-B_0+x))|_{x=B_0}\\
&= \int 1_A(z+x)\frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}}dz|_{x=B_0}\\
&= \int_A \frac{e^{-\frac{(y-x)^2}{2t}}}{\sqrt{2\pi t}}dy|_{x=B_0},
\end{aligned}$$

then the transition function is

$$p(t,x,y)=\frac{1}{\sqrt{2\pi t}}e^{-\frac{(y-x)^2}{2t}},$$

and the transition semigroup of Brownian motion is

$$P_t f(x) = E_x f(B_t) = \int \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} f(y) dy.$$

By a direct calculation, we get

$$\frac{\partial p(t,x,y)}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial y^2} \tag{1.4.6}$$

$$= \frac{1}{2}\frac{\partial^2 p}{\partial x^2}. \tag{1.4.7}$$

These two equations correspond to Kolmogorov backward and forward equations (1.2.5) and (1.2.6).

Next, let us consider the generator of the semigroup. For this purpose, we denote $C_u(R^1) := \{f : R^1 \to R^1$ bounded and uniformly continuous$\}$ for convenience. A direct calculation leads to

$$\frac{P_t f(x) - f(x)}{t} = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{f(x+\sqrt{t}z) - f(x)}{t} dz. \tag{1.4.8}$$

If we denote

$$C_u^2(R^1) := \{f \in C_u(R^1) : f'' \text{ is uniformly continuous and bounded}\},$$

then

$$\begin{aligned}\mathscr{A} f(x) &= \lim_{t\downarrow 0} \int \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \left\{ \frac{f'(x)}{\sqrt{t}} z + \frac{f''(x+\theta\sqrt{t}z)}{2t} tz^2 \right\} dz \\ &= \lim_{t\downarrow 0} \int \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \frac{1}{2} \left\{ f''(x+\theta\sqrt{t}z) - f''(x) \right\} dz + \frac{1}{2} f''(x) \\ &= \frac{1}{2} f''(x).\end{aligned}$$

Hence

$$\left\| \frac{P_t f - f}{t} - \frac{1}{2} f'' \right\| \to 0, \quad f \in \mathscr{D}(\mathscr{F}),$$

and

$$\mathscr{A} f = \frac{1}{2} f''.$$

Generally speaking, it is difficult to find out $\mathscr{D}(\mathscr{F})$. But it is usually sufficient to find out a density subset. There are many conclusions ensuring the Kolmogorov forward equation to be valid, which are closely related to (parabolic) PDEs. Interested readers may refer to monographs of diffusion process and partial differential equation, such as Ref. [100].

1.5 Poisson process

Poisson process is another important stochastic process with continuous time. It was first introduced by and named after the French mathematician Poisson. In this section, we will give the definition of Poisson process and discuss some of its basic mathematical properties.

Suppose we are given a probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{0\le t\le\infty}, P)$. Let T_n be a strictly increasing positive random variable sequence and $T_0 = 0$ a.s.

Definition 1.5.1. *Define $N_t = \sum_{n\ge1} \mathbb{1}_{t\ge T_n}$, taking values in $\mathbb{N}^+ := \mathbb{N} \cup \{\infty\}$, and $\mathbb{1}_{t\ge T_n} = 1$ if $t \ge T_n(\omega)$ and 0 otherwise. The process $N = \{N_t : 0 \le t \le \infty\}$ is called the counting process associated with $\{T_n; n \ge 1\}$.*

Define $T = \sup_n T_n$, then $[T_n, T_{n+1}) = \{N = n\}$ and $[T, \infty) = \{N = \infty\}$ and

$$[T_n, \infty) = \{N \ge n\} = \{(t, \omega) : N_t(\omega) > n\}.$$

It is immediate that $N_t - N_s := \sum_{n\ge1} \mathbb{1}_{s<T_n\le t}$, representing the number of the event arriving in $(s, t]$. Obviously, the counting process has left limit and right continuous sample path (i.e., cádlág) if $T = \infty$ a.s.

The counting process N is adapted if and only if the corresponding random variable $\{T_n\}_{n\ge1}$ is a stopping time. Indeed, if $\{T_n\}_{n\ge1}$ is a stopping time and $T_0 = 0$ a.s., then the event $\{N_t = n\} = \{\omega : T_n(\omega) \le t < T_{n+1}(\omega)\} \in \mathscr{F}_t$ for any n. Hence $N_t \in \mathscr{F}_t$ and N is adapted. On the contrary, if N is adapted, then $\{T_n \le t\} = \{N_t \ge n\} \in \mathscr{F}_t$ and hence $\{T_n\}_{n\ge1}$ is a stopping time.

Definition 1.5.2. *The adapted counting process N is called a Poisson process if $T = \infty$ a.s., and*
(1) *$N_t - N_s$ is independent of $\mathscr{F}_s$ for any $0 \le s < t < \infty$,*
(2) *$N_t - N_s$ and $N_v - N_u$ have identical distributions for any $0 \le s < t < \infty$, $0 \le u < v < \infty$ satisfying $t - s = v - u$.*

The following characterization is useful in practice.

Theorem 1.5.1. *Let N be a Poisson process, then there exists a $\lambda > 0$ such that*

$$P(N_t = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \tag{1.5.1}$$

for any $n = 0, 1, 2, \cdots$.

The parameter λ is the arrival rate of event. This theorem implies that N_t has Poisson distribution with parameter λt. The proof can be found in Ref. [171] and is omitted here. As a direct corollary, $E(N_t) = \lambda t$ and $Var(N_t) = \lambda t$.

What follows is a collection of properties of Poisson process.

(i) The Poisson process is stochastically continuous, i.e., for any real number δ such that $t \geq 0$ and $t + \delta \geq 0$, there holds

$$P(N_{t+\delta} \neq N_t) = 1 - P(N_{t+\delta} = N_t) = 1 - e^{-\lambda|\delta|} \leq \lambda|\delta|.$$

(ii) The sample path is nondecreasing on R^+ in probability 1, i.e., for any $0 \leq s < t$,

$$P(N_t - N_s \geq 0) = \sum_{k=0}^{\infty} P(N_t - N_s = k) = 1. \tag{1.5.2}$$

(iii) As $t \to \infty$,

$$P(N_t(\omega) \text{ is continuous on } [0,t]) = P(N_t(\omega) = N_0(\omega)) = e^{-\lambda t} \to 0.$$

Let $A_\infty = \{N_t(\omega) \text{ is continuous in } [0,\infty)\}$ and $A_t = \{N_s(\omega) \text{ is continuous in } [0,t]\}$, then $A_t \subset A_s$ for $t \geq s$ and hence

$$P(A_\infty) = P\left(\cap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} e^{-\lambda n} = 0.$$

(iv) For almost all sample paths, $\left|\lim_{t\to t_0+, s\to t_0-} N_t - N_s\right| = 1$ at every jump point.

(v) For any fixed $t_0 \geq 0$, almost all sample paths are continuous. Indeed if $t_0 > 0$, then $P(N(t_0 + \varepsilon, \omega) - N(t_0 - \varepsilon, \omega) > 0) = 1 - e^{-2\lambda\varepsilon} \to 0$ as $\varepsilon \to 0$. While if $t_0 = 0$, it is right continuous, i.e., $P(N(\tau,\omega) - N(0,\omega) > 0) = 1 - e^{-\lambda\tau} \to 0$ as $\tau \to 0$.

(vi) Define $\tau_n(\omega)$ to be the arrival time of the nth event of Poisson process N, $\tau_0 = 0$ and $\tau_n = \inf\{t : t > \tau_{n-1}, N_t = n\}$ for $n \geq 1$. Then $S_n(\omega) = \tau_n(\omega) - \tau_{n-1}(\omega)$ denotes the waiting time of the nth event. Obviously, $\{N_t \geq n\} = \{\tau_n \leq t\}$ and $\{N_t = n\} = \{\tau_n \leq t < \tau_{n+1}\} = \{\tau_n \leq t\} - \{\tau_{n=1} \leq t\}$. Hence the distribution function of τ_n is $P(\tau_n \leq t) = 0$ for $t < 0$ and $P(\tau_n \leq t) = P(N_t \geq n) = 1 - e^{-\lambda t}\sum_{k=0}^{n-1}(\lambda t)^k/k!$ for $t \geq 0$ and the probability density function of τ_n is $f_{\tau_n}(t) = \lambda(\lambda t)^{n-1}e^{-\lambda t}\mathbb{1}_{t\geq 0}/(n-1)!$.
In particular, when $n = 1$

$$P(S_1 \leq t) = P(\tau_1 \leq t) = (1 - e^{-\lambda t})\mathbb{1}_{t\geq 0},$$

i.e., $S_1 \sim E(\lambda)$ obeys the exponential distribution with parameter $\lambda > 0$. Generally, $\{S_n\}_{n\geq 1}$ are independent and obey the exponential distribution with parameter λ iff the counting process $N = \{N_t\}_{t\geq 0}$ is a Poisson process (see Ref. [171]).

(vii) Let $N = \{N_t\}_{t\in R^+}$ be a Poisson process with parameter λ, then $N_t - \lambda t$ and $(N_t - \lambda t)^2 - \lambda t$ are both martingales. Indeed, it suffices to note that $E(N_t - \lambda t - (N_s - \lambda s)|\mathscr{F}_s) = E(N_t - \lambda t - (N_s - \lambda s)) = 0$ for $0 \leq s < t < \infty$. The same result holds for $(N_t - \lambda t)^2 - \lambda t$.

1.6 Lévy process

Lévy process is a large class of stochastic processes including Brownian motion and Poisson process. In this section, we will give the definition, the Lévy–Itô theorem and the Lévy–Khintchine formula. Most of these contents and details can be found in Ref. [7].

1.6.1 Characteristic function and infinite divisibility

Consider the R^d-valued random variable X in a probability space $(\Omega, \mathscr{F}, P)$ with probability distribution P_X. The characteristic function is defined by

$$\phi_X(\xi) = E(e^{i(\xi,X)}) = \int_{R^d} e^{i(\xi,y)} P_X(dy), \quad \xi \in R^d.$$

Example 1.6.1 (Gauss random variable). *Random variance $X = (X_1, \cdots, X_d)$ is Gaussian or normal, denoted by $X \sim N(m, A)$, if it has probability density function*

$$f(x) = \frac{1}{\sqrt{(2\pi)^d \det A}} \exp\left\{-\frac{1}{2}(x-m, A^{-1}(x-m))\right\},$$

where $m \in R^d$ is the mean vector and A is strictly positive definite symmetric $d \times d$ covariance matrix. It is easy to calculate

$$\phi_X(\xi) = \exp\left\{i(m,\xi) - \frac{1}{2}(\xi, A\xi)\right\}.$$

Example 1.6.2 (Compound Poisson random variable). *Suppose that $\{Z(n)\}_{n\in\mathbb{N}}$ is a sequence of R^d-valued iid random variables having common law μ_Z and $N \sim \pi(\lambda)$ is a Poisson random variable independent of all the $Z(n)$. The compound Poisson random variable is defined as*

$$X = Z(1) + \cdots + Z(N).$$

Then the characteristic function of X is given by

$$\begin{aligned}
\phi_X(\xi) &= \sum_{n=0}^{\infty} E\left(\exp\left[i(\xi, Z(1) + \cdots + z(N))\right] | N = n\right) P(N = n) \\
&= \sum_{n=0}^{\infty} E\left(\exp\left[i(\xi, Z(1) + \cdots + Z(n))\right]\right) e^{-\lambda}\frac{\lambda^n}{n!} \\
&= e^{-\lambda}\sum_{0}^{\infty} \frac{[\lambda\phi_Z(\xi)]^n}{n!} = \exp\left[\lambda(\phi_Z(\xi) - 1)\right],
\end{aligned}$$

where $\phi_Z(\xi) = \int_{R^d} e^{i(\xi,y)}\mu_Z(dy)$.

Definition 1.6.1. *Let X be an R^d-valued random variable with distribution law μ_X. Then X is said to be* infinitely divisible *if for all $n \in \mathbb{N}$, there are iid random variables $Y_1^{(n)}, \cdots, Y_n^{(n)}$ such that $X \overset{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$, where $\overset{d}{=}$ means identically distributed.*

Let $\mathcal{M}_1$ be the set of Borel probability measure in R^d. Define the convolution of two probability measures $\mu_i \in \mathcal{M}_1$ $(i = 1, 2)$ as follows:

$$\mu_1 * \mu_2(A) = \int_{R^d} \mu_1(A - x)\mu_2(dx) \ \forall A \in \mathcal{B}(R^d),$$

where $A - x = \{y - x, y \in A\}$. It can be shown that $\mu_1 * \mu_2$ is a probability measure on R^d. If there is $\mu^{1/n} \in \mathcal{M}_1$ such that $\mu = \mu^{1/n} * \cdots * \mu^{1/n}$ (n times), then μ is said to have a convolution nth root $\mu^{1/n}$. The measure $\mu \in \mathcal{M}_1(R^d)$ is infinitely divisible if it has a convolution nth root in $\mathcal{M}_1$ for each $n \in \mathbb{N}$. It is not difficult to show that $\mu \in \mathcal{M}_1(R^d)$ is infinitely divisible if and only if for each n, there exists $\mu^{1/n} \in \mathcal{M}_1(R^d)$ such that $\phi_\mu(\xi) = [\phi_{\mu^{1/n}}(\xi)]^n$ for each $\xi \in R^d$.

Definition 1.6.2. *A Borel measure ν on $R^d \backslash \{0\}$ is called a Lévy measure if*

$$\int_{R^d \backslash \{0\}} (|y|^2 \wedge 1)\nu(dy) < \infty.$$

Theorem 1.6.1 (Lévy–Khintchine formula). *If there is a vector $b \in R^d$, a positive-definite symmetric matrix $d \times d$ matrix A and a Lévy measure ν on $R^d \backslash \{0\}$ such that for any $\xi \in R^d$,*

$$\begin{aligned}\phi_\mu(\xi) = \exp\Big\{ & i(b, \xi) - \frac{1}{2}(\xi, A\xi) \\ & + \int_{R^d \backslash \{0\}} [e^{i(\xi, y)} - 1 - i(\xi, y)\chi_{\hat{B}}(y)]\nu(dy)\Big\},\end{aligned} \tag{1.6.1}$$

then $\mu \in \mathcal{M}_1$ is infinitely divisible, where $\hat{B} = B_1(0)$. Conversely, any mapping of the form (1.6.1) *is the characteristic function of an infinitely divisible probability measurable on R^d.*

The proof can be found in Ref. [7] and is omitted here. There are three important examples of the Lévy–Khintchine formula, i.e., (1) Gauss random variable: b is the mean value, A is the covariance matrix and $\nu = 0$; (2) Poisson random variable: $b = 0, A = 0$ and $\nu = \lambda\delta_1$; and (3) compound Poisson random variable: $b = 0, A = 0$ and $\nu = \lambda\mu$, where $c > 0$ and μ is a probability measure on R^d.

1.6.2 Lévy process

Definition 1.6.3. *The stochastic process $X = \{X_t\}_{t \geq 0}$ is a Lévy process if*

(L1) *$X_0 = 0$ a.s.,*

(L2) *X has independent and stationary increments, i.e., for any $0 \le t_1 < t_2 < \cdots < t_n$, the random variables $X(t_1) - X(t_0), X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1})$ are independent and the distribution of $X(t) - X(s)$ is independent of $t - s$,*

(L3) *X is stochastically continuous, i.e., $\lim_{t\to s} P(|X_t - X_s| > a) = 0$ for any $a > 0$ and $s \ge 0$.*

One can show that any Lévy process has a càdlàg modification. By definition, any Lévy process X is infinitely divisible since for any $n \in \mathbb{N}$, $X_t = Y_1^{(n)}(t) + \cdots + Y_n^{(n)}(t)$, where $Y_k^{(n)}(t) = X_{\frac{kt}{n}} - X_{\frac{(k-1)t}{n}}$ are iid by (L2).

Denote by $\phi_{X_t}(\xi) = e^{\eta(t,\xi)}$ $(\forall t \ge 0, \xi \in R^d)$ the characteristic function of X_t and call $\eta(t, \cdot)$ the Lévy symbol. By definition of the Lévy process and infinite divisibility, it's not difficult to show $\eta(t,\xi) = t\eta(1,\xi)$ and $\eta(\xi) = \log\left(E(e^{i(\xi,X_1)})\right)$, where $\eta(\xi) = \eta(1,\xi)$ is the Lévy symbol of X_1.

Applying the Lévy–Khintchine formula to Lévy process $X = \{X_t\}_{t\ge 0}$, we have for any $t \ge 0$ and $u \in R^d$ that

$$E(e^{i(\xi,X_t)}) = \exp\left(t\left\{i(b,\xi) - \frac{1}{2}(\xi, A\xi) + \int_{R^d\setminus\{0\}} [e^{i(\xi,y)} - 1 - i(\xi,y)\chi_{\hat{B}}(y)]\nu(dy)\right\}\right).$$

Obviously, (b, A, ν) are characteristics of X_1.

The following are two examples.

Example 1.6.3 (Brownian motion with drift). *Let $b \in R^d$, $B(t)$ be an m-dimensional Brownian motion, A be a $d \times d$-dimensional positive-definite symmetric matrix and σ be a $d \times m$ matrix such that $\sigma\sigma^T = A$. The process $C = \{C(t)\}_{t\in R^+}$ with $C(t) = bt + \sigma B(t)$ is a Brownian motion in R^d. Then C is a Lévy process with Lévy symbol $\eta_C(\xi) = i(b,\xi) - \frac{1}{2}(\xi, A\xi)$.*

Example 1.6.4 (Compound Poisson process). *Let $N = (N(t); t \ge 0)$ be a Poisson process independent of $Z(n)$ with parameter λ. Define compound Poisson process $Y = \{Y(t)\}_{t\ge 0}$ with $Y(t) = Z(1) + \cdots + Z(N(t))$ and denote $Y(t) \sim \pi(\lambda t, \mu_Z)$. It's easy to show that Y is a Lévy process with Lévy symbol*

$$\eta_Y(\xi) = \int_{R^d} (e^{i(\xi,y)} - 1)\lambda\mu_Z(dy).$$

Now, we have seen that the first two parts in the Lévy–Khintchine formula correspond to Brownian motion with drift and compound Poisson process. As to the last part, we leave it to the next section. But we first introduce the compensated Poisson process. Let $N(t)$ be a Poisson process with parameter λ and $\tilde{N}(t) = N(t) - \lambda t$, then the process $\tilde{N} = \{\tilde{N}(t)\}_{t\ge 0}$ is called a compensated Poisson process.

1.6.3 Lévy–Itô decomposition

For a Lévy process $X = \{X(t)\}_{t\geq 0}$, we introduce the jump process $\Delta X = \{\Delta X(t)\}_{t\geq 0}$ defined by $\Delta X(t) = X(t)-X(t-)$, where $X(t-) = \lim_{s\uparrow t} X(s)$. It is easy to show that if X is increasing (a.s.) Lévy process and is such that $\Delta X(t)$ takes values in $\{0,1\}$, then X is a Poisson process. Obviously, the jump process ΔX is an adapted process, but generally not a Lévy process. Indeed, consider a Poisson process $N = \{N(t)\}_{t\geq 0}$, it's easy to verify that

$$P(\Delta N(t_2) - \Delta N(t_1) = 0|\Delta N(t_1) = 1) \neq P(\Delta N(t_2) - \Delta N(t_1) = 0),$$

for $0 \leq t_1 < t_2 < \cdots < \infty$. Hence, ΔN cannot have independent increments and is not a Lévy process.

Setting $0 \leq t < \infty$ and $A \in \mathcal{B}(R^d\backslash\{0\})$, we define

$$N(t,A) = \#\{0 \leq s \leq t : \Delta X(s) \in A\}.$$

Obviously, for any $\omega \in \Omega$ and $t \geq 0$, the function $A \to N(t,A)(\omega)$ is a counting measure on $\mathcal{B}(R^d\backslash\{0\})$ and $E(N(t,A))$ is a Borel measure on $\mathcal{B}(R^d\backslash\{0\})$. Denote $\mu(\cdot) = E(N(1,\cdot))$ and call it the intensity measure associated with X. If $A \in \mathcal{B}(R^d\backslash\{0\})$ is bounded below, i.e., 0 does not belong to $\bar{A}$, then for any $t \geq 0$, one has $N(t,A) < \infty$ a.s. (see Ref. [7]). It's worth noting that this is not necessarily true when A isn't bounded below, since it may have accumulation of infinitely many small jumps.

Let $(S,\mathcal{A})$ be a measurable space and $(\Omega,\mathscr{F},P)$ be a probability space.

Definition 1.6.4. *A random measure M on $(S,\mathcal{A})$ is a collection of random variables $(M(B), B \in \mathcal{A})$ such that:*
(1) $M(\emptyset) = 0$,
(2) *(σ-additivity) for any given sequence $\{A_n\}_{n\in\mathbb{N}}$ of mutually disjoint sets in $\mathcal{A}$, there holds $M(\cup_{n\in\mathbb{N}}) = \sum_{n\in\mathbb{N}} M(A_n)$ a.s.,*
(3) *for any mutually disjoint family $(B_1,\cdots,B_n)$ in $\mathcal{A}$, the random variable $M(B_1),\cdots,M(B_n)$ are mutually independent.*

In particular, we say M is a Poisson random measure if $M(B)$ has a Poisson distribution whenever $M(B) < \infty$.

Let U be a measurable space equipped with a σ-algebra $\mathcal{U}$, $S = R^+ \times U$ and $\mathcal{A} = \mathcal{B}(R^+) \otimes \mathcal{U}$. Let $p = \{p(t)\}_{t\geq 0}$ be a U-valued adapted process such that M is a Poisson random measure on S, where $M([0,t) \times A) = \#\{0 \leq s < t; p(s) \in A\}$ for each $t \geq 0$ and $A \in \mathcal{U}$. The process p is usually called the Poisson point process and M its associated Poisson random measure. When U is a topological space and $\mathcal{U}$ is its Borel σ-algebra, we define, for any $A \in \mathcal{U}$, a process $M_A = \{M_A(t)\}_{t\geq 0}$ by $M_A(t) = M([0,t) \times A)$. If there is $V \in \mathcal{U}$ such that M_A is a martingale whenever $\bar{A} \cap V = \emptyset$, then M is called a martingale-valued measure and V is called the associated forbidden set.

When $U = R^d\backslash\{0\}$ and $\mathcal{U}$ is its Borel σ-algebra. Then for a Lévy process $X = \{X(t)\}_{t\geq 0}$, ΔX is a Poisson point process and N is its Poisson random measure. It is easy to show that for each $t > 0$ and $\omega \in \Omega$, $N(t,\cdot)(\omega)$ is a counting measure on $\mathcal{B}(R^d\backslash\{0\})$ and for each A bounded below, $\{N(t,A)\}_{t\geq 0}$ is a Poisson process with intensity $\mu(A) = E(N(1,A))$. For any $t \geq 0$ and A bounded below, define the compensated Poisson random measure

$$\tilde{N}(t,A) = N(t,A) - t\mu(A).$$

Then $\{\tilde{N}(t,A)\}_{t\geq 0}$ is a martingale and so $\tilde{N}$ extends to a martingale-valued measure with forbidden set $\{0\}$.

Let f be a Borel measurable function from R^d to R^d and A be bounded below. Then for any $t > 0$ and $\omega \in \Omega$, we can define the Poisson integral of f as

$$\int_A f(x)N(t,dx)(\omega) = \sum_{x\in A} f(x)N(t,\{x\})(\omega). \tag{1.6.2}$$

For $t \geq 0$ and $\xi \in R^d$, the characteristic function is

$$E\left(\exp\left\{i\left(\xi, \int_A f(x)N(t,dx)\right)\right\}\right) = \exp\left\{t\int_A (e^{i(\xi,x)} - 1)\mu_f(dx)\right\},$$

where $\mu_f = \mu\circ f^{-1}$. If A_1 and A_2 are two disjoint sets and bounded below, then $\{\int_{A_1} f(x)N(t,dx)\}_{t\geq 0}$ and $\{\int_{A_1} f(x)N(t,dx)\}_{t\geq 0}$ are mutually independent [7, Theorem 2.4.6].

Let $f \in L^1(A,\mu_A)$, then we can definite the compensate Poisson integral of f as

$$\int_A f(x)\tilde{N}(t,dx) = \int_A f(x)N(t,dx) - t\int_A f(x)\mu(dx),$$

which is a martingale and whose characteristic function is given by

$$E\exp\left\{i\left(\xi, \int_A f(x)\tilde{N}(t,dx)\right)\right\} = \exp\left\{t\int_A \left[e^{i(\xi,y)} - 1 - i(\xi,y)\mu_f(dx)\right]\right\}.$$

Let $X = \{X(t)\}_{t\geq 0}$ be a Lévy process and $A = \{x : |x| \geq 1\}$. Obviously, A is bounded below. Consider the compound Poisson process $\{\int_{|x|\geq 1} xN(t,dx)\}_{t\geq 0}$ and define a new process $Y = \{Y(t)\}_{t\geq 0}$ as

$$Y(t) = X(t) - \int_{|x|\geq 1} xN(t,dx).$$

Then Y is also a Lévy process. We further define a new Lévy process $\hat{Y}$ as

$$\hat{Y}(t) = Y(t) - E(Y(t)).$$

It can be verified that $\hat{Y}$ is a càdlàg square-integrable martingale with zero expectation and has decomposition

$$\hat{Y}(t) = Y_{\mathrm{c}}(t) + Y_{\mathrm{d}}(t),$$

where Y_{c} and Y_{d} are independent Lévy processes, Y_{c} has continuous sample path and there exists a $d \times d$ positive-definite matrix A such that

$$E(e^{i(\xi, Y_{\mathrm{c}}(t))}) = e^{-t(\xi, A\xi)/2}$$

and

$$Y_{\mathrm{d}}(t) = \int_{|x|<1} x\tilde{N}(t, dx).$$

The continuous part Y_{c} can be proved to be a Brownian motion and the discontinuous part Y_{d} is the compensate sum of small jumps and

$$E(e^{i(\xi, Y_{\mathrm{d}}(t))}) = \exp\left\{t \int_{|x|<1} [e^{i(\xi,x)} - 1 - i(\xi, x)]\mu(dx)\right\} \quad \forall t \geq 0, \xi \in R^d.$$

Till now, we should understand the meaning of the last part in the Lévy–Khintchine formula. This leads to the following Lévy–Itô decomposition [7, 14].

Theorem 1.6.2 (Lévy-Itô decomposition). *Let $X = \{X(t)\}_{t\geq 0}$ be a Lévy process, then there exists $b \in R^d$, a Brownian motion B_A with covariance matrix A and an independent Poisson random measure N on $R^+ \times (R^d \backslash \{0\})$ such that for each $t \geq 0$,*

$$X(t) = bt + B_A(t) + \int_{|x|<1} x\tilde{N}(t, dx) + \int_{|x|\geq 1} xN(t, dx), \tag{1.6.3}$$

where $b = E\left(X(1) - \int_{|x|\geq 1} xN(1, dx)\right)$.

By independence,

$$E(e^{i(\xi, X(t))}) = E(e^{i(u, Y_{\mathrm{c}}(t))})E(e^{i(u, Y_{\mathrm{d}}(t))})E(e^{i(u, \int_{|x|\geq 1} xN(t,dx))}).$$

Recall that an adapted process X is a semimartingale if it has the decomposition

$$X(t) = X(0) + M(t) + C(t),$$

where $M = \{M(t)\}_{t\geq 0}$ is a local martingale and $C = \{C(t)\}_{t\geq 0}$ is an adapted process with bounded variation. A direct consequence of the Lévy–Itô decomposition is that any Lévy processes $X = \{X(t)\}_{t\geq 0}$ is a semimartingale. Indeed, for any $t \geq 0$, we have

$X(t) = M(t) + C(t)$ for $M(t) = B_A(t) + \int_{|x|<1} x\tilde{N}(t, dx)$ and $C(t) = bt + \int_{|x|\geq 1} xN(t, dx)$. Here M is a martingale and $Y(t) = \int_{|x|\geq 1} xN(t, dx)$ is a compound Poisson process and hence for any partition $\mathcal{P}$ of $[0, t]$,

$$\mathrm{Var}_{\mathcal{P}}(Y) \leq \sum_{0\leq s\leq t} |\Delta X(s)|\chi_{[1,\infty)}(\Delta X(s)) < \infty \quad a.s..$$

1.7 The fractional Brownian motion

Another important class of stochastic processes is the fractional Brownian motion. A fractional Brownian motion is defined initially by Kolmogorov in Hilbert space framework [158]. A fractional Brownian motion W^H with (Hurst) index $H \in (0, 1)$ is a centered Gaussian process with covariance function

$$E(W^H(t), W^H(s)) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \quad s, t \geq 0.$$

When $H = \frac{1}{2}$, it reduces to the standard Brownian motion. A fractional Brownian motion has stationary increment

$$E\left((W^H(t) - W^H(s))^2\right) = |t-s|^{2H},$$

and is H-self-similarity, i.e., for any $c > 0$,

$$\left\{\frac{1}{C^H} W^H(ct)\right\}_{t\geq 0} \overset{d}{=} \{W^H(t)\}_{t\geq 0},$$

where $\overset{d}{=}$ means identically distributed.

When $H \neq \frac{1}{2}$, W^H is neither a semimartingale nor a Markov process. Here are some properties of the fractional Brownian motion.

1. (Time homogeneity): For any $s > 0$, the process $\{W^H(t+s) - W^H(s)\}_{t>0}$ is a fractional Brownian motion with Hurst index H.
2. (Symmetry): The process$\{-W^H(t)\}_{t>0}$ is also a fractional Brownian motion with Hurst parameter H.
3. (Scaling): For any $c > 0$, the process $\{c^H W^H(t/c)\}_{t\geq 0}$ is a fractional Brownian motion with Hurst parameter H.
4. (Time reversibility): The process $\{X(t)\}_{t\geq 0}$ defined by $X(t) = t^{2H} W^H(1/t)$ with $X(0) = 0$ is also a fractional Brownian motion with Hurst parameter H.

Proposition 1.7.1. *For $H \in [0, 1]$, the sample path W^H of a fractional Brownian motion is a.s. α-Hölder continuous with $\alpha < H$.*

Several expressions of the fractional Brownian motion are listed below.

(1) **Mean motion expression:**

$$W^H(t) = \frac{1}{c_1(H)} \int_{\infty}^{\infty} \left[(t-x)_+^{H-\frac{1}{2}} - (-x)_+^{H-\frac{1}{2}} \right] M(dx), \quad t \in R,$$

where $c_1(H) = \sqrt{\int_0^\infty [(1+x)^{H-\frac{1}{2}} - x^{H-\frac{1}{2}}]^2 dx + \frac{1}{2H}}$, $M(dx)$ is a Gaussian random measure and $(a)_+ = a\chi_{[0,\infty)}(a)$.

(2) **Harmonic analysis expression:**

$$W^H(t) = \frac{1}{c_2(H)} \int_{-\infty}^{\infty} \frac{e^{ixt}-1}{ix} |x|^{\frac{1}{2}-H} M(dx), \quad t \in R, \tag{1.7.1}$$

where $c_2(H) = \sqrt{\frac{\pi}{H\Gamma(2H)\sin \pi H}}$.

(3) **Volterra expression:**

$$W^H(t) = \int_0^t K^H(t,s)dB(s), \quad t \ge 0, \tag{1.7.2}$$

where $K^H(t,s) = \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})} \Gamma(H-\frac{1}{2}, \frac{1}{2}-H, H+\frac{1}{2}, 1-\frac{t}{s})$, $s < t$, Γ is a Gauss hypergeometric function and $\{B(t)\}_{t\ge 0}$ is a Brownian motion. $K^H(t,s)$ can also be expressed as

$$K^H(t,s) = c_H \left[\frac{t^{H-\frac{1}{2}}}{s^{H-\frac{1}{2}}} (t-s)^{H-\frac{1}{2}} - \left(H - \frac{1}{2}\right) \int_0^t \frac{u^{H-\frac{1}{2}}}{s^{H-\frac{1}{2}}} (u-s)^{H-\frac{1}{2}} du \right],$$

where $c_H = \sqrt{\frac{\pi H(2H-1)}{\Gamma(2-2H)\Gamma(H+\frac{1}{2})^2 \sin(\pi(H-\frac{1}{2}))}}$.

2 The stochastic integral and Itô formula

This chapter introduces the stochastic integral, the Itô integral and the Itô formula. Especially, they are discussed in the infinite dimensional case in Section 2.3 for application in partial differential equations. Most of these materials in this chapter can be found in Refs [7, 75, 145, 211].

2.1 Stochastic integral

The stochastic integral $\int H\,dX_t$ of H with respect to a stochastic process $X_t = X_0 + M_t + A_t$ ($M_0 = A_0 = 0$) is defined in this section, where M_t is a locally square-integrable martingale and A_t is an adaptive càdlàg process with bounded variation on a compact set. We know that when A is of bounded variation and H is continuous, the integral $\int H\,dA$ is well defined.

As an example, consider the integral with compensated Poisson process $M_t = N_t - \lambda t$, which is obviously a real-valued martingale with bounded variation. For simplicity, we take H to be a bounded jointly measurable process. Hence,

$$\begin{aligned} I_t &= \int_0^t H_s dM_s = \int_0^t H_s d(N_s - \lambda s) \\ &= \int_0^t H_s dN_s - \lambda \int_0^t H_s ds \\ &= \sum_{n=1}^{\infty} H_{\tau_n} 1_{t \geq \tau_n} - \lambda \int_0^t H_s ds, \end{aligned}$$

where $(\tau_n)_{n\geq 1}$ are the arrival times of the Poisson process N_t. If, furthermore, H is a bounded and adapted process with continuous sample path, then by the dominated convergence theorem,

$$\begin{aligned} E(I_t - I_s|\mathscr{F}_s) =& E\left(\int_s^t H_u dM_u|\mathscr{F}_s\right) \\ =& E\left(\lim_{n\to\infty} \sum_{t_k, t_{k+1} \in \pi_n} H_{t_k}(M_{t_{k+1}} - M_{t_k})|\mathscr{F}_s\right) \\ =& \lim_{n\to\infty} \sum_{t_k, t_{k+1} \in \pi_n} E(E(H_{t_k}(M_{t_{k+1}} - M_{t_k})|\mathscr{F}_{t_k})|\mathscr{F}_s) = 0. \end{aligned}$$

This implies that the integral I is a martingale and the stochastic Stieltjes integral of an adapted and bounded continuous process with respect to a martingale is still a martingale.

But it is not clear when M_t is a square-integrable martingale, which does not necessarily has bounded variation on any finite interval. To tackle this problem, the following Banach–Steinhaus theorem is necessary, whose proof can be referred to, for example, Yosida [259].

Theorem 2.1.1. *Let X be a Banach space, Y be a normed linear space and $\{T_\alpha\}_{\alpha\in I}$ be a family of bounded linear operators from X to Y. If $\{T_\alpha x\}$ is bounded for any $x \in X$, then $\{T_\alpha\}$ is bounded.*

Consider a right-continuous function $x(t)$ on $[0,1]$ and let Π_n be a binary rational division on $[0,1]$ with $\lim_{n\to\infty}$ mesh $(\Pi_n) = 0$. It's nature to ask under what restrictions should be imposed on x to ensure the following summation:

$$S_n = \sum_{t_k, t_{k+1}\in\Pi_n} h(t_k)(x(t_{k+1}) - x(t_k)) \tag{2.1.1}$$

converge for any continuous function h as $n \to \infty$.

It follows from real analysis theory that when $x(t)$ has finite variation, S_n converges to the integration $\int_0^1 h(s)dx(s)$. But the following theorem shows that it is also necessary for x to be of finite variation.

Theorem 2.1.2. *If S_n converges for any continuous function $h(s)$, then x has finite variation.*

Proof. Let X be the Banach space of continuous functions equipped with maximum norm. For $h \in X$, let $T_n(h) = \sum_{t_k,t_{k+1}\in\Pi_n} h(t_k)(x(t_{k+1}) - x(t_k))$. Then for any fixed n, we construct $h \in X$ such that $h(t_k) = \mathrm{sign}(x(t_{k+1}) - x(t_k))$ and $\|h\| = 1$. For such h, $T_n(h) = \sum_{t_k,t_{k+1}\in\Pi_n} |x(t_{k+1}) - x(t_k)|$. Hence, $\|T_n\| \geq \sum_{t_k,t_{k+1}\in\Pi_n} |x(t_{k+1}) - x(t_k)|$ for all n and $\sup_n T_n$ is bigger than the total variation of x. On the other hand, for any $h \in X$, $\lim_{n\to\infty} T_n(h)$ exists and hence $\sup_n \|T_n(h)\| < \infty$. It follows from Banach–Steinhaus theorem that $\sup_n \|T_n\| < \infty$ and hence x has finite variation. ■

2.1.1 Itô integral

Now, we consider the simplest case when $B_s(\omega)$ is a standard Brown motion. We will define the corresponding stochastic integral

$$\int_0^t f(s,\omega)dB_s(\omega). \tag{2.1.2}$$

First, we take a partition of the interval $[0,t]$ to be $t_k = t_k^n = k \cdot 2^{-n}$ with $0 \leq k \cdot 2^{-n} \leq t$. It then seems reasonable to define

$$\int_0^t f(s,\omega)dB_s(\omega) = \sum_{j\geq 0} f(t_j^*, \omega)[B_{t_{j+1}} - B_{t_j}](\omega),$$

where $t_j^* \in [t_j, t_{j+1}]$. But the following example shows that different values of t_j^* lead to different results unlike the Riemann–Stieltjes integral. Let $f(t,\omega) = B_t(\omega)$, then choice of $t_j^* = t_j$ leads to I_1, while the choice of $t_j^* = t_{j+1}$ leads to I_2. Since $\{B_t\}$ has independent increments, it then follows that

$$E[I_1] = \sum_{j\geq 0} E[B_{t_j}(B_{t_{j+1}} - B_{t_j})] = 0$$

and

$$E[I_2] = \sum_{j\geq 0} E[B_{t_{j+1}}(B_{t_{j+1}} - B_{t_j})] = \sum_{j\geq 0} E[(B_{t_{j+1}} - B_{t_j})^2] = t.$$

The following two choices of t_j^* are the most common and useful ones.

Itô integral. Choose $t_j^* = t_j$ and the corresponding integral is denoted by $\int_0^t f(t,\omega)dB_t(\omega)$.

Stratonovitch integral. Choose $t_j^* = (t_j + t_{j+1})/2$ and the corresponding integral is denoted by $\int_0^t f(t,\omega) \circ dB_t(\omega)$.

In order to make integral (2.1.2) well defined, the integrand $f(t,\omega)$ must be confined to certain class of functions. For this, we first define $\mathscr{F}_t$ to be the σ-algebra generated by the random variables $\{B_s\}_{0\leq s\leq t}$ where $B_s(\omega)$ is an n-dimensional Brownian motion.

Definition 2.1.1. *Let $\mathcal{V} = \mathcal{V}(S,T)$ be a class of functions $f(t,\omega) : [0,\infty) \times \Omega \to R$ such that the mapping $(t,\omega) \to f(t,\omega)$ is $\mathcal{B} \times \mathscr{F}$ measurable, where $\mathcal{B}$ is a Borel σ-algebra on $[0,\infty)$, $f(t,\omega)$ is $\mathscr{F}_t$ adapted and $E\int_0^t f(s,\omega)^2 ds < \infty$.*

Let $f \in \mathcal{V}$ be a step function

$$f(t,\omega) = \sum_j e_j(\omega)\chi_{[t_j,t_{j+1})}(t),$$

then the Itô integral can be defined as

$$\int_0^t f(s,\omega)dB_s(\omega) = \sum_{j\geq 0} e_j(\omega)[B_{t_{j+1}} - B_{t_j}](\omega). \tag{2.1.3}$$

Lemma 2.1.1 (Itô isometry). *If $f(t,\omega) \in \mathcal{V}$ be a step function, then*

$$E\left[\left(\int_0^t f(s,\omega)dB_s(\omega)\right)^2\right] = E\left[\int_0^t f(s,\omega)^2 ds\right]. \tag{2.1.4}$$

Proof. Since $E[e_i e_j \Delta B_i \Delta B_j] = E[e_j^2] \cdot (t_{j+1} - t_j)$ for $i = j$ and 0 for $i \neq j$ due to the independence of $e_i e_j \Delta B_i$ and ΔB_j, it follows that

$$E\left[\left(\int_0^t f dB\right)^2\right] = \sum_{i,j} E[e_i e_j \Delta B_i \Delta B_j] = \sum_j E[e_j^2] \cdot (t_{j+1} - t_j) = E\left[\int_0^t f^2 ds\right],$$

where $\Delta B_j = B_{t_{j+1}} - B_{t_j}$. ■

For a general $f \in \mathcal{V}$, the Itô integral can be defined via approximation.

Step 1. Let $g \in \mathcal{V}$ be bounded and $g(\cdot, \omega)$ continuous for any ω. Define $\phi_n(t,\omega) = \sum_j g(t_j, \omega) \cdot \chi_{[t_j, t_{j+1})}(t)$ be a step function, then $\int_0^t (g - \phi_n)^2 ds \to 0$ as $n \to \infty$ for each ω by the continuity of $g(\cdot, \omega)$. Hence, $E[\int_0^t (g - \phi_n)^2 ds] \to 0$ as $n \to \infty$.

Step 2. If $h \in \mathcal{V}$ is bounded, then there exists a bounded function sequence $g_n \in \mathcal{V}$ such that $g_n(\cdot, \omega)$ is continuous for all ω and n, and

$$E\left[\int_0^t (h - \phi_n)^2 ds\right] \to 0.$$

Indeed, assume $|h(t,\omega)| \le M$ for all (t,ω). For any n, define the nonnegative continuous function ψ_n on R such that $\psi_n(x) = 0$ for $x \le -\frac{1}{n}$ and $x \ge 0$ and $\int_{-\infty}^{\infty} \psi_n(x) dx = 1$. Then define $g_n(t,\omega) = \int_0^t \psi_n(s-t) h(s,\omega) ds$, then $g_n(\cdot,\omega)$ is continuous for any ω and $|g_n(t,\omega)| \le M$. Since $h \in \mathcal{V}$ it follows that $g_n(t,\cdot)$ is still $\mathscr{F}_t$ measurable and moreover, since $\{\psi_n\}$ is an approximate identity, it follows that

$$\int_0^t (g_n(s,\omega) - h(s,\omega))^2 ds \to 0$$

as $n \to \infty$ for any ω. The conclusion then follows from bounded convergence theorem.

Step 3. Finally, let $f \in \mathcal{V}$. If we take $h_n(t,\omega) = -n$ for $f(t,\omega) < -n$, $h_n(t,\omega) = f(t,\omega)$ for $-n \le f(t,\omega) \le n$ and $h_n(t,\omega) = n$ if $f(t,\omega) > -n$, then $h_n \in \mathcal{V}$ is bounded for each n and

$$E\left[\int_0^t (f - h_n)^2 ds\right] \to 0, \quad \text{as } n \to \infty.$$

Definition 2.1.2 (Itô integral). *Let $f \in \mathcal{V}$, then Itô integral is defined by*

$$\int_0^t f(s,\omega) dB_s(\omega) = \lim_{n\to\infty} \int_0^t \phi_n(s,\omega) dB_s(\omega), \tag{2.1.5}$$

where the limit is taken in $L^2(P)$ and $\{\phi_n\} \subset \mathcal{V}$ is a sequence of step functions such that

$$E\left[\int_0^t (f(s,\omega) - \phi_n(s,\omega))^2 ds\right] \to 0, \quad \textit{as } n \to \infty.$$

We remark that from the definition, it follows that $E[\int_0^t f dB_s] = 0$ when $f \in \mathcal{V}$ and the Itô isometry also holds for $f \in \mathcal{V}$, i.e.,

$$E\left[\left(\int_0^t f(s,\omega)dB_s\right)^2\right] = E\left[\int_0^t f^2(s,\omega)ds\right] \quad \forall f \in \mathcal{V}.$$

Furthermore, there exists a time-continuous version of the Itô integral $\int_0^\tau f(s,\omega)dB_s(\omega)$ for $0 \le \tau \le t$, i.e., there exists a time-continuous stochastic process J_τ on $(\Omega, \mathscr{F}, P)$ such that $P(J_\tau = \int_0^\tau f dB) = 1$ for all $0 \le \tau \le t$. Hence, we always assume that the Itô integral is time continuous.

Theorem 2.1.3. *For any $T > 0$, let $f \in \mathcal{V}(0,T)$. Then the integral $M_t(\omega) = \int_0^t f(s,\omega)dB_s$ is a $\mathscr{F}_t$ martingale and the following martingale inequality holds:*

$$P\left(\sup_{0\le t\le T} |M_t| \ge \lambda\right) \le \frac{1}{\lambda^2} E\left[\int_0^T f(s,\omega)^2 ds\right], \quad \lambda, T > 0.$$

2.1.2 The stochastic integral in general case

In this section, we will consider the stochastic integral with respect to real-valued martingale measure M on $R^+ \times E$. Let

$$M((s,t],A] = M(t,A] - M(s,A],$$

where $0 \le s < t < \infty$ and $A \in \mathcal{B}(E)$. To define the stochastic integral, some restrictions on M are also needed. We assume (M1) $M(0,A) = 0$ a.s., (M2) $M((s,t],A)$ is independent of $\mathscr{F}_s$ and (M3) there exists a σ-finite measure ρ on $R^+ \times E$ such that $E[M(t,A)^2] = \rho(t,A)$ for all $0 \le s < t < \infty$ and $A \in \mathcal{B}$, where $\rho(t,A) = \rho((0,T],A)$ for brevity. The martingale–valued measures satisfying (M1)–(M3) are said to be of type $(2,\rho)$. The martingale–valued measure M is said to be continuous if the sample paths $t \mapsto M(t,A)(\omega)$ are continuous for a.a. $\omega \in \Omega$ and $A \in \mathcal{B}(E)$. In what follows, we always assume $\rho((0,t],A) = t\mu(A)$ for some σ-finite measure μ on E.

Let $F : [0,T] \times E \times \Omega \to R$ be a mapping such that $(x,\omega) \mapsto F(t,x,\omega)$ is $\mathcal{B}(E) \times \mathscr{F}_t$ measurable for any $0 \le t \le T$ and the mapping $t \mapsto F(t,x,\omega)$ is left continuous for any $x \in E$ and $\omega \in \Omega$. Fix $E \in \mathcal{B}(E)$ and $0 < T < \infty$, and let $\mathcal{P}$ be the smallest σ-algebra such that the mapping $F : [0,T] \times E \times \Omega \to R$ is measurable. We call such $\mathcal{P}$ the predictable σ-algebra and any $\mathcal{P}$-measurable mapping F is called predictable.

Fix $T > 0$ and define $\mathcal{H}_2(T,E)$ to be the linear space of all predictable mapping $F : [0,T] \times E \times \Omega \to R$ such that $\int_0^T \int_E E[|F(t,x)|^2]\rho(dt,dx) < \infty$. In such a space, we identify two mappings that coincide a.e. with respect to the measure $\rho \times P$. We define an inner product $\langle \cdot, \cdot \rangle_{T,\rho}$ to be

$$\langle F, G \rangle_{T,\rho} = \int_0^T \int_E E[F(t,x)G(t,x)]\rho(dt,dx),$$

which also introduces a norm in the usual way $\|F\|^2_{T,\rho} = \langle F, F\rangle_{T,\rho}$ and makes $\mathcal{H}_2(T, E)$ be a Hilbert space.

As before, we first consider stochastic integrals of simple functions. Let $\mathcal{S}(T, E)$ be the linear space of all *simple processes* in $\mathcal{H}_2(T, E)$, i.e.,

$$F = \sum_{j=1}^{m} \sum_{k=1}^{n} c_k F(t_j) \chi_{(t_j, t_{j+1}]} \chi_{A_k}$$

for some $0 \leq t_1 \leq t_2 \leq \cdots \leq t_{m+1} = T$ and disjoint Borel subsets $A_1, A_2, \cdots, A_n$ of E with $\mu(A_i) < \infty$, where $c_k \in R$ and $F(t_j)$ is a bounded $\mathcal{F}_{t_j}$-measurable random variable. Since F is left continuous and $\mathcal{B}(E) \otimes \mathcal{F}_t$ measurable, it is predictable. It can be proved that $\mathcal{S}(T, E)$ is dense in $\mathcal{H}_2(T, E)$. If we set $c_k F(t_j) = F_k(t_j)$, then we can rewrite F as

$$F = \sum_{j,k}^{m,n} F_k(t_j) \chi_{(t_j, t_{j+1}]} \chi_{A_k}. \tag{2.1.6}$$

Just like the Itô integral, for fixed $T > 0$ and $F \in \mathcal{S}(T, E)$ of the form (2.1.6) we define the stochastic integral with respect to a $(2, \rho)$-type martingale-valued measure M to be

$$I_T(F) = \int_0^T \int_E F(t, x) M(dt, dx) = \sum_{j,k=1}^{m,n} F_k(t_j) M((t_j, t_{j+1}], A_k). \tag{2.1.7}$$

It can be proved that for any $T \geq 0$ and $F \in \mathcal{S}(T, E)$ there holds

$$E[I_T(F)] = 0, \quad E[I_T(F)^2] = \|F\|^2_{T,\rho}. \tag{2.1.8}$$

Since I_T is a linear isometry from $\mathcal{S}(T, E)$ to $L^2(\Omega, \mathscr{F}, P)$ thanks to eq. (2.1.8), the stochastic integral I_T can be extended to integrand in $\mathcal{H}_2(T, E)$ and is still called *Itô integral*. By this extension, the Itô isometry formula (2.1.8) still holds for $F \in \mathcal{H}_2(T, E)$. One can also show that $\{I_t(F); t \geq 0\}$ is $\mathscr{F}_t$ adapted and $\{I_t(F); t \geq 0\}$ is a square-integrable martingale.

Finally, we extend the stochastic integral to a more general class of stochastic processes. Let $F : [0, T] \times E \times \Omega$ be predictable and $P(\int_0^T \int_E |F(t, x)|^2 \rho(dt, dx) < \infty) = 1$ and define $\mathcal{P}_2(T, E)$ to be the set of all equivalence classes of mapping F that coincide a.e. with respect to the measure $\rho \times P$.

It is obvious that $\mathcal{P}_2(T, E)$ is a linear space, $\mathcal{H}_2(T, E) \subseteq \mathcal{P}_2(T, E)$ and $\mathcal{S}(T, E)$ is dense in $\mathcal{P}_2(T, E)$ in the sense that for any $F \in \mathcal{P}_2(T, E)$, we can find a sequence $\{F_n\} \subset \mathcal{S}(T, E)$ such that

$$P\left(\lim_{n\to\infty} \int_0^T \int_E |F_n(t, x) - F(t, x)|^2 \rho(dt, dx) = 0\right) = 1.$$

For these simple functions $\{F_n\} \subset \mathcal{S}(T, E)$, one can define

$$I_{T,n} = \int_0^T \int_E F_n(t,x) M(dt,dx),$$

for $n \geq 0$ in the usual way. It is not difficult to prove that this sequence is a Cauchy sequence in probability and has a unique limit in probability. Denote this limit by $\hat{I}_T(F)$ and call it an extended stochastic integral. If

$$P\left(\int_0^t \int_E |F(t,x)|^2 \rho(dt,dx) < +\infty \right) = 1,$$

then $\{\hat{I}_t(F)\}_{t \geq 0}$ can also be regarded as a stochastic integral.

Generally speaking, the process $\{\hat{I}_t(F)\}_{t \geq 0}$ is no longer a martingale, but can be shown to be a local martingale and have a càdlàg correction.

2.1.3 Poisson stochastic integral

In this section, we will use the stochastic integral established to a Poisson random measure. First of all, let $E = \hat{B} \backslash \{0\}$, where $\hat{B}$ is the unit ball, N be a Poisson random measure of $R^+ \times (R^d - \{0\})$ with intensity ν and $\tilde{N}$ be the associated compensated Poisson random measure, which is a martingale measure. If $H = (H^1, \cdots, H^d) \in \mathcal{P}_2(T,E)$, then define $Z(t) = (Z^1(t), \cdots, Z^d(t))$ by

$$Z^i(T) = \int_0^T \int_{|x|<1} H^i(t,x) \tilde{N}(dt,dx).$$

Let A be a Borel set in $R^d \backslash \{0\}$ that is bounded below, and introduce the composite Poisson process $P = \{\int_A xN(t,dx)\}_{t \geq 0}$. Let K be predictable, then the Poisson stochastic integral can be extended to

$$\int_0^T \int_A K(t,x) N(dt,dx) = \sum_{0 \leq u \leq T} K(u, \Delta P(u)) \chi_A(\Delta P(u)). \tag{2.1.9}$$

In particular, when H is square integrable, then

$$\int_0^T \int_A H^i(t,x) \tilde{N}(dt,dx) = \int_0^T \int_A H^i(t,x) N(dt,dx) - \int_0^T \int_A H^i(t,x) \nu(dx) dt.$$

As an example, one can directly show that for each $t \geq 0$

$$\int_0^t N(s) d\tilde{N}(s) - \int_0^t N(s-) d\tilde{N}(s) = N(t). \tag{2.1.10}$$

Hence, the process $\{\int_0^t N(s) d\tilde{N}(s)\}_{t \geq 0}$ cannot be a local martingale, since here $N(s)$ is càglàd, but not predictable.

Let us still consider $E = \hat{B}\backslash\{0\}$. If we put all the stochastic integrals together, then for all $1 \le i \le d$, $1 \le j \le m$ and $t \ge 0$,

$$\begin{aligned} Y^i(t) =& Y^i(0) + \int_0^t G^i(s)ds + \int_0^t F^i_j dB^j(s) \\ &+ \int_0^t \int_{|x|<1} H^i(s,x)\tilde{N}(ds,dx) + \int_0^t \int_{|x|\ge 1} K^i(s,x)N(ds,dx), \end{aligned} \tag{2.1.11}$$

where $|G^i|^{\frac{1}{2}}, F^i_j \in \mathcal{H}_2(T), H^i \in \mathcal{H}_2(T,E)$ and K^i is predictable, B is an m-dimensional standard Brownian motion and N is an independent Poisson random measure on $R^+ \times (R^d\backslash\{0\})$ with compensator $\tilde{N}$ and intensity ν, which is assumed to be a Lévy measure. Thus the above defined R^d-valued stochastic process is called a *Lévy-type stochastic integral*. It is not difficult to illustrate that Y has càdlàg modifications and if $Y(0)$ is $\mathscr{F}_0$ measurable, then Y is adapted and a semi martingale. It can be rewritten in a more compact form

$$dY(t) = G(t)dt + F(t)dB(t) + H(t,x)\tilde{N}(dt,dx) + K(t,x)N(dt,dx).$$

Let X be a Lévy process with characteristic (b, a, ν) and Lévy–Itô decomposition

$$X(t) = bt + B_a(t) + \int_{|x|<1} x\tilde{N}(t,dx) + \int_{|x|\ge 1} xN(t,dx).$$

Let $L \in \mathcal{P}_2(t)$ for all $t \ge 0$, and choose $F^i_j = \sigma^i_j L$, $H^i = K^i = x^i L$ and $\sigma^T\sigma = a$ in eq. (2.1.11), then we can construct processes $Y = \{Y(t)\}_{t\ge 0}$ by

$$dY(t) = L(t)dX(t),$$

which is called a *Lévy stochastic integral*.

2.2 Itô formula

By the definition of Itô integral, one has

$$\begin{aligned} \int_0^t B_s dB_s =& \frac{1}{2}B_t^2 - \frac{1}{2}t, \\ \int_0^t B_s^2 dB_s =& \frac{1}{3}B_t^3 - \int_0^t B_s ds, \end{aligned}$$

where $B_0 = 0$. This indicates that the usual chain rule in calculus does not hold for stochastic integrals. In the following, we will introduce the Itô formula, which plays a vital role in stochastic integral theory. Let B_t be a one-dimensional Brownian motion in a given probability space $(\Omega, \mathscr{F}, P)$. A stochastic process X_t is called a Itô process if it has the form

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s, \tag{2.2.1}$$

where $v \in \mathcal{V}$ and u is $\mathscr{F}_t$ adapted.

Theorem 2.2.1. *Let X_t be a Itô process and $dX_t = udt + vdB_t$. Suppose $g \in C^2([0,\infty) \times R)$, then $Y_t = g(t, X_t)$ is still a Itô process and*

$$dY_t = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t) \cdot (dX_t)^2, \tag{2.2.2}$$

where $(dX_t)^2 = (dX_t) \cdot (dX_t)$ is calculated by the following principles:

$$dt \cdot dt = dt \cdot dB_t = dB_t \cdot dt = 0, \quad dB_t \cdot dB_t = dt.$$

Proof. The proof can be found in Ref. [192] and is omitted here. ∎

The Itô formula can be generalized to the n-dimensional case in an obvious way.

Theorem 2.2.2. *Let $dX_t = udt + vdB_t$ be an n-dimensional Itô process, $g(t,x) = (g_1(t,x), \cdots, g_p(t,x))$ a C^2 mapping from $[0,\infty) \times R^n$ to R^p, then $Y(t,\omega) = g(t, X_t)$ is also a Itô process and*

$$dY_k = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_i \frac{\partial g_k}{\partial x_i}(t,X)dX_i + \frac{1}{2}\frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j,$$

where $dB_i dB_j = \delta_{ij}dt$ and $dB_i dt = dtdB_i = 0$.

As an application of the Itô formula, one has the following integration by parts formula.

Corollary 2.2.1. *Let X_t, Y_t be two Itô processes,*

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dB_s, \quad Y_t = Y_0 + \int_0^t K'_s ds + \int_0^t H'_s dB_s,$$

then

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t,$$

where $\langle X, Y \rangle_t = \int_0^t H_s H'_s ds$.

Proof. By the Itô formula

$$
\begin{aligned}
(X_t + Y_t)^2 &= (X_0 + Y_0)^2 + 2\int_0^t (X_s + Y_s)d(X_s + Y_s) + \int_0^t (H_s + H_s')^2 ds,\\
X_t^2 &= X_0^2 + 2\int_0^t X_s dX_s + \int_0^t H_s^2 ds,\\
Y_t^2 &= Y_0^2 + 2\int_0^t Y_s dY_s + \int_0^t H_s'^2 ds.
\end{aligned}
$$

The conclusion is obtained by subtracting the last two equations from the first equation. ∎

Next, let $M(t)$ be the following Poisson stochastic integral:

$$
M^i(t) = M^i(0) + \int_0^t \int_A K^i(t,x)N(dt,dx), \tag{2.2.3}
$$

where $1 \le i \le d$, $t \ge 0$, A is bounded below and K^i is predictable.

Theorem 2.2.3. *Let M be the Poisson stochastic integral* (2.2.3)*, then for $f \in C(R^d)$ and any $t \ge 0$, one has in probability 1 that*

$$
f(M(t)) - f(M(0)) = \int_0^t \int_A [f(M(s-) + K(s,x)) - f(M(s-))]N(ds,dx).
$$

Proof. Let $Y(t) = \int_A xN(t,dx)$ and define the time of jump for Y as $T_0^A = 0$ and $T_n^A = \inf\{t > T_{n-1}^A : \Delta Y(t) \in A\}$ for all $n \in \mathbb{N}$. Then by definition of the stochastic integral, we have

$$
\begin{aligned}
f(M(t)) - f(M(0)) &= \sum_{0\le s\le t} f(M(s)) - f(M(s-))\\
&= \sum_{n=1}^{\infty} f(M(t\wedge T_n^A)) - f(M(t\wedge T_{n-1}^A))\\
&= \sum_{n=1}^{\infty} f\left(M(t\wedge T_n^A-) + K(t\wedge T_n^A, \Delta Y(t\wedge T_n^A))\right) - f(M(t\wedge T_{n-1}^A))\\
&= \int_0^t \int_A [f(M(s-) + K(s,x)) - f(M(s-))]N(ds,dx).
\end{aligned}
$$

The proof is completed. ∎

Next, consider the Leŷy-type stochastic integral

$$
Y^i(t) = Y^i(0) + Y_c^i(t) + \int_0^t \int_A K^i(s,x)N(ds,dx), \tag{2.2.4}
$$

where for any $1 \le i \le d$ and $t \ge 0$,

$$Y_c^i(t) = \int_0^t G^i(s)ds + \int_0^t F_j^i(s)dB^j(s).$$

Denote the quadratic variation process of Y_c^i by $\{[Y_c^i, Y_c^j](t)\}_{t\ge0}$, then

$$[Y_c^i, Y_c^j](t) = \sum_{k=1}^m \int_0^t F_k^i(s)F_k^j(s)ds.$$

Lemma 2.2.1. *If Y is a Lévy-type stochastic integral in the form* (2.2.4)*, then for any $f \in C^2(R^d)$ and $t \ge 0$, one has in probability 1 that*

$$f(Y(t)) - f(Y(0)) = \int_0^t \partial_i f(Y(s-))dY_c^i(s) + \frac{1}{2}\int_0^t \partial_i\partial_j f(Y(s-))d[Y_c^i, Y_c^j](s) + \int_0^t \int_A [f(Y(s-) + K(s,x)) - f(Y(s-))]N(ds, dx).$$

Proof. The proof can be referred to Ref. [7]. ■

Finally, let us consider the more general form of Lévy stochastic integral. Let Y satisfy

$$dY(t) = dY_c(t) + dY_d(t) = G(t)dt + F(t)dB(t) + \int_{|x|<1} H(t,x)\tilde{N}(dt, dx) + \int_{|x|\ge1} K(t,x)N(dt, dx), \tag{2.2.5}$$

where $1 \le i \le d$, $1 \le j \le m$, $t \ge 0$, $|G^i|^{\frac{1}{2}}, F_j^i \in \mathcal{P}_2(T)$, $H^i \in \mathcal{P}_2(T, E)$, K is predictable and $E = \hat{B}\backslash\{0\}$. Then the following Itô formula holds [7].

Theorem 2.2.4. *Let Y be the Lévy-type stochastic integral in the form* (2.2.5)*, then for any $f \in C^2(R^d)$ and $t \ge 0$, one has in probability 1 that*

$$f(Y(t)) - f(Y(0)) = \int_0^t \partial_i f(Y(s-))dY^i(s) + \frac{1}{2}\int_0^t \partial_i\partial_j f(Y(s-))d[Y_c^i, Y_c^j](s) + \sum_{0\le s\le t} [f(Y(s)) - f(Y(s-)) - \Delta Y^i(s)\partial_i f(Y(s-))].$$

In the final part, we consider the chain rule of the Stratonovitch integral defined at the beginning of Section 2.1.1. Let $M = \{M(t)\}_{t\ge0}$ with $M^i(t) = \int_0^t F_j^i(s)dB^j(s)$ and $G = (G^1, \cdots, G^d)$ such that $G_i F_j^i \in \mathcal{P}_2(t)$ for $1 \le j \le m, t \ge 0$. Then the *Stratonovitch integral* $\int_0^t G^i(s) \circ dM_i(s)$ is defined to be the limit in probability of the summation

$$\sum_{j=0}^{m(n)} \frac{G_i(t_{j+1}^{(n)}) + G_i(t_j^{(n)})}{2} [M^i(t_{j+1}^{(n)}) - M^i(t_j^{(n)})],$$

where $(\Pi_n, n \in \mathbb{N})$ is a partition of the interval $[0, T]$

$$\Pi_n = \{0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{m(n)}^{(n)} < t_{m(n)}^{(n)} = T\},$$

and $\lim_{n\to\infty} \max_{0\le j\le m(n)} |t_{j+1}^{(n)} - t_j^{(n)}| = 0$. Here, $\circ$ denotes that the integral is taken in the Stratonovitch sense. The Stratonovitch integral is related to the Itô integral via the formula

$$\int_0^t G^i(s) \circ dM_i(s) = \int_0^t G^i(s) dM_i(s) + \frac{1}{2}[G^i, M_i](t), \tag{2.2.6}$$

which in differential form can be written as

$$G^i(t) \circ dM_i(t) = G^i(t) dM_i(t) + \frac{1}{2} d[G^i, M_i](t). \tag{2.2.7}$$

The chain rule of Stratonovitch integral is stated in the following [7].

Theorem 2.2.5. *Let M be a Brownian integral and $f \in C^3(R^d)$, then for any $t \ge 0$, with probability 1 we have*

$$f(M(t)) - f(M(0)) = \int_0^t \partial_i f(M(s)) \circ dM^i(s). \tag{2.2.8}$$

Proof. By eq. (2.2.7), we have

$$\partial_i f(M(t)) \circ dM^i(t) = \partial_i f(M(t)) dM^i(t) + \frac{1}{2} d[\partial_i f(M(\cdot)), M^i](t)$$

and by Itô formula, for $1 \le i \le d$,

$$d\{\partial_i f(M(t))\} = \partial_j \partial_i f(M(t)) dM^j(t) + \frac{1}{2} \partial_j \partial_k \partial_i f(M(t)) d[M^j, M^k](t),$$

which gives

$$d[\partial_i f(M(\cdot)), M^i](t) = \partial_i \partial_j f(M(t)) d[M^i, M^j](t).$$

By using Itô formula again, we have

$$\begin{aligned}\int_0^t \partial_i f(M(s)) \circ dM^i(s) &= \partial_i f(M(s)) dM^i(s) + \frac{1}{2} \int_0^t \partial_i \partial_j f(M(s)) d[M^i, M^j](s) \\ &= f(M(t)) - f(M(0)).\end{aligned}$$

The proof is completed. ∎

2.3 The infinite-dimensional case

2.3.1 Q-Wiener process and the stochastic integral

Let H and U be two Hilbert spaces and $Q \in L(U)$ be a symmetric nonnegative operator. We first consider the case when $\mathrm{Tr}Q < \infty$. In this case, there exists a complete orthonormal system $\{e_k\}$ in U and nonnegative bounded sequence of nonnegative real numbers λ_k such that $Qe_k = \lambda_k e_k$ for all $k = 1, 2, \dots$.

Definition 2.3.1. *The U-valued stochastic process $W = \{W(t), t \geq 0\}$ is a Q-Wiener process if*
(1) $W(0) = 0$,
(2) *$W(t)$ has continuous trajectories,*
(3) *W has independent increments, and*
(4) $\mathscr{L}(W(t) - W(s)) = \mathcal{N}(0, (t-s)Q)$ *for all* $t \geq s \geq 0$.

If furthermore, for a given σ-algebra flow $\{\mathscr{F}_t\}_{t\geq 0}$, $W(t)$ is $\mathscr{F}_t$ measurable and $W(t+h) - W(t)$ is independent of $\mathscr{F}_t$ for all $h, t \geq 0$, then W is called a Q-Wiener process with respect to $\{\mathscr{F}_t\}_{t\geq 0}$.

Since W has independent increments, a Q-Wiener process W with $\mathrm{Tr}Q < \infty$ is necessarily a Gaussian process with zero mean and $\mathrm{Cov}[W(t)] = tQ$. Furthermore, there exists a sequence of real-valued, mutually independent standard Brownian motions $\beta_j(t)$, $j = 1, 2, \cdots$, on $(\Omega, \mathscr{F}, P)$ such that $W(t) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\beta_j(t)e_j$, which converges in $L^2(\Omega, P)$.

Fix $T > 0$ and let $L = L(U, H)$ be the space of bounded linear oprators. An L-valued stochastic process $\{\Phi(t)\}_{t\in[0,T]}$ is called an elementary process if there exists a sequence $0 = t_0 < t_1 < \cdots < t_k = T$ and a sequence of L-valued random variables $\Phi_0, \Phi_1, \cdots, \Phi_{k-1}$ such that Φ_m are $\mathscr{F}_{t_m}$ measurable and $\Phi(t) = \Phi_m$ for all $t \in (t_m, t_{m+1}]$ and $m = 0, 1, \cdots, k-1$. For such elementary process Φ, the stochastic integral with respect to W is defined naturally by

$$\mathcal{I}_W(\Phi) := \int_0^t \Phi(s)dW(s) = \sum_{m=0}^{k-1} \Phi_m(W_{t_{m+1}\wedge t} - W_{t_m\wedge t}). \tag{2.3.1}$$

Let $U_0 = Q^{1/2}(U)$. Then U_0 is a Hilbert space with the inner product

$$\langle f, g\rangle_0 = \sum_{k=1}^{\infty} \frac{1}{\lambda_k}\langle f, e_k\rangle\langle g, e_k\rangle = \langle Q^{-1/2}f, Q^{-1/2}g\rangle$$

and $\{g_j = \sqrt{\lambda_j}e_j\}$ is an orthonormal basis. Let $L_2^0 = L_2(U_0, H)$ be the space of Hilbert–Schmidt (HS) operators from U_0 to H with HS norm

$$\|\Psi\|_{L_2^0}^2 = \sum_{h,k=1}^{\infty} |\langle \Psi g_h, f_k\rangle|^2 = \sum_{h,k=1}^{\infty} \lambda_h |\langle \Psi e_h, f_k\rangle|^2 = \|\Psi Q^{1/2}\|^2 = \mathrm{Tr}[\Psi Q \Psi^*],$$

where $\{f_j\}$ is a standard orthonormal basis of H. Let $\Phi = \{\Phi(t)\}_{t\in[0,T]}$ be a measurable L_2^0-valued process and define

$$|||\Phi|||_t = \left\{ E\int_0^t \|\Phi(s)\|_{L_2^0}^2 ds \right\}^{1/2} \quad \forall t \in [0,T]. \tag{2.3.2}$$

Similar to the proof of Lemma 2.1.1, we can prove

Proposition 2.3.1 (Isometry). *Let Φ be an elementary process and $|||\Phi|||_T < \infty$, then the integral process $\mathcal{I}_W(\Phi)$ is a continuous square-integrable H-valued martingale on* $[0,T]$ *and*

$$E\left|\int_0^t \Phi(s)dW(s)\right|^2 = |||\Phi|||_t^2, \quad 0 \le t \le T. \tag{2.3.3}$$

This shows that the stochastic integral is an isometry from the space of all elementary processes equipped with norm $|||\cdot|||_T$ to the space $\mathscr{M}_T^2(H)$ of H-valued square-integrable martingale. Then our task is to extend the integral into space of general L_2^0-valued predictable process Φ such that $|||\Phi|||_T < \infty$. We have the following result.

Proposition 2.3.2. *Let Φ be an L_2^0-valued predictable process and $|||\Phi|||_T < \infty$, then there exists a sequence of elementary processes $\{\Phi_n\}$ such that $|||\Phi - \Phi_n|||_T \to 0$ as $n \to \infty$.*

The proof can be found in Ref. [75]. Note that all the L_2^0-valued predictable process with $|||\Phi|||_T < \infty$ form a Hilbert space, denoted by $\mathscr{N}_W^2(0,T;L_2^0)$ or $\mathscr{N}_W^2(0,T)$ or even $\mathscr{N}_W^2$ for short. Since the set of elementary processes is dense in $\mathscr{N}_W^2$, the definition can be extended to all elements in $\mathscr{N}_W^2$ and moreover $\mathcal{I}_W(\Phi)$ is a continuous square-integrable martingale.

Similar to the Itô integral in Section 2.1.2, the stochastic integral here defined can be extended to the L_2^0-valued predictable processes satisfying more weaker condition

$$P\left\{ \int_0^t \|\Phi(s)\|_{L_2^0}^2 ds < \infty \right\} = 1. \tag{2.3.4}$$

All such processes form a linear space denoted by $\mathscr{N}_W(0,T;L_2^0)$, or $\mathscr{N}_W(0,T)$ or $\mathscr{N}_W$.

To extend the stochastic integral, we use the stopping time technique. First, we note that if $\Phi \in \mathscr{N}_W^2(0,T;L_2^0)$ and τ be an $\mathscr{F}_t$-stopping time such that $P(\tau \le T) = 1$, then P-a.s.

$$\int_0^t I_{[0,\tau]}(s)\Phi(s)dW(s) = \mathcal{I}_W(\Phi)(\tau \wedge t) \quad \forall t \in [0,T]. \tag{2.3.5}$$

Define for $\mathscr{N}_W(0,T;L_2^0)$ the stopping time

$$\tau_n = \inf\left\{t \in [0,T] : \int_0^t \|\Phi(s)\|_{L_2^0}^2 ds \geq n\right\}, \quad \inf \emptyset = T.$$

Then $I_{[0,\tau_n)}\Phi \in \mathscr{N}_W^2(0,T;L_2^0)$ and $\mathcal{I}_W(I_{[0,\tau_n)}\Phi)$ is well defined. $\mathcal{I}_W(\Phi)$ is defined as the limit of this sequence in probability. Moreover,

$$\mathcal{I}_W(\Phi)(\tau_n \wedge t) = \mathcal{I}_W(I_{[0,\tau_n]}\Phi)(\tau_n \wedge t) = M_n(\tau_n \wedge t), \quad t \in [0,T], n = 1,2,\cdots,$$

where M_n is a H-valued continuous square-integrable martingale. This is referred to the *local martingale property* of stochastic integral.

Next, consider the cylindrical Wiener process. First of all, we are given a filtered probability space $(\Omega, \mathscr{F}, \{\mathcal{F}_t\}, P)$.

Definition 2.3.2. *Let $W : [0,\infty) \times U \to L^2(\Omega, \mathscr{F}, P)$ be a stochastic process on U such that (i) $E|W(t,x)|^2 = t|x|_U^2$ for all $t \geq 0$ and $x \in U$, and (ii) $\{W(t,x)\}_{t\geq 0}$ is a real-valued $\{\mathscr{F}_t\}$-adapted Wiener process for any $x \in U$, then W is called an $\mathscr{F}_t$-adapted cylindrical Wiener process on U.*

As a consequence of this definition, we have for all $t, s \geq 0$ and $x, y \in U$, that $E[W(t,x)W(s,y)] = (t\wedge s)\langle x,y\rangle_U$. This also implies that $E[W(t,e_n)W(s,e_n)] = t\wedge s$, where $\{e_n\}_{n\geq 1}$ is an orthonormal basis of U. In particular, if we set $\beta_n(t) := W(t,e_n)$, then $\beta_n(\cdot)$ is a sequence of independent standard real valued Wiener process. Furthermore, let H be another Hilbert space such that embedding $U \to H$ is HS, then $W(t) = \sum_n \beta_n(t)e_n$ converges in $L^2(\Omega, \mathscr{F}, P; H)$ for $t \geq 0$.

We will also call this process W a Q-Wiener process on U. Obviously, $Q = I$ (an infinite-dimensional identity matrix) and $\mathrm{Tr}Q = \infty$. Now, we consider the case of extended integral to $\mathrm{Tr}Q \leq \infty$. Let $W_N(t) = \sum_{j=1}^N \beta_j(t)e_j$ for all $t \in [0,T]$, where $\{e_j\}$ is an orthonormal basis of U. Suppose Φ is stochastically integrable with respect to the Q-Wiener process. It is easy to see that $\mathcal{I}_W(\Phi) = \mathcal{I}_{W_N}(\Phi) + \mathcal{I}_{W^N}(\Phi)$ and hence

$$E\|\mathcal{I}_W(\Phi)(T) - \mathcal{I}_{W_N}(\Phi)(T)\|^2 = E\int_0^T \|\Phi(s)(Q^N)^{1/2}\|_{L_2^0}^2 ds.$$

If $|||\Phi|||_T < \infty$, then as $N \to \infty$, we have

$$E\int_0^T \|\Phi(s)(Q^N)^{1/2}\|_{L_2^0}^2 ds \to 0.$$

By the martingale property of stochastic integral, we have

$$E \sup_{0\leq t\leq T} \|\Phi \cdot W(t) - \Phi \cdot W_N(t)\| \to 0, \quad N \to \infty,$$

from which we can choose a subsequence $\{\mathcal{I}_{W_{N_k}}(\Phi)\}$ converging P-a.s. and uniformly in $[0, T]$. Therefore, by taking limit of stochastic integrals w.r.t. a finite-dimensional Wiener process, one obtains the stochastic integral with respect to an infinite-dimensional (possible cylindrical) Wiener process, which is independent of the choice of the subsequence. The definition of the stochastic integral for $\Phi \in \mathcal{N}_W(0, T; L_2^0)$ can be obtained similarly by the localization method above. For general cylindrical Q-Wiener process W, the stochastic integral can also be defined via the limit process and we omit the details here.

First, we summarize the above discussion in the following theorem.

Theorem 2.3.1. *Let $\Phi \in \mathcal{N}_W^2(0, T; L_2^0)$, then the stochastic integral $\mathcal{I}_W(\Phi)$ is a continuous square-integrable martingale with quadratic variation*

$$\langle\langle \mathcal{I}_W(\Phi)(t)\rangle\rangle = \int_0^t (\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^* ds, \quad s, t \in [0, T]. \tag{2.3.6}$$

If $\Phi \in \mathcal{N}_W(0, T; L_2^0)$, then $\mathcal{I}_W(\Phi)$ is a local martingale.

The following propositions are given directly here without proof. The readers may refer to Ref. [75].

Proposition 2.3.3. *Let $\Phi_1, \Phi_2 \in \mathcal{N}_W^2(0, T; L_2^0)$, then*

$$E\mathcal{I}_W(\Phi_i) = 0, \quad E\|\mathcal{I}_W(\Phi_i)\|^2 < \infty, \quad t \in [0, T], i = 1, 2.$$

The correlation operator $V(t, s) = Cor(\mathcal{I}_W(\Phi_1), \mathcal{I}_W(\Phi_2))$ for all $t, s \in [0, T]$ is given by

$$V(t, s) = E\int_0^{t\wedge s} (\Phi_2(r)Q^{1/2})(\Phi_1(r)Q^{1/2})^* dr. \tag{2.3.7}$$

As a corollary, one has directly that

$$E\langle \mathcal{I}_W(\Phi_1), \mathcal{I}_W(\Phi_2)\rangle = E\int_0^{t\wedge s} \mathrm{Tr}[(\Phi_2(r)Q^{1/2})(\Phi_1(r)Q^{1/2})^*]dr. \tag{2.3.8}$$

Furthermore, if Φ_1, Φ_2 are $L(U, H)$-valued processes, this can be simplified as

$$E\langle \mathcal{I}_W(\Phi_1), \mathcal{I}_W(\Phi_2)\rangle = E\int_0^{t\wedge s} \mathrm{Tr}[(\Phi_2(r)Q\Phi_1(r)^*]dr.$$

Proposition 2.3.4. *Let $\Phi \in \mathcal{N}_W(0, T; L_2^0)$, then for any $a, b > 0$, there holds*

$$P\left(\sup_{t\in[0,T]} |\mathcal{I}_W(\Phi)(t)| > a\right) \le \frac{b}{a^2} + P\left(\int_0^T \|\Phi(t)\|_{L_2^0}^2 dt > b\right). \tag{2.3.9}$$

2.3.2 Itô formula

Now, let us discuss the Itô formula of stochastic integral. Assume that W is a U-valued Q-Wiener process. Let Φ be an L_2^0-valued process stochastically integrable in $[0,T]$, ϕ an H-valued predictable process Bochner integrable on $[0,T]$, P-a.s., and $X(0)$ an $\mathscr{F}_0$-measurable H-valued random variable. Then the process

$$X(t) = X(0) + \int_0^t \phi(s)ds + \int_0^t \Phi(s)dW(s), \quad t \in [0,T]$$

is well defined.

Theorem 2.3.2. *Let $F : [0,T] \times H \to R^1$ and its partial derivatives T_t, F_x and F_{xx} be uniformly continuous on bounded subsets of $[0,T] \times H$, then P-a.s., for any $t \in [0,T]$*

$$\begin{aligned} dF(t,X(t)) =& \langle F_x(t,X(t)), \Phi(t)dW(t)\rangle + \left\{F_t(t,X(t)) + \langle F_x(t,X(t)), \phi(t)\rangle\right\} dt \\ &+ \frac{1}{2} Tr\left[F_{xx}(t,X(t))(\Phi(t)Q^{1/2})(\Phi(t)Q^{1/2})^*\right] dt. \end{aligned}$$

This result can be proved first for elementary processes and then via a limiting process it is extended to a general stochastically integral process. For applications, we will consider the Burkholder-Davis-Gundy-type inequality for the stochastic integrals. First of all, let's state the martingale inequality without proof, see Theorem 1.3.3.

Proposition 2.3.5. *Let $E\int_0^T \|\Phi(s)\|^2_{L_2^0} ds < \infty$, then*

(1) *for any $p \geq 1$ and $\lambda > 0$,*

$$P\left(\sup_{t\leq T}\left|\int_0^t \Phi(s)dW(s)\right| \geq \lambda\right) \leq \frac{1}{\lambda^p} E|\int_0^T \Phi(s)dW(s)|^p,$$

(2) *for any $p > 1$,*

$$E\left(\sup_{t\leq T}\left|\int_0^t \Phi(s)dW(s)\right|^p\right) \leq \frac{p}{p-1} E|\int_0^T \Phi(s)dW(s)|^p.$$

Proposition 2.3.6. *For any $r \geq 1$ and any L_2^0-valued predictable process $\Phi(t)$, $t \in [0,T]$, there holds*

$$E\left(\sup_{s\in[0,t]}\left|\int_0^s \Phi(\sigma)dW(\sigma)\right|^{2r}\right) \leq c_r\, E\left(\int_0^t \|\Phi(\sigma)\|^2_{L_2^0} ds\right)^r. \tag{2.3.10}$$

Proof. When $r = 1$, the results follows from Itô's isometry. Let $r > 1$ and set $Z(t) = \int_0^t \Phi(\sigma)dW(\sigma)$, $f(x) = |x|^{2r}$, then

$$f_{xx}(x) = 4r(r-1)|x|^{2(r-2)}x \otimes x + 2r|x|^{2(r-1)}I,$$

and

$$\|f_{xx}(x)\| \le 2r(2r-1)|x|^{2(r-1)}.$$

This gives that

$$|\mathrm{Tr}\Phi^*(t)f_{xx}(Z(t))\Phi(t)Q| \le 2r(2r-1)|Z(t)|^{2(r-1)}\|\Phi(t)\|^2_{L_2^0}.$$

By Itô formula and the martingale property of the stochastic integral, we obtain

$$\begin{aligned} E|Z(t)|^{2r} &\le r(2r-1)E(\int_0^t |Z(s)|^{2(r-1)}\|\Phi(\sigma)\|^2_{L_2^0}d\sigma) \\ &\le r(2r-1)E(\sup_{s\in[0,t]}|Z(s)|^{2(r-1)}\int_0^t \|\Phi(\sigma)\|^2_{L_2^0}d\sigma). \end{aligned}$$

By Hölder inequality ($p = r/(r-1)$) and martingale inequality, one has

$$\begin{aligned} E|Z(t)|^{2r} \le& r(2r-1)\left[E\left(\sup_{s\in[0,t]}|Z(s)|^{2(r-1)p}\right)\right]^{1/p} \cdot \left[E\left(\int_0^t \|\Phi(\sigma)\|^2_{L_2^0}d\sigma\right)^r\right]^{1/r} \\ \le& c_r'\left[E|Z(t)|^{2r}\right]^{1-\frac{1}{r}} \cdot \left[E\left(\int_0^t \|\Phi(\sigma)\|^2_{L_2^0}ds\right)^r\right]^{\frac{1}{r}}. \end{aligned}$$

One completes the proof by applying once more the martingale inequality. ∎

Proposition 2.3.7. *For any $r \ge 1$, and any L_2^0-predictable process $\Phi(\cdot)$, it holds*

$$\sup_{s\in[0,t]} E\left|\int_0^s \Phi(\sigma)dW(\sigma)\right|^{2r} \le (r(2r-1))^r\left(\int_0^t \left(E\|\Phi(s)\|^{2r}_{L_2^0}\right)^{1/r} ds\right)^r.$$

Proof. The case $r = 1$ is straightforward. Let $r > 1$ and $Z(t)$ be as above, then

$$E|Z(t)|^{2r} \le r(2r-1)E\left(\int_0^t |Z(s)|^{2(r-1)}\|\Phi(s)\|^2_{L_2^0}ds\right).$$

Since from Hölder inequality that

$$E\left(|Z(s)|^{2(r-1)}\|\Phi(s)\|^2_{L_2^0}\right) \le \left(E|Z(s)|^{2r}\right)^{(r-1)/r}\left(E\|\Phi(s)\|^{2r}_{L_2^0}\right)^{1/r},$$

one has

$$
\begin{aligned}
E|Z(t)|^{2r} &\leq r(2r-1)\int_0^t \left(E|Z(s)|^{2r}\right)^{(r-1)/r}\left(E\|\Phi(s)\|_{L_2^0}^{2r}\right)^{1/r} ds \\
&\leq r(2r-1)\int_0^t \left(\sup_{u\in[0,s]} E|Z(u)|^{2r}\right)^{(r-1)/r}\left(E\|\Phi(s)\|_{L_2^0}^{2r}\right)^{1/r} ds.
\end{aligned}
$$

The right-hand side is monotonic in t, hence

$$
\begin{aligned}
\sup_{s\in[0,t]} E|Z(s)|^{2r} &\leq r(2r-1)\int_0^t \left(\sup_{u\in[0,s]} E|Z(u)|^{2r}\right)^{(r-1)/r}\left(E\|\Phi(s)\|_{L_2^0}^{2r}\right)^{1/r} ds \\
&\leq r(2r-1)\left(\sup_{s\in[0,t]} E|Z(s)|^{2r}\right)^{(r-1)/r}\int_0^t \left(E\|\Phi(s)\|_{L_2^0}^{2r}\right)^{1/r} ds.
\end{aligned}
$$

Hence, the proposition holds. ■

The following proposition concerns that Fubini theorem in stochastic integrals. Let $(E,\mathscr{E})$ be a measurable space, and Φ is a measurable mapping from $(\Omega_T\times E,\mathscr{P}_T\times\mathscr{B}(E))$ to $(L_2^0,\mathscr{B}(L_2^0))$.

Theorem 2.3.3. *Let μ be a finite-positive measure on $(E,\mathscr{E})$ and $\int_E |||\Phi(\cdot,\cdot,x)|||_T\mu(dx) < +\infty$, then P-a.s.*

$$
\int_E\left[\int_0^T \Phi(t,x)dW(t)\right]\mu(dx) = \int_0^T\left[\int_E \Phi(t,x)\mu(dx)\right]dW(t).
$$

The proof of the theorem can be referred to literature [75].

2.4 Nuclear operator and HS operator

In this section, we introduce some basic concepts of HS and nuclear operators. Let E, G be Banach spaces and $L(E,G)$ be the Banach space of all bounded linear operators from E to G equipped with the operator norm. We denote by E^* and G^* the dual spaces of E and G, respectively. A linear operator $T \in L(E,G)$ is called *nuclear* if it can be represented in the form

$$
Tx = \sum_{j=1}^{\infty} a_j\varphi_j(x) \quad \forall x \in E,
$$

where $\{a_j\}\subset G$ and $\{\varphi_j\}\subset E^*$ are such that $\sum_{j=1}^{\infty}\|a_j\|\cdot\|\varphi_j\| < +\infty$. We denote by $L_1(E,G)$ the space of all nuclear operator from E to G, which is a separable Banach space with the nuclear norm

$$\|T\|_1 = \inf\left\{\sum_{j=1}^{\infty}\|a_j\|\cdot\|\varphi_j\| : x = \sum_{j=1}^{\infty}a_j\varphi_j(x)\right\}.$$

When $G = E$, we write $L_1(E)$ instead of $L_1(E,E)$ for short. Let K be a Banach space, it's obvious that if $T \in L_1(E,G)$ and $S \in L(G,K)$ then $ST \in L_1(E,K)$ and $\|ST\|_1 \le \|S\|\|T\|_1$. Let H be a separable Hilbert space and $\{e_j\}$ be an orthonormal basis of H, then the trace of $T \in L_1(H)$ can be defined as $\mathrm{Tr}T = \sum_{j=1}^{\infty}\langle Te_j, e_j\rangle$. It can be shown that the trace is well defined for $T \in L_1(H)$ and is independent of the choice of the orthonormal basis $\{e_j\}$. Furthermore, for $T \in L_1(H)$, there holds $|\mathrm{Tr}T| \le \|T\|_1$. Hence, if $T \in L_1(H)$ and $S \in L(H)$, then $TS, ST \in L_1(H)$ and $\mathrm{Tr}TS = \mathrm{Tr}ST \le \|T\|_1\|S\|$.

Let E, F be two separable Hilbert spaces. A linear operator $T \in L(E,F)$ is called a HS operator if $\sum_{k=1}^{\infty}|Te_k|^2 < \infty$. It can be shown that the definition of HS operator is independent of the choice of basis $\{e_k\}$. The space of all HS operators $L_2(E,F)$ is a separable Hilbert space with the scalar product

$$\langle S, T\rangle_2 = \sum_{k=1}^{\infty}\langle Se_k, Te_k\rangle, \quad S, T \in L_2(E,F).$$

Denote by $\|T\|_2 = \langle T, T\rangle_2^{1/2}$ the corresponding HS norm. If $E = F$, we write $L_2(E,E) = L_2(E)$ for short. For any $b \in E$, $a, h \in F$, if we define $(b\otimes a)\cdot h = b\langle a, h\rangle$, then $\{f_j \otimes e_k\}_{j,k\in\mathbb{N}}$ constructs a set of complete orthonormal basis of $L_2(E,F)$.

Proposition 2.4.1. *Let E, H be two separable Hilbert spaces and $K(E,H)$ be the space of all compact operator from E to H, then $L_1(E,H) \subset L_2(E,H) \subset K(E,H)$.*

Proof. Let $T \in L_1(E,H)$, then $Tx = \sum_{j=1}^{\infty}a_j\varphi_j(x)$ for some $\{a_j\} \subset H$ and $\{\varphi_j\} \subset E^*$ such that $\sum_j \|a_j\|_H\|\varphi_j\|_{E^*} < \infty$. Let $\{e_k\}$ be an orthonormal basis of E, then

$$\begin{aligned}\sum_k |Te_k|_H^2 &= \sum_k \Big|\sum_n a_n\varphi_n(e_k)\Big|_H^2 \le \sum_k\sum_{n,l}|\langle a_n, a_l\rangle_H||\varphi_n(e_k)||\varphi_l(e_k)| \\ &\le \sum_{n,l}|a_n||a_l|\left(\sum_k \varphi_n^2(e_k)\right)^{1/2}\left(\sum_k \varphi_l^2(e_k)\right)^{1/2} \le \left(\sum_n |a_n||\varphi_n|\right)^2,\end{aligned}$$

which implies the first conclusion. To show the second one, we note that the HS norm is stronger than the operator norm, $K(E,H)$ is a closed set of $L(E,H)$ and each $T \in L_2(E,H)$ can be approximated in $L_2(E,H)$ by the sequence of finite rank operators $T_n = \sum_{k\le n} f_k \otimes T^*f_k$ (i.e., $T_nx = \sum_{k\le n}\langle Tx, f_k\rangle f_k$, $x \in E$), where $\{f_k\}$ is an orthonormal basis of H. Since finite rank operators are compact, the result follows. ∎

Proposition 2.4.2. *A nonnegative operator $T \in L(H)$ is nuclear if and only if for an orthonormal basis $\{e_k\}$ on H such that $\sum_{j=1}^{\infty}\langle Te_j, e_j\rangle < +\infty$. Moreover, in this case, $TrT = \|T\|_1$.*

Proof. We need only to show the sufficiency part, since $|\mathrm{Tr}T| \le \|T\|_1$. From Proposition 2.4.1, we know that T is compact. Let $\{\lambda_j, f_j\}$ be a sequence of eigenvalues and eigenvectors of T, then $Tx = \sum_{k=1}^{\infty} \lambda_k \langle x, f_k\rangle f_k$ for all $x \in H$. It follows that

$$\sum_{j=1}^{\infty}\langle Te_j, e_j\rangle = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\lambda_k|\langle f_j, e_k\rangle|^2 = \sum_{k=1}^{\infty}\lambda_k < +\infty.$$

Therefore, T is nuclear and $\mathrm{Tr}T = \sum_{k=1}^{\infty}\lambda_k$. In this case, $\mathrm{Tr}T = \|T\|_1$. ∎

Proposition 2.4.3. *Let E, F, G be the separable Hilbert spaces. If $T \in L_2(E,F), S \in L_2(F,G)$, then $ST \in L_1(E,G)$ and $\|ST\|_1 \le \|S\|_2\|T\|_2$.*

Proof. Let $\{f_k\}$ be an orthonormal basis of F, then

$$Tx = \sum_k \langle Tx, f_k\rangle f_k = \sum_k \langle x, T^* f_k\rangle f_k.$$

Hence, $STx = \sum_{k=1}^{\infty}\langle x, T^* f_k\rangle Sf_k$ for all $x \in E$ and by definition

$$\|ST\|_1 \le \sum_{k=1}^{\infty}|T^* f_k||Sf_k| \le \left(\sum_{k=1}^{\infty}|T^* f_k|^2\right)^{1/2}\left(\sum_{k=1}^{\infty}|Sf_k|^2\right)^{1/2}.$$

This completes the proof. ∎

3 OU processes and SDEs

In this chapter, we first introduce the Ornstein–Uhlenbeck (OU) processes and the stochastic convolution. Then we introduce some basic concepts of solutions to linear stochastic differential equations (SDE) and their properties. Finally, we consider the existence of solutions for a general nonlinear SDE based on fixed point theorem. Most of the materials come from Refs [75, 76].

3.1 Ornstein–Uhlenbeck processes

Consider the following stochastic partial differential equation:

$$dX_t = AX_t ds + BdW_t, \quad X_0 = \xi. \tag{3.1.1}$$

Suppose U, H are two Hilbert spaces, A is the infinitesimal generator [199] of a C_0 semigroup $\{S(t)\}_{t\geq 0}$ in H, $B \in L(U, H)$ is a bounded linear operator and ξ is a $\mathscr{F}_0$-measurable random variable in H. Let $\{W(t)\}_{t\geq 0}$ be a Q-Wiener process in probability space $(\Omega, \mathscr{F}, P)$ with covariance operator $Q \in L(U)$. If $\mathrm{Tr}Q < +\infty$, W is called a Q-Wiener process and when $\mathrm{Tr}Q = \infty$, it is called a cylindrical Wiener process. We also suppose that there is a set of complete orthonormal basis $\{e_k\}$ of U and a bounded nonnegative real sequence λ_k such that $Qe_k = \lambda_k e_k$ for $k = 1, 2, \cdots$, and a sequence of real-valued independent Brownian motions β_k such that

$$\langle W(t), u\rangle = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\langle e_k, u\rangle \beta_k, \quad u \in U, t \geq 0.$$

The noise is additive if B does not depend on X_t and multiplicative otherwise. We note that if W is a Q-Wiener process in U, then $W_1 = BW$ is a BQB^*-Wiener process in H. So without loss of generality we assume $U = H$.

If the H-valued $\mathscr{F}_t$-adapted stochastic process $X = \{X_t\}_{t\geq 0}$ is the solution of the equation, then the solution can be given by the Duhamel principle

$$X(t) = S(t)\xi + \int_0^t S(t-s)BdW(s), \quad t \geq 0, \tag{3.1.2}$$

where $S(t)$ is the solution operator generated by A. The process X_t is called an Ornstein–Uhlenbeck process and the stochastic integral

$$W_A(t) = \int_0^t S(t-s)BdW(s), \quad t \geq 0, \tag{3.1.3}$$

is called a stochastic convolution. Generally, if $\Phi(t) \in L_2(U,H), t \in [0,T]$ is a Hilber–Schmidt operator-valued adapted process such that the stochastic integral

$$W_A^{\Phi}(t) = \int_0^t S(t-s)\Phi(s)dW(s), \quad t \in [0,T]$$

is well defined, then W_A^{Φ} is also called the stochastic convolution. In particular, if $\Phi = B, t \in [0,T]$, then $W_A^{\Phi} = W_A$.

The following introduces some properties of the stochastic convolution. First of all, it can be proved that if $\|S(r)B\|_{L_2^0} \in L^2(0,T)$, then $\{W_A(t)\}_{t\geq 0}$ is a Guassian process, which is continuous in mean square and has a predictable version, and $\mathrm{Cov}W_A(t) = \int_0^t S(r)BQB^*S^*(r)dr$. Indeed, for fixed $0 \leq s \leq t \leq T$, then

$$W_A(t) - W_A(s) = \int_s^t S(t-r)BdW(r) + \int_0^s [S(t-r) - S(s-r)]BdW(r).$$

It then follows from the independence of these two integrals that

$$\begin{aligned} E|W_A(t) - W_A(s)|^2 = &\sum_{k=1}^{\infty} \lambda_k \int_0^{t-s} |S(r)Be_k|^2 dr \\ &+ \sum_{k=1}^{\infty} \lambda_k \int_0^s |(S(t-s+r) - S(r))Be_k|^2 dr. \end{aligned}$$

Then the continuity in mean square follows from the dominated convergence theorem and the Gaussian property follows from the properties of the stochastic integral.

Let $\alpha \in (0,1]$ and

$$Y_\alpha^{\Phi}(t) = \int_0^t (t-s)^{\alpha} S(t-s)\Phi(s)dW(s), \quad t \in [0,T]. \tag{3.1.4}$$

Then by stochastic Fubini theorem, it can be shown that if

$$\int_0^t (t-s)^{\alpha-1} \left[\int_0^s (s-\sigma)^{-2\alpha} E(\|S(t-\sigma)\Phi(\sigma)\|_2^2)d\sigma\right]^{1/2} ds < +\infty \tag{3.1.5}$$

holds for all $t \in [0,T]$, then

$$\int_0^t S(t-s)\Phi(s)dW(s) = \frac{\sin \alpha\pi}{\pi} \int_0^t (t-s)^{\alpha-1} S(t-s) Y_\alpha^{\Phi}(s)ds, \quad t \in [0,T]. \tag{3.1.6}$$

Indeed, the integral on the right-hand side is equal to

$$\begin{aligned}\int_0^t (t-s)^{\alpha-1}S(t-s)\int_0^s (s-\sigma)^{-\alpha}S(s-\sigma)\Phi(\sigma)dW(\sigma)ds \\ = \int_0^t\int_0^s (t-s)^{\alpha-1}(s-\sigma)^{-\alpha}S(t-\sigma)\Phi(\sigma)dW(\sigma)ds \\ = \int_0^t\int_\sigma^t (t-s)^{\alpha-1}(s-\sigma)^{-\alpha}dsS(t-\sigma)\Phi(\sigma)dW(\sigma) \\ = \frac{\pi}{\sin\alpha\pi}\int_0^t S(t-\sigma)\Phi(\sigma)dW(\sigma).\end{aligned}$$

The Hölder continuity of the stochastic convolution can then be proved. Assume that $S(t)$ be an analytic semigroup on H generated by A such that $\|S(t)\| \le Me^{-\omega t}$ for some positive constants M and ω and that there exists $\alpha \in (0,1/2)$ such that $\int_0^T t^{-2\alpha}\|S(t)B\|_{\mathrm{HS}}^2 dt < \infty$ for any $T > 0$, then for any $\gamma \in [0,\alpha)$, there is a $D((-A)^\gamma)$-valued Hölder-continuous modification of $W_A(t)$, with exponent smaller than $\alpha - \gamma$.

To show this, we first define for any $\alpha \in (0,1), \gamma \in [0,\alpha), p > 1$ the integral operator on the space $L^p(0,T;H)$

$$R_{\alpha,\gamma}\varphi(t)\int_0^t (t-\sigma)^{\alpha-1}(-A)^\gamma S(t-\sigma)\varphi(\sigma)d\sigma,\ t \in [0,T].$$

It can be shown that from Ref. [76, appendix, Prop. A.1.1.] that when $\gamma > 0, \alpha > \gamma + \frac{1}{p}$, then $R_{\alpha,\gamma}$ is a bounded operator from $L^p(0,T;H)$ to $C^{\alpha-\gamma-\frac{1}{p}}([0,T];D((-A)^\gamma))$ and when $\gamma = 0, \alpha > 1/p$, then for any $\delta \in (0,\alpha-\frac{1}{p})$, $R_{\alpha,\gamma}$ is a bounded operator from $L^p(0,T;H)$ to $C^\delta([0,T];H)$. Thanks to eq. (3.1.6), it then suffices to show that the process

$$Y_\alpha(s) = \int_0^s (s-\sigma)^{-\alpha}S(s-\sigma)BdW(\sigma),\ s \in [0,T]$$

has p-integrable trajectories. But this is obvious from the Burkholder–Davis–Gundy inequality

$$\begin{aligned}E\int_0^T |Y_\alpha(s)|^p ds &\le c_p\int_0^T \left(\int_0^s \|(s-\sigma)^{-\alpha}S(s-\sigma)B\|_2^2 d\sigma\right)^{p/2} ds \\ &\le c_p\int_0^T \left(\int_0^s \sigma^{-2\alpha}\|S(\sigma)B\|_2^2 d\sigma\right)^{p/2} ds < +\infty.\end{aligned}$$

Let us remark that similar results hold for $W_A^\Phi(t)$. Indeed, let $S(t), t \ge 0$ be an analytical semigroup on H generated by A such that $\|S(t)\| \le Me^{-\omega t}$ for all $t \ge 0$ and for some positive constants ω and M. Assume that there exists some $\alpha \in (0,1/2)$ and $C > 0$ such that $\int_0^T t^{-2\alpha}\|S(t)\|_{\mathrm{HS}}^2 dt < \infty$ for any $T > 0$ and P-a.s. $\|\Phi(t)\| \le C$ for all $t \ge 0$, then for any $\gamma \in [0,\alpha)$, there is a $D((-A)^\gamma)$-valued Hölder-continuous modification of W_A^Φ, with Hölder exponent smaller than $\alpha - \gamma$.

In the analytical case, if furthermore $\mathrm{Tr}Q < +\infty$, then one can show that for any $\alpha \in (0, 1/2)$, the process $W_A(\cdot)$ has α-Hölder-continuous trajectories and $(-A)^\alpha W_A(\cdot)$ is y-Hölder continuous for $y \in (0, \frac{1}{2} - \alpha)$ (see Ref. [75] for proof). In the critical case when $\alpha = 1/2$, the result does not hold. Indeed, an example given in Ref. [72] shows that if A is negative self-adjoint, then $(-A)^{1/2} W_A(t)$ may not have a pathwise-continuous version.

Finally, we show that when A generates a contraction semigroup and $\Phi \in \mathcal{N}_W^2(0, T; L_2^0)$, then W_A^Φ has continuous modifications and there is a constant C such that

$$E \sup_{s\in[0,t]} |W_A^\Phi(s)|^2 \le CE \int_0^t \|\Phi(s)\|_{L_2^0}^2 ds, \quad t \in [0, T]. \tag{3.1.7}$$

We now sketch the proof of this inequality. Let $X = W_A^\Phi$ for short, then X satisfies

$$dX(t) = AX(t)dt + \Phi(t)dW(t)$$

on the interval $[0, T]$. By Itô formula,

$$\begin{aligned} |X(t)|^2 &= 2\int_0^t \langle X(s), AX(s)\rangle ds + 2\int_0^t \langle X(s), \Phi(s)dW(s)\rangle + \int_0^t \|\Phi(s)\|_{L_2^0}^2 dt \\ &\le 2\int_0^t \langle X(s), \Phi(s)dW(s)\rangle + \int_0^t \|\Phi(s)\|_{L_2^0}^2 ds, \end{aligned}$$

since A generates a contraction semigroup. By taking supremum and then expectation, one has

$$E \sup_{s\in[0,t]} |X(s)|^2 \le 2E \sup_{s\in[0,t]} \left| \int_0^s \langle X(\sigma), \Phi(\sigma)dW(\sigma)\rangle \right| + E\int_0^t \|\Phi(s)\|_{L_2^0}^2 ds. \tag{3.1.8}$$

Let $M(t) = \int_0^t \langle X(s), \Phi(s)dW(s)\rangle$. Then from the martingale inequality

$$E \sup_{s\in[0,t]} |M(s)| \le 3E\left[\left(\int_0^t |(\Phi(s)Q^{1/2})^* X(s)|^2 ds \right)^{1/2} \right], \tag{3.1.9}$$

the first part of eq. (3.1.8) can be bounded by

$$\begin{aligned} 6E\left(\int_0^t |(\Phi(s)Q^{1/2})^* X(s)|^2 ds \right)^{1/2} &\le 6E\left(\int_0^t \|\Phi(s)\|_{L_2^0}^2 |X(s)|^2 ds \right)^{1/2} \\ &\le 6(E \sup_{s\in[0,t]} |X(s)|^2)^{1/2} \left(E\int_0^t \|\Phi(s)\|_{L_2^0}^2 ds \right)^{1/2}. \end{aligned}$$

It then follows that for any $\varepsilon > 0$,

$$E\sup_{s\in[0,t]}|X(s)|^2 \le 12\Big(\varepsilon E\sup_{s\in[0,t]}|X(s)|^2 + \frac{1}{\varepsilon}E\int_0^t \|\Phi(s)\|^2_{L_2^0}ds\Big) + E\int_0^t \|\Phi(s)\|^2_{L_2^0}ds.$$

Taking $\varepsilon = 1/24$, we get eq. (3.1.7). The above proof can be made rigorous by standard approximation arguments.

3.2 Linear SDEs

Consider the following linear stochastic differential equation:

$$dX(t) = [AX(t) + f(t)]dt + BdW(t), \quad X(0) = \xi, \tag{3.2.1}$$

where $A : D(A) \subset H \to H$, $B : U \to H$ are linear operators and f is an H-valued stochastic process. We assume A generates a strongly continuous semigroup $S(\cdot)$ on H, $B \in L(U, H)$, f is a predictable process which is Bochner integrable on any finite interval $[0, T]$ and finally the initial data ξ is $\mathscr{F}_0$ measurable.

Definition 3.2.1. *Under the above assumptions, we say that an H-valued predictable process $\{X(t)\}_{t\in[0,T]}$ is a strong solution of eq.* (3.2.1)*, if X takes values in $D(A)$, P_T-a.s., $\int_0^T |AX(s)|ds < \infty$, P-a.s. and for any $t \in [0, T]$*

$$X(t) = \xi + \int_0^t [AX(s) + f(s)]ds + \int_0^t BW(s), \quad P\text{-a.s.}.$$

We say that an H-valued predictable process $\{X(t)\}_{t\in[0,T]}$ is a weak solution of eq. (3.2.1)*, if the trajectories of $X(\cdot)$ are Bochner integrable, P a.s., and*

$$\begin{aligned}\langle X(t), \zeta\rangle = &\langle x, \zeta\rangle + \int_0^t [\langle X(s), A^*\zeta\rangle + \langle f(s), \zeta\rangle]ds \\ &+ \int_0^t \langle BW(s), \zeta\rangle, \quad P\text{-a.s.} \quad \forall \zeta \in D(A^*), \forall t \in [0, T].\end{aligned}$$

The definition of strong solution makes sense only if BW is an H-valued process, requiring $\mathrm{Tr}BQB^* < \infty$, while the definition of weak solution makes sense even for cylindrical Wiener processes, since the scalar process $\langle BW(t), \zeta\rangle$ ($t \in [0, T]$) is well defined.

Theorem 3.2.1. *Under the above assumptions on A, B and f, if $\|S(r)B\|_{L_2^0} \in L^2(0, T)$, then there exists a unique weak solution to eq.* (3.2.1)*, given by*

$$X(t) = S(t)\xi + \int_0^t S(t-s)f(s)ds + \int_0^t S(t-s)BdW(s), \quad t \in [0, T]. \tag{3.2.2}$$

Proof. By Duhamel principle and the superposition principle, it suffices to show that $W_A(t) = \int_0^t S(t-s)BdW(s)$ is a weak solution of equation

$$d\tilde{X} = A\tilde{X}dt + BdW, \quad \tilde{X}(0) = 0. \tag{3.2.3}$$

Fix $t \in [0,T]$ and let $\zeta \in D(A^*)$. From the stochastic Fubini theorem, we have

$$\begin{aligned}
\int_0^t \langle A^*\zeta, W_A(s)\rangle ds &= \int_0^t \left\langle A^*\zeta, \int_0^t S(s-r)BdW(r)\right\rangle ds \\
&= \int_0^t \left\langle \int_r^t B^*S^*(s-r)A^*\zeta ds, dW(r)\right\rangle \\
&= \int_0^t \left\langle \int_r^t \left(\frac{d}{ds}B^*S^*(s-r)\zeta\right) ds, dW(r)\right\rangle \\
&= \int_0^t \langle B^*S^*(t-r)\zeta - B^*\zeta, dW(r)\rangle \\
&= \langle \zeta, W_A(t)\rangle - \int_0^t \langle \zeta, BW(s)\rangle.
\end{aligned}$$

Hence, $W_A(\cdot)$ is a weak solution of eq. (3.2.3), completing the proof of existence.

Next, we prove uniqueness. First, we note that for any $t \in [0,T]$ and $\zeta(\cdot) \in C^1([0,T];D(A^*))$, we have

$$\langle W_A(t), \zeta(t)\rangle = \int_0^t \langle W_A(s), \zeta'(s) + A^*\zeta(s)\rangle ds + \int_0^t \langle \zeta(s), BdW(s)\rangle. \tag{3.2.4}$$

Now, if X is a weak solution to eq. (3.2.1) and $\zeta_0 \in D(A^*)$, then applying eq. (3.2.4) to $\zeta(s) = S^*(t-s)\zeta_0$ yields

$$\langle X(t), \zeta_0\rangle = \left\langle \int_0^t S(t-s)BdW(s), \zeta_0\right\rangle,$$

which implies $X = W_A$ since $D(A^*)$ is dense in H. Uniqueness follows.

It remains to show eq. (3.2.4). It suffices to consider a linearly dense subset of $C^1([0,T];D(A^*))$ of functions of the form $\zeta = \zeta_0\varphi(s), s \in [0,T]$, where $\varphi \in C^1([0,T])$ and $\zeta_0 \in D(A^*)$. Let

$$F_{\zeta_0}(t) = \int_0^t \langle X(s), A^*\zeta_0\rangle ds + \int_0^t \langle BdW(t), \zeta_0\rangle.$$

By Itô formula, one has

$$F_{\zeta_0}(t)\varphi(t) = \int_0^t \langle \zeta(s), BdW(s)\rangle + \int_0^t [\varphi(s)\langle X(s), A^*\zeta_0\rangle + \varphi'(s)\langle X(s), \zeta_0\rangle]ds.$$

Then eq. (3.2.4) follows since $F_{\zeta_0}(\cdot) = \langle X(\cdot), \zeta_0\rangle$, P-a.s. ∎

Then by the regularity of stochastic convolution, the weak solution is a mean square continuous predictable process. Indeed, more regularity can be obtained.

Theorem 3.2.2. *Assume $U = H, B = I$ and for some $\alpha > 0$ there holds*

$$\int_0^T t^{-\alpha} \|S(t)\|^2_{L_2^0} dt < +\infty. \tag{3.2.5}$$

Then the weak solution of eq. (3.2.1) *has continuous modifications.*

Proof. Without loss of generality, suppose $\xi = 0, f \equiv 0$. Then $X(t) = W_A(t)$. Furthermore, we assume $\mathrm{Tr}Q < \infty$. For fixed $\alpha \in (0, 1/2)$ and integer $m > 1/(2\alpha)$, the solution can be written as

$$W_A(t) = \frac{\sin \pi\alpha}{\pi} \int_0^t S(t-\sigma) \int_\sigma^t (t-s)^{\alpha-1}(s-\sigma)^{-\alpha} ds dW(\sigma).$$

By the stochastic Fubini theorem,

$$W_A(t) = \frac{\sin \pi\alpha}{\pi} \int_0^t S(t-s)(t-s)^{\alpha-1} Y(s) ds, \tag{3.2.6}$$

where

$$Y(s) = \int_0^s S(s-\sigma)(s-\sigma)^{-\alpha} dW(\sigma).$$

Define

$$z(t) = \frac{\sin \pi\alpha}{\pi} \int_0^t S(t-s)(t-s)^{\alpha-1} y(s) dW(\sigma), \quad t \in [0, T].$$

Then it's obvious that if $y(\cdot)$ is an H-valued continuous function, then $z(t)$ is also continuous. By Hölder inequality, there is a constant $C > 0$ depending on m, α, T and M such that

$$|z(t)|^{2m} \le C \int_0^T |y(s)|^{2m} ds, \quad t \in [0, T].$$

Hence,

$$\sup_{t\in[0,T]} |z(t)|^{2m} \le C \int_0^T |y(s)|^{2m} ds.$$

Therefore, if $y(\cdot) \in L^{2m}(0, T; H)$, then $z(\cdot)$ is continuous. Since the process $Y(\cdot)$ is Gaussian with covariance operator

$$\mathrm{Cov}(Y(s))\xi = \int_0^s (s-\sigma)^{-2\alpha} S(s-\sigma) Q S^*(s-\sigma) \xi ds \quad \forall \xi \in H,$$

there exists a constant $C_1 > 0$ such that $E|Y(s)|^{2m} \le C_1$ for all $s \in [0, t]$. Hence,

$$E\int_0^T |Y(s)|^{2m} ds \le C_1 T,$$

and eq. (3.2.6) defines a continuous version of $W_A(\cdot)$. ∎

Next, let us consider linear SDEs with multiplicative noise

$$dX(t) = [AX(t) + f(t)]dt + B(X(t))dW(t), \quad t \in [0, T], \tag{3.2.7}$$

with initial data $X(0) = \xi$, where $A : D(A) \subset H \to H$ is the infinitesimal generator of a strong continuous semigroup $S(\cdot)$, $B : D(B) \subset H \to L_2^0$ is a linear operator, ξ is an H-valued $\mathscr{F}_0$-measurable random variable, f is a predictable process with locally integrable trajectories. Let $\{g_j\}$ be an orthonormal basis of $U_0 = Q^{1/2}U$. Since for any $x \in D(B)$, $B(x)$ is a Hilbert–Schmidt operator from U_0 to H, then

$$\sum_{j=1}^{\infty} |B(x)g_j|^2 < \infty, \quad x \in D(B).$$

The operator $B_j x = B(x)g_j$ is linear and $B(x)u = \sum_{j=1}^{\infty} B_j x\langle u, g_j\rangle_{U_0}$ for any $x \in D(B)$, $u \in U_0$ and $j = 1, 2, \cdots$. Hence, if $W(t) = \sum_{j=1}^{\infty} \beta_j g_j$,

Definition 3.2.2. *An H-valued predictable process $X(t)$, $t \in [0, T]$ is the strong solution of eq.* (3.2.7), *if X takes values in $D(A) \cap D(B)$, P_T a.s., such that*

$$P\left(\int_0^T |X(s)| + |AX(s)|ds < \infty\right) = 1, \quad P\left(\int_0^T \|B(X(s))\|_{L_2^0}^2 ds < \infty\right) = 1,$$

and for any $t \in [0, T]$ and P-a.s.

$$X(t) = \xi + \int_0^t [AX(s) + f(s)]ds + \int_0^t B(X(s))W(s).$$

An H-valued predictable process $X(t)$, $t \in [0, T]$ is a weak solution of eq. (3.2.7), *if $X(\cdot)$ takes values in $D(B)$, P_T-a.s.,*

$$P\left(\int_0^T |X(s)|ds < \infty\right) = 1, \quad P\left(\int_0^T \|B(X(s))\|_{L_2^0}^2 ds < \infty\right) = 1, \tag{3.2.8}$$

and for all $t \in [0, T]$ and $\zeta \in D(A^*)$, there holds

$$\langle X(t), \zeta\rangle = \langle \xi, \zeta\rangle + \int_0^t [\langle X(s), A^*\zeta\rangle + \langle f(s), \zeta\rangle]ds + \int_0^t \langle BW(t), \zeta\rangle \quad P\text{-a.s.}$$

An H-valued predictable process $X(t)$, $t \in [0,T]$ is a mild solution of eq. (3.2.7), *if X takes value in $D(B)$, P-a.s.,* (3.2.8) *holds, and for any $t \in [0,T]$,*

$$X(t) = S(t)\xi + \int_0^t S(t-s)f(s)ds + \int_0^t S(t-s)B(X(s))dW(s).$$

As above, we let $A : D(A) \subset H \to H$ be the infinitesimal generator of the C_0 semigroup in H and define

$$W_A^{\Phi}(t) = \int_0^t S(t-s)\Phi(s)dW(s), \quad t \in [0,T], \Phi \in \mathscr{N}_W. \tag{3.2.9}$$

Then for any $\Phi \in \mathscr{N}_W$, the process $W_A^{\Phi}(\cdot)$ has a predictable version.

Obviously, a strong solution must be a weak solution and a weak solution must be a mild solution. Indeed, we have the following result.

Theorem 3.2.3. *Let $A : D(A) \subset H \to H$ be the infinitesimal generator of a C_0 semigroup $S(\cdot)$ in H. Then the strong solution of eq.* (3.2.7) *is also a weak solution, and a weak solution is also a mild solution. Conversely, if a mild solution X satisfies*

$$E\int_0^T \|B(X(s))\|_{L_2^0}^2 ds < +\infty,$$

then X is also the weak solution.

Next, let us simply discuss the existence of solutions. Only the case when B is bounded is considered.

Theorem 3.2.4. *Let A be the infinitesimal generator of a C_0 semigroup $S(\cdot)$ in H, $E|\xi|^2 < \infty$ and $B \in L(H, L_2^0)$. Then eq.* (3.2.7) *has a mild solution $X \in \mathscr{N}_W^2(0,T;H)$.*

Proof. Let $\mathscr{H}$ be the space of all H-valued predictable processes Y such that $|Y|_{\mathscr{H}} = \sup_{t\in[0,T]}(E|Y(t)|^2)^{1/2} < \infty$. For any Y, define

$$\begin{aligned}\mathscr{K}(Y)(t) &= S(t)\xi + \int_0^t S(t-s)f(s)ds + \int_0^t S(t-s)B(Y(s))dW(s)\\ \mathscr{K}_1(Y)(t) &= \int_0^t S(t-s)B(Y(s))dW(s), \quad t \in [0,T].\end{aligned}$$

By Hill–Yoside theorem, we can assume that $\|S(t)\| \le M$, $t \ge 0$, and then

$$\begin{aligned}|\mathscr{K}_1(Y)(t)|_{\mathscr{H}} &\le \sup_{t\in[0,T]} E\left(\int_0^t \|S(t-s)B(Y(s))\|_{L_2^0}^2 ds\right)^{1/2}\\ &\le M\|B\|_{L(H;L_2^0)}\sqrt{T}|Y|_{\mathscr{H}}.\end{aligned}$$

Let T be sufficiently small, then $\mathscr{K}$ is a contraction mapping and hence there is a unique fixed point. It is easy to illustrate that the fixed point is a solution of eq. (3.2.7). By the standard method of continuation, the case of general T can be treated. ■

By the regularity result in the previous sections, some existence result in the analytical case can be proved. See Da Prato and Zabczyk [75].

3.3 Nonlinear SDEs

In this section, we consider the more general nonlinear SDE,

$$dX(t) = (AX + F(X))dt + B(X)dW(t), \quad X(0) = \xi. \tag{3.3.1}$$

First, we assume the following Lipshitz conditions.

Assumption 3.3.1.

(i) *A is the infinitesimal generator of strongly continuous semigroup $S(t), t \geq 0$ in H;*

(ii) *$F : H \to H$ and there is a constant $c_0 > 0$ such that*

$$\begin{aligned} |F(x)| &\leq c_0(1 + |x|) \ \ \forall x \in H, \\ |F(x) - F(y)| &\leq c_0|x - y| \ \ \forall x, y \in H; \end{aligned}$$

(iii) *$B : H \to L(U, H)$ is strongly continuous (i.e., the mapping $x \mapsto B(x)u$ is continuous as a mapping from H to H, for any $u \in U$) such that for any $t > 0, x \in H, S(t)B(x)$ belongs to $L_2(U, H)$, and there is a locally square-integrable mapping $K : [0, \infty) \to [0, \infty)$ such that*

$$\begin{aligned} \|S(t)B(x)\|_{HS} &\leq K(t)(1 + |x|), \quad t > 0, x \in H, \\ \|S(t)B(x) - S(t)B(y)\|_{HS} &\leq K(t)|x - y|, \quad t > 0, x, y \in H. \end{aligned}$$

Definition 3.3.1. *An $\mathscr{F}_t$-adapted process $X(t), t \geq 0$ is a mild solution of eq.* (3.3.1) *if*

$$\begin{aligned} X(t) = S(t)\xi &+ \int_0^t S(t-s)F(X(s))ds \\ &+ \int_0^t S(t-s)B(X(s))dW(s), \quad t \in [0, T]. \end{aligned} \tag{3.3.2}$$

Let $\mathcal{H}_{p,T}$ be the Banach space of all H-valued predictable processes $Y(t)$ such that

$$\|Y\|_{p,T} = \sup_{t \in [0,T]} (E|Y(t)|^p)^{1/p} < +\infty.$$

Theorem 3.3.1. *Assume Assumption 3.3.1 holds and $p \geq 2$. Let the initial data $\xi \in \mathscr{F}_0$ and $E|\xi|^p < \infty$, then there is a unique mild solution X of eq.* (3.3.1) *in $\mathcal{H}_{p,T}$ and there is a constant C_T, independent of ξ, such that*

$$\sup_{t\in[0,T]} E|X(t)|^p \leq C_T(1+ E|\xi|^p). \tag{3.3.3}$$

Furthermore, if there is $\alpha \in (0, 1/2)$ such that $\int_0^1 s^{-2\alpha}K^2(s)ds < +\infty$, then $X(\cdot)$ is P-a.s. continuous, where K is defined in Assumption 3.3.1(iii).

Proof. For any $\xi \in L^p(\Omega; H)$ and $X \in \mathcal{H}_{p,T}$, define a process $Y = K(\xi, X)$ by

$$\begin{aligned} Y(t) = S(t)\xi &+ \int_0^t S(t-s)F(X(s))ds \\ &+ \int_0^t S(t-s)B(X(s))dW(s), \quad t \in [0, T]. \end{aligned} \tag{3.3.4}$$

By martingale inequality, $K(\xi, X) \in \mathcal{H}_{p,T}$ for any $X \in \mathcal{H}_{p,T}$. Let $M_T = \sup_{t\in[0,T]} \|S(t)\|$, then

$$\begin{aligned} E|Y(t)|^p &\leq 3^{p-1}\left\{\|S(t)\|^p\, E|\xi|^p + E\left[\left(\int_0^t |S(t-s)F(X(s))|ds\right)^p\right]\right. \\ &\quad \left. + E\left[\left|\int_0^t S(t-s)B(X(s))dW(s)\right|^p\right]\right\} \\ &\leq 3^{p-1}\left\{M_T^p\, E|\xi|^p + T^{p-1}M_T^p \int_0^t E|F(X(s))|^p ds\right. \\ &\quad \left. + c_p\left[\int_0^t (\,E\|S(t-s)B(X(s))\|_{\mathrm{HS}}^p)^{2/p} ds\right]^{p/2}\right\}. \end{aligned}$$

Moreover, we have

$$\int_0^t E|F(X(s))|^p ds \leq 2^{p-1}c_0^p \sup_{s\in[0,t]} (1+ E|X(s)|^p)t,$$

and

$$\begin{aligned} &\left[\int_0^t (\,E\|S(t-s)B(X(s))\|_{\mathrm{HS}}^p)^{2/p} ds\right]^{p/2} \\ &\quad \leq 2^{p-1}\left(\int_0^t K^2(t-s)(1+ E|X(s)|^p)^{2/p} ds\right)^{p/2} \\ &\quad \leq 2^{p-1}\left(\int_0^t K^2(t-s)ds\right)^{p/2} \sup_{s\in[0,t]} (1+E|X(s)|^p). \end{aligned}$$

Hence, there exist c_1, c_2, c_3 such that

$$\sup_{t\in[0,T]} E|X(s)|^p \le c_1 + c_2 E|\xi|^p + c_3 \sup_{t\in[0,T]} E|X(s)|^p. \tag{3.3.5}$$

Therefore, $Y \in \mathcal{H}_{p,T}$.

By the same method, if $X_1, X_2 \in \mathcal{H}_{p,T}$, $Y_1 = K(\xi, X_1)$, $Y_2 = K(\xi, X_2)$, then

$$\sup_{t\in[0,T]} E|Y_1(t) - Y_2(t)|^p \le c_3 \sup_{t\in[0,T]} E|X_1(t) - X_2(t)|^p.$$

Hence, we can choice a sufficiently small T, such that $c_3 < 1$. By the contraction mapping principle, there exists a unique solution of eq. (3.3.1) in $\mathcal{H}_{p,T}$. By iteration, it's not difficult to prove the existence and uniqueness of solution for the general $T > 0$. Consider equation in the interval $[0, \tilde{T}]$, $[\tilde{T}, 2\tilde{T}], \cdots$, where $c_3(\tilde{T}) < 1$. For such $\tilde{T}$, we have

$$\sup_{t\in[0,T]} E|X(t)|^p \le \frac{1}{1 - c_3(\tilde{T})}[c_1 + c_2\, E|\xi|^p].$$

Hence eq. (3.3.3) is proved.

The continuity can also be treated, see Da Prato and Zabczyk [75]. ■

4 Random attractors

In this chapter, the general determinate nonautonomous dynamical systems and general random dynamical systems are introduced. We will introduce the definition of random attractors and some existence results. The basic framework of random dynamical systems was established mainly by Crauel, Debussche and Flandoli [66, 68, 77] in the 1990s. See also the monograph of Arnold [5]. The materials in this chapter are mainly from their papers.

4.1 Determinate nonautonomous systems

Let (X, d) be a separable metric space, and $S(t, s) : X \to X$ $(-\infty < s \leq t < \infty)$ be a family of mappings such that (1) $S(t, r)S(r, s)x = S(t, s)x$ holds for any $s \leq r \leq t$ and $x \in X$ and (2) $S(t, s)$ is continuous in X for any s, t. In particular, $S(t, s)$ is closely connected with a nonautonomous differential equation, where $S(t, s)x$ denotes the state of the system at time t, started from the state x at time s. But we will not concern concrete differential equation models since our purpose of this chapter is only to introduce some abstract concepts.

First of all, we introduce several central concepts in the research of dynamical systems. Given $t \in R$, if for all bounded set $B \subset X$ there holds

$$d(S(t, s)B, K(t)) \to 0, \quad s \to -\infty, \tag{4.1.1}$$

then $K(t) \subset X$ is called an attracting set at time t, where $d(A, B)$ denotes the semidistance of subsets A, B in X defined as

$$d(A, B) = \sup_{x \in A} \inf_{y \in B} d(x, y).$$

The dynamical system $\{S(t, s)\}_{t \geq s}$ is called asymptotically compact, if for any time t there is a compact attracting set.

The ω-limit of a bounded set $B \subset X$ at time t is defined by

$$A(B, t) = \bigcap_{T < t} \overline{\bigcup_{s < T} S(t, s)B}.$$

If there is a compact attracting set $K(t)$ at time $t \in R$, then $A(B, t)$ is a nonempty subset contained in $K(t)$ and

$$\begin{aligned} A(B, t) = \{x \in X :\ & \text{there is } \{x_n\}_{n \in \mathbb{N}} \subset B \text{ and } \{s_n\}_{n \in \mathbb{N}} \subset R, \\ & \text{such that } s_n \to -\infty, \text{ and } S(t, s_n)x_n \to x\}. \end{aligned} \tag{4.1.2}$$

If there is a compact attracting set $K(t)$ at time t and we take $K(\tau) = S(\tau,t)K(t)$, then $K(\tau)$ is a compact attracting set for any $\tau > t$. If for any bounded set $B \subset X$, there is $s(B)$ such that $S(t,s)B \subset K(t)$ for all $s \leq s(B)$, then $K(t)$ is an absorbing set at the time t. The following property of ω-limit $A(B,t)$ is useful.

Lemma 4.1.1. *Given $t \in R$, if there is a compact attracting set $K(t)$ at the time t, then*

$$\lim_{s\to-\infty} d(S(t,s)B, A(B,t)) = 0.$$

Proof. If not, we suppose there exist $\varepsilon > 0$ and sequences $s_n \to -\infty$ and $x_n \in B$ such that

$$d(S(t,s_n)x_n, A(B,t)) > \varepsilon \quad \forall n > 0. \tag{4.1.3}$$

By definition (4.1.1), there is a sequence $y_n \in K(t)$, such that $d(S(t,s_n)x_n, y_n) \to 0$. Since $K(t)$ is compact, there is a subsequence n_k and $y \in K(t)$ such that $y_{n_k} \to y$ and $S(t,s_{n_k})x_{n_k} \to y$. Therefore, $y \in A(B,t)$. This contradicts to eq. (4.1.3). ∎

Theorem 4.1.1. *Given $t \in R$, suppose there is a compact attracting set $K(t)$. Then $A(t) := \overline{\bigcup_{B\subset X} A(B,t)}$ is a nonempty compact subset of $K(t)$ such that*

$$\lim_{s\to-\infty} d(S(t,s)B, A(t)) = 0.$$

Furthermore, it's minimal in the sense that if $\tilde{A}(t)$ is a close set that attracts all bounded set from $-\infty$, then $A(t) \subset \tilde{A}(t)$. Moreover, $A(\tau)(\tau > t)$ is invariant under the mapping $S(t,s)$, i.e., $S(\tau,r)A(r) = A(\tau)$ for all $\tau \geq r \geq t$.

Proof. First, we show minimality. Note that $\tilde{A}(t)$ is closed for all bounded set $B \subset X$ and attracts all bounded sets. Also from eq. (4.1.2), it follows that $A(B,t) \subset \tilde{A}(t)$. Hence $A(t) \subset \tilde{A}(t)$.

Next we show invariance. Given $x \in A(r)$ and let $x_n \in X$ and bounded set $B_n \subset X$ be such that $x_n \in A(B_n,r)$ and $x_n \to x$, then $S(\tau,r)x_n \to S(\tau,r)x$ and $A(\tau,r)x_n \in A(B_n,\tau)$. Hence, $S(\tau,r)x \in A(\tau)$ and $S(\tau,r)A(r) \subset A(\tau)$. On the other hand, let $y \in A(\tau)$, $B_n \subset X$ and $y_n \in A(B_n,t)$. Thanks to eq. (4.1.1), we can select subsequences $x_n^k \in B_n$ and $s_k \to -\infty$ such that $S(\tau,s_k)x_n^k \to y_n$ for all $k \to \infty$. Without loss of generality, we assume $s_k \leq r$. Since $d(S(r,s_k)x_n^k, K(r)) \to 0$, we can check that $S(r,s_k)x_n^k \to x_n \in K(r)$ up to a subsequence thanks to the compactness of $K(r)$. Then $x_n \in A(B_n,r)$ and $S(\tau,r)x_n \to y_n$. Since $A(B_n,r) \subset K(r)$, $x_n \to x \in K(r)$ up to a subsequence. This shows that $x \in A(r)$ and $S(\tau,r)x = y$. ∎

This leads to the following definition.

Definition 4.1.1. *$A(t)$ is the global attractor of the dynamical system $\{S(t,s)\}$ at time t.*

The following theorem will be often used.

Theorem 4.1.2. *Let $\{S(t,s)\}_{t\geq s}$ be asymptotically compact. Then for all $t \in R$, $A(t) \subset K(t)$ is the minimal closed set that is non-empty, compact and attracts all bounded sets from $-\infty$. Moreover, it is invariant in the sense that $S(t,s)A(s) = A(t)$ for all $t \geq s$.*

4.2 Stochastic dynamical systems

Let (X,d) be a complete separable metric space and $(\Omega, \mathscr{F}, P)$ be a probability space. Suppose that $S(t,s;\omega) : X \to X$ is a family of mappings with parameter ω for $-\infty < s \leq t < \infty$ satisfying

(1) $S(t,r;\omega)S(r,s;\omega)x = S(t,s;\omega)x$ for all $s \leq r \leq t$ and for any $x \in X$ and

(2) $S(t,s;\omega)$ is continuous in X for all $s \leq t$.

Definition 4.2.1. *Given $t \in R$ and $\omega \in \Omega$. We say that $K(t,\omega) \subset X$ is an* attracting set *if for any bounded set $B \subset X$, there holds $d(S(t,s;\omega)B, K(t,\omega)) \to 0$ as $s \to -\infty$. We say that $\{S(t,s;\omega)\}_{t\geq s;\omega\in\Omega}$ is* asymptotically compact *if there is a measurable set $\Omega_0 \subset \Omega$ with $P(\Omega_0) = 1$ such that for all $t \in R$ and $\omega \in \Omega_0$, there is a compact attracting set $K(t,\omega)$.*

It can be shown that if $\{S(t,s;\omega)\}_{t\geq s,\omega\in\Omega}$ is asymptotically compact, then for any $t \in R$, there is a measurable set $\Omega_t \subset \Omega$, $P(\Omega_t) = 1$, such that for any $\omega \in \Omega_t$, there is a compact attracting set $K(t,\omega)$ and vice versa. Indeed, let $\Omega_0 = \bigcap_{n\in\mathbb{N}} \Omega_{-n}$, then $P(\Omega_0) = 1$. For any n and $\omega \in \Omega_0$, there is a compact attracting set $K(-n,\omega)$ at time $t = -n$. Hence, there is a compact absorbing set at any time.

Similar to the nonautonomous systems, the random ω-limit set can be defined for bounded set $B \subset X$. For any bounded set $B \subset X$, define random ω-limit set as

$$A(B,t,\omega) = \bigcap_{T<t} \overline{\bigcup_{s<T} S(t,s;\omega)B} \tag{4.2.1}$$

and the set

$$A(t,\omega) = \overline{\bigcup_{B\subset X} A(B,t,\omega)} \tag{4.2.2}$$

is a random attractor.

Theorem 4.2.1. *Let $\{S(t,s,\omega)\}_{t\geq s,\omega\in\Omega}$ be asymptotically compact, then it holds for P-a.e. ω that for any $t \in R$, $A(t,\omega)$ is the nonempty compact subset of $K(t,\omega)$ that attracts all bounded set starting from $-\infty$, and it's the minimal set with such properties. Moreover, it is invariant in the sense that $S(t,s;\omega)A(s,\omega) = A(t,\omega)$ for all $s \leq t$.*

Next, we consider the measurability of the attractor. A family of closed subsets $A(\omega) \subset X$ for $\omega \in \Omega$ is measurable if the function $\omega \mapsto d(A(\omega), x)$ is measurable for all $x \in X$.

Proposition 4.2.1. *Let $\{S(t,s;\omega)\}_{t\geq s,\omega\in\Omega}$ be asymptotically compact and satisfy conditions (1), (2) and (3) for any $t \in R$ and $x \in X$, the mapping $(s,\omega) \mapsto S(t,s;\omega)x$ is measurable.*

Then for any $t \in R$ and all bounded set $B \subset X$, $A(B,t,\omega)$ and $A(t,\omega)$ are measurable with respect to the P-completion of $\mathscr{F}$.

Proof. It suffices to show the measurability of $\overline{\bigcup_{s<T} S(t,s;\omega)B}$ by definition of $A(t,\omega)$. Note

$$d(x, \overline{\bigcup_{s<T} S(t,s;\omega)B}) = \inf_{s<T} d(x, S(t,s;\omega)B), \quad x \in X. \tag{4.2.3}$$

It follows from Ref. [51, Theorem 23, p. 75] that the level set of $\inf_{s<T} d(x, S(t,s;\omega)B)$ is measurable w.r.t. the P-completion of $\mathscr{F}$. ∎

Next, let us consider the case with a shift on the probability space $(\Omega, \mathscr{F}, P)$. That is there is a measure-preserving transformations group θ_t, $t \in R$ on $(\Omega, \mathscr{F}, P)$ such that

$$S(t,s;\omega)x = S(t-s,0;\theta_s\omega)x, \quad P\text{-a.s.}, \tag{4.2.4}$$

for all $s < t$ and $x \in X$. In most of the applications, θ_t is defined as the shift

$$(\theta_t\omega)(s) = \omega(t+s) - \omega(t), \quad s,t \in R. \tag{4.2.5}$$

For such a group θ_t, we assume

(4) for any $s < t$ and $x \in X$, the mapping $\omega \mapsto S(t,s;\omega)x$ is measurable, and

(5) for any t and $x \in X$ and P-a.e. ω, the mapping $s \mapsto S(t,s;\omega)x$ is right continuous at any point.

Proposition 4.2.2. *Suppose (1), (2), (4), (5) and eq. (4.2.4) hold, and for P-a.e. ω, there is a compact attracting set $K(\omega)$ at the time $t = 0$, then the stochastic dynamical system $\{S(t,s;\omega)\}_{t\geq s,\omega\in\Omega}$ is asymptotically compact.*

Proof. It suffices to prove that there is a compact attracting set $K(t,\omega)$, P-a.s. for fixed $t \in R$. Let $\{x_n\} \subset X$ be dense and s_k be dense in $(-\infty, t)$. By assumption (4.2.4), except for a possible zero measure set $\Omega \backslash \Omega_0$, there holds for all $n, k \in \mathbb{N}$ that

$$S(t,s_k;\omega)x_n = S(t-s_k,0;\theta_{s_k}\omega)x_n = S(0,-t+s_k;\theta_t\omega)x_n.$$

Let Ω_1 be a set of full probability such that for any $\omega \in \Omega_1$, there is a compact attracting set $K(\omega)$ at the time $t = 0$. Let $\Omega_2 = \Omega_0 \bigcap \theta_{-t}\Omega_1$, then $P(\Omega_2) = 1$ and for all $\omega \in \Omega_2$, $\theta_t\omega \in \Omega_1$ and hence $K(\theta_t\omega)$ is well defined.

Fix $B \subset X$, bounded and let $\omega \in \Omega_2$, then from (2) and (5),

$$d(S(t,s;\omega)B, K(\theta_t\omega)) = \sup_{x_n \in B} \lim_{s_k \to s} d(S(t,s_k;\omega)x_n, K(\theta_t\omega)).$$

Similarly,

$$d(S(0,-t+s;\theta_t\omega)B, K(\theta_t\omega)) = \sup_{x_n \in B} \lim_{s_k \to s} d(S(0,-t+s_k;\theta_t\omega)x_n, K(\theta_t\omega)),$$
$$d(S(t,s,\ \omega)B, K(\theta_t\omega)) = d(S(0,-t+s;\theta_t\omega)B, K(\theta_t\omega)),$$

which imply that $d(S(t,s;\omega)B, K(\theta_t\omega)) \to 0$ as $s \to -\infty$. The result follows from the definition of compact attracting sets. ∎

The following are sufficient conditions to ensure (5) hold. Assume that

(6.a) for all s, $x \in X$ and P-a.e. ω, the mapping $t \mapsto S(t,s;\omega)x$ is continuous at $t = s$, and

(6.b) for any $s < t$ and P-a.e. ω, the mapping $x \mapsto S(t,s;\omega)x$ is continuous in X, uniformly in s on bounded sets

hold, then condition (5) holds.

Condition (5) can also be replaced with some more weaker conditions. For example, it can be replaced with

(7) for any t and $x \in X$, there exists a dense countable set $D(t,x) \subset (-\infty, t)$ such that for P-a.e. ω and $s < t$, there is a sequence $D(t,x) \ni s_n \to s$ such that $S(t,s_n;\omega)x \to S(t,s;\omega)x$.

Another condition that implies (7) is

(8) the mapping $t \mapsto S(t,s;\omega)x$ is continuous at $t = s$, uniformly in s on bounded intervals.

If the assumptions of Proposition 4.2.2 hold, then for a.a. $\omega \in \Omega$, the random attractor is well defined. It's a nonempty compact subset of X and is measurable w.r.t. ω. Denote $A(\omega) = A(0,\omega)$, then it is invariance in the sense that $S(t,s;\omega)A(\theta_s\omega) = A(\theta_t\omega)$.

Proposition 4.2.3. *Assume that the assumptions of Proposition 4.2.2 hold and $\theta_t (t \in R)$ be ergodic. Then there is a bounded set $B \subset X$ independent of ω such that $A(\omega)$ is the ω-limit of B at time $t = 0$. Moreover, $A(\omega)$ is the largest compact measurable set satisfying the property that if $\{\tilde{A}(\omega)\}_{\omega\in\Omega}$ is a measurable family of measurable compact subset such that for almost all ω, $S(t,s;\omega)\tilde{A}(\theta_s\omega) = \tilde{A}(\theta_t\omega)$, then $\tilde{A}(\omega) \subset A(\omega)$ for almost all ω.*

Proof. When $A(\omega)$ is defined, set

$$R(\omega) = \inf\{r \in R : \ A(\omega) \subset B(O,r)\}, \tag{4.2.6}$$

where $B(O, r)$ is a ball with radius r centered at $O \in X$. Otherwise, let $R(\omega) = 0$. The function $R(\omega)$ is obviously measurable, and by ergodicity of θ_t, for almost all ω, there exists a sequence $t_n \to \infty$ such that $R(\theta_{-t_n}\omega) \le R_0 + 1$, where $R_0 = \text{ess sup}_{\omega\in\Omega}\{R(\omega)\}$. Let $x \in A(\omega)$, by invariance we have for any $n \in \mathbb{N}$, there is $x_n \in A(\theta_{-t_n}\omega)$ such that $S(0, -t_n; \omega)x_n = x$. By $A(\theta_{-t_n}) \subset B(O, R_0 + 1)$ and eq. (4.1.2), we have $x \in A(B(O, R_0 + 1), 0, \omega)$, i.e., $A(\omega) \subset A(B(O, R_0 + 1), 0, \omega)$. The other inclusion is obvious; hence, the first part of the conclusion holds.

Consider a measurable family of compact sets $\{\tilde{A}(\omega)\}_{\omega\in\Omega}$, satisfying the invariance property. Similar arguments show that there is a bounded set B such that $\tilde{A}(\omega) \subset A(B, 0, \omega)$ holds for almost all ω. Then the result follows since $A(B, 0, \omega) \subset A(\omega)$. ∎

The above results are summarized in the following theorem:

Theorem 4.2.2. *Let the stochastic dynamical system $\{S(t, s; \omega)\}_{t\ge s,\omega\in\Omega}$ satisfy conditions (1), (2), (4) and (5) and assume that there is a measure-preserving transformation group $(\theta_t)_{t\in R}$ satisfying eq. (4.2.4), and for P-a.e. ω, there is a compact attracting set $K(\omega)$ at $t = 0$. For P-a.e., $\omega \in \Omega$, we set $A(\omega) = \overline{\bigcup_{B\subset X} A(B, \omega)}$ and*

$$A(B, \omega) = \bigcap_{T<0} \overline{\bigcup_{s<T} S(0, s; \omega)B}.$$

Then, the following conclusions hold for P-a.e. $\omega \in \Omega$:

(1) *$A(\omega)$ is a nonempty compact set of X and if X is connected, it is a connected subset of $K(\omega)$,*

(2) *the family $\{A(\omega)\}_{\omega\in\Omega}$ is measurable,*

(3) *$A(\omega)$ is invariant in the sense that $S(t, s; \omega)A(\theta_s\omega) = A(\theta_t\omega)$ for $s \le t$,*

(4) *it is the minimal closed set such that*

$$d(S(t, s; \omega), A(\theta_t\omega)) \to 0 \quad \textit{for} \quad s \to -\infty$$

for any $t \in R$ and $B \subset X$ bounded,

(5) *for any bounded set $B \subset X$, $d(S(t, s; \omega), A(\theta_t\omega)) \to 0$ in probability as $s \to \infty$.*

Moreover, if θ_t, $t \in R$ is ergodic, then

(6) *there is a bounded set $B \subset X$ such that $A(\omega) = A(B, \omega)$, and*

(7) *$A(\omega)$ is the largest compact set that is invariant.*

The connectness of $A(\omega)$ is proved in Ref. [68], and (5) is obvious since θ_t is measure preserving.

5 Applications

In this chapter, we consider several concrete stochastic partial differential equations (SPDEs) and their dynamics, including Ginzburg–Landau (GL) equation, stochastic damped forced Ostrovsky equation, some geophysical models and primitive equations. The interested readers may refer to the literature in the following sections.

5.1 Stochastic GL equation

In this section, the stochastic Ginzburg–Landau (SGL) equation (SGLE) with multiplicative noise and the associated random dynamical system (RDS) are considered. For details, the readers may refer to Refs [31, 125, 213, 257]. Precisely, let $D \subset R^n (n = 1, 2)$ be a bounded open set and ∂D be sufficiently smooth. We consider the following SGLE with multiplicative noise:

$$\begin{cases} du = (\lambda + i\alpha)\Delta u dt + \nu u dt - (k + i\beta)|u|^2 u dt + \sigma u dW(t), & x \in D, \\ u(t) = 0, \quad x \in \partial D, \end{cases} \tag{5.1.1}$$

with initial data $u(0) = u_0$, where $\lambda, \alpha, \nu, k, \beta, \sigma \in R, \lambda > |\alpha|, k > |\beta|, \sigma > 0$ are real parameters. The stochastic term $W(t) : \Omega \to R$ is a two-sided standard Wiener process and the unknown function u is a complex-valued function defined as $D \times R^+$.

We will let $L^2(D)$ and $H_0^1(D)$ denote the usual complex-valued Sobolev space on D, with L^2- and $H_0^1(D)$-inner product and norms being defined as

$$(u, v) = \Re \int_D u(x)\bar{v}(x)dx, \quad |u| = (u, u)^{1/2} \quad u, v \in L^2(D),$$

$$((u, v)) = \Re \sum_{i=1}^{n} \int_D D_i u D_i \bar{v} dx, \quad \|u\| = ((u, u))^{1/2} \quad u, v \in H_0^1(D).$$

The notations L^2 and H_0^1 used here are to distinguish the usual real-valued spaces L^2 and H_0^1. Denote $H = L^2(D), V = H_0^1(D), Au = -\Delta u$ and $f(u) = |u|^2 u$. The operator A is an isomorphism from $D(A) = V \cap H^2(D)$ onto H. Let $\{e_n\}$ be the orthonormal basis in H of the eigenvectors corresponding to eigenvalues $\lambda_n, \lambda_n > 0$ and $\lambda_n \to \infty$. For the first eigenvalue λ_1, we have $\lambda_1 |u|^2 \le \|u\|^2$. With these notations, the SGLE can be rewritten as the following abstract form:

$$du + (\lambda + i\alpha)Audt - \nu u dt + (k + i\beta)f(u)dt = \sigma u dW(t) \tag{5.1.2}$$

with initial data $u(0) = u_0$. When $\sigma = 0$, it is known [245] there is a unique solution

$$u \in C([0, T]; H) \cap L^2(0, T; V) \quad \forall T < \infty, u_0 \in H$$

and

$$u \in C([0,T];V) \cap L^2(0,T;D(A)) \ \ \forall T < \infty, u_0 \in V.$$

The following energy estimate can be obtained for the equation

$$\frac{1}{2}\frac{d}{dt}|u|^2 + \lambda\|u\|^2 - \nu|u|^2 + k|u|_{L^4}^4 = 0. \tag{5.1.3}$$

If $\nu < \lambda\lambda_1$, it implies $|u(u)| \leq |u_0| \exp\{(\nu - \lambda\lambda_1)t\}$, following that the steady solution $u \equiv 0$ is exponentially stable.

If $\nu = \lambda\lambda_1$, then by eq. (5.1.3), there is a constant C depending on u_0, k and the region D such that $|u(t)| \leq Ct^{-1/2}$; hence, $u \equiv 0$ is also asymptotically stable.

If $\nu > \lambda\lambda_1$, the dynamical system generated by GL equation has finite-dimensional global attractors.

From the above analysis, if $\nu \leq \lambda\lambda_1$, then all the trajectories converge to 0 as $t \to \infty$. Hence, the global attractor reduces to the steady solution $\{0\}$.

Now, let us consider the stochastic case. Introduce the process $z(t) = e^{-\sigma W(t)}$. Obviously, it satisfies the stochastic differential equation

$$dz(t) = \frac{1}{2}\sigma^2 z dt - \sigma z dW(t).$$

Hence, the process $v(t) = z(t)u(t)$ satisfies

$$dv(t) + (\lambda + i\alpha)Avdt - (\nu - \frac{\sigma^2}{2})vdt + (k + i\beta)zf(u)dt = 0.$$

Since $z(t)$ is a real-valued process, the above formula can be written as

$$dv(t) + (\lambda + i\alpha)Avdt - (\nu - \frac{\sigma^2}{2})vdt + (k + i\beta)z^{-2}f(v)dt = 0. \tag{5.1.4}$$

By Galerkin method and a priori estimate, it is not difficult to prove that for P-a.e. $\omega \in \Omega$, the equation has a unique strong solution

$$u \in C([s,t];H) \cap L^2(s,t;V) \ \ \forall s < t, u(s) \in H,$$

and

$$u \in C([s,t];V) \cap L^2(s,t;D(A)) \ \ \forall s < t, v(s) \in V.$$

Then $u(t) = \frac{v(t)}{z(t)} = e^{\sigma W(t)}v(t)$ is the solution of eq. (5.1.2).

Theorem 5.1.1. *Let $\nu < \lambda\lambda_1 + \frac{1}{2}\sigma^2$. Then there is a P-full set Ω_1 such that for all $\omega \in \Omega_1$, there exists $T(\omega) > 0$, such that for any $u_0 \in H$ and $t \geq T(\omega)$, there holds*

$$|u(t,\omega,u_0)| \leq |u_0| e^{\frac{1}{2}[\nu - \lambda\lambda_1 - \frac{1}{2}\sigma^2]t}. \tag{5.1.5}$$

Proof. Applying Itô formula to $|u(t)|^2$, we get

$$\begin{aligned}|u(t)|^2 = |u_0|^2 + 2\int_0^t \Big(-\lambda\|u(s)\|^2 + \Big(\nu + \frac{1}{2}\sigma^2\Big)|u(s)|^2 \\ - k|u(s)|_{L^4}^4\Big)ds + 2\int_0^t \sigma|u(s)|^2 dW(s).\end{aligned}$$

Applying Itô formula again to $\log|u(t)|^2$, we have

$$\begin{aligned}\log|u(t)|^2 &= \log|u_0|^2 - 2\int_0^t \frac{1}{|u(s)|^2}(\lambda\|u(s)\|^2 + k|u(s)|_{L^4}^4)ds \\ &\quad + (2\nu - \sigma^2)t + 2\sigma W(t) \\ &\leq \log|u_0|^2 + (2\nu - \lambda\lambda_2 - \sigma^2)t + 2\sigma W(t).\end{aligned}$$

Since $\lim_{t\to\infty} W(t)/t = 0$, P-a.s., there is a P-full set Ω_1 such that for $\omega \in \Omega_1$, there is a $T(\omega) > 0$ such that

$$\frac{2\sigma W(t)}{t} \leq (\lambda\lambda_1 + \frac{\sigma^2}{2} - \nu), \quad t \geq T(\omega).$$

Hence

$$\log|u(t)|^2 \leq \log|u_0|^2 + \Big(\nu - \lambda\lambda_2 - \frac{\sigma^2}{2}\Big)t. \tag{5.1.6}$$

The proof of the theorem is completed. ∎

5.1.1 The existence of random attractor

Now, let us consider the random attractor of the GL equation. Let $\Omega = \{\omega \in C(R,R) : \omega(0) = 0\}$, $\mathscr{F}$ be a Borel σ-algebra in Ω and P be a Wiener measure on $(\Omega, \mathscr{F})$. Denote $W(t,\omega) = \omega(t)$ and define

$$\theta_t\omega(s) = \omega(t+s) - \omega(t). \tag{5.1.7}$$

Then θ_t satisfies $\theta_t \circ \theta_s = \theta_{t+s}$, and hence $(\Omega, \mathscr{P}, \{\theta_t\}_{t\in R})$ is an ergodic metric dynamical system. Denote

$$u(t,\omega) = \psi(t,s;\omega)u_s, \quad v(t,\omega) = \phi(t,s;\omega)v_s,$$

where $u(t)$ is a solution of eq. (5.1.2) with initial data $u(s) = u_s$, and $v(t)$ is a solution of eq. (5.1.4) with initial data $v(s) = v_s$. Obviously, for $s \le r \le t$, there holds

$$\psi(t,s;\omega) = \psi(t,r;\omega)\psi(r,s;\omega).$$

By eq. (5.1.7), for any $s,t \in R^+, u_0 \in H$, there holds P-a.s.,

$$\psi(t+s,0;\omega)u_0 = \psi(t,0;\theta_s\omega)\psi(s,0;\omega)u_0.$$

Hence, the process $\varphi : R^+ \times \Omega \times H \to H$ defined by

$$\varphi(t,\omega)u_0 = \psi(t,0;\omega)u_0, \tag{5.1.8}$$

is an RDS on H over $(\Omega, \mathscr{F}, P, \{\theta_t\}_{t\in R})$. In order to prove the existence of a compact absorbing set, we first show

Lemma 5.1.1. *Given any ball $B(0,\rho)$ in H with radius ρ centered at 0, there exist random variables $r_t(\omega)$ and $t(\omega,\rho) \le -1$ such that for any $s \le t(\omega,\rho)$, $u_s \in B(0,\rho)$, $v_s = z(s)u_s$ and $-1 \le t \le 0$, we have $|\phi(t,s;\omega)v_s| \le r_t(\omega)$, P-a.s. Hence, $|\psi(0,s;\omega)u_s| \le r_0(\omega)$.*

Proof. Let $v(t) = v(t,s,v_s;\omega)$ be the solution of eq. (5.1.4) with initial value v_s, then

$$\begin{aligned}\frac{d}{dt}|v|^2 + \lambda\|v\|^2 &= -\lambda\|v\|^2 + (2\nu - \sigma^2)|v|^2 - 2kz^{-2}|v|_{L^4}^4 \\ &\le -(\lambda\lambda_1 + \sigma^2)|v|^2 + 2\nu|v|^2 - 2kz^{-2}|v|_{L^4}^4.\end{aligned} \tag{5.1.9}$$

Since $|v| \le |D|^{1/4}|v|_{L^4}$, we obtain

$$\begin{aligned}\frac{d}{dt}|v|^2 + \lambda\|v\|^2 &\le -(\lambda\lambda_1 + \sigma^2)|v|^2 + 2\nu|D|^{1/2}|v|_{L^4}^2 - 2kz^{-2}|v|_{L^4}^4 \\ &\le -(\lambda\lambda_1 + \sigma^2)|v|^2 + \nu^2|D|k^{-1}z^2 - kz^{-2}|v|_{L^4}^4 \\ &\le -(\lambda\lambda_1 + \sigma^2)|v|^2 + \nu^2|D|k^{-1}z^2.\end{aligned} \tag{5.1.10}$$

It follows that for $t \ge s$, we have

$$\begin{aligned}|v(t)|^2 &\le |v(s)|^2 e^{-(\lambda\lambda_1+\sigma^2)(t-s)} + \nu^2|D|k^{-1}\int_s^t e^{-(\lambda\lambda_1+\sigma^2)(t-\tau)}z^2(\tau)d\tau \\ &\le e^{(\lambda\lambda_1+\sigma^2)(s-t)}z^2(s)|u_s|^2 + \nu^2|D|k^{-1}\int_s^t e^{(\lambda\lambda_1+\sigma^2)(\tau-t)}z^2(\tau)d\tau.\end{aligned} \tag{5.1.11}$$

Since $e^{(\lambda\lambda_1+\sigma^2)s}z^2(s) = e^{(\lambda\lambda_1+\sigma^2)s-2\sigma W(s)} \to 0$, P-a.s. as $s \to \infty$, for $u_s \in B(0,\rho) \subset H$, there is a time $t(\omega,\rho) \le -1$, such that for $s \le t(\omega,\rho)$, it holds P-a.s. that $e^{(\lambda\lambda_1+\sigma^2)s}z^2(s)\rho^2 \le 1$. So we can choose positive variable $r_t(\omega)$,

$$r_t^2(\omega) = e^{-(\lambda\lambda_1+\sigma^2)t}(1 + \nu^2|D|k^{-1}\int_{-\infty}^{t} e^{(\lambda\lambda_1+\sigma^2)\tau}z^2(\tau)d\tau)$$

such that $|v(t)| \le r_t(\omega)$, P-a.s. ■

Note that $\nu < \lambda\lambda_1 + \sigma^2/2$, then by eq. (5.1.9), we have

$$|v(t)| \le |v(0)|e^{(\nu-\lambda\lambda_1-\sigma^2/2)t}.$$

Hence,

$$|u(t,\omega)| \le |u_0|e^{(\nu-\lambda\lambda_1-\sigma^2/2+\sigma[W(t)/t])t}.$$

Since $\lim_{t\to\infty}[W(t)/t] = 0$, P-a.s., there exists a $t(\omega)$ such that for all $t \ge t(\omega)$,

$$\sigma\frac{W(t)}{t} \le \frac{1}{2}(\lambda\lambda_1 + \sigma^2/2 - \nu).$$

Hence the same result as Theorem 5.1.1 holds. From the viewpoint of attractor, since for $s \le 0$,

$$|u(0)| \le |u_s|e^{-(\nu-\lambda\lambda_1-\sigma^2/2)s-\sigma W(s)},$$

we have $|u(0)| \to 0$ P-a.s. as $s \to -\infty$. So the attractor reduces to $\{0\}$.

Suppose $\nu \ge \lambda\lambda_1 + \sigma^2/2$ in the following. Lemma 5.1.1 shows there is an absorbing set $B(0, r_0(\omega))$ in H. In order to obtain the absorbing set in V, we assume $|\beta| \le k$.

First, integrating eq. (5.1.10) over $[-1, 0]$, we obtain

$$\int_{-1}^{0} \|v(s)\|^2 ds \le \frac{1}{\lambda}\left(|v(-1)|^2 + \nu^2|D|k^{-1}\int_{-1}^{0} z^2(s)ds\right).$$

Multiplying eq. (5.1.4) by $-\Delta\bar{v}$, integrating over D, and taking the real part, we have

$$\frac{d}{dt}\|v\|^2 + \lambda|\Delta v|^2 - \left(\nu - \frac{\sigma^2}{2}\right)\|v\|^2 = z^{-2}\Re(k + i\beta)\int f(v)\Delta\bar{v}dx. \tag{5.1.12}$$

By $|\beta| \le k$, the right-hand side (RHS) of eq. (5.1.12) is nonpositive:

$$\begin{aligned}\Re(k + i\beta)\int f(v)\Delta\bar{v}dx &= -\Re(k + i\beta)\int(|v|^2|\nabla v|^2 + v\nabla\bar{v}\nabla|v|^2)dx\\ &= -k\int |v|^2|\nabla v|^2 dx - \frac{k}{2}\int(\nabla|v|^2)^2 dx + \beta\Im\int v^2(\nabla\bar{v})^2 dx\\ &\le (|\beta| - k)\int |v|^2|\nabla v|^2 dx \le 0.\end{aligned}$$

So eq. (5.1.12) can be rewritten as

$$\frac{d}{dt}\|v\|^2 \leq -2\lambda|\Delta v|^2 + 2\left(v - \frac{\sigma^2}{2}\right)\|v\|^2 \leq 2\left(v - \lambda\lambda_1 - \frac{\sigma^2}{2}\right)\|v\|^2.$$

For any $s \in [-1, 0]$, it holds

$$\|v(0)\|^2 \leq \|v(s)\|^2 + 2\left(v - \lambda\lambda_1 - \frac{\sigma^2}{2}\right)\int_s^0 \|v(\tau)\|^2 d\tau.$$

Integrating over $[-1, 0]$, we get

$$\begin{aligned}\|v(0)\|^2 &\leq 2\left(\frac{1}{2} + v - \lambda\lambda_1 - \frac{\sigma^2}{2}\right)\int_{-1}^0 \|v(s)\|^2 ds \\ &\leq \frac{2}{\lambda}\left(\frac{1}{2} + v - \lambda\lambda_1 - \frac{\sigma^2}{2}\right)\left(|v(-1)|^2 + v^2|D|k^{-1}\int_{-1}^0 z^2(s)ds\right).\end{aligned} \tag{5.1.13}$$

Hence for given $\rho > 0$, there is a $T(\omega) \leq -1$ such that for $s \leq T(\omega)$ and $u_s \in B(0,\rho) \subset H$, one has

$$\|u(0)\|^2 = \|v(0)\|^2 \leq R_0^2(\omega), \quad P\text{-a.s.}, \tag{5.1.14}$$

where

$$\begin{aligned}R_0^2(\omega) =& \frac{2}{\lambda}\left(\frac{1}{2} + v - \lambda\lambda_1 - \frac{\sigma^2}{2}\right)\left(e^{\lambda\lambda_1+\sigma^2} + v^2|D|k^{-1}\right) \\ &\times\left(1 + \int_{-\infty}^0 e^{(\lambda\lambda_1+\sigma^2)\tau}z^2(\tau)d\tau + \int_{-1}^0 z^2(s)ds\right).\end{aligned} \tag{5.1.15}$$

Theorem 5.1.2. *Suppose $|\beta| \leq k$. The RDS associated with the SGLE has a global random attractor $\mathcal{A}(\omega)$. If $v < \lambda\lambda_1 + \frac{\sigma^2}{2}$, the random attractor reduces to $\{0\}$.*

Although the compact absorbing set obtained here guarantees the existence of a random attractor, generally speaking, the union of $\mathcal{A}(\omega)$ in ω is not compact. However, as $\sigma \to 0$, the $\mathcal{A}(\omega)$ may converge to the attractor of the corresponding deterministic system with probability 1. Since P is invariant under θ_t, such an asymptotic behavior with attraction property from 0 to ∞ can be obtained in a weaker convergence in probability. That is for any $\varepsilon > 0$ and all deterministic bounded set $B \subset H$, there holds

$$\lim_{t\to\infty} P(\mathrm{dist}(\varphi(t,\omega)B, \mathcal{A}(\theta_t\omega)) < \varepsilon) = 1.$$

Especially, if $v < \lambda\lambda_1 + \frac{\sigma^2}{2}$, the attraction as $t \to \infty$ holds not only in probability but also ω-wise.

5.1.2 Hausdorff dimension of random attractor

The theory of the dimension of attractor in the deterministic system can be generated to the stochastic case under some assumptions. The following result is due to Ref. [77].

Theorem 5.1.3. *Let $\mathcal{A}(\omega)$ be a compact measurable set which is invariant under a random map $S(\omega)$, $\omega \in \Omega$, for some ergodic metric dynamical system $(\Omega, \mathscr{F}, P, \{\theta_t\}_{t\in R})$. Assume*

i) *$S(\omega)$ is uniform differentiable a.s., on $\mathcal{A}(\omega)$, i.e. for every $u, u+h \in \mathcal{A}(\omega)$, there is a linear operator $DS(\omega,u) \in \mathcal{L}(H)$ such that*

$$|S(\omega)(u+h) - S(\omega)u - DS(\omega,u)h| \le \bar{k}(\omega)|h|^{1+\mu},$$

where $\mu > 0$, $\bar{k}(\omega)$ is a random variable satisfying $\bar{k}(\omega) \ge 1$ and $E(\log \bar{k}) < \infty$.

(ii) *there is a variable $\bar{\omega}_d(\omega)$ satisfying $E(\log(\bar{\omega}_d)) < 0$ such that $\omega_d(DS(\omega,u)) \le \bar{\omega}_d(\omega)$ for $u \in \mathcal{A}(\omega)$, where*

$$\omega_d(L) = \alpha_1(L)\cdots\alpha_d(L),$$
$$\alpha_i(L) = \sup_{F\subset H, dimF\le n-1} \inf_{\varphi\in F, |\varphi|=1} |L\varphi|, \qquad L \in \mathcal{L}(H)$$

(iii) *for $u \in \mathcal{A}(\omega)$ and some random variable $\bar{\alpha}_1(\omega) \ge 1$ satisfying $E(\log \bar{\alpha}_1) < \infty$, there holds $\alpha_1(DS(\omega,u)) \le \bar{\alpha}_1(\omega)$.*

Then the Hausdorff dimension $d_H(\mathcal{A}(\omega))$ of $\mathcal{A}(\omega)$ is less than d a.s.

By the theorem, the main task in the following is to verify the φ defined by eq. (5.1.8) is uniformly differentiable. Set $S(\omega) = \varphi(1,\omega)$, and $T(\omega)v_0 = \phi(1,0,v_0;\omega)$, then the random attractor $\mathcal{A}(\omega)$ is an invariant compact measurable set under S. Since $S(\omega) = e^{\sigma W(1)}T(\omega)$, it is easy to see that if $T(\omega)$ is uniformly differentiable a.s., then so it is with $S(\omega)$ with the Fréchet derivative $DS = e^{\sigma W(t)}DT$, where DT is the Fréchet derivative of T.

Lemma 5.1.2. *$T(\omega)$ is uniformly differentiable a.s. on $\mathcal{A}(\omega)$, i.e. for $v, v+h \in \mathcal{A}(\omega)$, there exists $DT(\omega,v) \in \mathcal{L}(H)$ such that P-a.s.,*

$$|T(\omega)(v+h) - T(\omega)v - DT(\omega,v)h| \le \bar{k}(\omega)|h|^{1+\mu},$$

where $\mu > 0, \bar{k}(\omega) \ge 1, E(\log \bar{k}(\omega)) < \infty$ and $DT(\omega,v_0)h = V(1)$, $V(t)$ solves the first variation equation of eq. (5.1.4)

$$dV/dt = L(t,v)V, \quad V(0) = h. \tag{5.1.16}$$

Here, $v(t) = \phi(t,0,v_0;\omega), L(t,v) = -(\lambda+i\alpha)A + (\nu - \sigma^2/2) - (k+i\beta)z^{-2}f'(v)$.

The proof of the lemma is given at the end of this section.

By eq. (5.1.16), we have

$$\frac{1}{2}\frac{d}{dt}|V|^2 = -\lambda\|v\|^2 + (\nu - \frac{\sigma^2}{2})|V|^2 \\ - \Re(k+i\beta)z^{-2}\int_D (|v|^2|V|^2 + 2v\bar{V}\Re(\bar{v}V))dx.$$

The last term at the RHS is nonpositive,

$$\begin{aligned} &-\Re(k+i\beta)z^{-2}\int (|v|^2|V|^2 + 2v\bar{V}\Re(\bar{v}V))dx \\ &= -k\int |v|^2|V|^2dx - 2\int \Re(\bar{v}V)\{k\Re(v\bar{V}) - \beta\Im(v\bar{V})\}dx \\ &\le -k\int |v|^2|V|^2dx + 2\beta\int \Im(v\bar{V})\Re(v\bar{V})dx \\ &\le (|\beta| - k)\int |v|^2|V|^2dx \le 0. \end{aligned}$$

Hence, $|V(t)|^2 \le |V(0)|e^{(\nu-\lambda\lambda_1-\sigma^2/2)t}$. Since $\alpha_1(DT(\omega,v))$ is equal to the operator norm of $DT(\omega,v) \in \mathcal{L}(H)$, by choosing $\bar{\alpha}_1(\omega) = \max\{1, e^{\sigma W(1)+\nu-\lambda\lambda_1-\sigma^2/2}\}$ it is not difficult to get $\alpha_1(DS(\omega,u)) \le \bar{\alpha}_1(\omega)$ and $E(\log\bar{\alpha}_1) < \infty$.

Noting

$$DT(\omega,v) = \exp\left\{\int_0^1 L(s,v(s))ds\right\},$$

and

$$DS(\omega,u) = \exp\left\{\sigma W(1) + \int_0^1 L(s,v(s))ds\right\},$$

one has from Ref. [245]

$$\omega_d(DS(\omega,u)) = \sup_{\xi_i\in H, |\xi_i|\le 1, i=1,\cdots,d}\left\{\exp\{\sigma W(1) + \int_0^1 \mathrm{Tr}(L(s,v(s))\circ Q_d(s))ds\right\},$$

where $Q_d(s)$ is the rectangular projector from H to $\mathrm{Span}\{V_1(s),\cdots,V_d(s)\}$, and $V_i(s)$ is the solution of eq. (5.1.16) with initial value $V(0) = \xi_i$.

Let $\varphi_i(s), i \in \mathbb{N}$ be an orthonormal basis of H such that $Q_d(s)H = \mathrm{Span}\{\varphi_1(s),\cdots,\varphi_d(s)\}$, then

$$\begin{aligned} \mathrm{Tr}(L(s,v(s))\circ Q_d(s)) &= \sum_{i=1}^d (L(s,v(s))\varphi_i(s),\varphi_i(s)) \\ &\le -\lambda\sum_{i=1}^d \|\varphi_i\|^2 + \left(\nu - \frac{\sigma^2}{2}\right)d \\ &\le -\lambda\sum_{i=1}^d \lambda_i^2 + \left(\nu - \frac{\sigma^2}{2}\right)d. \end{aligned}$$

Denoting $\bar{\omega}_d(\omega) = \exp\{\sigma W(1) - \lambda \sum_{i=1}^{d} \lambda_i^2 + (\nu - \frac{\sigma^2}{2})d\}$ and choosing a suitable d such that $\nu - \frac{\sigma^2}{2} < \frac{\lambda}{d} \sum_{i=1}^{d} \lambda_i^2$, then we have $\omega_d(DS) \le \bar{\omega}_d(\omega)$ and $E(\log(\bar{\omega}_d)) < 0$.

Hence, the following theorem holds.

Theorem 5.1.4. *If there exists a d such that $\nu \le \frac{\sigma^2}{2} + \frac{\lambda}{d} \sum_{i=1}^{d} \lambda_i^2$, then $d_H(\mathcal{A}(\omega)) < d$, P-a.s..*

Now, we prove Lemma 5.1.2 by three steps.

Step 1. Boundedness in L^p, $p \in \mathbb{Z}^+, 1 \le p \le 8$.

Lemma 5.1.3. *Let $\lambda \ge |\alpha|$ and $v(t)$ be the solution of eq.* (5.1.4)*, then for $p \in \mathbb{Z}^+, 1 \le p \le 4$, there is a random variable $I_{2p}(\omega)$ such that*

$$\int_0^1 |v(s)|_{L^{2p}}^{2p} ds \le I_{2p}(\omega), \tag{5.1.17}$$

and for any $m \ge 0$, $E(I_{2p}^m) < \infty$.

Proof. By the invariance of the random attractor $\mathcal{A}(\omega)$, for $v_0 \in \mathcal{A}(\omega)$, there exists a solution $v(t)$ of eq. (5.1.4) with $v(0) = v_0$ such that $v(t) \in \mathcal{A}(\theta_t \omega)$ for any $t \in R$. In order to obtain eq. (5.1.17), we first show that for $p \in \mathbb{Z}^+, 1 \le p \le 3$ and $r > 0$,

$$\int_{t-r}^{t} |v(s)|_{2p+2}^{2p+2} ds \le C \sup_{t-r \le s \le t} e^{-2\sigma W(s)} \int_{t-r-1}^{t} |v(s)|_{2p}^{2p} ds, \tag{5.1.18}$$

where C is a constant. Hereafter, we use $|\cdot|_p = |\cdot|_{L^p}$. Taking inner product of eq. (5.1.4) with $\bar{v}|v|^{2p-2}$, and using the estimate

$$\begin{aligned}
&\Re(\lambda + i\alpha) \int_D \Delta v \bar{v} |v|^{2p-2} dx \\
&= -\lambda p \int_D |\nabla v|^2 |v|^{2p-2} dx - (p-1)\Re(\lambda + i\alpha) \int_D (\bar{v} \nabla v)^2 |v|^{2p-4} dx \\
&\le -\lambda(p - \sqrt{2}p + \sqrt{2}) \int_D |\nabla v|^2 |v|^{2p-2} dx \le 0,
\end{aligned}$$

we obtain

$$\frac{1}{2p} \frac{d}{dt} |v|_{2p}^{2p} + e^{2\sigma W(t)} |v|_{2p+2}^{2p+2} \le \left(\nu - \frac{\sigma^2}{2}\right) |v|_{2p}^{2p}. \tag{5.1.19}$$

Integrating this from s to t leads to

$$\frac{1}{2p} |v(t)|_{2p}^{2p} \le \frac{1}{2p} |v(s)|_{2p}^{2p} + \left(\nu - \frac{\sigma^2}{2}\right) \int_s^t |v(\tau)|_{2p}^{2p} d\tau.$$

Then integrating on s from $t-1$ to t, we obtain

$$|v(t)|_{2p}^{2p} \le \int_{t-1}^{t} |v(s)|_{2p}^{2p} + 2p\left(\nu - \frac{\sigma^2}{2}\right)\int_{t-1}^{t} |v(s)|_{2p}^{2p} ds. \tag{5.1.20}$$

Integrating eq. (5.1.19) from $t-r$ to t, we get

$$\begin{aligned}
&\int_{t-r}^{t} e^{2\sigma W(s)} |v(s)|_{2p+2}^{2p+2} ds \\
&\le \frac{1}{2p}|v(t-r)|_{2p}^{2p} + \left(\nu - \frac{\sigma^2}{2}\right)\int_{t-r}^{t} |v(s)|_{2p}^{2p} ds \\
&\le \left(\frac{1}{2p} + \nu - \frac{\sigma^2}{2}\right)\int_{t-r-1}^{t-r} |v(s)|_{2p}^{2p} ds + \left(\nu - \frac{\sigma^2}{2}\right)\int_{t-r}^{t} |v(s)|_{2p}^{2p} ds \\
&\le \left(1 + 2\nu - \sigma^2\right)\int_{t-r-1}^{t} |v(s)|_{2p}^{2p} ds.
\end{aligned}$$

Hence, we have

$$\int_{t-r}^{t} |v(s)|_{2p+2}^{2p+2} ds \le (1 + 2\nu - \sigma^2) \sup_{t-r\le s\le t} e^{-2\sigma W(s)} \int_{t-r-1}^{t} |v(s)|_{2p}^{2p} ds.$$

Writing $S_i = \sup_{1-i\le s\le 1} e^{-2\sigma W(s)}$, we obtain

$$\begin{aligned}
\int_{0}^{1} |v(s)|_{2p+2}^{2p+2} ds &\le (1 + 2\nu - \sigma^2) S_1 \int_{-1}^{1} |v(s)|_{2p}^{2p} ds \\
&\le (1 + 2\nu - \sigma^2)^2 S_1 S_2 \int_{-2}^{1} |v(s)|_{2p-2}^{2p-2} ds \\
&\le (1 + 2\nu - \sigma^2)^p S_1 \cdots S_p \int_{-p}^{1} |v(s)|_{2}^{2} ds.
\end{aligned}$$

As $v(-3) \in \mathcal{A}(\theta_{-3}\omega)$, we get

$$|v(-3)| \le r_{-3}(\theta_{-3}\omega).$$

Finally, by eq. (5.1.11), the boundedness of $v(t)$ for $-3 \le t \le 1$ in H can be obtained. The proof is completed.

Step 2. Lipshitz property of the solution.

Let $v_i(t)$ $(i = 1, 2)$ be two solutions of eq. (5.1.4) with initial data $v_i(0) = v_i^0$. Setting $g(t) = v_1(t) - v_2(t)$, then $g(t)$ satisfies

$$\frac{dg}{dt} + (\lambda + i\alpha)Ag - \left(\nu - \frac{\sigma^2}{2}\right)g + (k + i\beta)z^{-2}(f(v_1) - f(v_2)) = 0.$$

Taking inner product with g, we have

$$\frac{1}{2}\frac{d}{dt}|g|^2 + \lambda\|g\|^2 - \left(\nu - \frac{\sigma^2}{2}\right)|g|^2 = -\Re(k+i\beta)z^{-2}\int_D (f(v_1) - f(v_2))(\bar{v}_1 - \bar{v}_2)dx.$$

Noting that RHS is bounded by $Cz^{-2}(|v_1|_6^3 + |v_2|_6^3)|g|^2$, then we get

$$|g(t)|^2 \le |g(0)|^2 \exp\left\{\left(\nu - \lambda\lambda_1 - \frac{\sigma^2}{2}\right)t + C\int_0^t z^{-2}(s)(|v_1(s)|_6^3 + |v_2(s)|_6^3)ds\right\}.$$

By eq. (5.1.17), one has

$$|g(1)|^2 \le |g(0)|^2 \exp\left\{\nu - \lambda\lambda_1 - \frac{\sigma^2}{2} + C\sup_{0\le s\le 1} z^{-2}(s)I_6(\omega)\right\}.$$

Finally, we get

$$|v_1(1) - v_2(1)| \le C(\omega)|v_1^0 - v_2^0|,$$

and $E(C(\omega)) < \infty$.

Step 3. Differentiability of $T(\omega)$.

Let $r(t) = v_1(t) - v_2(t) - V(t)$, where v_i $(i = 1, 2)$ are given as above and $V(t)$ satisfies the linear equation (5.1.16) with $L = L(t, v)$ and $h = v_1^0 - v_2^0$. Then $r(t)$ satisfies

$$\begin{aligned}\frac{dr}{dt} + (\lambda + i\alpha)Ar - \left(\nu - \frac{\sigma^2}{2}\right)r \\ = -(k+i\beta)z^{-2}(f(v_1) - f(v_2) - f'(v_2)(v_1 - v_2 - r)).\end{aligned}$$

Taking inner product about the formula with r, we obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|r|^2 + \lambda\|r\|^2 \le \left(\nu - \frac{\sigma^2}{2}\right)|r|^2 - \Re(k+i\beta)z^{-2}\int_D f'(v_2)|r|^2dx \\ - \Re(k+i\beta)z^{-2}\int_D (f(v_1) - f(v_2) - f'(v_2)(v_1 - v_2))\bar{r}dx.\end{aligned} \tag{5.1.21}$$

The second term on the RHS is nonpositive:

$$
\begin{aligned}
&- \Re(k+i\beta)z^{-2}\int_D f'(v_2)|r|^2dx \\
&= - \Re(k+i\beta)\int_D (|v_2|^2|r|^2 + 2v_2\bar{r}\Re(v_2\bar{r}))dx \\
&= - k\int |v_2|^2|r|^2dx - 2k\int (\Re(v_2\bar{r}))^2dx + 2\beta\int \Re(v_2\bar{r})\Im(v_2\bar{r})dx \\
&\le (|\beta|-k)\int |v_2|^2|r|^2dx \le 0.
\end{aligned}
$$

Now, we try to estimate the third term on the RHS of eq. (5.1.21). First, we observe

$$|f(v_1) - f(v_2) - f'(v_2)(v_1 - v_2)| \le C(|v_1|^2 + |v_2|^2)|v_1 - v_2|,$$

hence applying the Hölder inequality and the Sobolev embedding theorem leads to

$$
\begin{aligned}
-\Re(k+i\beta)z^{-2}&\int_D (f(v_1) - f(v_2) - f'(v_2)(v_1 - v_2))\bar{r}dx \\
&\le Cz^{-2}|(|v_1|^2 + |v_2|^2)(v_1 - v_2)|_s|r|_{s^*} \\
&\le Cz^{-4}|(|v_1|^2 + |v_2|^2)(v_1 - v_2)|_s^2 + \varepsilon\|r\|^2,
\end{aligned} \tag{5.1.22}
$$

where $\varepsilon > 0$, $s > 1$ and s^* is the conjugate index of s. Let

$$0 < \delta < 2/3, \quad 1 < s < 8/(6 + 3\delta).$$

It's easy to verify

$$1 < s < 2/(1+\delta), \quad s_1 := 2s(2-\delta)/(2 - s(1+\delta)) < 8.$$

Hence, there holds

$$
\begin{aligned}
|(|v_1|^2 + |v_2|^2)(v_1 - v_2)|_s^s &\le C\int_D (|v_1| + |v_2|)^{3s - s(1+\delta)}|v_1 - v_2|^{s(1+\delta)}dx \\
&\le C(|v_1|_{s_1}^{s(2-\delta)} + |v_2|_{s_1}^{s(2-\delta)})|v_1 - v_2|^{s(1+\delta)}.
\end{aligned}
$$

By eq. (5.1.21), there holds for sufficiently small ε that

$$\frac{1}{2}\frac{d}{dt}|r|^2 \le \left(\nu - \frac{\sigma^2}{2}\right)|r|^2 + Cz^{-4}(|v_1|_{s_1}^{2(2-\delta)} + |v_2|_{s_1}^{2(2-\delta)})|v_1 - v_2|^{2(1+\delta)}.$$

This yields

$$|r(1)|^2 \le C(\omega)\int_0^1 z^{-4}(|v_1|_{s_1}^{2(2-\delta)} + |v_2|_{s_1}^{2(2-\delta)})dsh^{2(1+\delta)}.$$

Denoting

$$\bar{k}_1^2(\omega) = C(\omega) \sup_{0\le t\le 1} z^{-4}(t) \int_0^1 (|v_1(s)|_{s_1}^{2(2-\delta)} + |v_2(s)|_{s_1}^{2(2-\delta)})ds,$$

and choosing $\bar{k}(\omega) = \max\{k_1(\omega), 1\}$ such that $E(\log \bar{k}(\omega)) < \infty$, we complete the proof of Lemma 5.1.2. ■

5.1.3 Generalized SGLE

Moreover, we consider the following stochastic generalized GL equation (see Refs [31, 125], for example):

$$\begin{aligned} du + \kappa u_x dt = (\chi u + (y_1 + iy_2)u_{xx} - (\beta_1 + i\beta_2)|u|^2 u - (\delta_1 + i\delta_2)|u|^4 u \\ - (v_1 + iv_2)|u|^2 u_x - (\mu_1 + i\mu_2)u^2 \bar{u}_x)dt + \Phi dW, \end{aligned} \tag{5.1.23}$$

with the following initial and boundary conditions:

$$u(0,t) = u(1,t) = 0, \quad u(x,t_0) = u_0(x). \tag{5.1.24}$$

W is a two-sided cylindricity Wiener process about time, which is a function valued in $L^2(0,1)$, defined on a complete probability space $(\Omega, \mathscr{F}, P)$. Or equivalently, if $\{\beta_i\}_{i\in\mathbb{Z}}$ are mutually independent Brown motions in $(\Omega, \mathscr{F}, P)$ and $\{e_i\}_{i\in\mathbb{Z}}$ is a set of orthonormal basis in $L^2(0,1)$, then

$$W(t) = \sum_i \beta_i e_i.$$

Φ is a Hilbert–Schmidt operator from $L^2(0,1)$ to some Hilbert space U.

Let z be the solution of the following linear equation:

$$dz = -(y_1 - iy_2)Azdt + \Phi dW, \;\; z(0) = 0. \tag{5.1.25}$$

Then $z \in C([0,\infty), V)$ and $v = u - z$ satisfies the following random differential equation:

$$\begin{aligned} \frac{dv}{dt} &= -\kappa v_x + \chi v - (y_1 + iy_2)Av - (\beta_1 + i\beta_2)|u|^2 u \\ &\quad - (\delta_1 + i\delta_2)|u|^4 u - (v_1 + iv_2)|u|^2 u_x \\ &\quad - (\mu_1 + i\mu_2)u^2 \bar{u}_x - \kappa z_x + \chi z, \\ v(t_0,\omega) &= v_0 = u_0 - z(t_0,\omega) =: \xi. \end{aligned} \tag{5.1.26}$$

It is not difficult to prove that, P-a.s.,

(i) for any $t_0 \in R, \xi \in V$ and any $T < \infty$, there exists a unique solution $v \in C([t_0, T); V) \cap L^2(t_0, T; V)$ of eq. (5.1.26),

(ii) if $\xi \in D(A)$, then $v \in C([t_0, T); V) \cap L^2(t_0, T; D(A))$,

(iii) for any $t \geq t_0$, the mapping $\xi \mapsto v(t)$ is continuous from $V \to V$.

Let $v(t, \omega; t_0, \xi)$ denote the solution of eq. (5.1.26), then

$$S(t, t_0; \omega)u_0 = v(t, \omega; t_0, u_0 - z(t_0, \omega)) + z(t, \omega)$$

defines an RDS.

In order to obtain the existence of random attractor, some a priori estimates are needed. It is shown in Ref. [125] that the following theorem holds.

Theorem 5.1.5. *Let $\kappa > 0, \gamma_1 > 0, \delta_1 > 0$, Φ be a Hilbert–Schmidt operator from $H \to D(A)$, by adaptive selected δ, the RDS generated by the generalized SGLE in V has a global random attractor $\mathcal{A}(\omega)$.*

5.2 Ergodicity for SGL with degenerate noise

In this section, we consider the ergodicity for the SGLE with cubic nonlinearity

$$du = \left[(\lambda + i\alpha)\Delta u + \gamma u - (\kappa + i\beta)|u|^{2\sigma} u\right] dt + \sum_{k=1}^{\infty} h_k(t, u(t)) dW_t^k, \tag{5.2.1}$$

where $\sigma = 1$, subjecting to the initial condition $u(0) = u_0 \in H^1$ and zero Dirichlet boundary conditions on a bounded smooth domain $D \subset R^3$.

In the following, we use $\langle u, v\rangle = \int_D u\bar{v}dx$ to denote the complex inner product in $L^2 = L^2(D)$ and $\langle\cdot,\cdot\rangle_1 = \langle\cdot,\cdot\rangle + \int_D \nabla u \nabla \bar{v} dx$. We use $|\cdot|_p$ to denote the L^p norm and $\|\cdot\|_s$ to denote the H^s Sobolev norm for $s \in R$. Of course, $|\cdot|_2$ and $\|\cdot\|_0$ coincide and is simply denoted as $\|\cdot\|$. $\mathcal{M}(H)$ = the probability measure space on H.

Let $\mathbf{h} = (h_1, \cdots, h_k, \cdots)$. We assume that the mapping

$$\mathbf{R}^+ \times \mathbf{C} \ni (t, u) \mapsto \mathbf{h}(t, u) \in l^2_{\mathbf{C}},$$

where $l^2_{\mathbb{C}}$ denotes the Hilbert space consisting of all sequences of square-summable complex numbers with standard norm $\|\cdot\|_{l^2}$. We now introduce some assumptions on $\mathbf{h}$ and the coefficients.

(**H1**) For any $T > 0$, there exists some constant $C = C_{T,\mathbf{h}} > 0$ such that for any $t \in [0, T]$, $x \in \mathbb{D}$, $u \in \mathbb{C}$, there holds

$$\|\mathbf{h}(t,u)\|_{l^2}^2 \le C|u|^2, \quad \|\partial_u \mathbf{h}(t,u)\|_{l^2} \le C_{\mathbf{h}}.$$

(**H2**) Assume $\gamma < \varrho_1 \lambda$, where ϱ_1 is the first eigenvalue of the Dirichlet problem

$$-\Delta u = \varrho u, \ \text{ in } D, \quad u = 0 \ \text{ on } \partial D.$$

(**H3**) $\beta^2 \le 3\kappa^2$.

(**H4**) There exists some constant $C = C_{T,\mathbf{h}} > 0$ such that

$$\|\partial_u \mathbf{h}(x,u) - \partial_v \mathbf{h}(x,v)\|_{l^2} \le C_{\mathbf{h}} \cdot |u - v|.$$

Lemma 5.2.1. *For any $T > 0$ and $u \in H^1$, there exists some constant C_T such that*

$$\|\boldsymbol{h}(t,u)\|_{L^2(l^2,H^1)}^2 \le C\|u\|_{H^1}^2. \tag{5.2.2}$$

Proof. Noticing that

$$\|\mathbf{h}(t,u)\|_{L^2(l^2,H^1)}^2 = \sum_{k=1}^{\infty} \int |h_k(t,u)|^2 dx + \sum_{k=1}^{\infty} \int |\nabla h_k(t,u)|^2 dx,$$

and

$$\nabla h_k(t,u) = \partial_u h_k(t,u) \nabla u,$$

eq. (5.2.2) is an immediate consequence of assumption **(H1)**. ■

Under assumption **(H1)**, we first show the existence and uniqueness of a strong solution for the 3D SGLE with H^1 initial data, which is stated in the following theorem.

Theorem 5.2.1. *Let* $\mathbf{h}$ *satisfy assumption* **(H1)**. *Then for any $u_0 \in H^1$, there exists a unique $u(t,x)$ such that*

1. $u \in L^2(\Omega, P; C([0,T]; H^1)) \cap L^2(\Omega, P; L^2(0,T; H^2))$ *for any $T > 0$ and*

$$\boldsymbol{E} \sup_{0 \le s \le t} \|u(t)\|_1^2 + \int_0^t \boldsymbol{E}\|u(s)\|_2^2 ds \le C(\|u_0\|_1) \quad \forall t \in [0,T] \tag{5.2.3}$$

2. *It satisfies the SGLE in the mild form for all $t \ge 0$,*

$$u(t) = u_0 + \int_0^t [Au(s) + N(u(s))]ds + \sum_{k=1}^{\infty} \int_0^t h_k(s, u(s)) dW_s^k, \quad P\text{-a.s.}, \tag{5.2.4}$$

where $Au = (\lambda + i\alpha)\Delta u$ and $N(u(s)) = -(\kappa + i\beta)|u(s)|^2 u(s) + \gamma u(s)$.

For fixed initial data $u_0 \in H^1$, we denote this unique solution by $u(t,u_0)$. Then, $\{u(t,u_0) : t \ge 0\}$ forms a strong Markov process with state space H^1. This leads us to the following concepts.

Let $\mathcal{P}_t$ be a Markov transition semigroup in the space $C_b(H^1)$, associated with the SGLE

$$\mathcal{P}_t\varphi(u_0) = \mathbf{E}\varphi(u(t,\cdot;0,u_0)), \quad t \geq 0, u_0 \in H^1, \varphi \in C_b(H^1).$$

Then its dual semigroup $\mathcal{P}_t^*$, defined on the probability measure space $\mathcal{M}(H^1)$, is given by

$$\int_{H^1} \varphi d(\mathcal{P}_t^*\mu) = \int_{H^1} \mathcal{P}_t\varphi d\mu \ \ \forall \varphi \in C_b(H^1), \forall \mu \in \mathcal{M}(H^1).$$

A measure $\mu \in \mathcal{M}(H^1)$ is called invariant provided $\mathcal{P}_t^*\mu = \mu$ for each $t \geq 0$.

We then turn to consider the case when the noise is degenerate, i.e., when only a few Fourier modes of the noise are nonzero. Precisely, for $N \in \mathbf{N}$, let $\Omega = C_0(\mathbf{R}^+, \mathbf{R}^N)$ be the space of all continuous functions with initial values 0, P the standard Wiener measure on $\mathcal{F} := \mathcal{B}(C_0(\mathbf{R}^+, \mathbf{R}^N))$. Then the coordinate process

$$W_t(\omega) := \omega(t), \quad \omega \in \Omega$$

is a standard Wiener process on $(\Omega, \mathcal{F}, P)$. In this case, the SGLE with degenerate noise has the form

$$du(t) = [(\lambda + i\alpha)\Delta u - (\kappa + i\beta)|u|^{2\sigma}u + \gamma u]dt + d\mathbf{w}_t, \tag{5.2.5}$$

where $\mathbf{w}_t = QW_t$ is the noise, and the linear map $Q : \mathbf{R}^N \to H^1$ is given by

$$Qe_i = q_i\mathbf{e}_i, \quad q_i > 0, i = 1, 2, \cdots, N,$$

where e_i is the canonical basis of $\mathbf{R}^N$ and $\mathbf{e}_i$ is orthonormal basis of H^1:

$$\Delta\mathbf{e}_i = -\lambda\mathbf{e}_i,$$

with $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N \to \infty$. Let $\mathcal{E}_0 = \mathrm{tr}QQ^* = \sum_{i=1}^N q_i^2/\lambda_i$ and $\mathcal{E}_1 = \mathrm{tr}QQ^* = \sum_{i=1}^N q_i^2$, then the quadratic variation of $\mathbf{w}_t$ in H and H^1 are given by $[\mathbf{w}]_0(t) = \mathcal{E}_0 t$ and $[\mathbf{w}]_1(t) = \mathcal{E}_1 t$.

For this equation, we have the following:

Theorem 5.2.2. *Assume* **(H1)–(H4)** *hold. Let* $\{\mathcal{P}_t\}$ *be the transition semigroup associated with eq.* (5.2.5)*, then for any sufficiently large N, there exists a unique invariant probability measure associated with* $\{\mathcal{P}_t\}_{t\geq 0}$.

In the proof, we need the following auxiliary proposition [126]. This proposition gives a sufficient condition for a transition semigroup to be an asymptotically strong Feller one.

Proposition 5.2.1. *Let t_n and δ_n be two positive sequences with t_n nondecreasing and δ converging to zero. A semigroup $\mathcal{P}_t$ on a Hilbert space $\mathcal{H}$ is asymptotically strong Feller if for all $\varphi : \mathcal{H} \to \mathbb{R}$ with $\|\varphi\|_\infty$ and $\|\nabla\varphi\|_\infty$ finite,*

$$|\nabla \mathcal{P}_{t_n}\varphi(x)| \leq C(\|x\|)(\|\varphi\|_\infty + \delta_n\|\nabla\varphi\|_\infty)$$

for all n, where $C : \mathbb{R}^+ \to \mathbb{R}$ *is a fixed nondecreasing function.*

5.2.1 Momentum estimate and pathwise uniqueness

In this section, we prove Theorem 5.2.1. We first show the existence of a martingale solution, and then pathwise uniqueness. Then by a theorem of Yamada–Watanabe, we indeed show the existence and uniqueness of a strong solution.

Roughly speaking, a strong solution is one that exists for a given probability space and given stochastic inputs while existence of a weak solution simply ensures that a solution exists on some probability space for some stochastic inputs having the specified distributional properties [145].

5.2.1.1 Higher order momentum estimates

We give some estimates of $\mathbf{E}\|u(t)\|_1^{2p}$ in this section. As can be seen later, these are key estimates when proving the asymptotically strong Feller property of the semigroup $\mathcal{P}_t$.

Lemma 5.2.2. *For all* $p = 1, 2, \cdots$, *integers, the following estimates hold:*

$$\boldsymbol{E}\|u(t)\|^{2p} \leq C(1 + \|u_0\|^{2p})(1 + t^p).$$

Proof. By Itû's formula and then taking real part, we have

$$\|u(t)\|^{2p} = \|u_0\|^{2p} + \sum_{i=1}^{4} I_i(t), \tag{5.2.6}$$

where

$$\begin{aligned}
I_1(t) &= 2p\int_0^t \|u(s)\|^{2(p-1)}\langle u(s), (\lambda + i\alpha)\Delta u(s)\rangle ds \\
&\quad + 2p\int_0^t \|u(s)\|^{2(p-1)}\langle u(s), yu(s)\rangle ds \\
&\quad - 2p\int_0^t \|u(s)\|^{2(p-1)}\langle u(s), (\kappa + i\beta)|u|^{2\sigma}u(s)\rangle ds, \\
I_2(t) &= 2p\sum_{k=1}^{\infty}\int_0^t \|u(s)\|^{2(p-1)}\langle u(s), h_k(u(s))\rangle dW_s^k, \\
I_3(t) &= p\sum_{k=1}^{\infty}\int_0^t \|u(s)\|^{2(p-1)}\|h_k\|^2 ds, \\
I_4(t) &= 2p(p-1)\sum_{k=1}^{\infty}\int_0^t \|u(s)\|^{2(p-2)}|\langle u(s), h_k(u(s))\rangle|^2 ds.
\end{aligned}$$

For I_1, it can be estimated that

$$I_1(t) \le 2p\lambda \int_0^t \|u(s)\|^{2(p-1)}\|u(s)\|_1^2 ds$$
$$+ 2p\gamma \int_0^t \|u(s)\|^{2p} ds - 2p\kappa \int_0^t \|u(s)\|^{2(p-1)}\||u|\|_{L^{2\sigma+2}}^{2\sigma+2} ds.$$

From assumptions **(H1–H2)**, I_3 and I_4 can be estimated that

$$I_3(t) + I_4(t) \le C_p \int_0^t \|u(s)\|^{2p} ds.$$

Taking expectations of eq. (5.2.6), we then have

$$\mathbf{E}\|u(t)\|^{2p} \le \|u_0\|^{2p} - 2p\lambda \int_0^t \mathbf{E}\|u(s)\|^{2(p-1)}\|u(s)\|_1^2 ds$$
$$+ C_p \int_0^t \mathbf{E}\|u(s)\|^{2p} ds - 2p\kappa \int_0^t \mathbf{E}\|u(s)\|^{2(p-1)}\||u|\|_{L^{2\sigma+2}}^{2\sigma+2} ds. \tag{5.2.7}$$

From this inequality, we can easily show the global but exponential growth estimate of the moment. However "exponential growth" is not enough for later purpose to prove ergodicity. Therefore, we use the following iterative method.

Note that from Hölder's inequality, we have

$$\|u\|^2 \le C\|u\|_{L^{2\sigma+2}}^2 \le \varepsilon\|u\|_{L^{2\sigma+2}}^{2\sigma+2} + C_\varepsilon \quad \forall \varepsilon > 0. \tag{5.2.8}$$

When $p = 1$, from eqs (5.2.7) and (5.2.8) (choosing ε sufficiently small)

$$\mathbf{E}\|u(t)\|^2 + \int_0^t \mathbf{E}\|u(s)\|_1^2 ds \le \|u_0\|^2 + C \cdot t \le C(1 + \|u_0\|^2)(1 + t). \tag{5.2.9}$$

When $p = 2$, using the result when $p = 1$, we have similarly

$$\begin{aligned}\mathbf{E}\|u(t)\|^4 &\le \|u_0\|^4 - 2p\lambda \int_0^t \mathbf{E}\|u(s)\|^2\|u(s)\|_1^2 ds \\ &\quad - \boldsymbol{\kappa} \int_0^t \mathbf{E}\|u(s)\|^2\|u(s)\|_{L^{2(\sigma+1)}}^{2(\sigma+1)} + C_\varepsilon \int_0^t \mathbf{E}\|u(s)\|^2 ds \\ &\le \|u_0\|^4 + C \int_0^t \mathbf{E}\|u(s)\|^2 ds \\ &\le C(1 + \|u_0\|^4)(1 + t^2).\end{aligned} \tag{5.2.10}$$

By induction, we have for general $p > 0$

$$\mathbf{E}\|u(t)\|^{2p} \le C(1 + \|u_0\|^{2p})(1 + t^p), \tag{5.2.11}$$

which finishes the proof of this lemma. ■

We now prove the following momentum estimates for $\|u(t)\|_1$.

Lemma 5.2.3. *Let **(H3)** hold, we have for all $p = 1, 2, \cdots$, integers that*

$$\boldsymbol{E}\|u(t)\|_1^{2p} \leq C(1 + \|u_0\|_1^{2p})(1 + t^{p+1}).$$

Proof. By Itô's formula, we have

$$\|u(t)\|_1^{2p} = \|u_0\|_1^{2p} + \sum_{i=1}^{4} Re(I_i(t)), \tag{5.2.12}$$

where

$$\begin{aligned} I_1(t) &= 2p \int_0^t \|u(s)\|_1^{2(p-1)} \langle u(s), (\lambda + i\alpha)\Delta u(s)\rangle_1 ds \\ &\quad + 2p \int_0^t \|u(s)\|_1^{2(p-1)} \langle u(s), \gamma u(s)\rangle_1 ds \\ &\quad - 2p \int_0^t \|u(s)\|_1^{2(p-1)} \langle u(s), (\kappa + i\beta)|u|^{2\sigma} u(s)\rangle_1 ds, \\ &=: I_1^{(1)}(t) + I_1^{(2)}(t) + I_1^{(3)}(t), \end{aligned}$$

and

$$\begin{aligned} I_2(t) &= 2p \sum_{k=1}^{\infty} \int_0^t \|u(s)\|_1^{2(p-1)} \langle u(s), h_k(u(s))\rangle_1 dW_s^k, \\ I_3(t) &= p \sum_{k=1}^{\infty} \int_0^t \|u(s)\|_1^{2(p-1)} \|h_k\|_1^2 ds, \\ I_4(t) &= 2p(p-1) \sum_{k=1}^{\infty} \int_0^t \|u(s)\|_1^{2(p-2)} |\langle u(s), h_k(u(s))\rangle_1|^2 ds. \end{aligned}$$

For the term $I_1^{(1)}(t)$ and $I_1^{(2)}(t)$, we have

$$\begin{aligned} Re(I_1^{(1)}(t)) &= -2p\lambda \int_0^t \|u(s)\|_1^{2(p-1)} \|\nabla u(s)\|_1^2 ds \\ Re(I_1^{(2)}(t)) &= 2p\gamma \int_0^t \|u(s)\|_1^{2p} ds. \end{aligned} \tag{5.2.13}$$

Recalling $\beta^2 \leq \kappa^2(\sigma^2 + 2\sigma)$ and noting that

$$\begin{aligned} & Re\langle u, (\kappa + i\beta)|u|^{2\sigma} u\rangle_1 \\ &= Re\langle \nabla u, (\kappa + i\beta)\nabla(|u|^{2\sigma} u)\rangle + Re\langle u, (\kappa + i\beta)|u|^{2\sigma} u\rangle \\ &\geq \kappa(\sigma + 1)\||u|^{\sigma}|\nabla u|\|^2 - Re(\kappa + i\beta) \int_{\mathcal{O}} |u|^{2(\sigma-1)}(u)^2(\nabla u^*)^2 dx + \kappa\|u\|_{L^{2(\sigma+1)}}^{2(\sigma+1)}, \end{aligned}$$

we have

$$I_1^{(3)}(t) \le -\kappa \|u\|_{L^{2(\sigma+1)}}^{2(\sigma+1)}.$$

For the last two terms, we have

$$I_3(t) + I_4(t) \le Cp(2p+1)\int_0^t \|u(s)\|_1^{2p} ds. \tag{5.2.14}$$

Taking expectations of eq. (5.2.12), using interpolation inequality and Young's inequality, we then have

$$\begin{aligned}\mathbf{E}\|u(t)\|_1^{2p} &\le \|u_0\|_1^{2p} - 2p\lambda \int_0^t \mathbf{E}\|u\|_1^{2(p-1)}\|u\|_2^2 ds + C_p \int_0^t \mathbf{E}\|u\|_1^{2p} ds \\ &\le \|u_0\|_1^{2p} - p\lambda \int_0^t \mathbf{E}\|u\|_1^{2(p-1)}\|u\|_2^2 ds + C_p \int_0^t \mathbf{E}\|u\|_1^{2(p-1)}\|u\|_0^2 ds \\ &\le \|u_0\|_1^{2p} - \frac{1}{2}p\lambda \int_0^t \mathbf{E}\|u\|_1^{2(p-1)}\|u\|_2^2 ds + C_p \int_0^t \mathbf{E}\|u\|_0^{2p} ds. \end{aligned} \tag{5.2.15}$$

Using the estimate in Lemma 5.2.2, we have

$$\mathbf{E}\|u(t)\|_1^{2p} \le C(1 + \|u_0\|_1^{2p})(1 + t^{p+1}).$$

We therefore end the proof. ∎

Corollary 5.2.1. *Under the above assumptions, there exists some constant C_T, such that*

$$\boldsymbol{E} \sup_{t\in[0,T]} \|u(t)\|_{H^1}^2 + \int_0^T \boldsymbol{E}\|u(s)\|_{H^2}^2 ds \le C_T.$$

Proof. Let

$$M(t) = 2\sum_{k=1}^{\infty} \int_0^t \langle u(s), h_k(u(s))\rangle_1 dW_s^k,$$

then, similar to the above lemma, we have

$$\|u(t)\|_1^2 + 2\lambda \int_0^t \|u\|_2^2 ds \le \|u_0\|_1^2 + C\int_0^t \|u\|_1^2 ds + M(t). \tag{5.2.16}$$

Taking expectations and using Gronwall's inequality, we obtain

$$\sup_{t\in[0,T]} \mathbf{E}\|u(t)\|_{H^1}^2 + \int_0^t \mathbf{E}\|u\|_{H^2}^2 ds \le C = C(T).$$

From eq. (5.2.16), using Burkholder–Davis–Gundy (BDG)'s inequality, we have from Lemma 5.2.1 and Young's inequality

$$\begin{aligned}\mathbf{E}\sup_{t\in[0,T]}\|u(t)\|_{H^1}^2 &\le C(T) + C\mathbf{E}\left(\int_0^T \|\mathbf{h}(u(s))\|_{L^2(l^2;L^0)}^2\|u(s)\|_{H^2}^2 ds\right)^{1/2}\\ &\le C(T) + \varepsilon\mathbf{E}\sup_{t\in[0,T]}\|\mathbf{h}(u(s))\|_{L^2(l^2;L^0)}^2 + C_\varepsilon\int_0^T \mathbf{E}\|u(s)\|_{H^2}^2 ds\\ &\le C(T,\varepsilon) + C_T\varepsilon\mathbf{E}\sup_{t\in[0,T]}\|(u(s))\|_{L^2}^2.\end{aligned}$$

Choosing $\varepsilon = 1/(2C_T)$, we then obtain

$$\mathbf{E}\sup_{t\in[0,T]}\|u(s)\|_{H^1}^2 \le C = C(T).$$

We complete the proof. ∎

5.2.1.2 Existence of martingale solutions

Definition 5.2.1. *We say that eq.* (5.2.1) *has a weak solution with initial law* $\nu \in \mathcal{M}(H^1)$ *if there exists a stochastic basis* $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\}_{t\ge 0})$, *an* H^1*-valued continuous* $\{\mathcal{F}_t\}$*-adapted stochastic process u and an infinite-dimensional sequence of independent standard* $\{\mathcal{F}_t\}$ *Brownian motions* $\{W^k(t) : t \ge 0, k \in \mathbb{N}\}$ *such that*

(i) $u(0)$ *has law* ν *in* H^1,

(ii) *for almost all* $\omega \in \Omega$ *and any* $T > 0$, $u(\cdot,\omega) \in L^2([0,T]; H^2)$,

(iii) *it satisfies the SGLE in the mild form*

$$u(t) = u_0 + \int_0^t [(\lambda + i\alpha)\Delta u(s) + N(u(s))]ds + \sum_{k=1}^{\infty}\int_0^t h_k(s,u(s))dW_s^k, \tag{5.2.17}$$

where $N(u(s)) = -(\kappa + i\beta)|u(s)|^{2\sigma}u(s) + \gamma u(s)$ *for all* $t \ge 0$, *P-a.s.*

This solution is denoted by $(\Omega, \mathcal{F}, P;\ \{\mathcal{F}_t\}_{t\ge 0}; W, u)$.

We first prove the following a priori estimates.

Lemma 5.2.4. *For any* $T > 0$, *there exists a positive constant* $C > 0$, *such that*

$$\sup_{t\in[0,T]}\boldsymbol{E}\|u(t)\|_0^2 + \int_0^T \boldsymbol{E}\|\nabla u(s)\|_0^2 ds \le C, \tag{5.2.18}$$

and

$$\sup_{t\in[0,T]} \boldsymbol{E}\|u(t)\|_1^2 + \int_0^T \boldsymbol{E}\|\nabla u(s)\|_2^2 ds \le C. \tag{5.2.19}$$

Proof. This lemma indeed has been proved in the previous section. ∎

With this lemma, by Galerkin's approximation and then the tightness method in Ref. [96] and Skorohod theorem in Ref. [140], we can show the existence of a martingale solution. Since what we care about most is ergodicity of invariant measures, and the details of existence of martingale solutions are omitted here for simplicity.

Remark 5.2.1. *It suffices to obtain the global existence of a martingale solution with the estimate in Lemma 5.2.4. However, this estimate is too weak to show the ergodicity for an invariant measure since it admits growing of solutions with time.*

5.2.1.3 Pathwise uniqueness

Generally there are two types of uniqueness in stochastic differential equations, namely *strong uniqueness* and *uniqueness in the sense of probability law* [145]. For our purpose, we only give the definition of pathwise uniqueness here.

Definition 5.2.2. *Given two weak solutions of eq.* (5.2.1) *defined on the same probability space together with the same Brownian motion*

$$(\Omega, \mathcal{F}, P;\ \{\mathcal{F}_t\}_{t\ge0}; W, u_1) \quad (\Omega, \mathcal{F}, P;\ \{\mathcal{F}_t\}_{t\ge0}; W, u_2),$$

if $P\{u_1(0) = u_2(0)\} = 1$, *then* $P\{u_1(t,\omega) = u_2(t,\omega), \forall t \ge 0\} = 1$.

We now prove the following strong uniqueness result. This is achieved by the technique of stopping times.

Theorem 5.2.3. *Let* $\sigma = 1$. *The strong uniqueness holds for the 3D cubic SGLE* (5.2.1) *with* H^1*-initial data.*

Proof. Let u_1 and u_2 be two weak solutions of eq. (5.2.1) defined on the same probability space together with the same Brownian motion and starting from the same initial data u_0. For any $T > 0$ and $R > 0$, we define the stopping time

$$\tau_R := \inf\{t \ge 0 : \|u_1(t,u_0)\|_1 \vee \|u_2(t,u_0)\|_1 \ge R\}.$$

Let $u(t) = u_1(t) - u_2(t)$, then by Itô's formula, we have

$$\begin{aligned}\|u(t)\|^2 = & -2\lambda\int_0^t \|\nabla u\|^2 ds + 2\Re\left[\int_0^t \langle N(u_1)-N(u_2), u\rangle ds\right] \\ & + 2\sum_{k=1}^{\infty}\int_0^t \langle h_k(t,u_1)-h_k(t,u_2), u\rangle dW_s^k \\ & + \sum_{k=1}^{\infty}\int_0^t \|h_k(s,u_1(s))-h_k(s,u_2(s))\|^2 ds \\ =: & \, I_1(t)+I_2(t)+I_3(t)+I_4(t).\end{aligned} \tag{5.2.20}$$

For I_1, we have

$$I_1(t\wedge\tau_R) = -2\lambda\int_0^{t\wedge\tau_R}\|u\|_1^2 ds + 2\lambda\int_0^{t\wedge\tau_R}\|u\|^2 ds. \tag{5.2.21}$$

For I_2, by direct calculations, it is easy to see that

$$\begin{aligned} I_2(t) &= 2\gamma\int_0^t \|u\|^2 ds - 2\Re(\kappa+i\beta)\int_0^t \langle |u_1|^2u_1 - |u_2|^2u_2, u(s)\rangle ds \\ &\le 2\gamma\int_0^t \|u\|^2 ds + C_{\kappa,\beta}\int_0^t \big\| |u|(|u_1|+|u_2|)\big\|^2 ds.\end{aligned}$$

By Hölder's inequality and Sobolev's embedding inequality, we have

$$\begin{aligned}\||u|(|u_1|+|u_2|)\|^2 &\le 2\|u\|_{L^3}^2(\|u_1\|_{L^6}^2+\|u_2\|_{L^6}^2) \\ &\le C\|u\|_1\|u\|(\|u_1\|_1^2+\|u_2\|_1^2) \\ &\le CR^2\|u\|_1\|u\|.\end{aligned}$$

Therefore, using Young's inequality again, we have

$$I_2(t\wedge\tau_R) \le \lambda\int_0^{t\wedge\tau_R}\|u\|_1^2 ds + C_R\int_0^{t\wedge\tau_R}\|u\|^2 ds. \tag{5.2.22}$$

Finally, for I_4, we have from assumption **(H1)** that

$$I_4(t\wedge\tau_R) \le C\int_0^{t\wedge\tau_R}\|u\|^2 ds. \tag{5.2.23}$$

Therefore, taking expectations of eq. (5.2.20) and using inequalities (5.2.21), (5.2.22), and (5.2.23), we have for any $t \in [0,T]$

$$\begin{aligned} & \mathbf{E}\|u(t\wedge\tau_R)\|^2 + \lambda\mathbf{E}\int_0^{t\wedge\tau_R}\|u(s)\|_1^2 ds \\ & \le C_R\mathbf{E}\int_0^{t\wedge\tau_R}\|u(s)\|^2 ds \le C_R\int_0^t \mathbf{E}\|u(s\wedge\tau_R)\|^2 ds.\end{aligned} \tag{5.2.24}$$

Applying Gronwall's inequality, we know that for any $t \in [0, T]$

$$\mathbf{E}\|u(t \wedge \tau_R)\|^2 = 0. \tag{5.2.25}$$

Since $\tau_R \to \infty$ as $R \to \infty$, uniqueness then follows by letting $R \to \infty$ and dominated convergence theorem. ∎

A result, named after Yamada and Watanabe, shows that weak existence plus pathwise uniqueness implies the existence of a unique strong solution, see Ref. [255]. An introduction of this can be found in Karatzas and Shreve [145]. Extensions are made by several authors, the interested readers may refer to Refs [91, 162]. Note that both weak existence of solutions and strong uniqueness have been obtained for the SGLE, we indeed have proven Theorem 5.2.1.

5.2.2 Invariant measures

In this section, we show the existence of an invariant measure following Krylov–Bogolyubov theorem [76]. For fixed $u_0 \in H^1$, we denote the unique solution in Theorem 5.2.1 by $u(t; u_0)$. Then $\{u(t, u_0) : t \geq 0, u_0 \in H^1\}$ forms a strong Markov process with state space H^1. For two initial data $u_{01}, u_{02} \in H^1$, we denote $u_i = u(t, u_{0i})$ the solutions starting from u_{0i}, $i = 1, 2$. In the following, we first show some stability result. Let $R > 0$, we define

$$\tau_R := \inf\{t \in [0, T] : \|u(t, u_{01})\|_1 \vee \|u(t, u_{02})\|_1 \geq R\}.$$

Lemma 5.2.5. *Let assumption* **(H1)** *hold, then there exists constant $C = C_{t,R}$ such that*

$$\mathbf{E}\|u(t \wedge \tau_R; u_{01}) - u(t \wedge \tau_R; u_{02})\|_1^2 \leq C_{t,R}\|u_{01} - u_{02}\|_1^2. \tag{5.2.26}$$

Proof. Let $w(t) = u(t, u_{01}) - u(t, u_{02})$. Multiplying the equation with $(I - \Delta)w$, using integration by parts formula and then taking the real part, we deduce

$$\begin{aligned} \|w(t \wedge \tau_R)\|_1^2 = \|w(0)\|_1^2 &+ 2\int_0^{t\wedge\tau_R} \langle Au_1(s) - Au_2(s), w(s)\rangle_1 ds \\ &+ 2\int_0^{t\wedge\tau_R} \langle N(u_1(s)) - N(u_2(s)), w(s)\rangle_1 ds \\ &+ 2\sum_{k=1}^{\infty}\int_0^{t\wedge\tau_R} \langle h_k(u_1(s)) - h_k(u_2(s)), w(s)\rangle_1 ds \\ &+ \sum_{k=1}^{\infty}\int_0^{t\wedge\tau_R} \|h_k(u_1(s)) - h_k(u_2(s))\|_1^2 ds \\ =: \|w(0)\|_1^2 &+ I_1(t) + I_2(t) + I_3(t) + I_4(t), \end{aligned} \tag{5.2.27}$$

where Au and $N(u)$ are as in Theorem 5.2.1.

For I_1, we have

$$
\begin{aligned}
I_1(t \wedge \tau_R) =& 2\lambda \int_0^{t\wedge\tau_R} \langle \Delta w(s), w(s)\rangle_1 ds \\
&= -2\lambda \int_0^{t\wedge\tau_R} \langle (I-\Delta)w(s), (I-\Delta)w(s)\rangle ds + 2\lambda \int_0^{t\wedge\tau_R} \langle w(s), (I-\Delta)w(s)\rangle ds \\
&= -2\lambda \int_0^{t\wedge\tau_R} \|w(s)\|_{H^2} ds + 2\lambda \int_0^{t\wedge\tau_R} \|w(s)\|_{H^1} ds
\end{aligned}
$$

For I_2, we have by mean value theorem

$$
\begin{aligned}
I_2(t \wedge \tau_R) &= 2\gamma \int_0^{t\wedge\tau_R} \langle w(s), (I-\Delta)w(s)\rangle_0 ds \\
&\quad - 2\Re(\kappa + i\beta) \int_0^{t\wedge\tau_R} \langle |u_1|^2 u_1 - |u_2|^2 u_2, (I-\Delta)w(s)\rangle ds \\
&\leq C \int_0^{t\wedge\tau_R} \|w(s)\|_1^2 ds + \frac{\lambda}{2} \int_0^{t\wedge\tau_R} \|w(s)\|_2^2 ds + C \int_0^{t\wedge\tau_R} \||w||U|^2\|^2 ds,
\end{aligned}
$$

where $U = |u_1| + |u_2|$. However, by the Sobolev embedding theorem,

$$
\begin{aligned}
C \int_0^{t\wedge\tau_R} \||w||U|^2\|^2 ds &\leq C \int_0^{t\wedge\tau_R} \|w\|_{L^\infty}^2 \|U\|_{L^4}^4 ds \\
&\leq C_R \int_0^{t\wedge\tau_R} \|w\|_2 \|w\|_0 ds \\
&\leq \frac{\lambda}{2} \int_0^{t\wedge\tau_R} \|w(s)\|_2^2 ds + C_R \int_0^{t\wedge\tau_R} \|w\|^2 ds. \qquad (5.2.28)
\end{aligned}
$$

Finally, for the last term I_4, we have

$$
\begin{aligned}
I_4(t \wedge \tau_R) &\leq \int_0^{t\wedge\tau_R} \|\nabla \mathbf{h}(s, u_1(s)) - \nabla \mathbf{h}(s, u_2(s))\|^2 ds \\
&\quad + \int_0^{t\wedge\tau_R} \|\mathbf{h}(s, u_1(s)) - \mathbf{h}(s, u_2(s))\|^2 ds =: J_1 + J_2.
\end{aligned}
$$

Recalling assumptions **(H1)** and **(H2)**, by chain rule, Sobolev's embedding and interpolation inequality, we have

$$
\begin{aligned}
&\|\nabla \mathbf{h}(s, u_1(s)) - \nabla \mathbf{h}(s, u_2(s))\|^2 \\
&\leq \int \|\partial_u \mathbf{h}(u_1) \cdot \nabla u_1 - \partial_u \mathbf{h}(u_1) \cdot \nabla u_2 + \partial_u \mathbf{h}(u_1) \cdot \nabla u_2 - \partial_u \mathbf{h}(u_2) \cdot \nabla u_2\|_{l^2}^2 dx \\
&\leq C_{\mathbf{h}}^2 \|\nabla w\|^2 + C_{\mathbf{h}}^2 \int |u_1 - u_2|^2 |\nabla u_2|^2 dx
\end{aligned}
$$

$$
\begin{aligned}
&\le C_{\mathbf{h}}^2\|\nabla w\|^2 + C_{\mathbf{h}}^2\|u_2\|_1^2|w|_{L^\infty}^2 \\
&\le C_{\mathbf{h}}^2\|w\|_1^2 + C_{\mathbf{h},R}\|u_2\|_1^2\|w\|_2\|w\|_1 \\
&\le \frac{\lambda}{2}\|w\|_2^2 + C_{\mathbf{h},R}\|w\|_1^2.
\end{aligned}
$$

Therefore,

$$
J_1 \le \int_0^{t\wedge\tau_R} \frac{\lambda}{2}\|w\|_2^2 + C_{\mathbf{h},R}\|w\|_1^2 ds.
$$

Note also that from assumption **(H1)**, J_2 can be bounded by

$$
J_2 \le C_{\mathbf{h}}^2 \int_0^{t\wedge\tau_R} \|w\|^2 ds. \tag{5.2.29}
$$

Therefore, collecting all the estimates for $I_1 - I_4$, and taking expectation of eq. (5.2.27), we get

$$
\begin{aligned}
\mathbf{E}\|w(t\wedge\tau_R)\|_1^2 &\le \|u_{01}-u_{02}\|_1^2 + C_{\mathbf{h},R}\int_0^{t\wedge\tau_R}\|w(s)\|_1^2 ds \\
&\le \|u_{01}-u_{02}\|_1^2 + C_{\mathbf{h},R}\int_0^{t}\mathbf{E}\|w(s\wedge\tau_R)\|_1^2 ds.
\end{aligned} \tag{5.2.30}
$$

By Gronwall's lemma, the stability estimate (5.2.26) holds. ■

Now we consider the transition semigroup associated with $\{u(t,u_0)\}$. Let $C_b(H^1)$ denote the set of all bounded and locally uniformly continuous functions on H^1. Clearly, under the norm

$$
\|\varphi\|_\infty := \sup_{u\in H^1}|\varphi(u)|,
$$

$C_b(H^1)$ is a Banach space. For $t > 0$, the semigroup $\mathcal{P}_t$ associated with $\{u(t,u_0) : t \ge 0, u_0 \in H^1\}$ is defined by

$$
(\mathcal{P}_t\varphi)(u_0) = \mathbf{E}(\varphi(u(t,u_0))), \quad \varphi \in C_b(H^1).
$$

Theorem 5.2.4. *Let $\sigma \le 1$ and assumption* **(H1)** *holds, for every $t > 0$, $\mathcal{P}_t$ maps $C_b(H^1)$ into itself, i.e., $\{\mathcal{P}_t\}_{t\ge 0}$ is a Feller semigroup on $C_b(H^1)$.*

Proof. Let $\varphi \in C_b(H^1)$ be given. By definition, $\mathcal{P}_t\varphi$ is bounded on H^1. To see that it is also locally uniformly continuous, it suffices to show that for any $\varepsilon > 0$, $t > 0$ and $m \in \mathbf{N}$, there exists $\delta > 0$ such that

$$\sup_{\substack{u_{01},u_{02}\in B_m\\ \|u_{01}-u_{02}\|\le\delta}} |\mathcal{P}_t\varphi(u_{01}) - \mathcal{P}_t\varphi(u_{02})| < \varepsilon, \tag{5.2.31}$$

where $B_m := \{u \in H^1 : \|u\|_1 \le m\}$. As before, we define

$$\tau_R = \inf\{t \ge 0 : \|u(t; u_{01})\|_1 \vee \|u(t; u_{02})\|_1 \ge R\}.$$

First, for solutions $u(t; u_{0i})(i = 1, 2)$, we know from eq. (5.2.3) that

$$\begin{aligned}
&\mathbf{E}|\varphi(u(t; u_{0i})) - \varphi(u(t \wedge \tau_R; u_{0i}))| \le 2\|\varphi\|_\infty P\{\tau_R < t\}\\
&\le 2\|\varphi\|_\infty \sup_{u_{0i}\in B_m} \mathbf{E}\left[\sup_{s\in[0,t]} \|u(s; u_{0i})\|^2\right] /R^2\\
&\le 2C_{t,m}\|\varphi\|_\infty/R^2.
\end{aligned} \tag{5.2.32}$$

Therefore, we can choose $R > m$ sufficiently large such that for $u_{01}, u_{02} \in B_m$

$$\mathbf{E}|\varphi(u(t; u_{0i})) - \varphi(u(t \wedge \tau_R; u_{0i}))| \le \varepsilon/4, \quad i = 1, 2. \tag{5.2.33}$$

Second, since φ is locally uniformly continuous, for the above fixed R, there exists $\delta_R > 0$ such that for any $u_1, u_2 \in B_m$ with $\|u_1 - u_2\|_{H^1} \le \delta_R$,

$$|\varphi(u_1) - \varphi(u_2)| \le \varepsilon/4.$$

Therefore, for $u_{01}, u_{02} \in B_m$ with $\|u_{01} - u_{02}\|^2_{H^1} \le \frac{\varepsilon\delta_R^2}{8C_{t,R}\|\varphi\|_\infty}$, by the stability result in Lemma 5.2.5 and Chebyshev inequality, we have

$$\begin{aligned}
&\mathbf{E}|\varphi(u(t \wedge \tau_R; u_{01})) - \varphi(u(t \wedge \tau_R; u_{02}))|\\
&= \int_{\Omega^+_{\delta_R}} |\varphi(u(t \wedge \tau_R; u_{01})) - \varphi(u(t \wedge \tau_R; u_{02}))|P(d\omega)\\
&\quad + \int_{\Omega^-_{\delta_R}} |\varphi(u(t \wedge \tau_R; u_{01})) - \varphi(u(t \wedge \tau_R; u_{02}))|P(d\omega)\\
&\le \varepsilon/4 + 2\|\varphi\|_\infty P\{\|u(t \wedge \tau_R; u_{01}) - u(t \wedge \tau_R; u_{01})\|_{H^1} > \delta_R\}\\
&\le \varepsilon/4 + 2\|\varphi\|_\infty \frac{\mathbf{E}\|u(t \wedge \tau_R; u_{01}) - u(t \wedge \tau_R; u_{01})\|^2_{H^1}}{\delta_R^2} \le \varepsilon/2,
\end{aligned} \tag{5.2.34}$$

where $\Omega^+_{\delta_R} := \{\omega : \|u(t \wedge \tau_R; u_{01}) - u(t \wedge \tau_R; u_{02})\|_{H^1} \ge \delta_R\}$ and $\Omega^-_{\delta_R} := \{\omega : \|u(t \wedge \tau_R; u_{01}) - u(t \wedge \tau_R; u_{02})\|_{H^1} < \delta_R\}$. Combining eqs (5.2.33) and (5.2.34), eq. (5.2.31) is proved by choosing $\delta = \frac{\sqrt{\varepsilon}\delta_R}{2\sqrt{2C_{t,R}\|\varphi\|_\infty}}$. ■

Theorem 5.2.5. *Let assumptions* **(H1)–(H4)** *hold, then there exists an invariant measure μ_* associated with the semigroup $\mathcal{P}_t$ such that for any $t \geq 0$ and $\varphi \in C_b(H^1)$*

$$\int_{H^1} \mathcal{P}_t\varphi(u)\mu_*(du) = \int_{H^1} \varphi(u)\mu_*(du).$$

Proof. Since $H^2 \hookrightarrow H^1$ compactly, by the classical Krylov–Bogolyubov theorem [76], to show the existence of an invariant measure it is enough to show that:

For any $\varepsilon > 0$, there exists $M > 0$ such that for all $T > 1$,

$$\frac{1}{T}\int_0^T P(\|u(s)\|_{H^2}^2 > M)ds < \varepsilon. \tag{5.2.35}$$

From eq. (5.2.9) and the first inequality of eq. (5.2.15), we know that

$$\begin{aligned} \mathbf{E}\|u(t)\|_1^2 &\leq \|u_0\|_1^2 - 2\lambda\int_0^t \mathbf{E}\|u\|_2^2 ds + C\int_0^t \mathbf{E}\|u\|_1^2 ds \\ &\leq \|u_0\|_{H^1}^2 - 2\lambda\int_0^t \mathbf{E}\|u\|_2^2 ds + C\cdot t. \end{aligned} \tag{5.2.36}$$

Therefore, there exists some constant C dependent on the parameters and the initial data u_0, but independent of t, such that for any $t \geq 1$,

$$\frac{1}{t}\int_0^t \mathbf{E}\|u(s)\|_2^2 ds \leq C. \tag{5.2.37}$$

By a standard argument of contradiction, eq. (5.2.35) is proved, which follows the existence of an invariant measure. ∎

5.2.3 Ergodicity

In this section, we prove Theorem 5.2.2. The following proposition that was proved in Hairer and Mattingly [126] will be useful in our analysis.

Proposition 5.2.2. *Let $\mathcal{P}_t$ be an asymptotically strong Feller Markov semigroup and there exists a point x such that $x \in \operatorname{supp}\mu$ for every invariant probability measure μ of $\mathcal{P}_t$, then there exists at most one invariant probability measure for $\mathcal{P}_t$.*

To apply this proposition to the SGLE, we divide the following proof into two parts. In the first one, we show that the semigroup $\mathcal{P}_t$ associated with the SGLE (5.2.5) with degenerate noise enjoys the asymptotically strong Feller property. In the second part, we show that 0 belongs to the support of any invariant measure for $\mathcal{P}_t$.

5.2.3.1 Asymptotically strong Feller property

To emphasize the dependence of the solutions on the noise realization, we denote $u(t,\omega;u_0)=\Phi_t(\omega,u_0)$. In other words, $\Phi_t : C([0,t];\mathbf{R}^N)\times H^1\to H^1$ is the solution map such that $\Phi_t(\omega,u_0)$ is a solution with initial data u_0 and noise realization ω.

Given $v\in L^2_{\mathrm{loc}}(\mathbf{R}^+,\mathbf{R}^N)$, the Malliavin derivative of the H^1-valued random variable $\Phi_t(\omega,u_0)$ with respect to ω in the direction v is given by

$$\mathfrak{D}^v u(t,\omega;u_0)=\lim_{\varepsilon\to 0}\frac{\Phi_t(\omega+\varepsilon V,u_0)-\Phi_t(\omega,u_0)}{\varepsilon},$$

where $V(t)=\int_0^t v(r)dr$ and the limit holds almost surely with respect to the Wiener measure. For $0\le s<t$, $\mathcal{J}_{s,t}\xi$ is the solution of the linearized equation

$$\partial_t\mathcal{J}_{s,t}\xi=(\lambda+i\alpha)\Delta\mathcal{J}_{s,t}\xi+\mathcal{N}(u(t,\omega;u_0),\mathcal{J}_{s,t}\xi),\quad \mathcal{J}_{s,s}\xi=\xi, \tag{5.2.38}$$

where $\mathcal{N}$ is linear with respect to the second argument and is given by

$$\mathcal{N}(\eta,\xi)=\gamma\xi-(\kappa+i\beta)\left[|\eta|^2\xi+2\Re(\bar{\eta}\cdot\xi)\eta\right]. \tag{5.2.39}$$

Note also that $\mathcal{J}_{s,t}$ enjoys the cocycle property $\mathcal{J}_{s,t}=\mathcal{J}_{s,r}\mathcal{J}_{r,t}$ for $r\in[s,t]$.

When $s=0$, we simply write $\mathcal{J}_t\xi=\mathcal{J}_{0,t}\xi$. It is not difficult to show that for every ω,

$$(\mathcal{J}_t\xi)(\omega)=\lim_{\varepsilon\downarrow 0}\frac{\Phi_t(\omega,u_0+\varepsilon\xi)-\Phi_t(\omega,u_0)}{\varepsilon}. \tag{5.2.40}$$

The proof of this limit is given at the end of this section.

Observe that v can be random and possibly nonadapted to the filtration generated by the increments of W. If we set $\mathcal{A}_t v=\mathfrak{D}^v u(t,\omega;u_0)$, then

$$\partial_t\mathcal{A}_t v=(\lambda+i\alpha)\Delta\mathcal{A}_t v+\mathcal{N}(u(t,\omega;u_0),\mathcal{A}_t v)+Qv(t),\quad \mathcal{A}_0 v=0. \tag{5.2.41}$$

Since $\mathcal{N}(\cdot,\cdot)$ is linear with respect to the second argument, by the formula of variation, $\mathcal{A}_t : L^2([0,t];\mathbf{R}^N)\to H^1$ is given by

$$\mathcal{A}_t v=\int_0^t \mathcal{J}_{s,t}Qv(s)ds.$$

Roughly speaking, $\mathcal{J}_t\xi$ is the perturbation of $u(t,\omega;u_0)$ caused by initial perturbation ξ of u_0, while $\mathcal{A}_t v$ is the perturbation at time t caused by an infinitesimal variation in the Wiener space over interval $[0,t]$. If we can find such a v such that they cause the same effect, i.e., $\mathcal{J}_t\xi=\mathcal{A}_t v$, then it can be shown in the spirit of Hairer and Mattingly [126] that $\mathcal{P}_t$ is strong Feller and ergodicity follows. However, in most cases with degenerate noise, such property does not hold, mainly due to the noninvertibility of the Malliavin matrix, see Ref. [76].

Therefore, we consider the difference $\rho(t) = \mathcal{J}_t\xi - \mathcal{A}_t v_{[0,t]}$ which we wish to drive to 0. Hereafter, $v_{[0,t]}$ is the restriction of v to the interval $[0,t]$. This is the idea of introducing *asymptotically strong Feller* in the paper [126].

From eqs (5.2.38) and (5.2.41), it is easy to see that $\rho(t)$ satisfies the equation

$$\partial_t\rho = (\lambda + i\alpha)\Delta\rho + \mathcal{N}(u(t,\omega;u_0),\rho) - Qv(t), \quad \rho(0) = \xi. \tag{5.2.42}$$

In the following, $\mathbb{H}^1_\ell$ denotes the finite-dimensional "low-frequency" subspace of H^1

$$H^1_\ell = \mathrm{span}\{e_1, \cdots, e_N\},$$

and H^1_h denotes the corresponding "high-frequency" subspace of H^1 such that the direct sum decomposition holds: $H^1 = H^1_\ell \oplus H^1_h$. This decomposition naturally associates the projecting operator $\pi_\ell : H^1 \to H^1_\ell$ with

$$\xi_\ell = \pi_\ell\xi := \sum_{i=1}^{N}\langle(-\Delta)\mathbf{e}_i, \xi\rangle e_i \in H^1_\ell \quad \forall \xi \in H^1.$$

We also use $\pi_h = I - \pi_\ell$ to denote the projection on the codimensional space H^1_h.

Before proving the main result, we give some lemmas.

Lemma 5.2.6. *For any $\eta, \xi \in H^1$, we have*

$$\begin{aligned}\langle\xi_h, \mathcal{N}(\eta, \xi_l + \xi_h)\rangle_1 \le{}& \frac{\lambda}{2}\|\Delta\xi_h\|^2 + C_0\left(\|\xi_h\|^2_{H^1} + \|\xi_\ell\|^2_{H^1}\right)\\ &\times\left(1 + \|\eta\|^2_{H^2} + \||\eta||\nabla\eta|\|^2_{L^2}\right),\end{aligned} \tag{5.2.43}$$

$$\|\pi_l\mathcal{N}(\eta,\xi)\|^2_{H^1} \le C_N\|\xi\|^2(1 + \|\eta\|^4_{L^4}), \tag{5.2.44}$$

where $\mathcal{N}(\cdot,\cdot)$ is given in eq. (5.2.39). *Here, C_0 does not depend on N.*

Proof. For eq. (5.2.43), we have

$$\langle\xi_h, \mathcal{N}(\eta, \xi_l + \xi_h)\rangle_{H^1} = I_1 + I_2 + I_3,$$

where

$$\begin{aligned}I_1 ={}& \gamma\langle\xi_h, \xi_l + \xi_h\rangle_1,\\ I_2 ={}& -(\kappa + i\beta)\langle\xi_h, |\eta|^2(\xi_l + \xi_h)\rangle_1,\\ I_3 ={}& -(\kappa + i\beta)\langle\xi_h, \Re(\bar\eta\cdot(\xi_l + \xi_h))\eta\rangle_1.\end{aligned}$$

For I_1, we have

$$I_1 \le \gamma\|\xi_h\|^2_{H^1}.$$

For I_2, we have

$$\begin{aligned} I_2 &= -(\kappa+i\beta)\langle\nabla\xi_h, |\eta|^2\nabla(\xi_\ell+\xi_h)\rangle_0 - (\kappa+i\beta)\langle\nabla\xi_h, \xi_h\nabla(|\eta|^2)\rangle_0 \\ &\le C\|\eta\|_{L^\infty}^2 \||\nabla\xi_h||\nabla\xi|\|_{L^1} + C\|\nabla\xi_h\|_{L^6}\|\xi\|_{L^3}\|\nabla(|\eta|^2)\|_{L^2} \\ &\le C\|\eta\|_{H^2}^2\left(\|\xi_h\|_{H^1}^2+\|\xi_\ell\|_{H^1}^2\right) + \frac{\lambda}{4}\|\xi_h\|_{H^2}^2 + C\|\xi\|_{H^1}^2\||\eta||\nabla\eta|\|_{L^2}^2. \end{aligned}$$

For I_3, it is easy to see that

$$I_3 \le I_2.$$

Adding these estimates together, we get eq. (5.2.43).

For eq. (5.2.44), since

$$\begin{aligned} \langle e_i, \xi\rangle_1 &= \langle\Delta e_i, \xi\rangle_0 \le C_{e_i}\|\xi\|, \\ \langle e_i, \pi_\ell(|\eta|^2\xi)\rangle_1 &\le C_{e_i}\|\eta\|_{L^4}^2\|\xi\|_{L^2}, \\ \langle e_i, \pi_\ell(\Re(\bar\eta\cdot\xi)\eta)\rangle &\le C_{e_i}\|\eta\|_{L^4}^2\|\xi\|_{L^2}, \end{aligned}$$

we have by adding them together

$$\|\pi_l\mathcal{N}(\eta,\xi)\|_{H^1}^2 = \sum_{i=1}^N \langle e_i, \mathcal{N}(\eta,\xi)\rangle_1^2 \le C_N\|\xi\|_{L^2}^2(1+\|\eta\|_{L^4}^4),$$

where the constant C_N depends on N. We complete the proof of eq. (5.2.44). ∎

Lemma 5.2.7. *For any $\ell > 0$, there exist constants $C_1, C_2 > 0$ such that for any $t > 0$ and $u_0 \in H^1$, there holds*

$$\boldsymbol{E}\exp\left\{\ell\int_0^t \|u(s)\|_{H^2}^2 + \||u||\nabla u|\|_{L^2}^2 ds\right\} \le \exp\left\{C_1\|u_0\|_{H^1}^2 + C_2 t\right\}.$$

Proof. Let $F_t := C(\|u(t)\|_{H^2}^2 + \||u||\nabla u|\|_{L^2}^2)$.

By Itô formula, we have

$$\begin{aligned} \|u(t)\|^2 = &\|u_0\|^2 - 2\lambda\int_0^t \|\nabla u(s)\|^2 ds + 2y\int_0^t \|u(s)\|^2 ds \\ &- 2\kappa\int_0^t \|u(s)\|_{L^4}^4 ds + 2\int_0^t \langle u(s), d\mathbf{w}_s\rangle + \mathcal{E}_0\cdot t. \end{aligned} \tag{5.2.45}$$

By Hölder's inequality and Young's inequality, we have

$$2y\|u(s)\|_{L^2}^2 - 2\kappa\|u(s)\|_{L^4}^4 \le C - \kappa\|u(s)\|_{L^2}^4.$$

Therefore, eq. (5.2.45) implies that

$$\|u(t)\|^2 = \|u_0\|^2 + \int_0^t \left(C - 2\lambda\|\nabla u(s)\|^2 - \kappa\|u(s)\|^4\right) ds + 2\int_0^t \langle u(s), d\mathbf{w}_s\rangle.$$

From Ref. [218, Lemma 6.2], it follows that for any $\ell > 0$

$$\mathbf{E}\exp\left\{\ell\int_0^t \|u(s)\|_{H^1}^2 ds\right\} \le \exp\left\{\ell\|u_0\|_{L^2}^2 + C_\ell t\right\}. \tag{5.2.46}$$

Here, the constant C_ℓ does not depend on N.

Again, by Itô's formula, we have (by similar estimates as in Lemma 5.2.3)

$$d\|u(t)\|_{H^1}^2 \le C(-F_t + \|u(t)\|_{H^1}^2)dt + 2\langle u(t), d\mathbf{w}\rangle_1 + \mathcal{E}_1 dt.$$

Integrating this identity, we have

$$\|u(t)\|_{H^1}^2 - \|u_0\|_{H^1}^2 \le -\int_0^t F_s ds + C\int_0^t \|u(s)\|_{H^1}^2 ds + 2\int_0^t \langle u(t), d\mathbf{w}\rangle_1 + \int_0^t \mathcal{E}_1 ds.$$

Multiplying this by ℓ and then taking exponential, we get

$$\begin{aligned} e^{\ell\|u(t)\|_{H^1}^2}\cdot e^{-\ell\|u_0\|_{H^1}^2} \le\, & e^{-\ell\int_0^t F_s ds}\cdot e^{C_\ell\int_0^t \|u(s)\|_{H^1}^2 ds} \\ & \cdot e^{-\ell\int_0^t \|u(s)\|_{H^1}^2 ds + 2\ell\int_0^t \langle u(t), d\mathbf{w}\rangle_1}\cdot e^{\mathcal{E}_1 t}. \end{aligned}$$

Let $M_s := -\ell\int_0^t \|u(s)\|_{H^1}^2 ds + 2\ell\int_0^t \langle u(t), d\mathbf{w}\rangle_1$. Take expectation and rearrange this inequality to get

$$\mathbf{E}e^{\ell\int_0^t F_s ds} \le \mathbf{E}\left[e^{-\ell\|u(t)\|_{H^1}^2}\cdot e^{\ell\|u_0\|_{H^1}^2}\cdot e^{C_\ell\int_0^t \|u(s)\|_{H^1}^2 ds}\cdot e^{M_s}\cdot e^{\mathcal{E}_1 t}\right].$$

By eq. (5.2.46), e^{M_s} is an exponential martingale and therefore

$$\mathbf{E}e^{5M_s} = \mathbf{E}e^{5M_0} = 1,$$

and

$$\mathbf{E}e^{-5\ell\|u(t)\|_{H^1}^2} \le 1,$$

we have from Hölder's inequality that

$$\mathbf{E}e^{\ell\int_0^t F_s ds} \le e^{\ell\|u_0\|_{H^1}^2}\cdot e^{\mathcal{E}_1 t}\left(\mathbf{E}e^{C_\ell\int_0^t \|u(s)\|_{H^1}^2 ds}\right)^{1/5}.$$

Recalling the estimate in eq. (5.2.46), we easily get

$$\mathbf{E}e^{\ell\int_0^t F_s ds} \leq e^{C_1\|u_0\|_{H^1}^2} \cdot e^{C_2 t},$$

where C_1 depends on l and C_2 depends on $\mathcal{E}_0, \mathcal{E}_1$ and ℓ. ■

Corollary 5.2.2. *For any $p \geq 1$ integers, we have*

$$\boldsymbol{E}\left[\int_0^t \|u(s)\|_{H^2}^2 + \||u||\nabla u|\|_{L^2}^2 ds\right]^p \leq p! \exp\{\|u_0\|^2 + Ct\}.$$

Proof. The result follows from Lemma 5.2.7 and the fundamental inequality $x^p \leq p! e^x$. ■

We are now in a proper position to prove the *asymptotically strong Feller* property for the semigroup $\mathcal{P}_t$.

Proposition 5.2.3. *Let $\{\mathcal{P}_t\}_{t\geq 0}$ be the semigroup associated with the SGLE* (5.2.5). *There exist some constant $N_* \in \mathbf{N}$ and constants $C, \delta > 0$ such that for any $t > 0$, $u_0 \in H^1$, and any Fréchet differentiable function φ on H^1 with $\|\varphi\|_\infty, \|\nabla\varphi\|_\infty < \infty$,*

$$\|\nabla\mathcal{P}_t\varphi(u_0)\| \leq C_0 \exp\{C_1\|u_0\|^2\}(\|\varphi\|_\infty + \|\nabla\varphi\|_\infty e^{-\delta t}). \tag{5.2.47}$$

Proof. For any $\xi \in H^1$ with $\|\xi\|_{H^1} = 1$. We define

$$\zeta_\ell(t) = \begin{cases} \xi_\ell \cdot (1 - \frac{t}{2\|\xi_\ell\|_{H^1}}), & t \in [0, 2\|\xi_\ell\|_{H^1}] \\ 0, & (2\|\xi_\ell\|_{H^1}, \infty). \end{cases}$$

For the high-frequency part, we let $\zeta_h(t)$ satisfy the equation

$$\partial_t\zeta_h = (\lambda + i\alpha)\Delta\zeta_h + \pi_h\mathcal{N}(u(t), \zeta_\ell + \zeta_h), \quad \zeta_h(0) = \xi_h. \tag{5.2.48}$$

Define

$$v(t) = Q^{-1}G_t, \tag{5.2.49}$$

where $G_t = \frac{\zeta_\ell(t)}{2\|\zeta_\ell(t)\|_{H^1}} + (\lambda + i\alpha)\Delta\zeta_\ell + \pi_\ell\mathcal{N}(u(t), \zeta_\ell + \zeta_h)$. It is immediate that $\zeta = \zeta_\ell + \zeta_h$ satisfy the equation

$$\partial_t\zeta(t) = -\frac{1}{2}\frac{\zeta_\ell(t)}{\|\zeta_\ell(t)\|_{H^1}} + (\lambda + i\alpha)\Delta\zeta_h(t) + \pi_h\mathcal{N}(u(t), \zeta_\ell(t) + \zeta_h(t)),$$

with the same initial data $\zeta(0) = \xi$. From eq. (5.2.42), it is clear that $\rho(t)$ and $\zeta(t)$ satisfy the same equation, with the same initial data $\rho(0) = \zeta(0) = \xi$; therefore, $\rho = \zeta$.

In the following, we will show that for given ξ and $v(t)$ defined as in eq. (5.2.49), $\rho(t)$ tends to 0 as $t \to \infty$. To get estimate (5.2.47), we give two preliminary estimates in the following two steps.

Step 1. *Show that there exist constants $\nu > 0$ and $C > 0$ such that*

$$\mathbf{E}\|\zeta(t)\|_{H^1} \leq Ce^{C\|u_0\|_{H^1}^2-\delta t}. \tag{5.2.50}$$

For the "low-frequency" part ζ_l, we have by definition $\|\zeta_l(t)\|_{H^1} \leq 1$ for $0 \leq t \leq 2$ and $\|\zeta_l(t)\|_{H^1} = 0$ for $t \geq 2$. In particular,

$$\mathbf{E}\|\zeta_l(t)\|_{H^1}^p \leq C. \tag{5.2.51}$$

For the "high-frequency" part ζ_h, we use eq. (5.2.48). By Lemma 5.2.6,

$$\begin{aligned}
\frac{d}{dt}\|\zeta_h(t)\|_{H^1}^2 =& -2\lambda\|\Delta\zeta_h(t)\|^2 + 2\langle\zeta_h(t), \pi_h\mathcal{N}(u(t), \zeta_l+\zeta_h)\rangle_{H^1} \\
\leq& -\lambda\|\Delta\zeta_h(t)\|^2 + 2C_0(1+F_t)\left(\|\zeta_h\|_{H^1}^2 + \|\zeta_\ell\|_{H^1}^2\right) \\
\leq&(-\lambda\lambda_N + 2C_0(1+F_t))\|\zeta_h(t)\|_{H^1}^2 + 2C_0(1+F_t)\|\zeta_\ell\|_{H^1}^2.
\end{aligned}$$

Noting that $\|\zeta_l(t)\|^2 = 0$ for all $t \geq 2$, we have by Gronwall's inequality that

$$\begin{aligned}
\|\zeta_h(t)\|_{H^1}^2 \leq&\|\zeta_h(0)\|_{H^1}^2 \exp\left\{-\lambda\lambda_N t + 2C_0\int_0^t(1+F_s)ds\right\} \\
&+ C\exp\left\{-\lambda\lambda_N(t-2) + 2C_0\int_0^t(1+F_s)ds\right\}\int_0^2\|\zeta_\ell(s)\|_{H^1}^2 ds.
\end{aligned} \tag{5.2.52}$$

Recall that $\lambda_N \to \infty$ as $N \to \infty$ and the fact that $\|\zeta_l(t)\|_{H^1} \leq 1$ for $0 \leq t \leq 2$. By Lemma 5.2.7, and Hölder's inequality, there exist constant $\delta > 0$ and N_* such that when $N \geq N_*$,

$$\mathbf{E}\|\zeta_h(t)\|_{H^1}^2 \leq Ce^{C\|u_0\|_{H^1}^2-\delta t}.$$

Indeed, from eq. (5.2.52) and Corollary 5.2.2, we know that for any $p \geq 1$, there exist constants C, δ and N_* such that for $N \geq N_*$,

$$\mathbf{E}\|\zeta_h(t)\|_{H^1}^p \leq C(1+\|\zeta_h(0)\|_{H^1}^p)e^{C\|u_0\|_{H^1}^2-\delta t}.$$

Combining this and the low-frequency estimate (5.2.51) together, we have the higher moment estimate

$$\mathbf{E}\|\zeta(t)\|_{H^1}^p \leq Ce^{C\|u_0\|^2-\delta t}. \tag{5.2.53}$$

In particular, eq. (5.2.50) holds when $p = 1$.

Step 2. *Show that $\int_0^\infty \mathbf{E}|v(t)|^2dt \leq Ce^{C\|u_0\|^2}$.*

For $v(t)$ in eq. (5.2.49), since

$$\|\Delta\zeta_\ell\|_{H^1}^2 \le \lambda_N^2 \|\zeta_\ell\|_{L^2}^2 \le C_N \|\zeta(t)\|_{L^2}^2,$$

we get from Lemma 5.2.6 that

$$\begin{aligned}\mathbf{E}|v(t)|^2 &\le C\mathbf{E}\|G_t\|_{H^1}^2 \\ &\le C_N\left\{1_{t\le 2} + \mathbf{E}\left[\|\zeta(t)\|_{L^2}^2(1+\|u(t)\|_{L^4}^4)\right]\right\} \\ &\le C_N\left\{1_{t\le 2} + \left(\mathbf{E}\|\zeta(t)\|^4\right)^{1/2}\left(1+\mathbf{E}\|u(t)\|_{L^4}^8\right)^{1/2}\right\}.\end{aligned}$$

From Lemma 5.2.3 with $p = 4$, $\mathbf{E}\|u(t)\|_{L^4}^8$ grows at most polynomially while from eq. (5.2.53) we know that $\mathbf{E}\|\zeta(t)\|^4$ decays exponentially; therefore,

$$\int_0^\infty \mathbf{E}|v(t)|^2 dt \le Ce^{C\|u_0\|^2}, \tag{5.2.54}$$

as is expected. From the proof, it seems necessary to get the higher momentum estimate of $\|u(t)\|_{H^1}$ in Lemma 5.2.3.

Finally, we turn to the proof of eq. (5.2.47). Let $\mathcal{P}_t$ and φ be as above. By chain rule and integration by parts formula, we have

$$\begin{aligned}\langle\nabla\mathcal{P}_t\varphi(u_0),\xi\rangle_{H^1} &= \mathbf{E}\langle\nabla(\varphi(u(t))),\xi\rangle_{H^1} = \mathbf{E}\langle(\nabla\varphi)(u(t)),\mathcal{J}_t\xi\rangle_{H^1} \\ &= \mathbf{E}\langle(\nabla\varphi)(u(t)),\mathcal{A}_t v_{[0,t]} + \rho(t)\rangle_{H^1} \\ &= \mathbf{E}\left[\mathfrak{D}^v(\varphi(u(t;u_0)))\right] + \mathbf{E}\langle(\nabla\varphi)(u(t)),\rho(t)\rangle_{H^1} \\ &= \mathbf{E}\left[\varphi(u(t;u_0))\int_0^t v(s)dW_s\right] + \mathbf{E}\langle(\nabla\varphi)(u(t)),\rho(t)\rangle_{H^1} \\ &\le \|\varphi\|_\infty\mathbf{E}\left|\int_0^t v(s)dW_s\right| + \|\nabla\varphi\|_\infty\mathbf{E}\|\rho(t)\|_{H^1}.\end{aligned}$$

Now, since $v_{[0,t]}$ is adapted to the Wiener path, we have by eq. (5.2.54)

$$\mathbf{E}\left|\int_0^t v(s)dW_s\right| \le \left(\int_0^t \mathbf{E}\|v(s)\|^2 ds\right)^{1/2} \le Ce^{C\|u_0\|^2}.$$

Then Estimate (5.2.47) then follows from estimate (5.2.50). ■

5.2.3.2 A support property

We first prove the following lemma. The idea can be found in Refs [76, 218].

Lemma 5.2.8. *Assume* **(H4)** *holds, then for any $r_1, r_2 > 0$, there exists some $T > 0$ such that*

$$\inf_{\|u_0\|_{H^1}\le r_1} P\{\omega : \|u(T,\omega;u_0)\|_{H^1} \le r_2\} > 0.$$

Proof. Set $v(t) = u(t) - \mathbf{w}(t)$, then

$$v'(t) = (\lambda + i\alpha)\Delta(v(t) + \mathbf{w}(t)) + \gamma(v + \mathbf{w}) - (\kappa + i\beta)|v + \mathbf{w}|^2(v + \mathbf{w}). \tag{5.2.55}$$

Let $T > 0$ and $\varepsilon > 0$, to be determined later. We assume that

$$\sup_{t\in[0,T]} \|\mathbf{w}(t,\omega)\|_{H^6} < \varepsilon. \tag{5.2.56}$$

By multiplying the equation with $v(t)$, and integrating by parts, we have

$$\begin{aligned}\frac{d}{dt}\|v(t)\|^2 &= -2\lambda\|\nabla v(t)\|^2 + 2\Re(\lambda + i\alpha)\langle\Delta\mathbf{w}(t), v(t)\rangle\\ &\quad + 2\gamma\|v(t)\|^2 + 2\gamma\Re\langle\mathbf{w}(t), v(t)\rangle\\ &\quad - 2\Re(\kappa + i\beta)\langle|v + \mathbf{w}|^2(v + \mathbf{w}), v(t)\rangle\\ &=: I_1 + I_2 + I_3.\end{aligned} \tag{5.2.57}$$

For I_1 and I_2, we have

$$\begin{aligned}I_1 + I_2 &\le -2\lambda\|\nabla v(t)\|^2 + 2\gamma\|v(t)\|^2 + C\varepsilon\|v(t)\|,\\ &\le -2\lambda\|\nabla v(t)\|^2 + 2\gamma\|v(t)\|^2 + C\varepsilon + C\varepsilon\|v(t)\|^2.\end{aligned}$$

For I_3, since

$$\begin{aligned}2Re(\kappa + i\beta)\langle|v + \mathbf{w}|^2(v + \mathbf{w}), \mathbf{w}\rangle &\le 2C\|\mathbf{w}\|_{L^\infty}\|v + \mathbf{w}\|_{L^3}^3\\ &\le C\varepsilon\|v\|_{L^3}^3 + C\varepsilon^4\\ &\le C\varepsilon\|\nabla v\|_{L^2}\|v\|_{L^2}^2 + C\varepsilon\\ &\le \delta\|\nabla v\|_{L^2}^2 + C\varepsilon\delta^{-1}\|v\|_{L^2}^4 + C\varepsilon\delta^{-1},\end{aligned}$$

it is estimated that

$$I_3 \le -2\kappa\|v + \mathbf{w}\|_{L^4}^4 + \delta\|\nabla v\|_{L^2}^2 + C\varepsilon\|v\|_{L^2}^4 + C\varepsilon.$$

Let $\tilde{\lambda} = \varrho_1(2\lambda - \delta) - 2\gamma$, from assumption **(H2)**, we can choose δ sufficiently small (e.g. $\delta = \lambda - \gamma\varrho_1^{-1}$) such that $\tilde{\lambda} > 0$. Fix such a δ, then from estimates for I_1, I_2, I_3, we obtain

$$\frac{d}{dt}\|v(t)\|^2 \le -\tilde{\lambda}\|v(t)\|^2 + C\varepsilon\|v(t)\|^4 + C\varepsilon.$$

Therefore, from Lemma 6.1 in Ref. [218], we have that for any $\varepsilon', h > 0$, there exist $T' > 0$ and ε small enough such that

$$\sup_{t\in[0,T']} \|v(t)\|_{L^2} \le 2r_1 \tag{5.2.58}$$

and

$$\sup_{t\in[T',T'+h]} \|v(t)\|_{L^2} < \varepsilon'. \tag{5.2.59}$$

Now we turn to the H^1 estimates. Taking inner product of eq. (5.2.55) with $-\Delta v(t)$ and using integrating by parts, we have

$$\begin{aligned}\frac{d}{dt}\|v(t)\|_{H^1} &= 2Re\langle(\lambda + i\alpha)\Delta v, v\rangle_{H^1} + 2Re\langle(\lambda + i\alpha)\Delta \mathbf{w}, v\rangle_{H^1}\\ &\quad + 2y\langle v, v\rangle_{H^1} + 2y\langle \mathbf{w}, v\rangle_{H^1}\\ &\quad - 2Re\langle(\kappa + i\beta)|v + \mathbf{w}|^2(v + \mathbf{w}), v + \mathbf{w}\rangle_{H^1}\\ &\quad + 2Re\langle(\kappa + i\beta)|v + \mathbf{w}|^2(v + \mathbf{w}), \mathbf{w}\rangle_{H^1}\\ &= J_1 + J_2 + J_3 + J_4.\end{aligned}$$

Similar to the previous estimates, we have

$$\begin{aligned}J_1 &\le -2\lambda\|\nabla v(t)\|_{H^1}^2 + C\varepsilon + C\varepsilon\|v(t)\|_{L^2}^2\\ J_2 &\le 2y\|v(t)\|_{H^1}^2 + C\varepsilon + C\varepsilon\|v(t)\|_{L^2}^2.\end{aligned}$$

For J_3, we have under assumption $\beta \le \sqrt{3}\kappa$,

$$\begin{aligned}J_3 &= -2\kappa\|v + \mathbf{w}\|_{L^4}^4 - 2\Re(\kappa + i\beta)\langle\nabla(|v + \mathbf{w}|^2(v + \mathbf{w})), \nabla(v + \mathbf{w})\rangle\\ &\le -2\kappa\|v + \mathbf{w}\|_{L^4}^4 - 4\kappa\||v + \mathbf{w}|\nabla(v + \mathbf{w})\|_{L^2}^2\\ &\quad + 2\Re(\kappa + i\beta)\int(v + \mathbf{w})^2(\nabla\overline{(v + \mathbf{w})})^2\\ &\le -2\kappa\|v + \mathbf{w}\|_{L^4}^4.\end{aligned}$$

Finally, for J_4, we have by interpolation

$$\begin{aligned}J_4(t) &= 2\Re\langle(\kappa + i\beta)|v + \mathbf{w}|^2(v + \mathbf{w}), (I - \Delta)\mathbf{w}\rangle\\ &\le \delta\|\nabla v\|_{L^2}^2 + C\varepsilon\|v\|_{L^2}^4 + C\varepsilon.\end{aligned}$$

Combining the above estimates yields that (choosing $\delta = \lambda - y\varrho_1^{-1}$)

$$\begin{aligned}\frac{d}{dt}\|v(t)\|_{H^1}^2 &\le -2\lambda\|\nabla v(t)\|_{H^1}^2 + 2y\|v(t)\|_{H^1}^2 + \delta\|v(t)\|_{H^1}^2 + C\varepsilon + C\varepsilon\|v(t)\|_{L^2}^4\\ &\le -\tilde{\lambda}\|v\|_{H^1}^2 + C\|v(t)\|_{L^2}^2 + C\varepsilon + C\varepsilon\|v(t)\|_{L^2}^4,\end{aligned}$$

where $\tilde{\lambda} = \lambda\varrho_1 - y > 0$. By Gronwall's inequality, we have

$$\|v(t)\|_{H^1}^2 \le e^{-\tilde{\lambda}(t-s)}\|v(s)\|_{H^1}^2 + \frac{C}{\tilde{\lambda}}\left\{C\sup_{r\in[s,t]}\|v(r)\|_{L^2}^2 + C\varepsilon\sup_{r\in[s,t]}\|v(r)\|_{L^2}^4 + C\varepsilon\right\}.$$

Letting $s = 0$ and $t = T'$, we have by eq. (5.2.58) that

$$\|v(T')\|_{H^1}^2 \le C(r_1^4 + 1).$$

Then letting $s = T'$ and $t = T' + h$, we have

$$\|v(T'+h)\|_{H^1}^2 \le Ce^{-\tilde{\lambda}h}(r_1^4+1) + C\{\sup_{r\in[T',T'+h]}\|v(r)\|_{L^2}^2 + C\varepsilon\sup_{r\in[T',T'+h]}\|v(r)\|_{L^2}^4 + C\varepsilon\},$$

which together with eq. (5.2.59) yields that there exist a T large enough and $\varepsilon \in (0,1)$ small enough such that

$$\|v(T)\|_{H^1} < \frac{r_2}{2}. \tag{5.2.60}$$

Taking eq. (5.2.56) into account, there exist a T large enough and $\varepsilon \in (0,1)$ small enough such that

$$\|u(T,\omega,u_0)\|_{H^1} < r_2. \tag{5.2.61}$$

Let

$$\Omega_\varepsilon = \{\omega : \sup_{t\in[0,T]}\|\mathbf{w}(t,\omega)\|_{H^6} < \varepsilon\},$$

then

$$\Omega_\varepsilon \subset \bigcap_{\|u_0\|_{H^1}\le r_1}\{\omega : \|u(T,\omega;u_0)\|_{H^1} \le r_2\}.$$

Since Ω_ε is an open set and $P\{\Omega_\varepsilon\} > 0$, the result follows. ∎

Now we prove the following:

Proposition 5.2.4. 0 *belongs to the support of any invariant measure of* $\{\mathcal{P}_t\}_{t\ge0}$.

Proof. For every invariant measure μ, we choose some $r_1 > 0$ such that

$$\mu(B_{r_1}) \ge 1/2.$$

By the definition of an invariant measure and the above lemma, for any $r_2 > 0$,

$$\begin{aligned}\mu(B_{r_2}) = \mathcal{P}_t^*\mu(B_{r_2}) &= \int_{H^1} \mathcal{P}_t(x, B_{r_2})\mu(dx)\\ &= \int_{H^1} \mathcal{P}_t \mathbf{1}_{B_{r_2}}(x)\mu(dx) \geq \int_{B_{r_1}} \mathcal{P}_t \mathbf{1}_{B_{r_2}}(x)\mu(dx)\\ &\geq \inf_{x\in B_{r_1}} \mathcal{P}_t \mathbf{1}_{B_{r_2}}(x) \cdot \mu(B_{r_1}) > 0,\end{aligned}$$

which implies that 0 belongs to the support of μ. ∎

5.2.3.3 Proof of the derivative flow equation (5.2.40)

Lemma 5.2.9. *For any $T > 0$, there exists a constant $C_T > 0$ such that for each ω and $u_0 \in H^1$*

$$\sup_{t\in[0,T]} \|u(t,\omega)\|_{H^1}^2 + \int_0^T \|u(t,\omega)\|_{H^2}^2 dt \leq C_T \left(1 + \|u_0\|_{H^1}^2 + \sup_{t\in[0,T]} \|\boldsymbol{w}(t,\omega)\|_{H^6}^4\right).$$

Proof. Consider estimate (5.2.57) in Lemma 5.2.8. For I_1 and I_2 we have a new estimate

$$I_1 + I_2 \leq -2\lambda\|\nabla v(t)\|_{L^2}^2 + C\|\Delta \mathbf{w}\|_{L^2}\|v(t)\|_{L^2} + (2y+1)\|v(t)\|_{L^2}^2 + C\|\mathbf{w}(t)\|_{L^2}^2.$$

For I_3, we have

$$\begin{aligned}I_3 &= -2\Re(\kappa + i\beta)\langle |v + \mathbf{w}|^2(v + \mathbf{w}), v(t)\rangle\\ &= -2\Re(\kappa + i\beta)\langle |v + \mathbf{w}|^2(v + \mathbf{w}), v + \mathbf{w}\rangle\\ &\quad - 2\Re(\kappa + i\beta)\langle |v + \mathbf{w}|^2(v + \mathbf{w}), \mathbf{w}\rangle =: I_{31} + I_{32}.\end{aligned}$$

Since

$$I_{31} \leq -2\kappa\|v + \mathbf{w}\|_{L^4}^4,$$

and

$$\begin{aligned}I_{32} &\leq c\|\mathbf{w}\|_{L^\infty}\|v + \mathbf{w}\|_{L^3}^3\\ &\leq c\|\mathbf{w}\|_{L^\infty}\|v + \mathbf{w}\|_{L^4}^3\\ &\leq \kappa\|v + \mathbf{w}\|_{L^4}^4 + c\|\mathbf{w}\|_{L^\infty}^4,\end{aligned}$$

we have

$$\frac{d}{dt}\|v(t)\|_{L^2}^2 \leq C\|v(t)\|_{L^2}^2 + C(\|\mathbf{w}\|_{H^2}^2 + \mathbf{w}\|_{H^2}^4).$$

By Gronwall's lemma, we have

$$\sup_{t\in[0,T]} \|v(t)\|_{L^2}^2 \le C_T \left(1+ \sup_{t\in[0,T]} (\|\mathbf{w}\|_{H^2}^2 + \|\mathbf{w}\|_{H^2}^4)\right).$$

The second step is to estimate the H^1-norm; however, since this is similar to the proof of Lemma 5.2.8, we omit the details for simplicity. ∎

For $\xi \in H^1$, let us consider a small perturbation of the initial value given by $u_\varepsilon(0) = u_0 + \varepsilon\xi$. The corresponding solution of eq. (5.2.5) is denoted by $u_\varepsilon(t)$.

Set

$$\xi_\varepsilon(t) = \left(u_\varepsilon(t) - u(t)\right)/\varepsilon.$$

Then $\xi_\varepsilon(t)$ satisfies

$$\begin{cases} \xi_\varepsilon'(t) = & (\lambda + i\alpha)\Delta\xi_\varepsilon(t) + y\xi_\varepsilon(t) - (\kappa + i\beta)|u_\varepsilon(t)|^2\xi_\varepsilon(t) \\ & -(\kappa + i\beta)(|u_\varepsilon|^2 - |u(t)|^2)u(t)/\varepsilon, \\ \xi_\varepsilon(0) = & \xi. \end{cases} \tag{5.2.62}$$

For this equation, we have

Lemma 5.2.10. *For any $T > 0$, there is a $C_T > 0$ such that for any $\varepsilon \in (0,1)$*

$$\sup_{t\in[0,T]} \|\xi_\varepsilon(t)\|_{H^1}^2 + \int_0^T \|\xi_\varepsilon(t)\|_{H^2}^2 dt \le C_T. \tag{5.2.63}$$

Proof. As in the proof of Lemma 5.2.5, we have

$$\begin{aligned} \frac{d}{dt}\|\xi_\varepsilon(t)\|_{H^1}^2 \le & -\lambda\|\xi_\varepsilon(t)\|_{H^2}^2 + C\|\xi_\varepsilon(t)\|_{H^1}^2 + C\|\xi_\varepsilon(t)\|_{L^6}^2\|u_\varepsilon(t)\|_{L^6}^4 \\ & + C\|\xi_\varepsilon(t)\|_{L^6}^2(\|u_\varepsilon(t)\|_{L^6}^4 + \|u(t)\|_{L^6}^4) \\ \le & -\lambda\|\xi_\varepsilon(t)\|_{H^2}^2 + C\left(1 + \|u_\varepsilon(t)\|_{L^6}^4 + \|u(t)\|_{L^6}^4\right)\|\xi_\varepsilon(t)\|_{H^1}^2, \end{aligned}$$

which together with Lemma 5.2.9 yields the desired estimate. ∎

Now, consider the difference

$$\Lambda_\varepsilon(t) = \xi_\varepsilon(t) - \mathcal{J}_t\xi. \tag{5.2.64}$$

It is not hard to see that $\Lambda_\varepsilon(t)$ satisfies

$$\Lambda_\varepsilon'(t) = (\lambda + i\alpha)\Delta\Lambda_\varepsilon(t) + y\Lambda_\varepsilon(t) - \sum_{i=1}^{5} F_i(t), \tag{5.2.65}$$

where

$$\begin{aligned}
F_1(t) =&(\kappa + i\beta)\left[|u_\varepsilon(t)|^2 - |u(t)|^2\right]\xi_\varepsilon(t),\\
F_2(t) =&(\kappa + i\beta)|u|^2\Lambda_\varepsilon(t),\\
F_3(t) =&-\varepsilon(\kappa + i\beta)|\xi_\varepsilon(t)|^2 u(t),\\
F_4(t) =&-(\kappa + i\beta)|u|^2\Lambda_\varepsilon(t),\\
F_5(t) =&-(\kappa + i\beta)\overline{\Lambda_\varepsilon(t)}(u(t))^2.
\end{aligned}$$

From eq. (5.2.65) and Young's inequality, we have

$$\frac{d}{dt}\|\Lambda_\varepsilon(t)\|_{H^1}^2 \le -\lambda\|\Lambda_\varepsilon(t)\|_{H^2}^2 + C\|\Lambda_\varepsilon(t)\|_{L^2}^2 + C\sum_{i=1}^{5}\|F_i(t)\|_{L^2}^2.$$

Now, we estimate the sum on the RHS.

For F_i, $i = 1, \cdots, 5$, by Hölder's inequality and Sobolev's embedding, we have

$$\begin{aligned}
\|F_1(t)\|_{L^2}^2 \le& C\varepsilon^2\int|\xi_\varepsilon(t)|^4(|u_\varepsilon(t)|^2 + |u(t)|^2)\\
\le& C\varepsilon^2\|\xi_\varepsilon(t)\|_{L^6}^4\left(\|u_\varepsilon(t)\|_{L^6}^2 + \|u(t)\|_{L^6}^2\right)\\
\le& C\varepsilon^2\|\xi_\varepsilon(t)\|_{H^1}^4\left(\|u_\varepsilon(t)\|_{H^1}^2 + \|u(t)\|_{H^1}^2\right),
\end{aligned}$$

$$\|F_2(t)\|_{L^2}^2 + \|F_4(t)\|_{L^2}^2 + \|F_5(t)\|_{L^2}^2 \le C\|\Lambda_\varepsilon(t)\|_{H^1}^2\|u(t)\|_{H^1}^4,$$

and

$$\begin{aligned}
\|F_3(t)\|_{L^2}^2 \le& C\varepsilon^2\|\xi_\varepsilon(t)\|_{L^6}^4\|u(t)\|_{L^6}^2\\
\le& C\varepsilon^2\|\xi_\varepsilon(t)\|_{H^1}^4\|u(t)\|_{H^1}^2.
\end{aligned}$$

Combining the above calculations with the estimates in Lemmas 5.2.9 and 5.2.10, we have

$$\frac{d}{dt}\|\Lambda_\varepsilon(t)\|_{H^1}^2 \le C_1\varepsilon^2 + C_2\|\Lambda_\varepsilon(t)\|_{H^1}^2, \tag{5.2.66}$$

which yields by Gronwall's inequality,

$$\|\Lambda_\varepsilon(t)\|_{H^1}^2 \le C_1te^{C_2t}\cdot\varepsilon^2.$$

As $\varepsilon \to 0$, we have

$$\|\Lambda_\varepsilon(t)\|_{H^1} \to 0,$$

which gives eq. (5.2.40).

5.2.4 Some remarks

In the final section, we would like to make several comments on the ergodicity for the SGLE with degenerate additive noise.

We would like to remark that the discussion in this section relies heavily on the approach proposed by Hairer and Mattingly [126], who proved the ergodicity of the stochastic Navier–Stokes equation (SNSE) with degenerate noise. In that paper, they developed two main tools: the *asymptotically strong Feller* property for the semigroup associated with the flow of an SPDE and the approximative integration by parts formula in the Malliavin calculus. They are rather powerful and indispensable techniques for ergodicity for an SPDE driven by degenerate stochastic forcing. As pointed out by Hairer and Mattingly [126], the asymptotically strong Feller property is much weaker than the strong Feller property, and many equations driven by degenerate noise have only the former one rather than the latter one [218].

But there are several differences between the analysis of the two-dimensional (2D) SNSE in Ref. [126] and the SGLE. For the 2D Navier–Stokes equation, uniqueness holds for L^2–solutions, and hence the solution for the SNSE generalizes a Markovian semigroup of the stochastic flow. However, for the three-dimensional (3D) GL equation, because of the lack of pathwise uniqueness of L^2–solutions, the flow does not generalize a semigroup and the analysis fails. Therefore, we consider H^1 solutions in this section.

Another difference is that although the nonlinear term $u \cdot \nabla u$ in SNSE consists of one order derivative, it is quadratic algebraically. But in the SGLE, the nonlinear term is cubic. To handle this nonlinear term when proving the asymptotically strong Feller property, estimates of higher order momentum seem to be indispensable, even if one considers the L^2 solution.

For ergodicity of SGL, see also Ref. [191] by coupling method.

5.3 Stochastic damped forced Ostrovsky equation

5.3.1 Introduction

This section is concerned with the asymptotic behavior of solutions to the following damped forced Ostrovsky equation with additive noise defined in the entire space $\mathbb{R}$:

$$du - (\beta u_{xxx} + \alpha D_x^{-1} u - (u^2)_x - \lambda u + f)dt = \sum_{i=1}^{m} h_i dw_i, \tag{5.3.1}$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, t > 0, \tag{5.3.2}$$

where $D_x^{-1} = \mathcal{F}_x^{-1} \frac{1}{i\xi} \mathcal{F}_x$, α, β, λ are real constants with $\beta \neq 0, \alpha > 0, \lambda > 0$, f is time independent, h_i ($i = 1, 2, \cdots, m$) are given functions defined on $\mathbb{R}$, and $\{w_i(t)\}_{i=1}^m$ are mutually independent two-sided Wiener processes on a probability space.

When $\lambda = 0, f = 0, h_i = 0 (1 \le i \le m)$, system (5.3.1) becomes the standard Ostrovsky equation:

$$u_t - \beta u_{xxx} - \alpha D_x^{-1} u + (u^2)_x = 0, \quad x \in \mathbb{R}, t > 0. \tag{5.3.3}$$

The Ostrovsky equation (5.3.3) was derived by Ostrovsky [193] as a model for the propagation of weakly nonlinear dispersive long surface and internal waves of small amplitude in a rotating fluid. Here, free surface $u(t,x)$ has been rendered nondimensional with respect to the constant depth h of liquid and gravitational acceleration g and the parameter $\alpha = n^2/c_0 > 0$ measures the effect of rotation, where the wave speed $c_0 = \sqrt{gh}$, and n is the local Coriolis parameter. The parameter β determines the type of dispersion, i.e., $\beta < 0$ (negative dispersion) for surface and internal waves in the ocean and surface waves in a shallow channel with an uneven bottom [10] and $\beta > 0$ (positive dispersion) for capillary waves on the surface of liquid or for oblique magneto-acoustic waves in plasma [102, 109]. Equation (5.3.1) models the situation when nonlinearity, dispersion, dissipation rotation and stochastic effects are taken into account at the same time.

The well-posedness of eqs (5.3.3)–(5.3.2) has been studied in Refs [133, 135, 173, 249] and the references therein. The well-posedness of eqs (5.3.1)–(5.3.2) without noise was obtained in Ref. [122] in Bourgain function spaces $\tilde{X}_{s,b}$ (see below for a precise definition of $\tilde{X}_{s,b}$) with $b > 1/2$. The Bourgain function spaces were introduced by Bourgain [29] for the well-posedness of the Korteweg–de Vries (KdV) equation. This method was developed by Kenig, Ponce and Vega [152] and Tao [242] to study the Cauchy problem for nonlinear dispersive wave equations. In this section, the existence and uniqueness of solutions of the stochastic damped forced Ostrovsky equation (5.3.1)–(5.3.2) are studied in the above Bourgain spaces $\tilde{X}_{s,b}$. Roughly speaking, the index b represents the smoothness in time. However, stochastic systems (5.3.1)–(5.3.2) have the same time regularity as Brownian motion with $b < 1/2$. Hence, when trying to apply this method, the spaces $\tilde{X}_{s,b}$ with $b < 1/2$ have to be encountered. The well-posedness for the stochastic KdV equation and the stochastic Camassa–Holm equation in Bourgain spaces $X_{s,b}$ with $b < 1/2$ were studied in Refs [27, 58].

Another interesting thing is the long-term behavior for systems (5.3.1)–(5.3.2). It is known that the long-term behavior of random systems is captured by a pullback random attractor, which was introduced in Refs [68, 98] as an extension of the attractors theory of deterministic systems in Refs [127, 217, 245]. In the case of bounded domains, the existence of random attractors for SPDEs has been studied extensively by many authors (see Refs [5, 57, 66, 98] and the references therein). When the domain is the entire space $\mathbb{R}^n$, the existence of random attractors was established recently in Ref. [36] and the references therein. The crucial idea for the proof is the asymptotic compactness and existence of bounded absorbing sets for these equations. And the asymptotic compactness is usually proved by a tail estimate. In this section, the random attractors for eqs (5.3.1)–(5.3.2) are obtained by the same idea as above. Instead of a tail estimate, the asymptotic compactness for eqs (5.3.1)–(5.3.2) is checked by splitting the solutions into a decaying part plus a regular part as in Refs [122, 245].

Note that the phase function of semigroup of eq. (5.3.1) has nonzero points, which makes a difference from that of the linear KdV equation and Kadomtsev–Petviashvili (KP) equation and also makes the problem much more difficult. Therefore, the Fourier restriction operators

$$P^N h = \frac{1}{2\pi}\int_{|\xi|\geq N} e^{ix\xi}\hat{h}(\xi)d\xi,\ P_N h = \frac{1}{2\pi}\int_{\varepsilon\leq|\xi|\leq N} e^{ix\xi}\hat{h}(\xi)d\xi \qquad \forall N\geq\varepsilon>0$$

are used to eliminate the singularity of the phase function and to split the solution in Section 5.3.3. For simplicity, denote $P_N h = \frac{1}{2\pi}\int_{|\xi|\leq N} e^{ix\xi}\hat{h}(\xi)d\xi$.

The section is organized as follows. In Section 5.3.2, the well-posedness of the stochastic damped forced Ostrovsky equation is proved. In Section 5.3.3, we first prove that the solutions for the equation are bounded, and then split the solutions into two parts, one uniformly bounded in $\tilde{H}^3(\mathbb{R})$ and the other decaying in $\tilde{L}^2(\mathbb{R})$. These estimates are necessary for proving the existence of bounded absorbing sets and the asymptotic compactness of the equation. In the last section, the asymptotic compactness of the solution operator and then the existence of a pullback random attractor are proved.

5.3.2 Well-posedness

In this section, the stochastic estimate and the bilinear estimates are proved, then the local well-posedness of eqs (5.3.1)–(5.3.2) is established by contraction mapping principle.

First, we give some notations. Set $a(x,s) = (u^2)_x + \lambda u - f$, then the mild solution of eqs (5.3.1)–(5.3.2) is

$$u(t) = S(t)u_0(x) - \int_0^t S(t-s)a(x,s)ds + \sum_{i=1}^m \int_0^t S(t-s)h_i dw_i(s), \tag{5.3.4}$$

where $S(t) = \mathcal{F}_x^{-1}e^{-itm(\xi)}$ with $m(\xi) = \beta\xi^3 + \frac{\alpha}{\xi}$.

The space $\tilde{L}^2(\mathbb{R})$ is defined as follows:

$$\tilde{L}^2(\mathbb{R}) = \left\{ h\in L^2(\mathbb{R}) : \mathcal{F}_x^{-1}\left(\frac{\hat{h}(\xi)}{\xi}\right)\in L^2(\mathbb{R})\right\},$$

with the norm

$$\|h\|_{\tilde{L}^2(\mathbb{R})} = \|h\|_{L^2(\mathbb{R})} + \|\mathcal{F}_x^{-1}(\frac{\hat{h}(\xi)}{\xi})\|_{L^2(\mathbb{R})}.$$

The corresponding Sobolev space $\tilde{H}^s(\mathbb{R})$ is defined in a similar way

$$\tilde{H}^s(\mathbb{R}) = \left\{ h \in H^s(\mathbb{R}) : \mathcal{F}_x^{-1}\left(\frac{\hat{h}(\xi)}{\xi}\right) \in H^s(\mathbb{R}) \right\},$$

with the norm

$$\|h\|_{\tilde{H}^s(\mathbb{R})} = \|h\|_{H^s(\mathbb{R})} + \|\mathcal{F}_x^{-1}\left(\frac{\hat{h}(\xi)}{\xi}\right)\|_{H^s(\mathbb{R})}.$$

The definition of Bourgain space is given as follows.

Definition 5.3.1. *For $s, b \in \mathbb{R}$, the space $X_{s,b}$ is the completion of the Schwartz function space on $\mathbb{R}^2$ with respect to the norm*

$$\|u\|_{X_{s,b}} = \|\langle \tau + m(\xi) \rangle^b \langle \xi \rangle^s \hat{u}(\xi, \tau)\|_{L^2_\xi L^2_\tau},$$

where $\langle \cdot \rangle = (1+|\cdot|)$. Similar to $\tilde{H}^s(\mathbb{R})$, the modified Bourgain function space $\tilde{X}_{s,b}$ is defined as follows:

$$\|u\|_{\tilde{X}_{s,b}} = \|u(t)\|_{X_{s,b}} + \|D_x^{-1} u(t)\|_{X_{s,b}}.$$

For $T_1, T_2 > 0$, $\tilde{X}_{s,b}^{[T_1,T_2]}$ is defined by the space restricted to $[T_1, T_2]$ of functions in $\tilde{X}_{s,b}$ with the norm

$$\|u\|_{\tilde{X}_{s,b}^{[T_1,T_2]}} = inf\{\|\tilde{u}\|_{\tilde{X}_{s,b}}, \tilde{u} \in \tilde{X}_{s,b}, u = \tilde{u}|_{[T_1,T_2]}\}.$$

Let $\tilde{X}_{s,b}^T$ be defined by the space restricted to $[0, T]$ of functions in $\tilde{X}_{s,b}$.

Let $\zeta \in C_c^\infty(\mathbb{R})$ be a decreasing function with $\zeta \equiv 1$ on $[0,1]$ and supp$\zeta \subseteq [-1,2]$, and $\zeta_\epsilon(t) = \zeta(t/\epsilon), \epsilon > 0$. Denote the hyperplane $\mathbf{S} = \{(\xi, \tau, \xi_1, \tau_1, \xi_2, \tau_2) : \xi = \xi_1 + \xi_2, \tau = \tau_1 + \tau_2\}$. Throughout this section, denote the integral

$$\int_D h(\xi, \tau, \xi_1, \tau_1, \xi_2, \tau_2) d\delta = \int_{D'} h(\xi, \tau, \xi_1, \tau_1, \xi - \xi_1, \tau - \tau_1) d\xi d\tau d\xi_1 d\tau_1,$$

where $D \subset \mathbf{S}$ and $D' = \{(\xi, \tau, \xi_1, \tau_1) : (\xi, \tau, \xi_1, \tau_1, \xi_2, \tau_2)\} \in D$.

Now, we get the stochastic estimate and the bilinear estimates. Hereafter, C denotes a positive constant, whose value may change from one place to another.

Proposition 5.3.1. *Let* $0 < b < 1/2$, ζ *be defined as above, and* $h_i \in \tilde{L}^2(\mathbb{R})$ $(i = 1, 2, \cdots, m)$. *Then* $\phi = \sum_{i=1}^{m} \int_0^t S(t-s)h_i dw_i(s)$ *satisfies* $\zeta\phi \in L^2(\Omega; \tilde{X}_{0,b})$ *and*

$$\boldsymbol{E}(\|\zeta\phi\|^2_{\tilde{X}_{0,b}}) \le C \sum_{i=1}^{m} \|h_i\|^2_{\tilde{L}^2(\mathbb{R})},$$

where C is a constant depending on b, $\|\zeta\|_{\tilde{H}^b(\mathbb{R})}$, $\||t|^{1/2}\zeta\|_{\tilde{L}^2(\mathbb{R})}$, $\||t|^{1/2}\zeta\|_{L^\infty(\mathbb{R})}$.

Proof. The method in Ref. [27] can be applied here with little modification, so we omit it. ■

In order to obtain the bilinear estimates, we give some notations and lemmas. Denote

$$\sigma = \tau + m(\xi),\ \ \sigma_j = \tau_j + m(\xi_j), j = 1, 2,\ \ \mathcal{F}F_\kappa(\xi, \tau) = \frac{h(\xi, \tau)}{\langle\sigma\rangle^\kappa}.$$

Lemma 5.3.1. ([122]). *Let* $\kappa > 3/8$, $\rho > \frac{3(q-2)}{4q}$, $2 \le q \le 6$. *Then*

$$\|D_x^{\frac{1}{8}} F_\kappa\|_{L_t^4 L_x^4} \le C\|h\|_{L_\xi^2 L_\tau^2},$$
$$\|F_\rho\|_{L_t^q L_x^q} \le C\|h\|_{L_\xi^2 L_\tau^2}.$$

Lemma 5.3.2. ([152]). *If* $\frac{1}{2} < \ell < 1$, *then*

$$\int_{\mathbb{R}} \frac{dx}{(1+|x-\alpha|)^\ell (1+|x-\beta|)^\ell} \le \frac{C}{(1+|\alpha-\beta|)^{2\ell-1}}. \tag{5.3.5}$$

Proposition 5.3.2. *Let* $1/2 < b < 5/8$ *and* $3/8 < b' < 1/2$. *Assume that the Fourier transform* $\mathcal{F}u$ *of* u *is supported in* $\{(\xi, \tau) : |\xi| \ge N\}, N > 0$. *Then*

$$\|\partial_x(u_1 u_2)\|_{\tilde{X}_{0,b-1}} \le \frac{C}{N^{\frac{1}{8}}} \|u_1\|_{\tilde{X}_{0,b'}} \|u_2\|_{\tilde{X}_{0,b'}}. \tag{5.3.6}$$

Remark 5.3.1. *The similar result in Ref. [122] is* $1/2 < b < 9/16$ *and* $b' > 1/2$, *while* $b' < 1/2$ *is needed in our case. Fortunately, we can prove Proposition 3.2 with litter modification of the method used in Theorem 4.1 in Ref. [122].*

Proof. Because the proof of the bilinear estimate for $\|D_x^{-1}\partial_x(u_1u_2)\|_{X_{0,b-1}}$ can be obtained in a similar way, we only prove the bilinear estimate for $\|\partial_x(u_1u_2)\|_{X_{0,b-1}}$. By duality, it suffices to show

$$\int_{\mathbf{S}} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \mathcal{F}u_1 \mathcal{F}u_2 d\delta \le \frac{C}{N^{\frac{1}{8}}} \|h\|_{L^2_\xi L^2_\tau} \|f_1\|_{L^2_\xi L^2_\tau} \|f_2\|_{L^2_\xi L^2_\tau},$$

for $h \in L^2(\mathbb{R}), h \ge 0, f_j = \langle\sigma_j\rangle^{b'} \mathcal{F}u_j, j = 1, 2$. Let

$$\mathcal{F}F^j_\kappa(\xi,\tau) = \frac{f_j(\xi,\tau)}{\langle\sigma_j\rangle^\kappa},\ j = 1, 2.$$

In order to bound the integral, we split the domain of integration into two pieces. By symmetry, it suffices to estimate the integral in the domain $|\xi_1| \le |\xi_2|$ and so that $|\xi| \le 2|\xi_2|$.

Case 1. $|\xi| \le 4a$. Set $D_1 = \{(\xi,\tau,\xi_1,\tau_1,\xi_2,\tau_2) \in S : N \le 2a \le |\xi_1| \le |\xi_2|, |\xi| \le 4a\}$. Since $b < \frac{5}{8}$ and $b' > \frac{3}{8}$, by Lemma 5.3.1, we have

$$\begin{aligned}\int_{D_1} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta &\le \frac{C}{N^{\frac{1}{8}}} \int_{D_1} \frac{h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{|\xi_1|^{\frac{1}{8}} f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\ &\le \frac{C}{N^{\frac{1}{8}}} \|F_{1-b}\|_{L^2_x L^2_t} \|D_x^{\frac{1}{8}} F^1_{b'}\|_{L^4_x L^4_t} \|F^2_{b'}\|_{L^4_x L^4_t} \\ &\le \frac{C}{N^{\frac{1}{8}}} \|h\|_{L^2_\xi L^2_\tau} \|f_1\|_{L^2_\xi L^2_\tau} \|f_2\|_{L^2_\xi L^2_\tau}.\end{aligned} \tag{5.3.7}$$

Case 2. $|\xi| > 4a$. Set D_2 for the corresponding region. Notice that

$$\sigma - \sigma_1 - \sigma_2 = 3\beta\xi\xi_1\xi_2 \left(1 - \alpha\frac{\xi_1^2 + \xi^2 + \xi_1\xi_2}{3\beta(\xi\xi_1\xi_2)^2}\right),$$

which implies that if $|\xi_2| \ge |\xi_1| \ge 2a \ge N, |\xi| \ge 4a \ge 2N$, then

$$\max(|\sigma|, |\sigma_1|, |\sigma_2|) \ge C|\xi\xi_1\xi_2|.$$

Then one of the following cases always occurs:

$$\text{(a) } |\sigma| \ge |\xi||\xi_1||\xi_2|;\quad \text{(b) } |\sigma_1| \ge |\xi||\xi_1||\xi_2|;\quad \text{(c) } |\sigma_2| \ge |\xi||\xi_1||\xi_2|.$$

If (a) holds, let D_{21} be the corresponding region. Since $b < \frac{5}{8}$ and $b' > \frac{3}{8}$, by Lemma 5.3.1, we have

$$\begin{aligned}\int_{D_{21}} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta &\le \int_{D_{21}} \frac{|\xi| h(\xi,\tau)}{(|\xi||\xi_1||\xi_2|)^{1-b}} \frac{f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\ &\le \int_{D_{21}} |\xi|^b h(\xi,\tau) \frac{|\xi_1|^{\frac{1}{8}} f_1(\xi_1,\tau_1)}{|\xi_1|^{\frac{9}{8}-b}\langle\sigma_1\rangle^{b'}} \frac{|\xi_2|^{\frac{1}{8}} f_2(\xi_2,\tau_2)}{|\xi_2|^{\frac{9}{8}-b}\langle\sigma_2\rangle^{b'}} d\delta\end{aligned}$$

$$
\begin{aligned}
&\le \frac{C}{N^{\frac{9}{4}-3b}} \int_{D_{21}} h(\xi,\tau) \frac{|\xi_1|^{\frac{1}{8}} f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{|\xi_2|^{\frac{1}{8}} f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\
&\le \frac{C}{N^{\frac{9}{4}-3b}} \|F_0\|_{L_x^2L_t^2} \|D_x^{\frac{1}{8}} F_{b'}^1\|_{L_x^4L_t^4} \|D_x^{\frac{1}{8}} F_{b'}^2\|_{L_x^4L_t^4} \\
&\le \frac{C}{N^{\frac{9}{4}-3b}} \|h\|_{L_\xi^2L_\tau^2} \|f_1\|_{L_\xi^2L_\tau^2} \|f_2\|_{L_\xi^2L_\tau^2}.
\end{aligned} \tag{5.3.8}
$$

If (b) holds, set D_{22} for the corresponding region. Since $b < \frac{5}{8}$ and $b' > \frac{3}{8}$, by Lemma 5.3.1, we have

$$
\begin{aligned}
\int_{D_{22}} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta &\le \int_{D_{22}} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{(|\xi||\xi_1||\xi_2|)^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\
&\le C \int_{D_{22}} |\xi|^{\frac{7}{8}-b'} \frac{|\xi|^{\frac{1}{8}} h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{|\xi_1|^{b'}} \frac{|\xi_2|^{\frac{1}{8}} f_2(\xi_2,\tau_2)}{|\xi_2|^{\frac{1}{8}+b'}\langle\sigma_2\rangle^{b'}} d\delta \\
&\le \frac{C}{N^{3b'-\frac{3}{4}}} \int_{D_{22}} \frac{|\xi|^{\frac{1}{8}} h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} f_1(\xi_1,\tau_1) \frac{|\xi_2|^{\frac{1}{8}} f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\
&\le \frac{C}{N^{3b'-\frac{3}{4}}} \|D_x^{\frac{1}{8}} F_{1-b}\|_{L_x^4L_t^4} \|F_0^1\|_{L_x^2L_t^2} \|D_x^{\frac{1}{8}} F_{b'}^2\|_{L_x^4L_t^4} \\
&\le \frac{C}{N^{3b'-\frac{3}{4}}} \|h\|_{L_\xi^2L_\tau^2} \|f_1\|_{L_\xi^2L_\tau^2} \|f_2\|_{L_\xi^2L_\tau^2}.
\end{aligned} \tag{5.3.9}
$$

The argument for (c) is similar to (b).

By eqs (5.3.7)–(5.3.9), we have

$$
\|\partial_x(u_1u_2)\|_{\tilde{X}_{0,b-1}} \le \frac{C}{N^\varsigma} \|u_1\|_{\tilde{X}_{0,b'}} \|u_2\|_{\tilde{X}_{0,b'}},
$$

where $\varsigma = \min\{\frac{1}{8}, \frac{9}{4} - 3b, 3b' - \frac{3}{4}\}$. Since $1/2 < b < 5/8$ and $3/8 < b' < 1/2$, it follows that $\varsigma = \frac{1}{8}$. So the proof is completed. ∎

Proposition 5.3.3. *Let $\frac{1}{2} < b < \frac{9}{16}$, $\frac{25}{64} < b' < \frac{1}{2}$. Then we have*

$$
\|\partial_x(u_1u_2)\|_{\tilde{X}_{0,b-1}} \le C\|u_1\|_{\tilde{X}_{0,b'}} \|u_2\|_{\tilde{X}_{0,b'}}. \tag{5.3.10}
$$

Proof. Similar to Proposition 5.3.2, we only prove the bilinear estimate for $\|\partial_x(u_1u_2)\|_{X_{0,b-1}}$. By symmetry, it suffices to estimate the integral in the domain $|\xi_1| \le |\xi_2|$. Denote $\mathcal{F}F_\kappa^j, f_j, j = 1, 2$ as in Proposition 5.3.2. By duality, it suffices to show that

$$
\int_{\mathbf{S}} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \mathcal{F}u_1 \mathcal{F}u_2 d\delta \le C\|h\|_{L_\xi^2L_\tau^2} \|f_1\|_{L_\xi^2L_\tau^2} \|f_2\|_{L_\xi^2L_\tau^2}.
$$

Case 1. $|\xi| \leq 16$. Set E_1 for the corresponding region. For $b < 1$ and $b' > 1/3$, by Lemma 5.3.1, we have

$$\begin{aligned}\int_{E_1} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta &\leq C\|F_{1-b}\|_{L_x^2L_t^2}\|F_{b'}^1\|_{L_x^4L_t^4}\|F_{b'}^2\|_{L_x^4L_t^4} \\ &\leq C\|h\|_{L_\xi^2L_\tau^2}\|f_1\|_{L_\xi^2L_\tau^2}\|f_2\|_{L_\xi^2L_\tau^2}.\end{aligned}$$

Case 2. $|\xi_2| \geq |\xi_1| \geq 4$, and $|\xi| \geq 16$. Notice that

$$\sigma - \sigma_1 - \sigma_2 = 3\beta\xi\xi_1\xi_2 \left(1 - \alpha \frac{\xi_1^2 + \xi^2 + \xi_1\xi_2}{3\beta(\xi\xi_1\xi_2)^2}\right),$$

which implies that if $|\xi| \geq 2$, $|\xi_1| \geq 2$ and $|\xi_2| \geq 2$, then

$$\max(|\sigma|, |\sigma_1|, |\sigma_2|) \geq C|\xi\xi_1\xi_2|.$$

Then one of the following cases always occurs:

$$\text{(a) } |\sigma| \geq |\xi||\xi_1||\xi_2|; \quad \text{(b) } |\sigma_1| \geq |\xi||\xi_1||\xi_2|; \quad \text{(c) } |\sigma_2| \geq |\xi||\xi_1||\xi_2|.$$

If (a) holds, set E_{21} for the corresponding region. Since $|\xi| \leq 2|\xi_2|$, by Lemma 5.3.1, we have

$$\begin{aligned}\int_{E_{21}} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta &\leq \int_{E_{21}} \frac{|\xi| h(\xi,\tau)}{(|\xi||\xi_1||\xi_2|)^{1-b}} \frac{f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\ &\leq C\int_{E_{21}} h(\xi,\tau) \frac{|\xi_1|^{b-1} f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{|\xi_2|^{2b-1} f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\ &\leq C\|F_0\|_{L_x^2L_t^2}\|D_x^{\frac{1}{8}} F_{b'}^1\|_{L_x^4L_t^4}\|D_x^{\frac{1}{8}} F_{b'}^2\|_{L_x^4L_t^4} \\ &\leq C\|h\|_{L_\xi^2L_\tau^2}\|f_1\|_{L_\xi^2L_\tau^2}\|f_2\|_{L_\xi^2L_\tau^2},\end{aligned}$$

for $2b - 1 < \frac{1}{8}$, i.e. $b < \frac{9}{16}$ and $b' > \frac{3}{8}$.

If (b) holds, set E_{22} for the corresponding region. By Lemma 5.3.1, we have

$$\begin{aligned}\int_{E_{22}} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{\langle\sigma_1\rangle^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta &\leq \int_{E_{22}} \frac{|\xi| h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} \frac{f_1(\xi_1,\tau_1)}{(|\xi||\xi_1||\xi_2|)^{b'}} \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\ &\leq C\int_{E_{22}} \frac{|\xi|^{1-b'} h(\xi,\tau)}{\langle\sigma\rangle^{1-b}} f_1(\xi_1,\tau_1) \frac{f_2(\xi_2,\tau_2)}{\langle\sigma_2\rangle^{b'}} d\delta \\ &\leq C\|D_x^{\frac{1}{8}} F_{1-b}\|_{L_x^4L_t^4}\|F_0^1\|_{L_x^2L_t^2}\|F_{b'}^2\|_{L_x^4L_t^4} \\ &\leq C\|h\|_{L_\xi^2L_\tau^2}\|f_1\|_{L_\xi^2L_\tau^2}\|f_2\|_{L_\xi^2L_\tau^2},\end{aligned}$$

for $\frac{1}{3} < b' < \frac{1}{2}$, $\frac{1}{2} < b < \frac{5}{8}$.

The argument for (c) is similar to (b).

Case 3. $|\xi_1| \le 4$ or $|\xi_2| \le 4$, and $|\xi| \ge 16$.

Using Cauchy–Schwarz inequality and Fubini's theorem, it follows that

$$\begin{aligned}
&\| \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\langle\sigma\rangle^{b-1}|\xi|}{\langle\sigma_2\rangle^{b'}\langle\sigma_1\rangle^{b'}} f_1(\xi_1,\tau_1) f_2(\xi_2,\tau_2) d\xi_1 d\tau_1 \|_{L^2_\xi L^2_\tau} \\
&\le C \| \frac{|\xi|}{\langle\sigma\rangle^{1-b}} \left(\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{\langle\sigma_1\rangle^{2b'}\langle\sigma_2\rangle^{2b'}} d\xi_1 d\tau_1 \right)^{1/2} \|_{L^\infty_\xi L^\infty_\tau} \|u_1\|_{X_{0,b'}} \|u_2\|_{X_{0,b'}} .
\end{aligned} \tag{5.3.11}$$

Let $q := \frac{\xi_1}{\xi}$ and

$$\begin{aligned}
z_1 := \sigma - \sigma_1 - \sigma_2 &= 3\beta\xi\xi_1\xi_2 + \frac{\alpha}{\xi}\left[1 - \frac{\xi^3}{\xi\xi_1\xi_2}\right] \\
&= 3\beta\xi^3 q(1-q) + \frac{\alpha}{\xi}\left[1 - \frac{1}{q(1-q)}\right].
\end{aligned}$$

Then

$$\frac{dz_1}{d\xi_1} = 3\beta\xi^2(1-2q)\left[1 + \frac{\alpha}{3\xi^4(q-q^2)^2}\right]. \tag{5.3.12}$$

Noticing that $|q| = |\frac{\xi_1}{\xi}| < \frac{1}{4}$ and $\alpha > 0$, which implies from eq. (5.3.12) that $|\frac{dz_1}{d\xi_1}| \ge C\xi^2$, we conclude by Lemma 5.3.2 that

$$\begin{aligned}
&C\frac{|\xi|}{\langle\sigma\rangle^{1-b}} \left(\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{\langle\sigma_2\rangle^{2b'}\langle\sigma_1\rangle^{2b'}} d\xi_1 d\tau_1 \right)^{1/2} \\
&\le C\frac{|\xi|}{\langle\sigma\rangle^{1-b}} \left(\int_{-4}^{4} \frac{d\xi_1}{(1+|z_1-\sigma|)^{4b'-1}} \right)^{1/2} \\
&\le C\frac{|\xi|}{\langle\sigma\rangle^{1-b}} \frac{1}{|\xi|} \left(\int_{|z_1|\le 3|\sigma|} \frac{dz_1}{(1+|z_1-\sigma|)^{4b'-1}} \right)^{1/2} \\
&\le \frac{C}{\langle\sigma\rangle^{4b'-b-1}} \le C.
\end{aligned} \tag{5.3.13}$$

Estimates (5.3.11) and (5.3.13) complete the proof. ∎

Based on Propositions 5.3.1 and 5.3.3, the well-posedness of eqs (5.3.1)–(5.3.2) can be proved by the following lemma and contraction mapping principle.

Lemma 5.3.3. (see [27, 153]). *Let* $0 < a, b < 1/2$, $s \in \mathbb{R}$, $u_0 \in \tilde{H}^s(\mathbb{R})$, $h \in \tilde{X}^T_{s,c}$. *Then we have*

$$\|S(t)u_0\|_{\tilde{X}^T_{s,b}} \leq C\|u_0\|_{\tilde{H}^s},$$
$$\|\int_0^t S(t-s)h(s)ds\|_{\tilde{X}^T_{s,b}} \leq CT^{1-a-b}\|h\|_{\tilde{X}_{s,-a}}.$$

Theorem 5.3.1. *Let* $\lambda \in \mathbb{R}$, $f \in \tilde{L}^2(\mathbb{R})$, $h_i \in \tilde{L}^2(\mathbb{R})$ $(i = 1, 2, \cdots, m)$ *and* $u_0 \in \tilde{L}^2(\mathbb{R})$. *Then, for P-a.e.* $\omega \in \Omega$, *there is a* $T > 0$ *and a unique solution* $u(t)$ *of eqs (5.3.1)–(5.3.2) on* $[0, T]$, *which satisfies*

$$u \in C([0, T]; \tilde{L}^2(\mathbb{R})) \cap \tilde{X}^T_{0,b}.$$

Proof. From Proposition 5.3.1, we get

$$\|\phi\|_{\tilde{X}^T_{0,b}} \leq C\|\zeta\phi\|_{\tilde{X}_{0,b}}, \tag{5.3.14}$$

where ζ is given as before and $T \in [0, 1]$ for P-a.e. $\omega \in \Omega$. We fix $\omega \in \Omega$ such that eq. (5.3.14) and $u_0(\omega, \cdot) \in \tilde{L}^2(\mathbb{R})$ hold. Set

$$z(t) = U(t)u_0, \quad v(t) = u(t) - \phi(t) - z(t).$$

Then eq. (5.3.4) can be rewritten in terms of

$$\begin{aligned} v(t) = -\int_0^t S(t-s)[2(v+\phi+z)(v+\phi+z)_x \\ +\lambda(v+\phi+z) - f]ds. \end{aligned} \tag{5.3.15}$$

Let us introduce the complete metric space

$$B^T_R = \{v \in \tilde{X}^T_{0,b},\ \|v\|_{\tilde{X}^T_{0,b}} \leq R\},$$

with $R = \|\phi\|_{\tilde{X}^T_{0,b}} + \|u_0\|_{\tilde{L}^2(\mathbb{R})}$. We set

$$\Gamma v(t) = -\int_0^t S(t-s)[2(v+\phi+z)(v+\phi+z)_x + \lambda(v+\phi+z) - f]ds.$$

We will show that Γ is a contraction mapping in B^T_R, provided that T is chosen sufficiently small. With this aim in view, let $v, v_1, v_2 \in B^T_R$ be adapted processes. Noticing Propositions 5.3.1 and 5.3.3 and Lemma 5.3.3, we easily get

$$\|\Gamma v\|_{\tilde{X}^T_{0,b}} \leq CT^{1-a-b}(R^2 + \|\phi\|^2_{\tilde{X}^T_{0,b}} + \|u_0\|^2_{\tilde{L}^2(\mathbb{R})}),$$
$$\|\Gamma v_1 - \Gamma v_2\|_{\tilde{X}^T_{0,b}} \leq CT^{1-a-b}(R + \|\phi\|_{\tilde{X}^T_{0,b}} + \|u_0\|_{\tilde{L}^2(\mathbb{R})})\|v_1 - v_2\|_{\tilde{X}^T_{0,b}}.$$

Define the stopping time T by

$$T = \inf\{t > 0, 2Ct^{1-a-b}R \le 1/2\}.$$

Then Γ maps B_R^T in $\tilde{X}_{0,b}^T$ into itself, and

$$\|\Gamma v_1 - \Gamma v_2\|_{\tilde{X}_{0,b}^T} \le \frac{1}{2}\|v_1 - v_2\|_{\tilde{X}_{0,b}^T}.$$

Thus, the contraction mapping principle implies that there exists a unique solution u in $\tilde{X}_{0,b}^T$ on $[0, T]$ to eq. (5.3.15). It remains to show that the solution $u = z + v + \phi \in \tilde{X}_{0,c}^T + \tilde{X}_{0,b}^T$ in $C([0, T], \tilde{L}^2(\mathbb{R}))$ (note that here $b < 1/2, c > 1/2$). Since $c > 1/2$, we have $z \in C([0, T], \tilde{L}^2(\mathbb{R}))$ by the Sobolev imbedding theorem in time. Since $S(\cdot)$ is a unitary group in $\tilde{L}^2(\mathbb{R})$, we have ϕ as a continuous modification with values in $\tilde{L}^2(\mathbb{R})$ similar to Theorem 6.10 in Ref. [75].

Let $\tilde{u}$ be any prolongation of u in $\tilde{X}_{0,c} + \tilde{X}_{0,b}$, by Proposition 5.3.3, $\partial_x\tilde{u}^2 \in \tilde{X}_{0,-a}$ with $-\frac{1}{2} < a < 0$. It follows that (see Ref. [153])

$$\|\varphi_T \int_0^t S(t-s)\partial_x\tilde{u}^2 ds\|_{\tilde{X}_{0,1-a}} \le C\|\partial_x(\tilde{u}^2)\|_{\tilde{X}_{0,-a}}.$$

Since $1 - a > 1/2$, $\tilde{u} \in \tilde{X}_{0,1-a} \subset C([0, T], \tilde{L}^2(\mathbb{R}))$, where φ_T is a cutoff function defined by $\varphi \in C_0^\infty(\mathbb{R})$ with $\varphi = 1$ on $[0, 1]$, and $\varphi = 0$ on $t \le -1, t \ge 2$. Denote $\varphi_\delta(\cdot) = \varphi(\delta^{-1}(\cdot))$ for some $\delta \in \mathbb{R}$. This ends the proof of Theorem 5.3.1. ∎

Let $\Omega = \{\omega = (\omega_1, \omega_2, \cdots, \omega_m) \in C(\mathbb{R}; \mathbb{R}^m), \omega(0) = 0\}$, $\mathcal{F}$ be the Borel σ-algebra induced by the compact open topology of Ω, and P be the corresponding Wiener measure on Ω. In the sequel, we consider the probability space $(\Omega, \mathcal{F}, P)$. Then, we identify ω with

$$(w_1(t), w_2(t), \cdots, w_m(t)) = \omega(t), \ t \in \mathbb{R}.$$

Define the time shift by

$$\theta_t(\omega_s) = \omega(s + t) - \omega(t), \ t, s \in \mathbb{R}.$$

Then $(\Omega, \mathcal{F}, P, (\theta_t)_{t\in\mathbb{R}})$ is a metric dynamical system.

We now associate a continuous RDS with the stochastic damped forced Ostrovsky equation over $(\Omega, \mathcal{F}, P, (\theta_t)_{t\in\mathbb{R}})$. To this end, we introduce an auxiliary Ornstein–Uhlenbeck process, which enables us to change the stochastic equation (5.3.1) to a deterministic equation depending on a random parameter. Denote

$$\alpha_i = 2\|h_i\|_{\tilde{H}^{\frac{3}{2}}(\mathbb{R})}. \tag{5.3.16}$$

Then choose a sufficiently large κ such that

$$\kappa > \frac{8(\sum_{i=1}^{m} \alpha_i)^2}{\lambda^2}. \tag{5.3.17}$$

For each $i = 1, 2, \cdots, m$, let ρ_i be the stationary solution of the one-dimensional Itô equation

$$d\rho_i + \kappa\rho_i dt = dw_i, \tag{5.3.18}$$

the solution of which is called an Ornstein–Uhlenbeck process. We have

$$\rho_i(t) = \rho_i(\theta_t\omega_i) = -\kappa \int_{-\infty}^{0} e^{\kappa s}(\theta_t\omega_i)(\tau)d\tau, \ \ t \in \mathbb{R}.$$

For this solution, the random variable $|\rho_i(\omega_i)|$ is tempered, and $\rho_i(\theta_t\omega_i)$ is P-a.s. continuous. Setting $\rho(\theta_t\omega) = \Sigma_{i=1}^{m} h_i\rho_i(\theta_t\omega_i)$, by eq. (5.3.18), we obtain

$$d\rho + \kappa\rho dt = \sum_{i=1}^{m} h_i dw_i.$$

We now make the change $\eta(t) = u(t) - \rho(\theta_t\omega)$. Then η satisfies the following equation which depends on a random parameter:

$$\eta_t + (\eta^2)_x - \beta\eta_{xxx} - \alpha D_x^{-1}\eta + \lambda\eta = f + g, \tag{5.3.19}$$

where

$$g = (\kappa - \lambda)\rho(\theta_t\omega) + \beta\rho_{xxx}(\theta_t\omega) + \alpha D_x^{-1}\rho(\theta_t\omega) - (\rho(\theta_t\omega)^2)_x - 2(\rho(\theta_t\omega)\eta)_x.$$

Now, we give the global well-posedness of eq. (5.3.19) as follows.

Theorem 5.3.2. *Let $\eta_0 \in \tilde{L}^2(\mathbb{R})$. Then for P-a.e. $\omega \in \Omega$, there exists a unique solution $\eta(\cdot, \omega, \eta_0) \in C([0, +\infty); \tilde{L}^2(\mathbb{R}))$ of eq. (5.3.19) with initial value $\eta(0, \omega, \eta_0) = \eta_0$.*

Proof. The proof proceeds by a priori estimate. By the similar proof as in Ref. [122], one can see that for P-a.e. $\omega \in \Omega$ the above theorem holds. The detailed proof is omitted here. ■

By Theorem 5.3.2, we see that there is a continuous mapping from $\tilde{L}^2(\mathbb{R})$ into itself: $\eta_0 \to \eta(t, \omega, \eta_0)$, where $\eta(t, \omega, \eta_0)$ is the solution of eq. (5.3.19) with initial value η_0. Let $u(t, \omega, u_0) = \eta(t, \omega, u_0 - \rho(\omega)) + \rho(\theta_t\omega)$. Then the process u is the solution of problems (5.3.1)–(5.3.2). We can now define an RDS $\psi(t, \omega)$ in $\tilde{L}^2(\mathbb{R})$ by setting

$$\psi(t, \omega, u_0) = u(t, \omega, u_0) = \eta(t, \omega; u_0 - \rho(\omega)) + \rho(\theta_t\omega), \ \ t \geq 0.$$

Therefore, ψ is a continuous RDS associated with the stochastic damped forced Ostrovsky equation.

5.3.3 Uniform estimates of solutions

In this section, we derive uniform estimates on the solutions of eqs (5.3.1)–(5.3.2) when $t \to \infty$ with the purpose of proving the existence of a bounded random absorbing set and the asymptotic compactness of the RDS associated with the equation.

From now on, we always assume that $\mathcal{D}$ is the collection of all tempered subsets of $\tilde{L}^2(\mathbb{R})$ with respect to $(\Omega, \mathcal{F}, P, (\theta_t)_{\mathbb{R}})$.

5.3.3.1 Random absorbing set in $\mathcal{D}$.

We first derive the following uniform estimates on η in $\tilde{L}^2(\mathbb{R})$.

Lemma 5.3.4. *Assume $f \in \tilde{L}^2(\mathbb{R})$, $h_i \in \tilde{H}^3(\mathbb{R})$, $h_i^2 \in \tilde{H}^1(\mathbb{R})$ $(i = 1, 2, \cdots, m)$ and eq. (5.3.17) holds. Let $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ and $\eta_0(\omega) \in B(\omega)$. Then for P-a.e. $\omega \in \Omega$, there is $T = T(B, \omega) > 0$ such that for all $t \geq T$,*

$$\|\eta(t, \theta_{-t}\omega, \eta_0(\theta_{-t}\omega))\|_{\tilde{L}^2(\mathbb{R})} \leq r(\omega),$$

where $r(\omega)$ is a positive random function satisfying

$$e^{-\frac{1}{2}\lambda t} r(\theta_{-t}\omega) \to 0 \text{ as } t \to \infty. \tag{5.3.20}$$

Proof. Multiplying eq. (5.3.19) by η and $D_x^{-2}\eta$, then integrating it over $\mathbb{R}$, respectively, we can obtain,

$$\frac{d}{dt}\|\eta\|^2_{\tilde{L}^2(\mathbb{R})} + 2\lambda\|\eta\|^2_{\tilde{L}^2(\mathbb{R})} = 2(f + g, \eta)_{\tilde{L}^2(\mathbb{R})}. \tag{5.3.21}$$

Using integration by parts and Young's inequality, we have

$$\begin{aligned}
&2(f, \eta)_{\tilde{L}^2(\mathbb{R})} \leq \frac{\lambda}{5}\|\eta\|^2_{\tilde{L}^2(\mathbb{R})} + \frac{5}{\lambda}\|f\|^2_{\tilde{L}^2(\mathbb{R})},\\
&2((\kappa - \lambda)\rho(\theta_t\omega), \eta)_{\tilde{L}^2(\mathbb{R})} \leq \frac{\lambda}{5}\|\eta\|^2_{\tilde{L}^2(\mathbb{R})} + \frac{5(\kappa-\lambda)^2}{\lambda}\|\rho(\theta_t\omega)\|^2_{\tilde{L}^2(\mathbb{R})},\\
&2(\beta\rho_{xxx}(\theta_t\omega), \eta)_{\tilde{L}^2(\mathbb{R})} \leq \frac{\lambda}{5}\|\eta\|^2_{\tilde{L}^2(\mathbb{R})} + \frac{5\beta^2}{\lambda}\|\rho_{xxx}(\theta_t\omega)\|^2_{\tilde{L}^2(\mathbb{R})},\\
&2(\alpha D_x^{-1}\rho(\theta_t\omega), \eta)_{\tilde{L}^2(\mathbb{R})} \leq \frac{\lambda}{5}\|\eta\|_{\tilde{L}^2(\mathbb{R})} + \frac{5\alpha^2}{\lambda}\|D_x^{-1}\rho(\theta_t\omega)\|^2_{\tilde{L}^2(\mathbb{R})},\\
&2((\rho(\theta_t\omega)^2)_x, \eta)_{\tilde{L}^2(\mathbb{R})} \leq \frac{\lambda}{5}\|\eta\|^2_{\tilde{L}^2(\mathbb{R})} + \frac{5}{\lambda}\|\rho(\theta_t\omega)^2\|^2_{\tilde{H}^1(\mathbb{R})}.
\end{aligned} \tag{5.3.22}$$

Using integration by parts, it follows that

$$4((\rho(\theta_t\omega)\eta)_x, \eta)_{\tilde{L}^2(\mathbb{R})} = 2(\rho(\theta_t\omega)_x, \eta^2)_{\tilde{L}^2(\mathbb{R})} \leq 2\|\rho(\theta_t\omega)_x\|_{L^\infty}\|\eta\|^2_{\tilde{L}^2(\mathbb{R})}$$
$$\leq 2\sum_{i=1}^m \|h_i\|_{\tilde{H}^{3/2}(\mathbb{R})}|\rho_i(\theta_t\omega)|\|\eta\|^2_{\tilde{L}^2(\mathbb{R})}. \tag{5.3.23}$$

By eqs (5.3.21)–(5.3.23) and (5.3.16), we obtain

$$\frac{d}{dt}\|\eta\|^2_{\tilde{L}^2(\mathbb{R})} + \left(\lambda - \sum_{i=1}^m \alpha_i|\rho_i(\theta_t\omega)|\right)\|\eta\|^2_{\tilde{L}^2(\mathbb{R})} \leq \frac{5}{\lambda}\|f\|^2_{\tilde{L}^2(\mathbb{R})} + h(\theta_t\omega), \tag{5.3.24}$$

where

$$h(\theta_t\omega) = \frac{(\kappa-\lambda)^2}{\lambda}\|\rho(\theta_t\omega)\|^2_{\tilde{L}^2(\mathbb{R})}$$
$$+ \frac{5}{\lambda}\left(\beta^2\|\rho_{xxx}(\theta_t\omega)\|^2_{\tilde{L}^2(\mathbb{R})} + \alpha^2\|D_x^{-1}\rho(\theta_t\omega)\|^2_{\tilde{L}^2(\mathbb{R})} + \|\rho(\theta_t\omega)^2\|^2_{\tilde{H}^1(\mathbb{R})}\right).$$

Since $f \in \tilde{L}^2(\mathbb{R})$, $h_i \in \tilde{H}^3(\mathbb{R})$, $h_i^2 \in \tilde{H}^1(\mathbb{R})(i = 1, 2, \cdots, m)$, we have

$$\frac{5}{\lambda}\|f\|^2_{\tilde{L}^2(\mathbb{R})} + h(\theta_t\omega) \leq C\left(1 + \sum_{i=1}^m(|\rho_i(\theta_t\omega)|^2 + |\rho_i(\theta_t\omega)|^4)\right) := r_0(\theta_t\omega). \tag{5.3.25}$$

Applying Gronwall's lemma to eq. (5.3.24) and by eq. (5.3.25), we find that, for all $s \geq 0$,

$$\|\eta(s,\omega,\eta_0(\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \leq e^{-\lambda s+\sum_{i=1}^m \alpha_i\int_0^s|\rho_i(\theta_\tau\omega)|d\tau}\|\eta_0(\omega)\|^2_{\tilde{L}^2(\mathbb{R})}$$
$$+ \int_0^s e^{-\lambda(s-\sigma)+\sum_{i=1}^m \alpha_i\int_\sigma^s|\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma. \tag{5.3.26}$$

Replace ω by $\theta_{-t}\omega$ with $t \geq 0$ in eq. (5.3.26) to get that, for any $s \geq 0$ and $t \geq 0$,

$$\|\eta(s,\theta_{-t}\omega,\eta_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \leq e^{-\lambda s+\sum_{i=1}^m \alpha_i\int_0^s|\rho_i(\theta_{\tau-t}\omega)|d\tau}\|\eta_0(\theta_{-t}\omega)\|^2_{\tilde{L}^2(\mathbb{R})}$$
$$+ \int_0^s e^{-\lambda(s-\sigma)+\sum_{i=1}^m \alpha_i\int_\sigma^s|\rho_i(\theta_{\tau-t}\omega)|d\tau} r_0(\theta_{\sigma-t}\omega)d\sigma$$
$$= e^{-\lambda s+\sum_{i=1}^m \alpha_i\int_{-t}^{s-t}|\rho_i(\theta_\tau\omega)|d\tau}\|\eta_0(\theta_{-t}\omega)\|^2_{\tilde{L}^2(\mathbb{R})}$$
$$+ \int_{-t}^{s-t} e^{-\lambda(s-\sigma-t)+\sum_{i=1}^m \alpha_i\int_\sigma^{s-t}|\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma. \tag{5.3.27}$$

By eq. (5.3.27), it follows that for all $t \geq 0$,

$$\begin{aligned}\|\eta(t,\theta_{-t}\omega,\eta_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} &\le e^{-\lambda t+\sum_{i=1}^m \alpha_i \int_{-t}^0 |\rho_i(\theta_\tau\omega)|d\tau}\|\eta_0(\theta_{-t}\omega)\|^2_{\tilde{L}^2(\mathbb{R})}\\ &\quad+\int_{-t}^0 e^{\lambda\sigma+\sum_{i=1}^m \alpha_i \int_\sigma^0 |\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma.\end{aligned} \tag{5.3.28}$$

Note that $|\rho_i(\theta_s\omega)|(i = 1, 2, \cdots, m)$ is stationary and ergodic [66]. Then it follows from the ergodic theorem that

$$\lim_{t\to\infty}\frac{1}{t}\int_{-t}^0 |\rho_i(\theta_s\omega)|ds = E(|\rho_i(\omega)|). \tag{5.3.29}$$

Combining eq. (5.3.29) with the following inequality:

$$E(|\rho_i(\omega)|) \le (E(|\rho_i(\omega)|^2))^{1/2} \le \frac{1}{\sqrt{2\kappa}},$$

we obtain

$$\lim_{t\to\infty}\frac{1}{t}\int_{-t}^0 |\rho_i(\theta_s\omega)|ds < \frac{1}{\sqrt{2\kappa}}. \tag{5.3.30}$$

By eqs (5.3.17) and (5.3.30), there is $T_0(\omega) > 0$ such that for all $t > T_0(\omega)$,

$$\sum_{i=1}^m \alpha_i \int_{-t}^0 |\rho_i(\theta_s\omega)|ds < \frac{2\sum_{i=1}^m \alpha_i t}{\sqrt{2\kappa}} < \frac{1}{2}\lambda t. \tag{5.3.31}$$

By eqs (5.3.28) and (5.3.31), we find that, for all $t > T_0(\omega)$,

$$\begin{aligned}\|\eta(t,\theta_{-t}\omega,\eta_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} &\le e^{-\frac{1}{2}\lambda t}\|\eta_0(\theta_{-t}\omega)\|^2_{\tilde{L}^2(\mathbb{R})}\\ &\quad+\int_{-t}^0 e^{\lambda\sigma+\sum_{i=1}^m \alpha_i \int_\sigma^0 |\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma.\end{aligned} \tag{5.3.32}$$

Since $|\rho_i(\theta_\sigma\omega)|$ is tempered, the following integral is convergent:

$$r_1(\omega) = \int_{-\infty}^0 e^{\lambda\sigma+\sum_{i=1}^m \alpha_i \int_\sigma^0 |\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma. \tag{5.3.33}$$

Hence,

$$\|\eta(t,\theta_{-t}\omega,\eta_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \le e^{-\frac{1}{2}\lambda t}\|\eta_0(\theta_{-t}\omega)\|^2_{\tilde{L}^2(\mathbb{R})} + r_1(\omega). \tag{5.3.34}$$

By assumption, $\{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ is tempered and hence we have

$$e^{-\frac{1}{4}\lambda t}\|\eta_0(\theta_{-t}\omega)\|^2_{\tilde{L}^2(\mathbb{R})} \to 0 \text{ as } t \to \infty,$$

from which and eq. (5.3.34), it follows that there is $T = T(B, \omega) > 0$ such that for all $t \geq T$,

$$\|\eta(s, \theta_{-t}\omega, \eta_0(\theta_{-t}\omega))\|_{\tilde{L}^2(\mathbb{R})} \leq \sqrt{2r_1(\omega)} := r(\omega). \tag{5.3.35}$$

Next, we prove $r(\omega)$ satisfies eq. (5.3.20). Replacing ω by $\theta_{-t}\omega$ in eq. (5.3.33) we obtain that

$$\begin{aligned} r_1(\theta_{-t}\omega) &= \int_{-\infty}^{0} e^{\lambda\sigma + \sum_{i=1}^m \alpha_i \int_\sigma^0 |\rho_i(\theta_{\tau-t}\omega)|d\tau} r_0(\theta_{\sigma-t}\omega)d\sigma \\ &= \int_{-\infty}^{-t} e^{\lambda(\sigma+t) + \sum_{i=1}^m \alpha_i \int_\sigma^{-t} |\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma \\ &\leq e^{\frac{1}{2}\lambda t} \int_{-\infty}^{-t} e^{\frac{1}{2}\lambda\sigma + \sum_{i=1}^m \alpha_i \int_\sigma^0 |\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma \\ &\leq e^{\frac{1}{2}\lambda t} \int_{-\infty}^{0} e^{\frac{1}{2}\lambda\sigma + \sum_{i=1}^m \alpha_i \int_\sigma^0 |\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma. \end{aligned} \tag{5.3.36}$$

By eq. (5.3.31), we have

$$\begin{aligned} &e^{-\frac{1}{2}\lambda t} r(\theta_{-t}\omega) = e^{-\frac{1}{2}\lambda t}\sqrt{2r_1(\theta_{-t}\omega)} \\ &\leq e^{-\frac{1}{4}\lambda t}\Big(\int_{-\infty}^{0} e^{\frac{1}{2}\lambda\sigma + \sum_{i=1}^m \alpha_i \int_\sigma^0 |\rho_i(\theta_\tau\omega)|d\tau} r_0(\theta_\sigma\omega)d\sigma\Big)^{1/2} \to 0, \text{ as } t \to \infty, \end{aligned}$$

which along with eq. (5.3.35) completes the proof. ∎

The next lemma shows that ψ has a random absorbing set in $\mathcal{D}$.

Lemma 5.3.5. *Assume $f \in \tilde{L}^2(\mathbb{R}), h_i \in \tilde{H}^3(\mathbb{R}), h_i^2 \in \tilde{H}^1(\mathbb{R})(i = 1, 2, \cdots, m)$ and eq. (5.3.17) holds. Then there exists $\{K(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ such that $\{K(\omega)\}_{\omega\in\Omega}$ is a random absorbing set for ψ in $\mathcal{D}$, that is, for any $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ and P-a.e. $\omega \in \Omega$, there is $T_1 > 0$ such that*

$$\psi(t, \theta_{-t}\omega, B(\theta_{-t}\omega)) \subset K(\omega) \ \forall t \geq T_1.$$

Proof. Let $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ and define

$$\begin{aligned} \tilde{B}(\omega) = \{\eta \in \tilde{L}^2(\mathbb{R}) : \|\eta\|_{\tilde{L}^2(\mathbb{R})} \leq \|u(\omega)\|_{\tilde{L}^2(\mathbb{R})} + \|\rho(\omega)\|_{\tilde{L}^2(\mathbb{R})}, \\ u(\omega) \in B(\omega)\}. \end{aligned} \tag{5.3.37}$$

We claim that if $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$, then $\tilde{B} = \{\tilde{B}(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$.

Note that $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ implies that

$$\lim_{t\to\infty} e^{-\frac{1}{2}\lambda t} d(B(\theta_{-t}\omega)) = 0. \tag{5.3.38}$$

Since $\rho(\omega)$ is tempered, eqs (5.3.37)–(5.3.38) imply that

$$\begin{aligned}\lim_{t\to\infty} e^{-\frac{1}{2}\lambda t} d(\tilde{B}(\theta_{-t}\omega)) &\le \lim_{t\to\infty} e^{-\frac{1}{2}\lambda t} d(B(\theta_{-t}\omega)) + \lim_{t\to\infty} e^{-\frac{1}{2}\lambda t}\|\rho(\theta_{-t}\omega)\|_{\tilde{L}^2} \\ &= 0, \end{aligned} \tag{5.3.39}$$

which shows that $\tilde{B} = \{\tilde{B}(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$. Then by Lemma 5.3.4, for P-a.e. $\omega \in \Omega$, if $\eta_0(\omega) \in \tilde{B}(\omega)$, there is $T_1 = T_1(\tilde{B}, \omega)$ such that for all $t \ge T_1$,

$$\|\eta(t, \theta_{-t}\omega, \eta_0(\theta_{-t}\omega))\|_{\tilde{L}^2(\mathbb{R})} \le r(\omega), \tag{5.3.40}$$

where $r(\omega)$ is a positive random function satisfying

$$\lim_{t\to\infty} e^{-\frac{1}{2}\lambda t} r(\theta_{-t}\omega) = 0. \tag{5.3.41}$$

Denote

$$K(\omega) = \{u \in \tilde{L}^2(\mathbb{R}) : \|u\|_{\tilde{L}^2(\mathbb{R})} \le r(\omega) + \|\rho(\omega)\|_{\tilde{L}^2(\mathbb{R})}\}. \tag{5.3.42}$$

Then, by eq. (5.3.41), we have

$$\begin{aligned}\lim_{t\to\infty} e^{-\frac{1}{2}\lambda t} d(K(\theta_{-t}\omega)) &\le \lim_{t\to\infty} e^{-\frac{1}{2}\lambda t} r(\theta_{-t}\omega) + \lim_{t\to\infty} e^{-\frac{1}{2}\lambda t}\|\rho(\theta_{-t}\omega)\|_{\tilde{L}^2(\mathbb{R})} \\ &= 0, \end{aligned} \tag{5.3.43}$$

which implies that $K = \{K(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$. We now show that K is also an absorbing set of ψ in $\mathcal{D}$. Given $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ and $u_0(\omega) \in B(\omega)$, since $\psi(t, \omega, u_0(\omega)) = \eta(t, \omega, u_0(\omega) - \rho(\omega)) + \rho(\theta_t\omega)$, by eq. (5.3.40), we get that, for all $t \ge T_1$,

$$\begin{aligned}\|\psi(t, \theta_{-t}\omega, \eta_0(\theta_{-t}\omega))\|_{\tilde{L}^2(\mathbb{R})} &\le \|\eta(t, \theta_{-t}\omega, \eta_0(\theta_{-t}\omega))\|_{\tilde{L}^2(\mathbb{R})} + \|\rho(\omega)\|_{\tilde{L}^2(\mathbb{R})} \\ &\le r(\omega) + \|\rho(\omega)\|_{\tilde{L}^2(\mathbb{R})}, \end{aligned} \tag{5.3.44}$$

which implies that

$$\psi(t, \theta_{-t}\omega, \eta_0(\theta_{-t}\omega)) \subset K(\omega) \quad \forall t \ge T_1, \tag{5.3.45}$$

hence $K = \{K(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ is a closed absorbing set of ψ in $\mathcal{D}$, which completes the proof. ■

To obtain the random attractor for eqs (5.3.1)–(5.3.2), we need to establish the existence of a compact random absorbing set in $\tilde{L}^2(\mathbb{R})$. This will be done as in Ref. [122]. That is, we split the solution η in eq. (5.3.19) into a decaying part v and a regular part δ, i.e.

$\eta = v + \delta$. To determine v and δ, we need to split eq. (5.3.19):

$$\delta_t + (\delta^2)_x - \beta\delta_{xxx} - \alpha D_x^{-1}\delta + \lambda\delta = f - 2P_N((\delta v)_x + v\, v_x) + P_N g, \tag{5.3.46}$$

$$v_t + P^N((v^2)_x) - \beta v_{xxx} - \alpha D_x^{-1} v + \lambda v = -2P^N((\delta v)_x) + P^N g, \tag{5.3.47}$$

with the initial date

$$\delta_0 = \delta(t=0) = P_N(u_0 + \rho(\omega)), \;\; v_0 = v(t=0) = P^N(u_0 + \rho(\omega)). \tag{5.3.48}$$

5.3.3.2 Well-posedness and decay of the v part of the solution

Use $\delta = \eta - v$ to rewrite eq. (5.3.47) as

$$v_t - P^N((v^2)_x) - \beta v_{xxx} - \alpha D_x^{-1} v + \lambda v = -2P^N((\eta v)_x) + P^N g. \tag{5.3.49}$$

The local well-posedness of eqs (5.3.49)–(5.3.48) can be proved in an analogous way to that for eqs (5.3.1)–(5.3.2). Furthermore, we may consider eq. (5.3.49) starting from any time $t_0 \geq 0$, i.e.

$$v_{t_0} = v(t_0, \omega, v_0(\omega)) \in \tilde{L}^2(\mathbb{R}). \tag{5.3.50}$$

Then, by the fixed argument, we find a solution $\tilde{v}$ of the equation

$$\tilde{v} = \zeta_T(t)S(t)v_{t_0} + \zeta_T(t)\int_0^t S(t-s)\{P^N(\tilde{v}^2)_x - \lambda\tilde{v} - 2P^N\partial_x(\eta\tilde{v}) - P^N g\}ds.$$

Applying the estimates from Section 5.3.2, we find that for T sufficiently small,

$$\|\tilde{v}\|_{\tilde{X}_{0,b'}} \leq C\|v_0\|_{\tilde{L}^2(\mathbb{R})},$$

for $0 < b' < 1/2$. Since $\tilde{v}$ coincides with the solution v of eqs (5.3.49)–(5.3.48) locally in time around the origin, we see that

$$\|v\|_{\tilde{X}_{0,b'}^{[-T,T]}} \leq C\|v_{t_0}\|_{\tilde{L}^2(\mathbb{R})}. \tag{5.3.51}$$

We may repeat the argument above for an interval centered at a different initial time t_0 (in the interval of definition of v) to obtain

$$\|v\|_{\tilde{X}_{0,b'}^{[t_0-T,t_0+T]}} \leq C\|v(t_0)\|_{\tilde{L}^2(\mathbb{R})}, \tag{5.3.52}$$

where $T = T(\|v(t_0)\|_{\tilde{L}^2(\mathbb{R})}, \|u(t_0)\|_{\tilde{L}^2(\mathbb{R})}, \|\rho\|_{\tilde{L}^2(\mathbb{R})}, \lambda)$ is small.

The following lemma gives the decay of v, obtained by the bilinear estimate (5.3.6).

Lemma 5.3.6. *Assume $f \in \tilde{L}^2(\mathbb{R}), h_i \in \tilde{H}^3(\mathbb{R}), h_i^2 \in \tilde{H}^1(\mathbb{R})$ $(i = 1, 2, \cdots, m)$. Let $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ and $u_0(\omega) \in B(\omega)$. Then for every $\varepsilon > 0$ and P-a.e. $\omega \in \Omega$, there exist $T_1 > 0$ and sufficient large N, the solution v of eqs (5.3.49) and (5.3.48) satisfies, for all $t \geq T_1$,*

$$\|v(t, \theta_{-t}\omega, v_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \leq \varepsilon.$$

Proof. Similar to eq. (5.3.21), we have

$$\frac{d}{dt}\|v\|^2_{\tilde{L}^2(\mathbb{R})} + 2\lambda\|v\|^2_{\tilde{L}^2(\mathbb{R})} = 2(-P^N((\eta v)_x) + P^N g, v)_{\tilde{L}^2(\mathbb{R})}. \tag{5.3.53}$$

By the bilinear estimates (5.3.6), Hölder's inequality and Young's inequality, we have

$$\begin{aligned} ((\eta v)_x), 2v)_{\tilde{L}^2(\mathbb{R})} &= -(\eta, (v^2)_x)_{\tilde{L}^2(\mathbb{R})} \\ &\leq \|\eta\|_{\tilde{X}_{0,1-b}}\|(v^2)_x\|_{\tilde{X}_{0,b-1}} \\ &\leq \frac{C}{N^{\frac{1}{8}}}\|\eta\|_{\tilde{X}_{0,b'}}\|v\|^2_{\tilde{X}_{0,b'}}, \end{aligned} \tag{5.3.54}$$

$$\begin{aligned} (2(\rho\eta)_x, 2v)_{\tilde{L}^2(\mathbb{R})} &\leq C\|v\|_{\tilde{X}_{0,1-b}}\|(\rho\eta)_x\|_{\tilde{X}_{0,b-1}} \\ &\leq \frac{C}{N^{\frac{1}{8}}}\|v\|_{\tilde{X}_{0,1-b}}\|\rho\|_{\tilde{X}_{0,b'}}\|\eta\|_{\tilde{X}_{0,b'}} \\ &\leq \frac{C}{N^{\frac{1}{4}}}\|\eta\|^2_{\tilde{X}_{0,b'}}\|\rho\|^2_{\tilde{X}_{0,b'}} + \frac{\lambda}{4}\|v\|^2_{\tilde{L}^2(\mathbb{R})}, \end{aligned} \tag{5.3.55}$$

$$((\rho^2)_x, v)_{\tilde{L}^2(\mathbb{R})} \leq \frac{C}{N^{\frac{1}{4}}}\|\rho\|^4_{\tilde{X}_{0,b'}} + \frac{\lambda}{4}\|v\|^2_{\tilde{L}^2(\mathbb{R})}, \tag{5.3.56}$$

$$\begin{aligned} &(P^N[(\kappa-\lambda)\rho + \beta\rho_{xxx} + \alpha D_x^{-1}\rho], 2v)_{\tilde{L}^2(\mathbb{R})} \\ &\leq C\|P^N(\rho + \rho_{xxx} + D_x^{-1}\rho)\|^2_{\tilde{L}^2(\mathbb{R})} + \frac{\lambda}{2}\|v\|^2_{\tilde{L}^2(\mathbb{R})}. \end{aligned} \tag{5.3.57}$$

From eqs (5.3.53) to (5.3.57), we obtain

$$\frac{d}{dt}\|v\|^2_{\tilde{L}^2(\mathbb{R})} + 2\lambda\|v\|^2_{\tilde{L}^2(\mathbb{R})} = g_1 + C\|P^N(\rho + \rho_{xxx} + P^N D_x^{-1}\rho)\|^2_{\tilde{L}^2(\mathbb{R})}, \tag{5.3.58}$$

where

$$g_1 = \frac{C}{N^{\frac{1}{8}}}\|\eta\|_{\tilde{X}_{0,b'}}\|v\|^2_{\tilde{X}_{0,b'}} + \frac{C}{N^{\frac{1}{4}}}\|\eta\|^2_{\tilde{X}_{0,b'}}\|\rho\|^2_{\tilde{X}_{0,b'}} + \frac{C}{N^{\frac{1}{4}}}\|\rho\|^4_{\tilde{X}_{0,b'}}.$$

Integrating eq. (5.3.58) from t_0 to t, we obtain, for all $t \geq t_0$

$$\begin{aligned}\|v(t,\omega,v_0(\omega)\|^2_{\tilde{L}^2(\mathbb{R})} \leq \|v(t_0,\omega,v_0(\omega)\|^2_{\tilde{L}^2(\mathbb{R})}e^{-2\lambda(t-t_0)} \\ + \int_{t_0}^{t} e^{-2\lambda(t-s)}[g_1(\omega) + C\|P^N(\rho+\rho_{xxx}+D_x^{-1}\rho)\|^2_{\tilde{L}^2(\mathbb{R})}]ds.\end{aligned}$$

Replacing ω by $\theta_{-t}\omega$ in the above, we find that, for all $t \geq 0$,

$$\begin{aligned}&\|v(t,\theta_{-t}\omega,v_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \\ &\leq \|v(t_0,\theta_{-t}\omega,v_0(\theta_{-t}\omega)\|^2_{\tilde{L}^2(\mathbb{R})}e^{2\lambda(t-t_0)} + \int_{t_0}^{t} e^{-2\lambda(t-s)}[g_1(\theta_{-t}\omega) \\ &\quad + C\|P^N(\rho(\theta_{s-t}\omega)+\rho_{xxx}(\theta_{s-t}\omega)+D_x^{-1}\rho(\theta_{s-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})}]ds.\end{aligned} \tag{5.3.59}$$

Similar to eq. (5.3.35), we have

$$\|v(t_0,\theta_{-t}\omega,v_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \leq 2r_1(\omega). \tag{5.3.60}$$

By eq. (5.3.60), we find that, given $\varepsilon > 0$, there is $T_1 = T_1(B,\omega,\varepsilon) > 0$ such that for all $t \geq T_1$,

$$\|v(t_0,\theta_{-t}\omega,v_0(\theta_{-t}\omega)\|^2_{\tilde{L}^2(\mathbb{R})}e^{-2\lambda(t-t_0)} \leq \varepsilon. \tag{5.3.61}$$

Using eq. (5.3.52) with $t_0 = s$, for $t \in [t_0, t_0+\epsilon]$, we obtain

$$\begin{aligned}\int_{t_0}^{t} e^{-2\lambda(t-s)}g_1(\theta_{-t}\omega)ds \leq \frac{C}{N^{\frac{1}{8}}}\int_{t_0}^{t} e^{-\lambda(t-s)}\big[\|\eta\|_{\tilde{X}_{0,b'}}\|v(s,\theta_{-t}\omega,v_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \\ + \|\rho\|^2_{\tilde{X}_{0,b'}}(\|\eta(s,\theta_{-t}\omega,v_0(\theta_{-t}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} + \|\rho\|^2_{\tilde{X}_{0,b'}})\big]ds.\end{aligned} \tag{5.3.62}$$

Similar to Proposition 5.3.1, we have

$$\mathbf{E}\|\rho\|^2_{\tilde{X}_{0,b'}} \leq C\sum_{i=1}^{m}\|h_i\|^2_{\tilde{L}^2(\mathbb{R})},$$

hence, for P-a.e. $\omega \in \Omega$, it follows

$$\|\rho\|_{\tilde{X}_{0,b'}} \leq C\sum_{i=1}^{m}\|h_i\|_{\tilde{L}^2(\mathbb{R})}. \tag{5.3.63}$$

Similar to eq. (5.3.52), we have

$$\|\eta\|_{\tilde{X}^{[t_0,t]}_{0,b'}} \leq C\|\eta(t_0)\|_{\tilde{L}^2(\mathbb{R})}. \tag{5.3.64}$$

By eqs (5.3.62)–(5.3.64), given $\varepsilon > 0$, for N large enough, we have

$$\int_{t_0}^{t} e^{-2\lambda(t-s)} g_1(\theta_{-t}\omega)ds < \varepsilon. \tag{5.3.65}$$

Since $h_i \in \tilde{H}^3(\mathbb{R})$, there is $N_1 > 0$ such that for all $N \geq N_1$,

$$\sum_{i=1}^{m} \int_{|x|\geq N} g_2 dx \leq \varepsilon, \tag{5.3.66}$$

where

$$g_2 = 2|h_i|^2 + 2|\mathcal{F}_x^{-1}\left(\frac{\hat{h}_i(\xi)}{\xi}\right)|^2 + |\mathcal{F}_x^{-1}(\xi^2\hat{h}_i(\xi))|^2 + 2|\mathcal{F}_x^{-1}\left(\frac{\hat{h}_i(\xi^2)}{\xi}\right)|^2.$$

By eq. (5.3.66), the last term on the RHS of eq. (5.3.59) satisfies

$$\begin{aligned}
&\int_{t_0}^{t} e^{-2\lambda(t-s)} C\|P^N(\rho(\theta_{s-t}\omega) + \rho_{xxx}(\theta_{s-t}\omega) + D_x^{-1}\rho(\theta_{s-t}\omega))\|_{\tilde{L}^2(\mathbb{R})}^2 ds \\
&\leq Cm^2 \int_{t_0}^{t} e^{-2\lambda(t-s)} \sum_{i=1}^{m} \int_{|x|\geq N} g_2 |\rho_i(\theta_{s-t}\omega)| dxds \\
&\leq \varepsilon C \int_{t_0}^{t} e^{-2\lambda(t-s)} |\rho_i(\theta_{s-t}\omega)| ds = \varepsilon C \int_{t_0-t}^{0} e^{\lambda s} |\rho_i(\theta_s\omega)| ds \\
&\leq \varepsilon C \int_{-\infty}^{0} e^{2\lambda s} |\rho_i(\theta_s\omega)| ds := \varepsilon C r_2(\omega).
\end{aligned} \tag{5.3.67}$$

This completes the proof. ■

Lemma 5.3.6 implies that the solution v is globally defined and decays to zero in $\tilde{L}^2(\mathbb{R})$.

5.3.3.3 Regularity of the δ part of the solution

Since η and v are defined globally in time, so is $\delta = \eta - v$. Next, we prove an $\tilde{H}^3(\mathbb{R})$ bound for $\delta = P_N\delta + P^N\delta$.

Using eq. (5.3.34), we first observe that

$$\begin{aligned}
&\limsup_{t\to\infty} \|P_N\delta(t, \theta_{-t}\omega, P_N\delta_0(\theta_{-t}\omega))\|_{\tilde{H}^3(\mathbb{R})}^2 \\
&\leq \limsup_{t\to\infty} N^3 \|\delta(t, \theta_{-t}\omega, P_N\delta_0(\theta_{-t}\omega))\|_{\tilde{L}^2(\mathbb{R})}^2 \\
&\leq \limsup_{t\to\infty} N^3 \|\eta(t, \theta_{-t}\omega, \eta_0(\theta_{-t}\omega))\|_{\tilde{L}^2(\mathbb{R})} \leq C(1 + r(\omega))N^3.
\end{aligned} \tag{5.3.68}$$

Hence, we focus on an $\tilde{H}^3(\mathbb{R})$ estimate for $P^N\delta = y$. From eqs (5.3.46), (5.3.48) and $P^N(P_N) = 0$, it shows that y is the unique solution to the following equation:

$$y_t + 2P^N(P_N\delta + y)(P_N\delta + y)_x - \beta y_{xxx} - \alpha D_x^{-1}y + \lambda y = P^N f, \tag{5.3.69}$$

$$y(0) = 0. \tag{5.3.70}$$

We can obtain that the $\tilde{H}^3(\mathbb{R})$ bound for y is equivalent to prove an $\tilde{L}^2(\mathbb{R})$ estimate on $y' = y_t$, which solves

$$y'_t + P^N((P_N\delta + y)y')_x - \beta y'_{xxx} - \alpha D_x^{-1}y' + \lambda y' = -P^N((P_N\delta + y)(P_N\delta)_t)_x, \tag{5.3.71}$$

$$y'(0) = P^N f - P^N(P_N\delta(0)(P_N\delta(0))_x). \tag{5.3.72}$$

Similar to the arguments in Lemmas 5.3.4 and 5.3.6, we can obtain

$$\|y'\|^2_{\tilde{L}^2(\mathbb{R})} \le C(N,\omega)\ \forall t > 0. \tag{5.3.73}$$

Therefore, there exists $C > 0$ such that

$$\|\delta(t,\theta_{-t}\omega,\delta_0(\theta_{-t}\omega))\|^2_{\tilde{H}^3(\mathbb{R})} \le C(N,\omega)\ \forall t > 0. \tag{5.3.74}$$

The following energy equation for δ will be used in the next section:

$$\begin{aligned}\|\delta(t)\|^2_{\tilde{L}^2(\mathbb{R})} =& \|\delta(t_0)\|^2_{\tilde{L}^2(\mathbb{R})} e^{-2\lambda(t-t_0)} + 2\int_{t_0}^t e^{-2\lambda(t-s)}(f,\delta)_{\tilde{L}^2(\mathbb{R})}ds \\ &- 4\int_{t_0}^t e^{-2\lambda(t-s)}(P_N((\delta v)_x + -_x),\delta)_{\tilde{L}^2(\mathbb{R})}ds \\ &+ 2\int_{t_0}^t e^{-2\lambda(t-s)}(g,\delta)_{\tilde{L}^2(\mathbb{R})}ds.\end{aligned} \tag{5.3.75}$$

5.3.4 Asymptotic compactness and random attractors

In this section, we prove the existence of a $\mathcal{D}$-random attractor for the RDS ψ associated with the stochastic damped forced Ostrovsky eqs (5.3.1)–(5.3.2) on $\mathbb{R}$. It follows from Lemma 5.3.5 that ψ has a closed random absorbing set in $\mathcal{D}$, which along with the $\mathcal{D}$-pullback asymptotic compactness will imply the existence of a unique $\mathcal{D}$-random attractor. The $\mathcal{D}$-pullback asymptotic compactness of ψ is given below.

Lemma 5.3.7. *Assume that $f \in \tilde{L}^2(\mathbb{R})$. Let $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$ and $u_{0,n}(\theta_{-t_n}\omega) \in B(\theta_{-t_n}\omega)$. Then the RDS ψ is $\mathcal{D}$-pullback asymptotically compact in $\tilde{L}^2(\mathbb{R})$, that is, for P-a.e. $\omega \in \Omega$, the sequence $\{\psi(t_n,\theta_{-t_n}\omega,u_{0,n}(\theta_{-t_n}\omega))\}$ has a convergent subsequence in $\tilde{L}^2(\mathbb{R})$ provided $t_n \to \infty$.*

Proof. Let $B = \{B(\omega)\}_{\omega\in\Omega} \in \mathcal{D}$, $u_{0,n}(\theta_{-t_n}\omega) \in B(\theta_{-t_n}\omega)$ and $t_n \to \infty$. For P-a.e. $\omega \in \Omega$, we have

$$\{\delta_n(t_n + \cdot, \theta_{-(t_n+\cdot)}\omega, \delta_{0,n}(\theta_{-(t_n+\cdot)}\omega))\}_n \\ \text{is bounded in } C([-T, T]; \tilde{H}^3(\mathbb{R})), \tag{5.3.76}$$

and, for the time derivative

$$\{\partial_t\delta_n(t_n + \cdot, \theta_{-(t_n+\cdot)}\omega, \delta_{0,n}(\theta_{-(t_n+\cdot)}\omega))\}_n \\ \text{is bounded in } C([-T, T]; \tilde{L}^2(\mathbb{R})), \tag{5.3.77}$$

for each $T > 0$ (and starting with sufficiently large n so that $t_n - T \geq 0$). By Arzela–Ascoli theorem, there is a subsequence of $\{\delta_n(t_n + \cdot, \theta_{-(t_n+\cdot)}\omega, \delta_{0,n}(\theta_{-(t_n+\cdot)}\omega))\}_n$ such that

$$\delta_m(t_m + \cdot, \theta_{-(t_m+\cdot)}\omega, \delta_{0,m}(\theta_{-(t_m+\cdot)}\omega)) \to \bar{\eta}(\cdot) \text{ strongly in } C([-T, T]; \tilde{H}^s_{\mathrm{loc}}(\mathbb{R})), \\ \text{weakly star in } L^\infty([-T, T]; \tilde{H}^3(\mathbb{R})), \tag{5.3.78}$$

for all $s \in [1, 3)$. From eq. (5.3.74), we find that

$$\delta_m(t_m + t, \theta_{-(t_m+t)}\omega, \delta_{0,m}(\theta_{-(t_m+t)}\omega)) \to \bar{\eta}(t), \\ \text{weakly in } \tilde{H}^3(\mathbb{R}) \ \forall t \in \mathbb{R}. \tag{5.3.79}$$

Thus, to prove

$$\delta_m(t_m, \theta_{-t_m}\omega, \delta_{0,m}(\theta_{-t_m}\omega)) \to \bar{\eta}(0) \text{ in } \tilde{L}^2(\mathbb{R}), \tag{5.3.80}$$

we need only to show that for P-a.e. $\omega \in \Omega$,

$$\limsup_{m\to+\infty} \|\delta_m(t_m, \theta_{-t_m}\omega, \delta_{0,m}(\theta_{-t_m}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \leq \|\bar{\eta}(0)\|^2_{\tilde{L}^2(\mathbb{R})}. \tag{5.3.81}$$

By Lemma 5.3.6, we find that

$$\|v_n(t_n + t, \theta_{-(t_n+t)}\omega, v_{0,n}(\theta_{-(t_n+t)}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \to 0 \\ \text{uniformly for } t \geq -T \ \forall T > 0. \tag{5.3.82}$$

By eqs (5.3.78) and (5.3.82), taking the limit in the weak formulation of the equation for δ_n, we can prove that $\bar{\eta}$ is a solution of eq. (5.3.19) and satisfies the energy eq. (5.3.21).

Integrating eq. (5.3.21) with $\eta = \bar{\eta}$ from $-T$ to 0, we obtain

$$\begin{aligned} &\|\bar{\eta}(0)\|^2_{\tilde{L}^2(\mathbb{R})} - \|\bar{\eta}(-T)\|^2_{\tilde{L}^2(\mathbb{R})} e^{-2\lambda T} \\ &= 2\int_0^T e^{-2\lambda(T-s)} (\bar{g}, \bar{\eta}(-T+s))_{\tilde{L}^2(\mathbb{R})}, \end{aligned} \tag{5.3.83}$$

where

$$\bar{g} = f + (\kappa - \lambda)\rho + \beta\rho_{xxx} + \alpha D_x^{-1}\rho - (\rho^2)_x - 2(\rho\bar{\eta}(-T+s))_x. \tag{5.3.84}$$

We now rewrite the energy equation (5.3.75) for $\delta_n(t_n)$ with $t = t_n$ and $t_0 = t_n - T$:

$$\begin{aligned} \|\delta_n(t_n)\|^2_{\tilde{L}^2(\mathbb{R})} = \|\delta_n(t_n - T)\|^2_{\tilde{L}^2(\mathbb{R})} e^{-2\lambda T} &+ 2\int_0^T e^{-2\lambda(T-s)} (f, \delta_n)_{\tilde{L}^2(\mathbb{R})} ds \\ &- 4\int_0^T e^{-2\lambda(T-s)} (P_N((\delta v)_x + v v_x), \delta_n)_{\tilde{L}^2(\mathbb{R})} ds \\ &+ 2\int_0^T e^{-2\lambda(T-s)} (g, \delta_n)_{\tilde{L}^2(\mathbb{R})} ds, \end{aligned} \tag{5.3.85}$$

where for notational simplicity, we omitted the argument $t_n - T + s$ of the functions inside the time integrals. By using the uniform boundedness of δ_n in $\tilde{H}^3(\mathbb{R})$, decay estimate of v_n in Lemma 5.3.6 and weak star convergence of δ, we obtain

$$\begin{aligned} &\limsup_{m\to+\infty} \|\delta_m(t_m)\|^2_{\tilde{L}^2(\mathbb{R})} \\ &\leq Ce^{-2\lambda T} + 2\int_0^T e^{-2\lambda(T-s)} (\bar{g}, \bar{\eta}(-T+s))_{\tilde{L}^2(\mathbb{R})}. \end{aligned} \tag{5.3.86}$$

By substituting eq. (5.3.83), we obtain

$$\begin{aligned} &\limsup_{m\to+\infty} \|\delta_m(t_m)\|^2_{\tilde{L}^2(\mathbb{R})} \\ &\leq Ce^{-2\lambda T} + \|\bar{\eta}(0)\|^2_{\tilde{L}^2(\mathbb{R})} - \|\bar{\eta}(-T)\|^2_{\tilde{L}^2(\mathbb{R})} e^{-2\lambda T}, \end{aligned} \tag{5.3.87}$$

for any $T > 0$. Let $T \to +\infty$, we find

$$\limsup_{m\to+\infty} \|\delta_m(t_m, \theta_{-t_m}\omega, \delta_{0,m}(\theta_{-t_m}\omega))\|^2_{\tilde{L}^2(\mathbb{R})} \leq C\|\bar{\eta}(0)\|^2_{\tilde{L}^2(\mathbb{R})}. \tag{5.3.88}$$

So we conclude that as $m \to \infty$

$$\begin{aligned} \eta_m(t_m, \theta_{-t_m}\omega, \eta_{0,m}(\eta_{-t_m}\omega)) &= \delta_m(t_m, \theta_{-t_m}\omega, \delta_{0,m}(\theta_{-t_m}\omega)) \\ &\quad + v_m(t_m, \theta_{-t_m}\omega, v_{0,m}(\theta_{-t_m}\omega)) \\ &\to \bar{\eta}(0), \ \text{ strongly in } \tilde{L}^2(\mathbb{R}). \end{aligned} \tag{5.3.89}$$

Since $\psi(t,\omega,u_0(\omega)) = \eta(t,\omega,u_0(\omega)-\rho(\omega)) + \rho(\theta_t\omega)$, by eq. (5.3.89), it shows that ψ is $\mathcal{D}$-pullback asymptotically compact in $\tilde{L}^2(\mathbb{R})$. ■

We are now in a position to present our main result as follows.

Theorem 5.3.3. *Assume that $f \in \tilde{L}^2(\mathbb{R})$. Then the RDS ψ has a unique $\mathcal{D}$-random attractor in $\tilde{L}^2(\mathbb{R})$.*

Proof. Notice that ψ has a closed random absorbing set $\{K(\omega)\}_{\omega\in\Omega}$ in $\mathcal{D}$ by Lemma 5.3.5 and is $\mathcal{D}$-pullback asymptotically compact in $\tilde{L}^2(\mathbb{R})$ by Lemma 5.3.7. Hence, the existence of a unique $\mathcal{D}$-random attractor for ψ is obtained immediately. ■

5.4 Simplified quasi-geostrophic model

In order to study the oceanic dynamics and its influence on global climate, we focus on the dynamic equations that describe the motion of the ocean. One of the pioneers of Meteorology is V. Bjerkness who pointed out that the weather forecast can be viewed as a set of initial boundary value problems on mathematical physics. Taking the Boussinesq approximation and the hydrostatic balance into account, the primitive equation of large-scale ocean [216] is derived from the full Boussinesq system.

Due to the complexity of the primitive equations, difficult to study both in theory and in numerical, Charney and Philips [56] proposed a simplified quasi-geostrophic model, which is an approximation of rotating shallow-water equations for the small Rossby number. Hereinafter, we will refer the quasi-geostrophic model as the QG model. This section will concentrate on the QG equation and the associated dynamical system with stochastic external force and the reader may refer to Refs [32, 83, 120] for more information. We consider the following 2D stochastic QG on a regular enough bounded domain $D \subset \mathbb{R}^2$,

$$\left(\frac{\partial}{\partial t} + \frac{\partial\psi}{\partial x}\frac{\partial}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial}{\partial x}\right)(\Delta\psi - F\psi + \beta_0 y) = \frac{1}{R_e}\Delta^2\psi - \frac{r}{2}\Delta\psi + f(x,y,t), \tag{5.4.1}$$

where ψ is the stream function, $\frac{1}{R_e}\Delta^2\psi$ is the viscous term, $-\frac{r}{2}\Delta\psi$ is the friction, F is the Froude number ($F \approx O(1)$), R_e is the Reynolds number ($R_e \geq 10^2$), β_0 is a positive constant ($\beta_0 \approx O(10^{-1})$), r is the Ekman dissipation constant ($r \approx O(1)$) and $f(x,y,t) = -\frac{dW}{dt}$ is a Gaussian random field, which is white noise in time and subjects to the restrictions imposed later.

Let $A = -\frac{1}{R_e}\Delta$, then $A : L^2(D) \to L^2(D)$ is defined in $D(A) = H^2(D) \cap H_0^1(D)$. Here $L^2(D), H^2(D)$ and $H_0^1(D)$ are the usual Sobolev spaces with $|\cdot|_p$ being the norm of $L^p(D)$ and $\|\cdot\|$ being the norm of $H_0^1(D)$. Then the operator A is positive, self-adjoint with the compact operator inverse operator A^{-1}. We denote by $0 < \lambda_1 < \lambda_2 \leq \cdots$ the eigenvalues of A and by $e_1, e_2, \cdots$ the corresponding eigenvectors that form a complete orthogonal basis of $H_0^1(D)$. Remark that for any $u \in H_0^1(D)$, we have inequality $\frac{1}{R_e}\|u\|^2 \geq \lambda_1|u|_2^2$. Finally, we note $\{e^{-tA}\}_{t\geq 0}$ is semigroup generated by operator $-A$ on $L^2(D)$.

Let the stochastic process W be a two-side Wiener process

$$W(t) = \sum_{i=1}^{\infty} \mu_i \omega_i(t) e_i,$$

where $\omega_1, \omega_2, \cdots$ is a collection of independent standard Brownian motions on a probability space $(\Omega, \mathscr{F}, P)$, and the coefficients μ_i satisfy the condition: there exists $\beta_1 > 0$, such that

$$\sum_{i=1}^{\infty} \frac{\mu_i^2}{\lambda_i^{1/2-2\beta_1}} < \infty.$$

Consider the following boundary conditions:

$$\begin{aligned} \psi(x,y,t) &= 0, \quad \partial D, \\ \Delta\psi(x,y,t) &= 0, \quad \partial D. \end{aligned} \tag{5.4.2}$$

For any $u \in L^2(D)$, by solving the elliptic equation with Dirichlet boundary condition

$$\begin{aligned} F\psi - \Delta\psi &= u, \\ \psi(x,y,t)|_{\partial D} &= 0, \end{aligned}$$

we get $\psi = (FI - \Delta)^{-1}u = B(u)$. Using the elliptic regularity theory, $B : L^2(D) \to H_0^1(D) \cap H^2(D)$ is established. Therefore, eq. (5.4.1) can be rewritten as

$$u_t + J(\psi, u) - \beta_0\psi_x = \frac{1}{R_e}\Delta u + \left(\frac{F}{R_e} - \frac{r}{2}\right)u - F\left(\frac{F}{R_e} - \frac{r}{2}\right)\psi + \frac{dW}{dt}, \tag{5.4.3}$$

where $\psi = B(u)$, J is the Jacobian operator defined by $J(\psi, u) = \frac{\partial\psi}{\partial x}\frac{\partial u}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial u}{\partial x}$. Defining $G(u) = -J(\psi, u) + \beta_0\psi_x + \left(\frac{F}{R_e} - \frac{r}{2}\right)u - F\left(\frac{F}{R_e} - \frac{r}{2}\right)\psi$, the problem is transformed into

$$u_t - \frac{1}{R_e}\Delta u = G(u) + \frac{dW}{dt}, \tag{5.4.4}$$

with the boundary condition and initial condition,

$$\begin{aligned} u|_{\partial D} &= 0, \\ u(x,y,0) &= u_0. \end{aligned} \tag{5.4.5}$$

We note that dynamical systems (5.4.1)–(5.4.2) correspond to the problem of eqs (5.4.4)–(5.4.5).

5.4.1 The existence and uniqueness of solution

Equations (5.4.4)–(5.4.5) are rewritten as the following abstract form:

$$\begin{cases} du = -Audt + G(u)dt + dW, \\ u(0) = u_0. \end{cases} \tag{5.4.6}$$

Consider the linear equation

$$\begin{cases} du = -Audt + dW, \\ u(0) = u_0. \end{cases}$$

The solution of the above linear equation is unique and can be expressed as

$$W_A(t) = \int_0^t e^{-A(t-s)} dW(s).$$

Here, $W_A(t)$ has continuous version that takes value in $D(A^{1/4+\beta})$, $\beta < \beta_1$ for P-a.e. $\omega \in \Omega$. Particularly, it has a continuous version in

$$C_0(D) := \{u : u \in C(D), u \text{ is compactly supported in } D\}.$$

Let

$$v(t) = u(t) - W_A(t), \quad t \geq 0,$$

then u satisfies eq. (5.4.6) if and only if v satisfies the equation

$$\begin{cases} \frac{dv}{dt} + Avdt = G(v(t) + W_A(t)), \\ v(0) = u_0. \end{cases} \tag{5.4.7}$$

This will be written in the following integral form:

$$v(t) = e^{-At} u_0 + \int_0^t e^{-A(t-s)} G(v + W_A) ds, \tag{5.4.8}$$

and we say the solution v of eq. (5.4.8) is the mild solution of eq. (5.4.7).

5.4.1.1 Local existence

In the following, we will employ the Banach fixed point theorem to prove that there exists a $T > 0$ such that eq. (5.4.8) has solutions in $C([0, T]; L^2(D))$, for any $m > 0$, we define

$$\Sigma(m, T) = \{v \in C([0, T]; L^2(D)) : |v(t)|_2 \leq m, \forall t \in [0, T]\}.$$

Lemma 5.4.1. *If for P-a.e.* $\omega \in \Omega$, $u_0 \in L^2(D)$ *and* $m > |u_0|_2$, *then there exists a* T *such that the integral eq.* (5.4.8) *has a unique solution in* $\Sigma(m,T)$, *and for P-a.e.* $\omega \in \Omega, v \in C((0,T]; H^\alpha(D))$ *with* $0 \le \alpha < \frac{1}{2}$, *here* $H^\alpha(D)$ *is the usual Sobolev space.*

Proof. First, we review some properties of the semigroup e^{-tA} [199]:

$$\begin{aligned}
&e^{-tA}A^\alpha = A^\alpha e^{-tA},\\
&|A^\alpha e^{-tA}u|_2 \le \frac{c}{t^\alpha}|u|_2,\\
&|e^{-tA}u|_2 \le c|u|_2.
\end{aligned}$$

Hereinafter, C represents the positive constant, which may vary from one line to another one. The definition of fractional-order differential operator A^α can be found in Ref. [199]. For fixed $\omega \in \Omega$, define

$$Mv(t) = e^{-tA}u_0 + \int_0^t e^{(t-s)(-A)}G(v(s) + W_A(s))ds,$$

then

$$|Mv(t)|_2 = \sup_{\varphi \in L^2(D), |\varphi|_2 = 1} |\langle Mv, \varphi\rangle|,$$

where $\langle\cdot,\cdot\rangle$ is the inner product in $L^2(D)$ and

$$\langle Mv, \varphi\rangle = \langle e^{-tA}u_0, \varphi\rangle + \int_0^t \langle e^{(t-s)(-A)}G(v(s) + W_A(s)), \varphi\rangle ds.$$

Assume that $\varphi \in C_0^\infty(D)$, $v + W_A \in H_0^1(D)$, which can be achieved by the fact that C_0^∞ is dense in $L^2(D)$, and here C_0^∞ denotes a collection of infinitely differentiable functions with compact support in D.

Let $\psi = B(v + W_A)$, and

$$\begin{aligned}
J =&\langle e^{(t-s)(-A)}G(v(s) + W_A(s)), \varphi\rangle\\
=&\int_D e^{(t-s)(-A)}\left\{\frac{\partial\psi}{\partial x}\frac{\partial(v+W_A)}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial(v+W_A)}{\partial x}\right.\\
&\left. + \beta_0(B(v+W_A))_x + \left(\frac{F}{R_e} - \frac{r}{2}\right)(v+W_A) - F\left(\frac{F}{R_e} - \frac{r}{2}\right)\psi\right\}\varphi\\
=&\int_D e^{(t-s)(-A)}\frac{\partial\psi}{\partial x}\frac{\partial(v+W_A)}{\partial y}\varphi + \int_D e^{(t-s)(-A)}\frac{\partial\psi}{\partial y}\frac{\partial(v+W_A)}{\partial x}\varphi\\
&+ \int_D e^{(t-s)(-A)}\left(\frac{F}{R_e} - \frac{r}{2}\right)(v+W_A)\varphi - \int_D e^{(t-s)(-A)}F\left(\frac{F}{R_e} - \frac{r}{2}\right)\psi\varphi\\
&+ \beta_0\int_D e^{(t-s)(-A)}(B(v+W_A))_x\varphi = J_1 + J_2 + J_3 + J_4 + J_5.
\end{aligned} \tag{5.4.9}$$

Using integration by parts and Hölder's inequality, we get

$$\begin{aligned}
|J_1| &= |\int_D e^{-(t-s)A}\frac{\partial\psi}{\partial x}\frac{\partial(v+W_A)}{\partial y}\varphi| \\
&= |\int_D (e^{-(t-s)A}\frac{\partial\psi}{\partial x}\varphi)_y(v+W_A)| \\
&\le \left(\int_D |(e^{-(t-s)A}\frac{\partial\psi}{\partial x}\varphi)_y|^2\right)^{1/2}|v+W_A|_2 \\
&\le \left(\int_D |(e^{-(t-s)A}\varphi)_y\frac{\partial\psi}{\partial x}+e^{-(t-s)A}\varphi\frac{\partial^2\psi}{\partial x\partial y}|^2\right)^{1/2}|v+W_A|_2 \\
&\le c\left[\left(\int_D |(e^{-(t-s)A}\varphi)_y\frac{\partial\psi}{\partial x}|^2\right)^{1/2}+|e^{-(t-s)A}\varphi|_\infty|v+W_A|\right]|v+W_A|_2,
\end{aligned} \tag{5.4.10}$$

where we used $\|\psi\|_{H^2}\le c|v+W_A|_2$ and $\|\cdot\|_{H^q}$ denotes the usual Sobolev norm of $H^q(D)$.

By Hölder's inequality, Sobolev's embedding theorem, Gagliardo–Nirenberg's inequality and Poincaré's inequality, the first summand on the RHS of inequality (5.4.10) has

$$\begin{aligned}
\left(\int_D |(e^{-(t-s)A}\varphi)_y\frac{\partial\psi}{\partial x}|^2\right)^{1/2} &\le \left(\int_D |(e^{-(t-s)A}\varphi)_y|^4\right)^{1/4}\left(\int_D |\frac{\partial\psi}{\partial x}|^4\right)^{1/4} \\
&\le c\|(e^{-(t-s)A}\varphi)_y\|_{H^{1/2}}\|\psi_x\|_{H^1} \\
&\le c|e^{-(t-s)A}\varphi\|_{H^{3/2}}|v+W_A|_2 \\
&\le c|A^{3/4}e^{-(t-s)A}\varphi_y\|_2|v+W_A|_2 \\
&\le c(t-s)^{-3/4}|\varphi|_2|v+W_A|_2.
\end{aligned} \tag{5.4.11}$$

Similarly, one can get

$$\begin{aligned}
|e^{-(t-s)A}\varphi)|_\infty|v+W_A|_2^2 &\le c|e^{-(t-s)A}\varphi\|_{H^{1+\varepsilon_0}}|v+W_A|_2^2 \\
&\le c|a^{\frac{1}{2}+\varepsilon_0}e^{-(t-s)A}\varphi\|_2|v+W_A|_2^2 \\
&\le c(t-s)^{-(\frac{1}{2}+\varepsilon_0)}|\varphi|_2|v+W_A|_2^2,
\end{aligned} \tag{5.4.12}$$

where ε_0 is a positive constant. Applying eqs (5.4.10)–(5.4.12), it is not difficult to get

$$|J_1|\le c(t-s)^{-3/4}|\varphi|_2|v+W_A|_2^2+c(t-s)^{-(\frac{1}{2}+\varepsilon_0)}|\varphi|_2|v+W_A|_2^2. \tag{5.4.13}$$

Similarly, we obtain

$$|J_2|\le c(t-s)^{-3/4}|\varphi|_2|v+W_A|_2^2+c(t-s)^{-\frac{1}{2}+\varepsilon_0}|\varphi|_2|v+W_A|_2^2, \tag{5.4.14}$$

$$\begin{aligned}|J_3| =&|\frac{F}{R_e}-\frac{r}{2}||\int_D e^{-(t-s)A}(v+W_A)\varphi| \\ \le& c|e^{-(t-s)A}(v+W_A)|_2|\varphi|_2 \le c|v+W_A|_2|\varphi|_2,\end{aligned} \tag{5.4.15}$$

$$\begin{aligned}|J_4| =&|F\left(\frac{F}{R_e}-\frac{r}{2}\right)||\int_D e^{-(t-s)A}\psi\varphi| \\ \le& c|e^{-(t-s)A}\psi|_2|\varphi|_2 \\ \le& c|\psi|_2|\varphi|_2 \le c|v+W_A|_2|\varphi|_2,\end{aligned} \tag{5.4.16}$$

$$\begin{aligned}|J_5| =&\beta_0|\int_D e^{-(t-s)A}(B(v+W_A))_x\varphi| \\ \le& \beta_0|\int_D e^{-(t-s)A}(B(v+W_A))_x|_2|\varphi|_2 \\ \le& c|(B(v+W_A))_x|_2|\varphi|_2 \le c|v+W_A|_2|\varphi|_2.\end{aligned} \tag{5.4.17}$$

Thus, by eq. (5.4.9) and the above estimates of J_i, we obtain

$$|Mv(t)|_2 \le |u_0|_2 + c(t^{1/4}+t^{\frac{1}{2}-\varepsilon_0})|v+W_A|_2^2 + ct|v+W_A|_2. \tag{5.4.18}$$

Obviously, for any $m > |u_0|_2$, there exists a $T_1 > 0$, such that $Mv \in \Sigma(m, T_1)$.

For any $v_1, v_2 \in \Sigma(m, T_1)$,

$$Mv_1 - Mv_2 = \int_0^t e^{-(t-s)A}(G(v_1+W_A(s)) - G(v_2+W_A(s)))ds.$$

Similar to eq. (5.4.18), one can get

$$\begin{aligned}|Mv_1 - Mv_2|_2 \le& c(t^{1/4}+t^{\frac{1}{2}-\varepsilon_0})\sup_{0\le s\le t}(|v_1(s)+W_A(s)|_2 + |v_2(s)+W_A(s)|_2) \\ &\times \sup_{0\le s\le t}|v_1(s)-v_2(s)|_2 + ct\sup_{0\le s\le t}|v_1(s)-v_2(s)|_2.\end{aligned}$$

Thus, one can choose an appropriate $T_2 > 0$ such that M is a contraction map.

Let $T = \min\{T_1, T_2\}$. By the Banach fixed point theorem, eq. (5.4.8) has a unique solution in $\Sigma(m, T)$. Noting for v, a similar result as eq. (5.4.18) enables us to obtain $v \in C((0, T]; H^\alpha(D))$, $0 \le \alpha < 1/2$. In fact, we can prove

$$|A^{\alpha/2}Mv|_2 \le t^{\frac{\alpha}{2}}|u_0|_2 + c(t^{\frac{1}{4}-\frac{\alpha}{2}} + t^{\frac{1}{2}-\varepsilon_0-\frac{\alpha}{2}})|v+W_A|_2^2 + ct^{1-\frac{\alpha}{2}}|v+W_A|_2.$$

∎

5.4.1.2 Global existence

In order to establish the existence of global solution, we need to make some prior estimates.

Lemma 5.4.2. *If for P-a.s. $\omega \in \Omega$, $v \in C([0,T]; L^2(D))$ is the solution of eq.* (5.4.8)*, then*

$$|v(t)|_2^2 \le (|u_0|_2^2 + cv_\infty^4 + cv_\infty^2)e^{c(v_\infty^2+1)t},$$

where $v_\infty = \sup_{0\le t\le T} |W_A(t)|_\infty$.

Proof. Fix $\omega \in \Omega$. Choose $\{u_0^n\} \subset C_0^\infty(D)$ such that $u_0^n \to u_0$ in $L^2(D)$ and let $\{W_A^n\}$ be a set of sufficiently smooth stochastic processes so that

$$W_A^n(t) = \int_0^t e^{-(t-s)A} dW^n(s) \to W_A(t) \text{ in } C([0,T_0]\times D) \text{ for a.s. } \omega \in \Omega.$$

From the proof of Lemma 5.4.1, we see that there exists $v^n \in C([0,T^n]; L^2(D))$, such that $T^n \to T$ and $v^n \to v$ is strong convergence in $C([0,T]; L^2(D))$, where v is a weak solution of eq. (5.4.8). Then we have

$$\frac{\partial v^n}{\partial t} + Av^n = G(v^n + W_A^n(t)),\ v^n(0) = u_0^n. \tag{5.4.19}$$

Doing inner product with v^n, we get

$$\frac{1}{2}\frac{d}{dt}|v^n|_2^2 + \frac{1}{R_e}\|v^n\|^2 = \int_D G(v^n + W_A^n(d))v^n. \tag{5.4.20}$$

By

$$\int_D J(B(v^n), v^n)v^n = 0,\quad \int_D J(B(W_A^n), v^n)v^n = 0,$$

we have

$$\begin{aligned}
&|\int_D G(v^n + W_A^n(d))v^n| \\
=&|\int_D \Big(-J(B(v^n + W_A^n(t)), v^n + W_A^n) + \beta_0(B(v^n + W_A^n))_x \\
&+\left(\frac{F}{R_e} - \frac{r}{2}\right)\cdot(v^n + W_A^n) - F\left(\frac{F}{R_e} - \frac{r}{2}\right)B(v^n + W_A^n)\Big)\cdot v^n| \\
\le&|\int_D J(B(W_A^n), W_A^n)v^n| + |\int_D J(B(v^n), W_A^n)v^n| \\
&+|\int_D \left(\frac{F}{R_e} - \frac{r}{2}\right)\cdot(v^n + W_A^n)v^n| + |\int_D F\left(\frac{F}{R_e} - \frac{r}{2}\right)B(v^n + W_A^n)v^n| \\
&+\beta_0|\int_D (B(v^n + W_A^n))_x v^n| =: I_1 + I_2 + I_3 + I_4 + I_5.
\end{aligned} \tag{5.4.21}$$

Now we estimate $I_i, 1 \le i \le 5$. Using integration by parts, Hölder's inequality and Young's inequality, we get

$$\begin{aligned}
I_1 &= |\int_D \frac{\partial B(W_A^n)}{\partial x}\frac{\partial W_A^n}{\partial y}v^n - \frac{\partial B(W_A^n)}{\partial y}\frac{\partial W_A^n}{\partial x}v^n| \\
&= |\int_D \frac{\partial B(W_A^n)}{\partial x}W_A^n\frac{\partial v^n}{\partial y} - \frac{\partial B(W_A^n)}{\partial y}W_A^n\frac{\partial v^n}{\partial x}| \\
&\le |W_A^n|_\infty|\frac{\partial B(W_A^n)}{\partial x}|_2|\frac{\partial v^n}{\partial y}|_2 + |W_A^n|_\infty|\frac{\partial B(W_A^n)}{\partial y}|_2|\frac{\partial v^n}{\partial x}|_2 \\
&\le c|W_A^n|_\infty^2(|\frac{\partial B(W_A^n)}{\partial x}|_2^2 + |\frac{\partial B(W_A^n)}{\partial y}|_2^2)| + \varepsilon\|v^n\|_2 \\
&\le c|W_A^n|_\infty^2|W_A^n|_2^2 + \varepsilon\|v^n\|^2 \le c|W_A^n|_\infty^4 + \varepsilon\|v^n\|^2,
\end{aligned}$$

where $\varepsilon > 0$ is a positive constant. Similar to the estimate of I_1, we can also obtain

$$I_2 \le c|W_A^n|_\infty^2|v^n|_2^2 + \varepsilon\|v^n\|^2, \tag{5.4.22}$$

$$I_3 \le c|v^n + W_A^n|_2|v^n|_2 \le c|v^n|_2^2 + |W_A^n|_\infty^2, \tag{5.4.23}$$

$$\begin{aligned}
I_4 &\le c|B(v^n + W_A^n)|_2|v^n|_2 \\
&\le c|v^n + W_A^n|_2|v^n|_2 \le c|v^n|_2^2 + |W_A^n|_\infty^2,
\end{aligned} \tag{5.4.24}$$

$$\begin{aligned}
I_5 &\le c|(B(v^n + W_A^n))_x|_2|v^n|_2 \\
&\le c|v^n + W_A^n|_2|v^n|_2 \le c|v^n|_2^2 + |W_A^n|_\infty^2.
\end{aligned} \tag{5.4.25}$$

Substituting the estimates of $I_i, 1 \le i \le 5$ into eq. (5.4.20), we get

$$\frac{1}{2}\frac{d}{dt}|v^n|_2^2 + \frac{1}{R_e}\|v^n\|^2 \le 2\varepsilon\|v^n\|_2^2 + c(1 + |W_A^n|_\infty^2)|v^n|_2^2 + c|W_A^n|_\infty^4 + 3|W_A^n|_\infty^2.$$

Choosing a sufficiently small ε, such that

$$\frac{d}{dt}|v^n|_2^2 + \frac{1}{R_e}\|v^n\|^2 \le c(1 + |W_A^n|_\infty^2)|v^n|_2^2 + c|W_A^n|_\infty^4 + 3|W_A^n|_\infty^2,$$

and applying the Gronwall inequality, we obatin

$$|v^n|_2^2 \le |u_0^n|_2^2 e^{\int_0^t c(1+|W_A^n|_\infty^2)ds} + \int_0^t (c|W_A^n|_\infty^4 + 3|W_A^n|_\infty^2)e^{\int_s^t c(1+|W_A^n|_\infty^2)d\tau}ds. \tag{5.4.26}$$

Letting $n \to \infty$, we get

$$\begin{aligned}
|v(t)|_2^2 &\le |u_0|_2^2 e^{\int_0^t c(1+|W_A^n|_\infty^2)ds} + \int_0^t (c|W_A^n|_\infty^4 + 3|W_A^n|_\infty^2)e^{\int_s^t c(1+|W_A^n|_\infty^2)d\tau}ds \\
&\le |u_0|_2^2 e^{c(v_\infty^2+1)t} + c(cv_\infty^4 + 3v_\infty^2)e^{c(v_\infty^2+1)t} \\
&\le (|u_0|_2^2 + cv_\infty^4 + cv_\infty^2)e^{c(v_\infty^2+1)t}.
\end{aligned} \tag{5.4.27}$$

This completes the proof. ∎

Theorem 5.4.1. *If for P-a.s., $\omega \in \Omega, u_0 \in L^2(D)$, then eq.* (5.4.7) *has a unique global solution $v(x, y, t)$ and for P-a.s. $\omega \in \Omega$ and $T > 0$, $v \in C((0, T]; H^\alpha(D))$ with* $(0 \le \alpha < 1/2)$.

5.4.2 Existence of random attractors

For any $\alpha > 0$, let

$$z(t) = W_A^\alpha(t) = \int_{-\infty}^{t} e^{-(t-s)(A+\alpha)} dW(s),$$

where $W_A^\alpha(t)$ is the weak solution of the following initial value problem:

$$\begin{cases} dz = -(A+\alpha)zdt + dW(t), \\ z(0) = \int_{-\infty}^{0} e^{s(A+\alpha)} dW(s). \end{cases}$$

Obviously, if u is the weak solution of the following initial boundary value problem,

$$(P_1) \quad \begin{cases} du - \frac{1}{R_e}\Delta u dt = G(u)dt + dW(t), \\ u(s,\omega) = u_s \\ u|_{\partial D} = 0, \end{cases}$$

then $v(t) = u(t) - z(t)$ is the solution of the following problem:

$$(P_2) \quad \begin{cases} dv = -Av + G(v+z)dt + \alpha z, \\ v(s,\omega) = u_s - z(s) \\ v|_{\partial D} = 0. \end{cases} \tag{5.4.28}$$

5.4.2.1 Well-posedness and regularity of solution with problem (P_2)

In order to investigate the asymptotic behavior of solutions for the problem (P_1), we must study higher regularity about v.

Theorem 5.4.2.

(i) *For $T > s$, $u_s \in L^2(D)$ for P-a.e. $\omega \in \Omega$, the problem (P_2) has a unique solution $v \in C(s, T; L^2(D)) \cap L^2(s, T; H_0^1(D))$ in the weak sense for P-a.s. $\omega \in \Omega$. For any $\beta < \beta_1$, the solution $u = v + z$ of the problem (P_1) satisfies $u \in C(s, T; L^2(D)) \cap L^2(s, T; D(A^{\min\{\frac{1}{4}+\beta, \frac{1}{2}\}}))$ a.s., here β_1 has been given in the definition of the stochastic process $W(t)$.*

(ii) *For some $\theta \in (0, 2\beta_1) \cap (0, \frac{1}{2}]$ and $u_s \in D(A^\theta)$ for P-a.s. $\omega \in \Omega$, then $v \in C(s, T; D(A^\theta)) \cap L^2(s, T; D(A^{\frac{1}{2}+\theta}))$, for P-a.s. $\omega \in \Omega$.*

Theorem 5.4.2 can be proved using the classic Faedo–Galerkin method (see [174]). As this method is standard, we only give the key estimates in the following.

– **Energy estimates of** v

In the sequel, $\omega \in \Omega$ is fixed. By choosing v as a test function in eq. (5.4.28), we get

$$\frac{1}{2}\frac{d|v|_2^2}{dt} + \frac{1}{R_e}\|v\|^2 = \int_D G(v+z)v + \alpha \int_D zv, \tag{5.4.29}$$

where

$$\begin{aligned}\int_D G(v+z)v = &- \int_D [J(B(v+z), v+z)v + \beta_0 (B(v+z))_x v] \\ &+ \int_D \left[\left(\frac{F}{R_e} - \frac{r}{2}\right)(v+z)v - F\left(\frac{F}{R_e} - \frac{r}{2}\right) B(v+z)v\right].\end{aligned} \tag{5.4.30}$$

Using integration by parts, Hölder's inequality and Young's inequality, we get

$$\begin{aligned}&- \int_D [J(B(v+z), v+z)v \\ &\quad = \int_D J(B(v+z), v+z)z \\ &\quad = \int_D J(B(v+z), v)z \\ &\quad \le |(B(v+z))_x|_4 |v_y|_2 |z|_4 + |(B(v+z))_y|_4 |v_x|_2 |z|_4 \\ &\quad \le c|v+z|_2||v_y|_2|z|_4 + c|v+z|_2|v_x|_2|z|_4 \\ &\quad \le c|v+z|_2^2||z|_4^2 + \varepsilon\|v\|^2 \\ &\quad \le c|v|_2^2|z|_4^2 + |z|_2^2|z|_4^2 + \varepsilon\|v\|^2 \\ &\quad \le c|v|_2^2|z|_4^2 + c|z|_4^4 + 2\varepsilon\|v\|^2\end{aligned} \tag{5.4.31}$$

and

$$|\beta_0 \int_D (B(z))_x v| \le c|(B(z))_x|_2 |v|_2 \le c|z|_2|v|_2 \le c|z|_2^2 + \varepsilon|v|_2^2. \tag{5.4.32}$$

Similarly, we also obtain

$$|\left(\frac{F}{R_e} - \frac{r}{2}\right)\int_D zv| \le c|z|_2^2 + \varepsilon|v|_2^2, \tag{5.4.33}$$

$$|-F\left(\frac{F}{R_e} - \frac{r}{2}\right)\int_D vB(z)| \le c|z|_2^2 + \varepsilon|v|_2^2. \tag{5.4.34}$$

By the definition of the operator B and the hypothesis of the parameter β_0, F, R_e, r, we get

$$\beta_0 \int_D (B(v))_x v + \left(\frac{F}{R_e} - \frac{r}{2}\right)\int_D v^2 - F\left(\frac{F}{R_e} - \frac{r}{2}\right)\int_D vB(v) \le 0. \tag{5.4.35}$$

Using eqs (5.4.29)–(5.4.35), and choosing a sufficiently small $\varepsilon > 0$, we obtain

$$\frac{d|v|_2^2}{dt} + \frac{1}{R_e}\|v\|^2 \le c|z|_4^4|v|_2^2 + c|z|_2^2 + c|z|_4^4. \tag{5.4.36}$$

By Poincaré inequality $\frac{1}{R_e}\|v\|^2 \ge \lambda_1|v|_2^2$,

$$\frac{d|v|_2^2}{dt} \le \left(c|z|_4^4 - \frac{\lambda_1}{R_e}\right)|v|_2^2 + c|z|_2^2 + c|z|_4^4.$$

Thus for $t \in [s, T]$, an application of Gronwall's inequality leads to

$$\begin{aligned}|v(t)|_2^2 \le\ & e^{\int_s^t(-\lambda_1 + c|z(\tau)|_4^4)d\tau}|v(s)|_2^2 \\ & + \int_s^t e^{\int_\sigma^t(-\lambda_1 + c|z(\tau)|_4^4)d\tau}(c|z|_2^2 + c|z|_4^4)d\sigma.\end{aligned} \tag{5.4.37}$$

Finally, we integrate eq. (5.4.36) about t on $[t_1, t_2](\subset [s, T])$ and obtain

$$\frac{1}{R_e}\int_{t_1}^{t_2}\|v(\tau)\|^2 d\tau \le |v(t_1)|_2^2 + \int_{t_1}^{t_2}(c|z|_4^4|v|_2^2 + c|z|_2^2 + c|z|_4^4)d\tau. \tag{5.4.38}$$

In order to obtain the energy estimates of $A^\theta v$, we first give the following lemma (see Ref. [95]).

Lemma 5.4.3. *For any two real-valued functions $f, g \in H^{\frac{1}{2}+\theta}(D)$, $0 < \theta < 1/2$,*

$$\|fg\|_{H^{2\theta}} \le \|f\|_{H^{\frac{1}{2}+\theta}}\|g\|_{H^{\frac{1}{2}+\theta}}.$$

– **Energy estimates of $A^\theta v$**

Assume $v \in C(s, T; D(A^\theta)) \cap L^2(s, T; D(A^{\frac{1}{2}+\theta}))$ is a solution of the problem (P_2). Since the proof of the case $\theta = 1/2$ is classic (see [245]), we consider only the case $0 < \theta < 1/2$.

First of all, we have

$$\frac{1}{2}\frac{d|A^\theta v|_2^2}{dt} + \frac{1}{R_e}|A^{\frac{1}{2}+\theta}v|^2 = \int_D A^\theta G(v+z)A^\theta v + \alpha\int_D A^\theta z A^\theta v. \tag{5.4.39}$$

Using the definition of the operator A^θ, the interpolation inequality and Lemma 5.4.3, we get

$$
\begin{aligned}
-\int_D [A^\theta J(B(v+z), v+z)A^\theta v
&= -\int_D J(B(v+z), v+z)A^{2\theta} z \\
&= -\int_D \left(\frac{\partial B(v+z)}{\partial x}\frac{\partial (v+z)}{\partial y} - \frac{\partial B(v+z)}{\partial y}\frac{\partial (v+z)}{\partial x}\right) A^{2\theta} z \\
&= \int_D \left(\frac{\partial B(v+z)}{\partial x}\frac{\partial A^{2\theta} v}{\partial y} - \frac{\partial B(v+z)}{\partial y}\frac{\partial A^{2\theta} v}{\partial x}\right)(v+z) \\
&\le c\|\frac{\partial A^{2\theta} v}{\partial y}\|_{H^{-2\theta}}\|\frac{\partial B(v+z)}{\partial x}(v+z)\|_{H^{2\theta}} \\
&\quad + c\|\frac{\partial A^{2\theta} v}{\partial x}\|_{H^{-2\theta}}\|\frac{\partial B(v+z)}{\partial y}(v+z)\|_{H^{2\theta}} \\
&\le c\|A^{2\theta} v\|_{H^{1-2\theta}}\left(\|\frac{\partial B(v+z)}{\partial x}(v+z)\|_{H^{2\theta}} + \|\frac{\partial B(v+z)}{\partial y}(v+z)\|_{H^{2\theta}}\right) \\
&\le c\|A^{2\theta} v\|_{H^{1-2\theta}}\left(\|\frac{\partial B(v+z)}{\partial x}\|_{H^{\frac{1}{2}+\theta}} + \|\frac{\partial B(v+z)}{\partial y}\|_{H^{\frac{1}{2}+\theta}}\right)\|v+z\|_{H^{\frac{1}{2}+\theta}} \\
&\le c|A^{\frac{1}{2}+\theta} v|_2 |A^{\frac{1}{2}\theta}(v+z)|_2 |A^{\frac{1}{4}+\frac{\theta}{2}}(v+z)|_2 \\
&\le \varepsilon|A^{\frac{1}{2}+\theta} v|_2^2 + c|A^{\frac{1}{2}\theta}(v+z)|_2^2 |A^{\frac{1}{4}+\frac{\theta}{2}}(v+z)|_2^2 \\
&\le \varepsilon|A^{\frac{1}{2}+\theta} v|_2^2 + c(|A^{\frac{1}{2}\theta} v|_2^2 + |A^{\frac{1}{2}\theta} z|_2^2)(|A^{\frac{1}{4}+\frac{\theta}{2}} v|_2^2 + |A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^2) \\
&\le \varepsilon|A^{\frac{1}{2}+\theta} v|_2^2 + c(|A^{\theta} v|_2|v|_2 + |A^{\frac{1}{2}\theta} z|_2^2)(|A^{\frac{1}{4}+\frac{\theta}{2}} v|_2^2 + |A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^2) \\
&\le \varepsilon|A^{\frac{1}{2}+\theta} v|_2^2 + c|A^{\theta} v|_2|v|_2|A^{\frac{1}{4}+\frac{\theta}{2}} v|_2^2 \\
&\quad + c|A^{\theta} v|_2|v|_2|A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^2 + |A^{\frac{1}{2}\theta} z|_2^2|A^{\frac{1}{4}+\frac{\theta}{2}} v|_2^2 + |A^{\frac{1}{2}\theta} z|_2^2|A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^2 \\
&\le \varepsilon|A^{\frac{1}{2}+\theta} v|_2^2 + c|A^{\theta} v|_2^2|v|_2^2 + c|A^{\frac{1}{4}+\frac{\theta}{2}} v|_2^4 + |A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^4 + c|A^{\frac{\theta}{2}} z|_2^4 \\
&\le \varepsilon|A^{\frac{1}{2}+\theta} v|_2^2 + c|A^{\theta} v|_2^2|v|_2^2 + c|A^{\frac{1}{4}+\frac{\theta}{2}} v|_2^4 + c|A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^4.
\end{aligned}
$$

By interpolation inequality and Young's inequality, we have

$$
\begin{aligned}
|A^{\frac{1}{4}+\frac{\theta}{2}} v|_2^4 &\le c|A^{\frac{1}{4}} v|_2^2 |A^{\frac{1}{4}+\theta} v|_2^2 \\
&\le c|v|_2|A^{\frac{1}{2}} v|_2|A^{\theta} v|_2|A^{\frac{1}{2}+\theta} v|_2 \\
&\le \varepsilon|A^{\frac{1}{2}+\theta} v|_2^2 + c|v|_2^2|A^{\frac{1}{2}} v|_2^2|A^{\theta} v|_2^2.
\end{aligned}
\tag{5.4.40}
$$

The combination of these two inequalities yields

$$
\begin{aligned}
-\int_D A^\theta J(B(v+z), v+z)A^\theta v \le 2\varepsilon|A^{\frac{1}{2}+\theta} v|_2^2 + c|A^{\theta} v|_2^2|v|_2^2 \\
+ c|v|_2^2|A^{\frac{1}{2}} v|_2^2|A^{\theta} v|_2^2 + c|A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^4.
\end{aligned}
\tag{5.4.41}
$$

By Hölder's inequality and Cauchy–Schwarz's inequality, we obtain

$$\begin{aligned}\beta_0 \int_D A^\theta (B(v+z))_x A^\theta v &\le c|A^\theta (B(v+z))_x|_2 |A^\theta v|_2 \\ &\le c|A^\theta (B(v+z))_x|_2 + 2 + c|A^\theta v|_2^2 \\ &\le c|v+z|_2^2 + c|A^\theta v|_2^2 \\ &\le c|A^\theta v|_2^2 + c|z|_2^2.\end{aligned} \tag{5.4.42}$$

Similarly,

$$\begin{aligned}\left(\frac{F}{R_e} - \frac{r}{2}\right) \int_D A^\theta (v+z) A^\theta v - F\left(\frac{F}{R_e} - \frac{r}{2}\right) \int_D A^\theta B(v+z) A^\theta v \\ \le \varepsilon |A^{\frac{1}{2}+\theta} v|_2^2 + c|A^\theta v|_2^2 + c|z|_2^2.\end{aligned} \tag{5.4.43}$$

By eqs (5.4.39), (5.4.41)–(5.4.43), we get

$$\begin{aligned}\frac{1}{2}\frac{d|A^\theta v|_2^2}{dt} + \frac{1}{R_e}|A^{\frac{1}{2}+\theta} v|^2 &\le 3\varepsilon |A^{\frac{1}{2}+\theta} v|_2^2 + c|v|_2^2 |A^{\frac{1}{2}} v|_2^2 |A^\theta v|_2^2 \\ &\quad + c|A^\theta v|_2^2 |v|_2^2 + c|A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^4 + c|A^\theta v|_2^2 + c|z|_2^2.\end{aligned}$$

By choosing a sufficiently small $\varepsilon > 0$, we obtain

$$\begin{aligned}\frac{d|A^\theta v|_2^2}{dt} + \frac{1}{R_e}|A^{\frac{1}{2}+\theta} v|^2 &\le c|v|_2^2 |A^{\frac{1}{2}} v|_2^2 |A^\theta v|_2^2 + c|A^\theta v|_2^2 |v|_2^2 \\ &\quad + c|A^{\frac{1}{4}+\frac{\theta}{2}} v|_2^4 + c|A^\theta v|_2^2 + c|z|_2^2.\end{aligned} \tag{5.4.44}$$

By Gronwall's inequality, for any $t \in [s, T]$, we have

$$\begin{aligned}|A^\theta v(t)|_2^2 &\le e^{\int_s^t (c|v|_2^2 |A^{1/2} v|_2^2 + c|v|_2^2 + c) d\tau} |A^\theta v(s)|_2^2 \\ &\quad + \int_s^t e^{\int_\sigma^t (c|v|_2^2 |A^{1/2} v|_2^2 + c|v|_2^2 + c) d\tau} (c|z|_2^2 + c|A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^2) d\sigma.\end{aligned} \tag{5.4.45}$$

Integrating eq. (5.4.44) about t on $[t_1, t_2] \subset [s, T]$, we have

$$\begin{aligned}\frac{1}{R_e}\int_{t_1}^{t_2} |A^{\frac{1}{2}+\theta} v|^2 d\tau &\le |A^\theta v(t_1)|_2^2 + \int_{t_1}^{t_2} c(|v|_2^2 |A^{\frac{1}{2}} v|_2^2 |A^\theta v|_2^2 \\ &\quad + |A^\theta v|_2^2 |v|_2^2 + |A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^4 + |A^\theta v|_2^2 + |z|_2^2).\end{aligned}$$

5.4.2.2 Dissipative properties in $L^2(D)$

Next, we study the dissipative properties of the dynamical systems associated with eq. (5.4.1) combining with the boundary conditions (5.4.2) and obtain the existence of random attractor. The basic knowledge of random attractor can be seen in Chapter 4.

Theorem 5.4.3. *The 2D QG equation with a general white noise*

$$du - \frac{1}{R_e}\Delta u dt = G(u)dt + dW(t), \quad u|_{\partial D} = 0,$$

has a random attractor in $L^2(D)$.

Proof. By the ergodicity of the process z and Sobolev's embedding $D(A^{\frac{1}{4}+\beta}) \hookrightarrow L^4$ (see Ref. [76]), we get

$$\lim_{s_0\to-\infty} \frac{1}{-s_0} \int_{s_0}^{1} |z(\tau)|_4^4 d\tau = E(|z(0)|_4^4).$$

The results from Ref. [95] show that $E(|z(0)|_4^4)$ is arbitrary small provided α is sufficiently large. Thus

$$\lim_{s\to-\infty} \frac{1}{-s} \int_{s_0}^{1} (-\lambda_1 + c|z(\tau)|_4^4 d\tau = -\lambda_1 + cE(|z(0)|_4^4) \le -\frac{\lambda_1}{2}.$$

This implies the existence of $S_0(\omega)$ such that $s < S_0(\omega)$,

$$\int_s^0 (-\lambda_1 + c|z|_4^4)d\tau \le -\frac{\lambda_1}{4}(-s). \tag{5.4.46}$$

As $|z(s)|_2^2$ and $|z(s)|_4^4$ have at most polynomial growth as $s \to -\infty$ (see Ref. [95]), eqs (5.4.37) and (5.4.46) imply that there exists a random variable $r_1(\omega)$, almost surely, such that

$$|v(t)|_2^2 \le r_1(\omega), \quad \text{for any } t \in [-1, 0], \quad a.s., \tag{5.4.47}$$

where u_s is in a bounded set in $L^2(D)$.

Let $t_1 = -1$, $t_2 = 0$. By eqs (5.4.38) and (5.4.47), there exists an almost surely finite random variable $r_2(\omega)$ such that

$$\int_{-1}^{0} \|v(\tau)\|^2 d\tau \le r_2(\omega) \text{ a.s.} \tag{5.4.48}$$

By eq. (5.4.45), for $t = 0$, $s \in [-1, 0]$, we have

$$\begin{aligned} |A^\theta v(0)|_2^2 &\le e^{\int_s^0 (c|v|_2^2 |A^{1/2}v|_2^2 + c)d\tau} |A^\theta v(s)|_2^2 \\ &\quad + \int_s^0 e^{\int_\sigma^0 (c|v|_2^2 |A^{1/2}v|_2^2 + c)d\tau} (c|z|_2^2 + c|A^{\frac{1}{4}+\frac{\theta}{2}} z|_2^2) d\sigma. \end{aligned} \tag{5.4.49}$$

Integrating eq. (5.4.49) about s on $[-1,0]$, we get

$$|A^\theta v(0)|_2^2 \le e^{\int_{-1}^0 (c|v|_2^2|A^{1/2}v|_2^2+c)d\tau} \int_{-1}^0 |A^\theta v(s)|_2^2 ds \\ + \int_{-1}^0 e^{\int_\sigma^0 (c|v|_2^2|A^{1/2}v|_2^2+c)d\tau}(c|z|_2^2 + c|A^{\frac{1}{4}+\frac{\theta}{2}}z|_2^2)d\sigma. \tag{5.4.50}$$

Combining eq. (5.4.47) with eq. (5.4.48) and using eq. (5.4.50), we obtain an almost surely finite random variable $r_3(\omega)$ such that

$$|A^\theta v(0)|_2^2 \le r_3(\omega), \quad \text{a.s.}$$

Set $S(t,s,\omega)u_s = u(t,\omega)$. Without loss of generality, let $\Omega = \{\omega : \omega \in C(R, D(A^{-\frac{1}{4}+\beta_1})), \omega(0) = 0\}$, P be a Wiener measure, $W(t,\omega) = \omega(t), t \in R, \omega \in \Omega$, then we can define a measure-preserving transformation $\{\theta_t\}_{t\in R}$ on Ω, such that $\{\theta_s\omega(t)\} = \omega(t+s) - \omega(s)$ for $t,s \in R$ and $\omega \in \Omega$. Let $K(\omega)$ be a ball of radius $r_3(\omega) + |A^\theta z(0,\omega)|_2^2$ in $D(A^\theta)$. $K(\omega)$ is a compactly attracting set at $t = 0$ (by compact embedding $H^{2\theta}(D) \hookrightarrow L^2(D)$). Thus by an application of general theorem in Chapter 4, we get the existence of random attractors. ∎

5.5 Stochastic primitive equations

The mathematical study of the primitive equations originated from a series of articles by Lions, Temam and Wang in the early 1990s [175–177]. They defined the notions of weak and strong solutions and also proved the existence of weak solutions. Existence of strong solutions (local in time) and their uniqueness were obtained in Refs [115, 246]. Hu et al. [131] studied the local existence of strong solutions to the primitive equations under the small depth hypothesis. Cao and Titi [44] developed a delicate approach to prove that the L^6-norm of the fluctuation $\widetilde{v}$ of horizontal velocity is bounded, and obtained the global well-posedness for the 3D viscous primitive equations. Another different proof of this result was given by Kobelkov [155, 156]. The existence of the attractor was obtained in Ref. [139]. In Ref. [161], existence and uniqueness for different, physically relevant boundary conditions are established with a third method (different from both [44, 155]), which deals with the pressure terms directly in the equations. For a general reference on the current research of the (deterministic) mathematical theory for the primitive equations, we can refer to Ref. [207]. Moreover, the deterministic 2D primitive equations were studied also in Refs [205, 206].

The addition of white-noise-driven terms to the basic governing equations for a physical system is natural for both practical and theoretical applications. Stochastic solutions of the 2D primitive equations of the ocean and atmosphere with an additive noise have been studied in Ref. [90]. In Refs [103, 119], weak and strong random attractors have been obtained for 3D stochastic primitive equations with additive noise, respectively. The existence and uniqueness of solutions for 2D stochastic primitive equations with multiplicative noise have been discussed in Ref. [113], where they ignored the coupling with the temperature and salinity equations. There have also been other recent works on the stochastic 2D and 3D primitive equations [78, 112], in both works a coupling with temperature and salinity equations as well as physically relevant boundary conditions are considered.

5.5.1 Stochastic 2D primitive equations with Lévy noise

First, we introduce some definitions and basic properties of Wiener processes and Lévy processes. For further details, one can see Ref. [75] or [204], for example.

In this section, $W(t)$ are independent Wiener processes defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, taking values in Hilbert space H, with linear symmetric positive covariant operators Q. We assume that Q is trace class (and hence compact [75]), i.e. $\mathrm{tr}(Q) < \infty$. As in Ref. [84], let $H_0 = Q^{\frac{1}{2}}H$. Then H_0 is a Hilbert space with the scalar product

$$(u, \psi)_0 = (Q^{-\frac{1}{2}}u, Q^{-\frac{1}{2}}\psi)\ \forall u, \psi \in H_0$$

together with the induced norm $|\cdot|_0 = \sqrt{(\cdot,\cdot)_0}$. Let L_Q be the space of linear operators S such that $SQ^{\frac{1}{2}}$ is a Hilbert–Schmidt operator (and thus a compact operator [75]) from H to H. The norm in the space L_Q is defined by $|S|^2_{L_Q} = \mathrm{tr}(SQS^*)$, where S^* is the adjoint operator of S.

If $X = X(t)$ is a Lévy process, the jump of $X(t)$ is given by $\triangle X(t) = X(t) - X(t-)$. Let $Z \in \mathcal{B}(H)$, we define

$$N(t, Z) = N(t, Z, w) = \sum_{0<s\le t} \chi_Z(\triangle X(s)).$$

$N(t, Z)$ is called the Poisson random measure of $X(t)$. We denote $\widetilde{N}(dt, dZ) = N(dt, dZ) - dt\lambda(dz')$, which is called compensated Poisson random measure, where $\lambda(dz')$ is a σ-finite measure on Z.

Definition 5.5.1. *Let $I = [a, b]$ be an interval in $\mathbb{R}^+$. A mapping $g : I \to \mathbb{R}^d$ is said to be càdlàg if, for all $t \in [a, b]$, g has a left limit at t and g is right continuous at t. Let $\mathcal{D}([0, T], H)$ be the space of all càdlàg paths from $[0, T]$ into H.*

Definition 5.5.2. *Let E and F be separable Banach spaces, let $F_t := \mathcal{B}(\mathbb{R}^+ \times E) \otimes \mathcal{F}_t$ be the product σ-algebra generated by the semi-ring $\mathcal{B}(\mathbb{R}^+ \times E) \times \mathcal{F}_t$ of the product sets $Z \times F$, $Z \in \mathcal{B}(\mathbb{R}^+ \times E)$, $F \in \mathcal{F}_t$ (where $\mathcal{F}_t$ is the filtration of the process $X(t)$). Let $T > 0$ and*

$$\mathbb{H}(Z) = \Big\{ g : \mathbb{R}^+ \times Z \times \Omega \to F, \textit{such that } g \textit{ is } F_T/\mathcal{B}(F)$$
$$\textit{measurable and } g(t, z', w) \textit{ is } \mathcal{F}_t\textit{-adapted } \forall z' \in Z, \quad \forall t \in (0, T] \Big\}.$$

Let $p \geq 1$,

$$\mathbb{H}^p_\lambda([0,T] \times Z; F) = \Big\{ g \in \mathbb{H}(Z) : \int_0^T \int_Z \mathbb{E}\|g(t, z', w)\|^p_F \lambda(dz')dt < \infty \Big\}.$$

Definition 5.5.3. *An* $\mathbb{R}^d$*-valued stochastic process* $Y(t)$ *is a Lévy-type stochastic integral if it can be written in the following form, for each* $1 \leq i \leq d$, $1 \leq j \leq m$, *we have* $|G^i|^{1/2}$, $F^i_j \in L^2[0,T]$, $H^i \in \mathbb{H}^2_\lambda([0,T] \times Z; E)$ *and* $K(t)$ *is predictable:*

$$\begin{aligned} dY^i(t) &= G^i(t)dt + F^i_j(t)dB^j(t) + \int_{|z'|<1} H^i(t,z')\widetilde{N}(dt, dz') \\ &\quad + \int_{|z'|\geq 1} K^i(t, z')N(dt, dz'). \end{aligned} \tag{5.5.1}$$

Let

$$dY_c(t) = G^i(t)dt + F^i_j(t)dB^j(t),$$

and the discontinuous part of Y,

$$dY_d(t) = \int_{|z'|<1} H^i(t,z')\widetilde{N}(dt, dz') + \int_{|z'|\geq 1} K^i(t,z')N(dt, dz').$$

Then we have Itô's theorem for Lévy-type stochastic integral of the form (5.5.1).

Lemma 5.5.1 (Itô's theorem [7]). *If* $Y(t)$ *is a Lévy-type stochastic integral of the form* (5.5.1), *then for each* $f \in C^2(\mathbb{R}^d)$, *we have*

$$\begin{aligned} & f(Y(t)) - f(Y(0)) \\ &\quad = \int_0^t \partial_i f(Y(s-))dY^i_c(s) + \frac{1}{2}\int_0^t \partial_i\partial_j f(Y(s-))d[Y^i_c, Y^j_c](s) \\ &\quad + \int_0^t \int_{|z'|\geq 1} [f(Y(s-) + K(s,z')) - f(Y(s-))]N(ds, dz') \\ &\quad + \int_0^t \int_{|z'|\leq 1} [f(Y(s-) + H(s,z')) - f(Y(s-))]\widetilde{N}(ds, dz') \\ &\quad + \int_0^t \int_{|z'|\leq 1} [f(Y(s-) + H(s,z')) - f(Y(s-)) \\ &\quad - H^i(s,z')\partial_i f(Y(s-))]\lambda(dz')ds. \end{aligned} \tag{5.5.2}$$

Lemma 5.5.2 (BDG inequality [204]). *For every $p \geq 1$, there is constant C_p such that for any real-valued square-integrable càdlàg martingale M with $M(0) = 0$, and for any $T > 0$,*

$$C_p^{-1}\boldsymbol{E}[M,M]_T^{p/2} \leq \sup_{0\leq t\leq T}|M(t)|^p \leq C_p\boldsymbol{E}[M,M]_T^{p/2}. \tag{5.5.3}$$

The 2D primitive equations can be formally derived from the full 3D system under the assumption of invariance with respect to the second horizontal variable y as in Ref. [113], since primitive equations is a large-scale model, one may neglect the effect of small scale and intermediate scale in its modeling. Such fluctuations can be caused by internal instability processes, as well as by external forcing. As usual, the atmospheric forcing field should be regarded as random. As a result, we arrive at the following stochastic evolution system:

$$\begin{aligned} du &= [\nu_1\Delta u - u\partial_x u - w\partial_z u - \partial_x p + f]dt \\ &\quad + \sqrt{\varepsilon}\sigma(t,u)dW(t) + \varepsilon\int_Z g(u,z')\widetilde{N}(dt,dz'), \end{aligned} \tag{5.5.4}$$

$$\partial_z p = 0, \tag{5.5.5}$$

$$\partial_x u = -\partial_z w, \tag{5.5.6}$$

with velocity $u = u(t,x,z) \in \mathbb{R}$, pressure p, $(x,z) \in \mathcal{M} = [0,l] \times [-h,0]$ and $t > 0$. Here Δ is the Laplacian operator, without loss of generality in this section, we take ν_1 to be 1, noting eq. (5.5.5), p does not depend on variable z. In the above formulation, we have ignored the coupling with the temperature and salinity equations in order to focus main attention on the difficulties from nonlinear terms in eq. (5.5.4) (see Ref. [113]).

We partition the boundary into the top $\Gamma_u = \{z = 0\}$, the bottom $\Gamma_b = \{z = -h\}$ and the sides $\Gamma_s = \{x = 0\} \cup \{x = l\}$. In this section, we consider the following boundary conditions:

$$\begin{aligned} &\text{on } \Gamma_u: \ \partial_z u = 0, \ w = 0, \\ &\text{on } \Gamma_b: \ \partial_z u = 0, \ w = 0, \\ &\text{on } \Gamma_s: \ \ u = 0. \end{aligned}$$

Due to eq. (5.5.6), we have that

$$w(x,z,t) = -\int_{-h}^{z}\partial_x u(x,\xi,t)d\xi. \tag{5.5.7}$$

We define the function spaces H and V as follows:

$$H = \left\{v \in L^2(\mathcal{M}) \mid \int_{-h}^{0} v dz = 0 \ \right\}, \tag{5.5.8}$$

$$V = \left\{v \in H^1(\mathcal{M}) \mid \int_{-h}^{0} v dz = 0, \ \ v|_{\Gamma_s} = 0\right\}. \tag{5.5.9}$$

These spaces are endowed with the L^2 and H^1 norms, which we, respectively, denote by $|\cdot|$ and $\|\cdot\|$. The inner products and norms on H, V are given by

$$(v, v_1) = \int_{\mathcal{M}} v v_1 dxdz, \quad ((v, v_1)) = \int_{\mathcal{M}} \nabla v \nabla v_1 dxdz,$$

and

$$|v| = (v, v)^{\frac{1}{2}}, \quad \|v\| = ((v, v))^{\frac{1}{2}},$$

where $v_1, v \in V$. Let V' be the dual space of V. We have the dense and continuous embedding $V \hookrightarrow H = H' \hookrightarrow V'$ and denote by $\langle u, \psi \rangle$ the duality between $u \in V$ and $u \in V'$.

Consider an unbounded linear operator $A : D(A) \to H$ with $D(A) = V \cap H^2(\mathcal{M})$ and define

$$\langle Au, v \rangle = ((u, v)) \; \forall u, v \in D(A).$$

The Laplace operator A is self-adjoint, positive, with compact self-adjoint inverses, which maps V to V'. Next we address the nonlinear term. In accordance with eq. (5.5.7) we take

$$\mathcal{W}(v) := -\int_{-h}^{z} \partial_x v(x, \widetilde{z}) d\widetilde{z} \tag{5.5.10}$$

and let

$$B(u, v) := u \partial_x v + \mathcal{W}(u) \partial_z v, \tag{5.5.11}$$

where $u, v \in V$.

Define the bilinear operator $B(u, v) : V \times V \to V'$ according to

$$\langle B(u, v), w \rangle = b(u, v, w),$$

where

$$b(u, v, w) = \int_{\mathcal{M}} (u \partial_x v w + \mathcal{W}(u) \partial_z v w) d\mathcal{M}.$$

In the sequel, when no confusion arises, we denote by C a constant that may change from one line to the next one.

Lemma 5.5.3 (Estimates for *b* and *B* (see [104, 113])). *The trilinear forms b and B have the following properties. There exists a constant $C > 0$ such that*

$$|b(u,v,w)| \le C\Big(|u|^{\frac{1}{2}}\|u\|^{\frac{1}{2}}\|v\||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}} + |\partial_x u||\partial_z v||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}}\Big),$$
$$u,v,w \in V, \tag{5.5.12}$$
$$b(u,v,v) = 0, \qquad u,v,w \in V, \tag{5.5.13}$$
$$\langle B(u,u), \partial_{zz}u\rangle = 0, \qquad u \in V. \tag{5.5.14}$$

Note that the above formulation is equivalent to projecting eqs (5.5.4)–(5.5.6) from $L^2(\mathcal{M})$ into the space $H(\mathcal{M})$ and thus the pressure term $p(x,t)$ is absent. With these notations, the above primitive equations can be rewritten as

$$\begin{aligned} du + [Au + B(u,u)]dt &= fdt + \sqrt{\varepsilon}\sigma(t,u)dW(t) \\ &\quad + \varepsilon\int_Z g(u,z')\widetilde{N}(dt,dz'). \end{aligned} \tag{5.5.15}$$

In this section, we assume σ and g satisfy the following hypotheses of joint continuity, Lipschitz condition and linear growth.

Assumption A. There exist positive constants K and L such that

(A.1) $\sigma \in C\big([0,T]\times V; L_Q(H_0,H)\big)$,

(A.2) $|\sigma(t,u)|^2_{L_Q} + \int_Z |g(u,z')|^2_H \lambda(dz') \le K(1+|u|^2)$, $g \in \mathbb{H}^2_\lambda([0,T]\times Z; H)$,

(A.3) $\int_Z |g(u,z')|^4_H \lambda(dz') \le K(1+|u|^4)$, $g \in \mathbb{H}^4_\lambda([0,T]\times Z; H)$,

(A.4) $|\sigma(t,u)-\sigma(t,\psi)|^2_{L_Q} + \int_Z |g(u,z')-g(\psi,z')|^2_H \lambda(dz') \le L\|u-\psi\|^2$, $\forall u,\psi \in V$.

To obtain well-posedness of solution, we have to give the additional assumptions on the σ, g:

Assumption B. There exists a positive constant K such that

(B.1) $|\partial_z\sigma(t,u)|^2_{L_Q} + \int_Z |\partial_z g(u,z')|^2_H \lambda(dz')$
$\le K(1+|\partial_z u|^2))$, $\partial_z g \in \mathbb{H}^2_\lambda([0,T]\times Z; H)$,

(B.2) $\int_Z |\partial_z g(u,z')|^4_H \lambda(dz') \le K(1+|\partial_z u|^4)$, $\partial_z g \in \mathbb{H}^4_\lambda([0,T]\times Z; H)$.

Example 5.5.1. *We assume $W(t) = \sum_{k=1}^\infty \lambda_k^{1/2}\beta_k(t)e_k \in H$, where $\beta_k(t)$ is a collection of independent standard Brownian motions, $\{e_k\}$ is an orthonormal basis in H consisting of eigen-elements of Q, with $Qe_k = \lambda_k e_k$.*

– *(Additive Noise) As in Ref. [113], we suppose the noise term does not depend on the solution u and can be written in the expansion:*

$$\begin{aligned} \sigma(t,u)dW(t) &= \sum_{k=1}^\infty \langle h(t,\omega), e_k\rangle \lambda_k^{-1/2}\lambda_k^{1/2} e_k d\beta_k(t) \\ &= \sum_{k=1}^\infty h_k(t,\omega)d\beta_k(t), \quad g(u,z') = g(z'), \end{aligned}$$

where $h_k(t,\omega) = \langle h(t,\omega), e_k\rangle e_k$. *Taking*

$$K = 2\max\Big\{\sup_{t,\omega}\sum_{k=1}^{\infty}|h_k(t,\omega)|_H^2, \sup_{t,\omega}\sum_{k=1}^{\infty}|\partial_z h_k(t,\omega)\big)|_H^2, \int_Z |g(z')|_H^2\lambda(dz'),$$
$$\int_Z |\partial_z g(z')|_H^2\lambda(dz'), \int_Z |g(z')|_H^4\lambda(dz'), \int_Z |\partial_z g(z')|_H^4\lambda(dz')\Big\},$$

and $L > 0$, *Assumptions A and B about Lipschitz condition and linear growth are sastified.*

– *(Independently forced models) Given a uniformly bounded sequence* $a_k(t,\omega) \in L^\infty([0,T]\times\Omega)$, *we assume that*

$$\sigma(t,u)dW(t) = \sum_{k=1}^{\infty} a_k(t,\omega)\langle u, e_k\rangle\lambda_k^{-1/2}\lambda_k^{1/2}e_k d\beta_k(t) = \sum_{k=1}^{\infty} a_k(t,\omega)u_k d\beta_k(t),$$
$$g(u,z') = u.$$

Taking

$$K = L = 2\max\Big\{\sup_{t,\omega,k}|a_k(t,\omega)|, \int_Z \lambda(dz') \equiv \boldsymbol{D}(Z)\Big\},$$

Assumptions A and B about Lipschitz condition and linear growth are satisfied.

5.5.1.1 Well-posedness

Let $X := \mathcal{D}\big([0,T];H\big)\cap L^2\big((0,T);V\big)$ denote the Banach space with the norm defined by

$$\|u\|_X = \Big\{\sup_{0\le s\le T}|u(s)|^2 + \int_0^T \|u(s)\|^2 ds\Big\}^{\frac{1}{2}}. \tag{5.5.16}$$

Recall that an $\mathcal{F}_t$-predictable stochastic process $u(t,\omega)$ is called the weak solution for stochastic primitive problem (5.5.15) on $[0,T]$ if u is in

$$L^2(\Omega, \mathcal{D}([0,T];H)\cap L^2((0,T);V)),$$

and satisfies

$$\langle u(t),\psi\rangle - \langle \xi,\psi\rangle + \int_0^t [\langle u(s), A\psi\rangle + \big\langle B(u(s)),\psi\big\rangle]ds = \int_0^t \langle f,\psi\rangle ds$$
$$+ \sqrt{\varepsilon}\int_0^t \big\langle \sigma(u(s))dW(s),\psi\big\rangle + \varepsilon\int_0^t\int_Z \langle g(u,z'),\psi\rangle \widetilde{N}(ds,dz'), \quad \text{a.s.,} \tag{5.5.17}$$

for all $\psi \in D(A)$ and all $t\in[0,T]$.

Remark 5.5.1. *We give a variational definition for solutions of the systems (5.5.15) in PDEs sense, note that weak refers to the spatial–temporal regularity of the solutions, and this solution is a strong one in the probabilistic meaning because of the stochastic basis given in advance. This is in contrast to the theory of weak solutions in probability sense considered for many nonlinear systems where the probability space is constructed as part of the solution. See Ref. [75].*

The main result of this section is the following theorem.

Theorem 5.5.1 (Well-posedness and a priori bounds). *There exists* $\varepsilon_0 := \varepsilon_0(K, L, T) > 0$*, such that the following existence and uniqueness are true for* $0 \le \varepsilon \le \varepsilon_0$*. Let the initial datum* ξ *satisfy* $\mathbf{E}\,|\xi|^4 < \infty$, $\mathbf{E}\,|\partial_z \xi|^4 < \infty$, and $f, \partial_z f \in L^4(\Omega; L^2(0,T;H))$, then there exists a unique weak solution u of the stochastic primitive problem (5.5.15) with initial condition $u(0) = \xi$. Furthermore, there exists a constant

$$C := C\Big(K, L, T, |f|_{L^4(\Omega;L^2(0,T;H))}, |\partial_z f|_{L^4(\Omega;L^2(0,T;H))}\Big)$$

such that for $\varepsilon \in [0, \varepsilon_0]$,

$$E\|u\|_X^2 \le C\big(1 + E|\xi|^2\big),\ E\Big(\sup_{0\le t\le T} |u(t)|^4 + \Big(\int_0^T \|u(t)\|^2\,dt\Big)^2\Big) \le C\,\big(1 + E|\xi|^4\big), \tag{5.5.18}$$

and satisfy the additional regularity

$$\partial_z u(t) \in L^4(\Omega, L^\infty(0,T;H) \cap L^2(0,T;V)). \tag{5.5.19}$$

Remark 5.5.2. *We should point out that we do not study the path regularity of the solution in this section. We study the adapted process* $u(t,x,\omega)$ *with regularity* $u \in L^2(\Omega, \mathcal{D}([0,T];H) \cap L^2((0,T);V))$*. The related paper [82] studied the path regularity.*

For $u \in V$, define

$$E(u) = -Au - B(u) + f. \tag{5.5.20}$$

We first obtain monotonicity property of E.

Lemma 5.5.4. *Assume that* $u, \psi \in V$*, we have*

$$\big\langle E(u) - E(\psi), u - \psi\big\rangle + \frac{1}{2}\|u-\psi\|^2 \le C|u-\psi|\,\|u-\psi\|\,\|\psi\| + C(1 + |\partial_z\psi|^4)|u-\psi|^2. \tag{5.5.21}$$

Proof. Set $U = u - \psi$, we deduce

$$\begin{aligned}\big\langle E(u) - E(\psi), U\big\rangle &= -\big\langle A(u) - A(\psi), U\big\rangle - \big\langle B(u) - B(\psi), U\big\rangle \\ &\equiv I_1 + I_2.\end{aligned}$$

Integrating by parts, Lemma 5.5.3, Hölder's inequality, Ladyzhenskaya's inequality for 2D domain and Young's inequality imply

$$\begin{aligned}
I_1 &= -\|u-\psi\|^2,\\
I_2 &= b(u-\psi, u-\psi, v)\\
&\leq |u-\psi|_{L^4}\|\psi\||u-\psi|_{L^4} + |\partial_x(u-\psi)||\partial_z\psi||u-\psi|^{\frac{1}{2}}\|u-\psi\|^{\frac{1}{2}}\\
&\leq C|u-\psi|\|u-\psi\|\|\psi\| + C\|u-\psi\|^{\frac{3}{2}}|u-\psi|^{\frac{1}{2}}|\partial_z\psi|\\
&\leq C|u-\psi|\|u-\psi\|\|\psi\| + \frac{1}{2}\|u-\psi\|^2 + C|\partial_z\psi|^4|u-\psi|^2.
\end{aligned}$$

Combining I_1, I_2, we end the proof. ■

Remark 5.5.3 (Outline of the proof for Theorem 5.5.1). *We first introduce the Galerkin systems associated with the original equation and establish some uniform a priori estimates (Propositions 5.5.1 and 5.5.2). These estimates together with the local monotonicity property (Lemma 5.5.4) and weak convergence methods play a fundamental role in proving the existence and uniqueness of the weak solution. It has appeared that the problem of global existence of solutions for the 2D stochastic primitive equations might be harder than 2D Navier–Stokes equations because the nonlinear term* $B(u,u)$ *is more complex. We deduce that the estimate of* $B(u,u)$ *satisfies eq.* (5.5.62). *To obtain eq.* (5.5.62), *we should establish the uniform bounds for* u_n, $\partial_z u_n$ *in* $L^p(\Omega; L^2(0,T;V)\cap L^\infty(0,T;H))$, $p=2,4$. *However, since we concern Lévy noise here, the estimates* u_n, $\partial_z u_n$ *in* $L^4(\Omega; L^2(0,T;V)\cap L^\infty(0,T;H))$ *are not immediately obtained by Itô's formula. Thus, we first estimate in* $L^2(\Omega; L^2(0,T;V)\cap L^\infty(0,T;H))$*(Proposition 5.5.1). Second, by more hypothesis (**A.3**), (**B.2**), after a series of estimates, we finally obtain the estimates* $u_n, \partial_z u_n$ *in*

$$L^4(\Omega; L^2(0,T;V)\cap L^\infty(0,T;H))$$

(Proposition 5.5.2).

We now introduce the Galerkin systems associated with the original equation and establish some uniform a priori estimates. For any $n \geq 1$, let $H_n = \operatorname{span}\{e_1, \cdots, e_n\} \subset \operatorname{Dom}(A)$ and $P_n : H \to H_n$ denote the orthogonal projection onto H_n. Suppose that the H-valued Wiener process W with covariance operator Q is such that

$$P_n Q^{\frac{1}{2}} = Q^{\frac{1}{2}} P_n, \quad n \geq 1,$$

which is true if $Qh = \sum_{n\geq 1}\lambda_n e_n$ with trace $\sum_{n\geq 1}\lambda_n < \infty$. Then for $H_0 = Q^{\frac{1}{2}}H$ and $(u,\psi)_0 = (Q^{-\frac{1}{2}}u, Q^{-\frac{1}{2}}\psi)$ with $u, \psi \in H_0$, we see that $P_n : H_0 \to H_0 \cap H_n$ is a contraction in both of the H and H_0 norms. Let $W_n = P_n W$, $\sigma_n = P_n\sigma$.

Consider the following stochastic ordinary differential equation on the n-dimensional space H_n defined by

$$d(u_n, \psi) = \left[\langle P_n E(u_n), \psi\rangle\right]dt + \sqrt{\varepsilon}\,\langle \sigma_n(u_n)dW_n, \psi\rangle \\ + \varepsilon \int_Z \langle g_n(u_n, z'), \psi\rangle \widetilde{N}(dt, dz'), \tag{5.5.22}$$

for $\psi \in H_n$ and $u_n(0) = P_n\xi$.

We note that formulation (5.5.22) allows one to treat u_n as a process in R^n. Thus one can apply the finite-dimensional Itô calculus to the Galerkin systems above.

Note that for $\psi \in V$, the map $u_n(t) = \sum_{k=1}^n d_n^k(t)e_k \in H_n \bigcap V \mapsto \langle -Au_n + f + \nabla p_s, \psi\rangle$ is globally Lipschitz, while using Lemma 5.5.3 the map B is locally Lipschitz. Furthermore, assumptions A and B imply that the map $u_n \to \langle \sigma_n(u_n)dW_n, \psi\rangle + \varepsilon \int_Z \langle g_n(u_n, z'), \psi\rangle \widetilde{N}(dt, dz')$ is globally Lipschitz. Hence, according to standard theory for ordinary differential equations, there exists a maximal càdlàg process $u_n \in H_n \bigcap V$ to (5.5.22), i.e., a stopping time $\tau_n^\varepsilon \le T$ such that eq. (5.5.22) holds for $t < \tau_n^\varepsilon$ and as $t \uparrow \tau_n^\varepsilon < T$, $\|\phi_n^\varepsilon(t)\|_{H_n \bigcap V} \to \infty$. Thus, the following definition for stopping time τ_N is reasonable.

For every $N > 0$, set

$$\tau_N = \inf\{t : |u_n(t)| \ge N\} \wedge \inf\{t : |\partial_z u_n(t)| \ge N\} \wedge T. \tag{5.5.23}$$

The following proposition provides the (global) existence and uniqueness of approximate solutions and also their uniform (a priori) estimates. This is the main preliminary step in the proof of Theorem 5.5.1.

Proposition 5.5.1. *There exists $\varepsilon_1 := \varepsilon_1(K, T)$ such that for $0 \le \varepsilon \le \varepsilon_1$, the following result holds. Let σ, g satisfy (**A.1**) and (**A.2**), $f \in L^2(\Omega; L^2(0, T; H))$ and $\xi \in L^2(\Omega, H)$. Then eq.* (5.5.22) *has a solution with a modification $u_n \in \mathcal{D}([0, T], H_n)$ and satisfies*

$$\sup_n \mathbf{E}\Big(\sup_{0\le t\le T} |u_n(t)|^2 + \int_0^T \|u_n(s)\|^2 ds\Big) \\ \le C(K, T, \mathbf{E}|f|^2_{L^2(\Omega;L^2(0,T;H))})\big(1 + \mathbf{E}|\xi|^2\big). \tag{5.5.24}$$

*Moreover, if σ, g satisfy (**A.3**), $f \in L^4(\Omega; L^2(0, T; H))$ and $\xi \in L^4(\Omega, H)$, we have*

$$\sup_n \mathbf{E}\Big(\sup_{0\le s\le T} |u_n|^4\Big) + \sup_n \mathbf{E}\Big(\int_0^T \|u_n(r)\|^2 dr\Big)^2 \\ \le C(K, T, \mathbf{E}|f|^4_{L^2(\Omega;L^2(0,T;H))})\big(1 + \mathbf{E}|\xi|^4\big). \tag{5.5.25}$$

Proof. Itô's formula yields that for $t \in [0, T]$ and τ_N defined by eq. (5.5.23),

$$|u_n(t \wedge \tau_N)|^2 = |P_n\xi|^2 + 2\sqrt{\varepsilon} \int_0^{t\wedge\tau_N} \left\langle \sigma_n(u_n(s))dW_n(s), u_n(s)\right\rangle \tag{5.5.26}$$

$$+ 2\int_0^{t\wedge\tau_N} \left\langle E(u_n(s)), u_n(s)\right\rangle ds + \varepsilon \int_0^{t\wedge\tau_N} \int_Z |g_n(u(s-), z')|^2 N(ds, dz')$$

$$+ \varepsilon \int_0^{t\wedge\tau_N} |\sigma_n(u_n(s))\, P_n|_{L_Q}^2\, ds$$

$$+ 2\varepsilon \int_0^{t\wedge\tau_N} \int_Z \langle u_n(s-), g_n(u(s-), z')\rangle \widetilde{N}(ds, dz'). \tag{5.5.27}$$

Using Lemma 5.5.3, this yields for $t \in [0, T]$,

$$|u_n(t \wedge \tau_N)|^2 + 2\int_0^{t\wedge\tau_N} \|u_n(r)\|^2\, dr \leq |P_n\xi|^2 + \sum_{1\leq j\leq 4} T_j(t), \tag{5.5.28}$$

where

$$\begin{aligned}
T_1(t) &= 2 \int_0^{t\wedge\tau_N} |\left\langle u_n, f\right\rangle| dr, \\
T_2(t) &= 2\sqrt{\varepsilon} \left| \int_0^{t\wedge\tau_N} \left\langle \sigma_n(u_n(r))\, dW_n(r), u_n(r)\right\rangle \right|, \\
T_3(t) &= \varepsilon \int_0^{t\wedge\tau_N} |\sigma_n(u_n(r))\, P_n|_{L_Q}^2\, dr \\
&\quad + \varepsilon \int_0^{t\wedge\tau_N} \int_Z |g_n(u(s-), z')|^2 N(ds, dz'), \\
T_4(t) &= 2\varepsilon \int_0^{t\wedge\tau_N} \int_Z \langle u_n(s-), g_n(u(s-), z')\rangle \widetilde{N}(ds, dz').
\end{aligned}$$

The Hölder's inequality and Young's inequality imply that

$$T_1(t) = 2\int_0^{t\wedge\tau_N} [\langle u_n(r), f\rangle] dr \leq \varsigma \int_0^{t\wedge\tau_N} \|u_n(r)\|^2 dr + \frac{1}{\varsigma}\Big(|f|^2_{L^2(0,t\wedge\tau_N;H)}\Big). \tag{5.5.29}$$

We deduce that

$$T_3(t) = \varepsilon \int_0^{t\wedge\tau_N} \int_Z |g_n(u(s-), z')|^2 \lambda(dz') ds + \varepsilon \int_0^{t\wedge\tau_N} \int_Z |g_n(u(s-), z')|^2 \widetilde{N}(ds, dz') + \varepsilon \int_0^{t\wedge\tau_N} |\sigma_n(u_n(r))\, P_n|_{L_Q}^2\, dr. \tag{5.5.30}$$

By Hölder's inequality, g_n is strong 2-integrable w.r.t $\widetilde{N}(dt,dz)$, then $|g_n|^2$ is strong 1-integrable w.r.t $\widetilde{N}(dt,dz)$. Using Theorem 4.12 in Ref. [225] on $|g_n|^2$ and the hypothesis (**A.2**), we obtain

$$\begin{aligned}
\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|T_3(s)|\Big) &\le \varepsilon\mathbf{E}\int_0^{t\wedge\tau_N}\int_Z |g_n(u(s-),z')|^2\lambda(dz')ds \\
&+ \varepsilon\mathbf{E}\int_0^{t\wedge\tau_N}\int_Z |g_n(u(s-),z')|^2\widetilde{N}(ds,dz') + \varepsilon\mathbf{E}\int_0^{t\wedge\tau_N}|\sigma_n(u_n(r))\,P_n|^2_{L_Q}\,dr \\
&\le 3\varepsilon\mathbf{E}\int_0^{t\wedge\tau_N}\int_Z |g_n(u(s-),z')|^2\lambda(dz')ds + \varepsilon\mathbf{E}\int_0^{t\wedge\tau_N}|\sigma_n(u_n(r))\,P_n|^2_{L_Q}\,dr \\
&\le \varepsilon KT + \varepsilon K\mathbf{E}\int_0^{t\wedge\tau_N}\|u_n(s)\|^2 dr.
\end{aligned} \tag{5.5.31}$$

By the BDG inequality, (**A.2**) and Schwarz's inequality, we get that for $t\in[0,T]$,

$$\begin{aligned}
\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|T_2(s)|\Big) &\le 2\sqrt{2\varepsilon}\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}|u_n(r)|^2\,|\sigma_n(u_n(r))\,P_n|^2_{L_Q}\,dr\Big\}^{\frac12} \\
&\le 2\sqrt{2\varepsilon K}\mathbf{E}\Big[(\sup_{0\le s\le t\wedge\tau_N}|u_n|^2)^{1/2}\Big(\int_0^{t\wedge\tau_N}(1+\|u_n(s)\|^2)ds\Big)^{1/2}\Big] \\
&\le \frac14\mathbf{E}\sup_{0\le s\le t\wedge\tau_N}|u_n|^2 + 8\varepsilon K\mathbf{E}\int_0^{t\wedge\tau_N}\|u_n(t)\|^2 dt + 8\varepsilon KT.
\end{aligned} \tag{5.5.32}$$

Now applying BDG inequality, condition (**A.2**) and Young's inequality to the term T_4, we get

$$\begin{aligned}
\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|T_4(s)|\Big) &\le 2\sqrt{2}\varepsilon\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}\int_Z|\langle u_n(s),g_n(u(s),z')\rangle|^2\lambda(dz')ds\Big\}^{\frac12} \\
&\le 2\sqrt{2}\varepsilon\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}\int_Z|u_n(s)|^2|g_n(u(s),z')|^2\lambda(dz')ds\Big\}^{\frac12} \\
&\le 2\sqrt{2K}\varepsilon\mathbf{E}\Big[(\sup_{0\le s\le t\wedge\tau_N}|u_n|^2)^{1/2}\Big(\int_0^{t\wedge\tau_N}(1+\|u_n(s)\|^2)ds\Big)^{1/2}\Big] \\
&\le \frac14\mathbf{E}\sup_{0\le s\le t\wedge\tau_N}|u_n|^2 + 8\varepsilon^2K\mathbf{E}\int_0^{t\wedge\tau_N}\|u_n(t)\|^2 dt + 8\varepsilon^2 KT.
\end{aligned} \tag{5.5.33}$$

Taking supremum up to time $t\wedge\tau_N$ before taking the expectation in eq. (5.5.28) and using eqs (5.5.29)–(5.5.33), we get (without loss of generality, we let $\varepsilon<1$).

$$\begin{aligned}
&\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|u_n|^2\Big) + 2(1-\varsigma-17\varepsilon K)\mathbf{E}\int_0^{t\wedge\tau_N}\|u_n(r)\|^2 dr \\
&\qquad\le 2\mathbf{E}|\xi|^2 + 2C\mathbf{E}|f|^2_{L^2(0,t\wedge\tau_N;H)} + 34\varepsilon KT.
\end{aligned} \tag{5.5.34}$$

Taking $T\wedge\tau_N\to T$ a.s. as $N\to\infty$, and $\varsigma=1/2$, we get for $0<\varepsilon\le\frac{1}{34K}$,

$$\mathbf{E}\Big(\sup_{0\le s\le T}|u_n|^2\Big)+\mathbf{E}\int_0^T\|u_n(r)\|^2dr\le C\Big(\mathbf{E}|\xi|^2,\mathbf{E}|f|^2_{L^2(0,T;H)},K,T\Big). \tag{5.5.35}$$

Now we estimate eq. (5.5.25). By eq. (5.5.28), we get

$$|u_n(t\wedge\tau_N)|^4+4\Big(\int_0^{t\wedge\tau_N}\|u_n(r)\|^2\,dr\Big)^2\le\ 2|P_n\xi|^4+8\sum_{1\le j\le 4}T_j^2(t). \tag{5.5.36}$$

By eq. (5.5.29), we deduce that

$$T_1^2(t)\le 2\varsigma^2\Big(\int_0^{t\wedge\tau_N}\|u_n(r)\|^2dr\Big)^2+\frac{2}{\varsigma^2}\Big(|f|^2_{L^2(0,t\wedge\tau_N;H)}\Big)^2. \tag{5.5.37}$$

By eq. (5.5.30), we get

$$\begin{aligned}\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|T_3^2(s)|\Big)\ &\le 3\varepsilon^2\mathbf{E}\Big(\int_0^{t\wedge\tau_N}\int_Z|g_n(u(s-),z')|^2\lambda(dz')ds\Big)^2\\&\quad+3\varepsilon^2\mathbf{E}\Big(\int_0^{t\wedge\tau_N}\int_Z|g_n(u(s-),z')|^2\widetilde{N}(ds,dz')\Big)^2\\&\quad+3\varepsilon^2\,\mathbf{E}\Big(\int_0^{t\wedge\tau_N}|\sigma_n(u_n(r))\,P_n|^2_{L_Q}\,dr\Big)^2\\&\le\sum_{i=1}^3U_i.\end{aligned} \tag{5.5.38}$$

Applying Hölder's inequality, we have

$$U_1\le 3\varepsilon^2CT^2\mathbf{E}\int_0^{t\wedge\tau_N}\int_Z|g_n(u(s-),z')|^4\lambda(dz')ds, \tag{5.5.39}$$

$$U_3\le 3\varepsilon^2CT\mathbf{E}\int_0^{t\wedge\tau_N}|\sigma_n(u_n(r))\,P_n|^4_{L_Q}\,dr. \tag{5.5.40}$$

By BDG inequality, we obtain

$$U_2\le 3C\varepsilon^2\mathbf{E}\int_0^{t\wedge\tau_N}\int_Z|g_n(u(s-),z')|^4\lambda(dz')ds, \tag{5.5.41}$$

Recalling the assumption (**A.3**), we obtain

$$\begin{aligned}\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|T_3^2(s)|\Big)&\le 3C(K)\varepsilon^2\mathbf{E}\int_0^{t\wedge\tau_N}(1+|u_n|^4)ds\\&\le 3C(K,T)\varepsilon^2+3C(K)\varepsilon^2\mathbf{E}\int_0^{t\wedge\tau_N}\sup_{0\le r\le s}|u_n(s)|^4ds.\end{aligned} \tag{5.5.42}$$

By the BDG inequality, (**A.2**) and Schwarz's inequality, we obtain

$$
\begin{aligned}
\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|T_2^2(s)|\Big) &\le C\sqrt{\varepsilon}\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}|u_n(r)|^2\,|\sigma_n(u_n(r))\,P_n|_{L_Q}^2\,dr\Big\}\\
&\le C\sqrt{\varepsilon}\mathbf{E}\Big[(\sup_{0\le s\le t\wedge\tau_N}|u_n|^2)\int_0^{t\wedge\tau_N}(1+\|u_n(s)\|^2)ds\Big]\\
&\le \frac{1}{40}\mathbf{E}\sup_{0\le s\le t\wedge\tau_N}|u_n|^4\\
&\quad + C(K)\varepsilon\mathbf{E}\Big(\int_0^{t\wedge\tau_N}\|u_n(t)\|^2dt\Big)^2 + C(K,T)\varepsilon.
\end{aligned}\tag{5.5.43}
$$

Now again applying BDG inequality, (**A.2**) and Young's inequality, we get

$$
\begin{aligned}
\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|T_5^2(s)|\Big) &\le C\varepsilon\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}\int_Z|\langle u_n(s),g_n(u(s),z')\rangle|^2\lambda(dz')ds\Big\}\\
&\le C\varepsilon\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}\int_Z|u_n(s)|^2|g_n(u(s),z')|^2\lambda(dz')ds\Big\}\\
&\le C(K)\varepsilon\mathbf{E}\Big[(\sup_{0\le s\le t\wedge\tau_N}|u_n|^2)\int_0^{t\wedge\tau_N}(1+\|u_n(s)\|^2)ds\Big]\\
&\le \frac{1}{40}\mathbf{E}\sup_{0\le s\le t\wedge\tau_N}|u_n|^4\\
&\quad + C(K)\varepsilon\mathbf{E}\Big(\int_0^{t\wedge\tau_N}\|u_n(t)\|^2dt\Big)^2 + C(K,T)\varepsilon.
\end{aligned}\tag{5.5.44}
$$

As we obtain eq. (5.5.35), let $N\to\infty$, by Gronwall's inequality and eqs (5.5.37)–(5.5.43) and (5.5.44), we get

$$
\begin{aligned}
&\sup_n\mathbf{E}\Big(\sup_{0\le s\le T}|u_n|^4\Big)+\sup_n\mathbf{E}\Big(\int_0^T\|u_n(r)\|^2dr\Big)^2\\
&\qquad\le C(K,T,\mathbf{E}|f|^4_{L^2(\Omega;L^2(0,T;H))})\big(1+\mathbf{E}|\xi|^4\big).
\end{aligned}\tag{5.5.45}
$$

This completes the proof of Proposition 5.5.1. ∎

Proposition 5.5.2. *There exists $\varepsilon_2 := \varepsilon_2(K,T)$ such that for $0\le\varepsilon\le\varepsilon_2$ the following result holds. Let $\partial_z f\in L^2(\Omega;L^2(0,T;H))$ and $\partial_z\xi\in L^2(\Omega,H)$. Then we have*

$$
\begin{aligned}
&\sup_n \boldsymbol{E}\Big(\sup_{0\le t\le T}|\partial_z u_n(t)|^2+\int_0^T\|\partial_z u_n(s)\|^2ds\Big)\\
&\qquad\le C(K,T,|\partial_z f|_{L^2(\Omega;L^2(0,T;H))})\big(1+\boldsymbol{E}|\partial_z\xi|^2\big).
\end{aligned}\tag{5.5.46}
$$

*Moreover, if σ, g satisfy (**B.2**), $\partial_z f\in L^4(\Omega;L^2(0,T;H))$ and $\partial_z\xi\in L^4(\Omega,H)$, we have*

$$\sup_n \boldsymbol{E}\Big(\sup_{0\le s\le T}|\partial_z u_n|^4\Big)+\sup_n \boldsymbol{E}\Big(\int_0^T \|\partial_z u_n(r)\|^2 dr\Big)^2$$
$$\le C\Big(\boldsymbol{E}|\partial_z f|^4_{L^2(0,T;H)}, K, T\Big)(1+\boldsymbol{E}|\partial_z \xi|^4). \tag{5.5.47}$$

Proof. We apply Itô formula to $|\partial_z u_n|^2$ for $t\in[0,T]$ and τ_N defined by eq. (5.5.23),

$$|\partial_z u_n(t\wedge\tau_N)|^2 + 2\int_0^{t\wedge\tau_N}\left[\|\partial_z u_n(r)\|^2+\|\partial_z\theta_n(r)\|^2\right]dr$$
$$\le \ |P_n\partial_z\xi|^2+\sum_{1\le j\le 5} J_j(t), \tag{5.5.48}$$

where

$$\begin{aligned}
J_1(t) &= 2\int_0^{t\wedge\tau_N}|\langle\partial_{zz}u_n, B(u_n)\rangle|dr,\\
J_2(t) &= 2\int_0^{t\wedge\tau_N}|\langle\partial_{zz}u_n, f\rangle|dr,\\
J_3(t) &= 2\sqrt{\varepsilon}\,\Big|\int_0^{t\wedge\tau_N}\langle\sigma_n(u_n(r))\,dW_n(r), \partial_{zz}u_n(r)\rangle\Big|,\\
J_4(t) &= \ \varepsilon\int_0^{t\wedge\tau_N}|\partial_z\sigma_n(u_n(r))\,P_n|^2_{L_Q}\,dr\\
&\quad+\varepsilon\int_0^{t\wedge\tau_N}\int_Z|\partial_z g_n(u(s-),z')|^2 N(ds,dz'),\\
J_5(t) &= 2\varepsilon\int_0^{t\wedge\tau_N}\int_Z\langle\partial_z u_n(s-), \partial_z g_n(u(s-),z')\rangle\widetilde{N}(ds,dz').
\end{aligned}$$

Noting that eq. (5.5.14), we have

$$J_1(t)=2\int_0^{t\wedge\tau_N}\Big|b\Big(u_n(r),u_n(r),\partial_{zz}u_n(r)\Big)\Big|dr=0,$$

As we obtain estimate (5.5.29), we have

$$J_2\le\varsigma\int_0^{t\wedge\tau_N}\|\partial_z u_n(r)\|^2dr+\frac{1}{\varsigma}\Big(|\partial_z f|^2_{L^2(0,t\wedge\tau_N;H)}\Big). \tag{5.5.49}$$

Then we estimate J_4, we deduce that

$$\begin{aligned}
J_4(t) &= \varepsilon\int_0^{t\wedge\tau_N}\int_Z|\partial_z g_n(u(s-),z')|^2\lambda(dz')ds\\
&\quad+\varepsilon\int_0^{t\wedge\tau_N}\int_Z|\partial_z g_n(u(s-),z')|^2\widetilde{N}(ds,dz')\\
&\quad+\varepsilon\int_0^{t\wedge\tau_N}|\partial_z\sigma_n(u_n(r))\,P_n|^2_{L_Q}\,dr.
\end{aligned} \tag{5.5.50}$$

As we obtain eq. (5.5.30), by hypothesis (**B.1**), we deduce

$$\begin{aligned}\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|J_4(s)|\Big) &\le 3\varepsilon\mathbf{E}\int_0^{t\wedge\tau_N}\int_Z|\partial_z g_n(u(s-),z^{'})|^2\lambda(dz^{'})ds\\ &\quad+\varepsilon\mathbf{E}\int_0^{t\wedge\tau_N}|\partial_z\sigma_n(u_n(r))\,P_n|_{L_Q}^2\,dr\\ &\le \varepsilon KT+\varepsilon K\mathbf{E}\int_0^{t\wedge\tau_N}\|\partial_z u_n(s)\|^2dr. \end{aligned}\tag{5.5.51}$$

Similarly, we have

$$\begin{aligned}\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|J_3(s)|\Big) &\le 2\sqrt{2\varepsilon}\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}|\partial_z u_n(r)|^2\,|\partial_z\sigma_n(u_n(r))\,P_n|_{L_Q}^2\,dr\Big\}^{\frac{1}{2}}\\ &\le\frac{1}{4}\mathbf{E}\sup_{0\le s\le t\wedge\tau_N}|\partial_z u_n|^2\\ &\quad+8\varepsilon K\mathbf{E}\int_0^{t\wedge\tau_N}\|\partial_z u_n(t)\|^2dt+8\varepsilon KT\end{aligned}\tag{5.5.52}$$

and

$$\begin{aligned}&\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|J_5(s)|\Big)\\ &\le\ 2\sqrt{2}\varepsilon\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}\int_Z|\langle\partial_{zz}u_n(s),g_n(u(s),z^{'})\rangle|^2\lambda(dz^{'})ds\Big\}^{\frac{1}{2}}\\ &\le 2\sqrt{2}\varepsilon\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}\int_Z|\partial_z u_n(s)|^2|\partial_z g_n(u(s),z^{'})|^2\lambda(dz^{'})ds\Big\}^{\frac{1}{2}}\\ &\le\frac{1}{4}\mathbf{E}\sup_{0\le s\le t\wedge\tau_N}|\partial_z u_n|^2+8\varepsilon^2K\mathbf{E}\int_0^{t\wedge\tau_N}\|\partial_z u_n(t)\|^2dt+8\varepsilon^2KT.\end{aligned}\tag{5.5.53}$$

Letting $N\to\infty$, by Gronwall's inequality and eqs (5.5.49)–(5.5.53), we get

$$\begin{aligned}&\sup_n\mathbf{E}\Big(\sup_{0\le s\le T}|\partial_z u_n|^2\Big)+\sup_n\mathbf{E}\int_0^T\|\partial_z u_n(r)\|^2dr\\ &\quad\le C(K,T,|\partial_z f|_{L^2(\Omega;L^2(0,T;H))})\big(1+\mathbf{E}|\partial_z\xi|^2\big).\end{aligned}\tag{5.5.54}$$

Similarly, by eq. (5.5.48), we get

$$\begin{aligned}&|\partial_z u_n(t\wedge\tau_N)|^4+4\Big(\int_0^{t\wedge\tau_N}\|\partial_z u_n(r)\|^2\,dr\Big)^2\\ &\quad\le\ 2|P_n\partial_z\xi|^4+10\sum_{1\le i\le 5}J_i^2(t).\end{aligned}\tag{5.5.55}$$

By eq. (5.5.49), we deduce that

$$J_2^2(t) \le 2\varsigma^2\Big(\int_0^{t\wedge\tau_N} \|\partial_z u_n(r)\|^2 dr\Big)^2 + \frac{2}{\varsigma^2}\Big(|\partial_z f|^2_{L^2(0,t\wedge\tau_N;H)}\Big)^2. \tag{5.5.56}$$

As we obtain eqs (5.5.42)–(5.5.44), by (**B.2**), we obtain

$$\begin{aligned}
\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|J_4^2(s)|\Big) &\le 3\varepsilon^2\mathbf{E}\Big(\int_0^{t\wedge\tau_N}\int_Z |\partial_z g_n(u(s-),z')|^2\lambda(dz')ds\Big)^2\\
&\quad + 3\varepsilon^2\mathbf{E}\Big(\int_0^{t\wedge\tau_N}\int_Z |\partial_z g_n(u(s-),z')|^2\widetilde{N}(ds,dz')\Big)^2\\
&\quad + 3\varepsilon^2\,\mathbf{E}\Big(\int_0^{t\wedge\tau_N}|\partial_z\sigma_n(u_n(r))\,P_n|^2_{L_Q}\,dr\Big)^2\\
&\le 3C(K)\varepsilon^2\mathbf{E}\int_0^{t\wedge\tau_N}(1+|\partial_z u_n|^4)ds\\
&\le 3C(K,T)\varepsilon^2 + 3C(K)\varepsilon^2\mathbf{E}\int_0^{t\wedge\tau_N}\sup_{0\le r\le s}|\partial_z u_n(s)|^4 ds,
\end{aligned} \tag{5.5.57}$$

$$\begin{aligned}
\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|J_3^2(s)|\Big) &\le C\sqrt{\varepsilon}\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}|\partial_z u_n(r)|^2\,|\partial_z\sigma_n(u_n(r))\,P_n|^2_{L_Q}\,dr\Big\}\\
&\le \frac{1}{40}\mathbf{E}\sup_{0\le s\le t\wedge\tau_N}|\partial_z u_n|^4\\
&\quad + C(K)\varepsilon\mathbf{E}\Big(\int_0^{t\wedge\tau_N}\|\partial_z u_n(t)\|^2 dt\Big)^2 + C(K,T)\varepsilon
\end{aligned} \tag{5.5.58}$$

and

$$\begin{aligned}
\mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N}|J_5^2(s)|\Big) &\le C\varepsilon\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N}\int_Z |\langle\partial_z u_n(s),\partial_z g_n(u(s),z')\rangle|^2\lambda(dz')ds\Big\}\\
&\le \frac{1}{40}\mathbf{E}\sup_{0\le s\le t\wedge\tau_N}|\partial_z u_n|^4\\
&\quad + C(K)\varepsilon\mathbf{E}\Big(\int_0^{t\wedge\tau_N}\|\partial_z u_n(t)\|^2 dt\Big)^2 + C(K,T)\varepsilon.
\end{aligned} \tag{5.5.59}$$

As we obtain eq. (5.5.35), let $N\to\infty$, by Gronwall's inequality and from eqs (5.5.56) to (5.5.59), we get

$$\begin{aligned}
&\mathbf{E}\Big(\sup_{0\le s\le T}|\partial_z u_n|^4\Big) + \mathbf{E}\Big(\int_0^T\|\partial_z u_n(r)\|^2 dr\Big)^2\\
&\quad\le C\Big(\mathbf{E}|\partial_z f|^4_{L^2(0,T;H)},K,T\Big)(1+\mathbf{E}|\partial_z\xi|^4).
\end{aligned} \tag{5.5.60}$$

∎

Due to Ladyzhenskaya's inequality for 2D domain, we now have the following bound in $L^4(\mathcal{M})$.

Proposition 5.5.3. *Let $\xi \in L^4(\Omega, H)$, then there exists a constant*

$$C_2 := C_2(K, T, M, |f|_{L^4(\Omega;L^2(0,T;H))}),$$

such that

$$\sup_n \boldsymbol{E} \int_0^T |u_n^\varepsilon(s)|_{L^4}^4 ds \le C_2(1 + \boldsymbol{E}|\xi|^4). \tag{5.5.61}$$

Proof of Theorem 5.5.1. Let $\Omega_T = [0, T] \times \Omega$ be endowed with the product measure $ds \otimes d\mathbb{P}$ on $\mathcal{B}([0, T]) \otimes \mathcal{F}$. Using a priori estimates in Propositions 5.5.1 and 5.5.2 and Banach–Alaoglu theorem, we deduce the existence of a subsequence of Galerkin elements u_n and processes $u \in L^2(\Omega_T, V) \cap L^4(\Omega_T, L^4(\mathcal{M})) \cap L^4(\Omega, L^\infty([0, T], H))$, $\partial_z u \in L^2(\Omega_T, V) \cap L^4(\Omega, L^\infty([0, T], H))$, $E \in L^2(\Omega_T, V')$, $S \in L^2(\Omega_T, L_Q)$ and $G \in \mathbb{H}^2_\lambda([0, T] \times Z, H)$, for which the following limits hold:
- (i) $u_n \to u$ weakly in $L^2(\Omega_T, V)$, $\partial_z u_n \to \partial_z u$ weakly in $L^2(\Omega_T, V)$,
- (ii) $u_n \to u$ weakly in $L^4(\Omega_T, L^4(D))$,
- (iii) u_n is weak star converging to u in $L^4(\Omega, L^\infty([0, T], H))$,
  $\partial_z u_n$ is weak star converging to $\partial_z u$ in $L^4(\Omega, L^\infty([0, T], H))$,
- (iv) $E(u_n) \to E$ weakly in $L^2(\Omega_T, V')$,
- (v) $\sigma_n(u_n)P_n \to S$ weakly in $L^2(\Omega_T, L_Q)$,
- (vi) $g_n(u_n) \to G$ weakly in $\mathbb{H}^2_\lambda([0, T] \times Z, H)$.

Indeed, (i)–(iii) are straightforward consequences of Propositions 5.5.1–5.5.3. Furthermore, by an application of eq. (5.5.12),

$$\begin{aligned}&\mathbf{E} \int_0^T |P_n B(u_n(t))|_{V'}^2 dt \\ &\quad \le C\mathbf{E}\Big[\sup_{t\in[0,T]} (|u_n(t)|^4 + |\partial_z u_n(t)|^4) + \Big(\int_0^T \|u_n(t)\|^2 dt\Big)^2\Big] \quad < \infty.\end{aligned} \tag{5.5.62}$$

Hence, $E(u_n(t))$ has a subsequence converging weakly in $L^2(\Omega_T, V')$ to $E(t)$.
Since diffusion coefficient has the linear growth property and u_n is bounded in $L^2([0, T], V)$ uniformly in n, the last two statements hold.

Then u has the Itô differential

$$du(t) = \xi dt + \sqrt{\varepsilon} S(t) dW(t) + E(t)dt + \varepsilon \int_Z G(t)\widetilde{N}(dt, dz'), \tag{5.5.63}$$

weakly in $L^2(\Omega_T, V')$.

In eq. (5.5.63) we still have to prove that $ds \otimes d\mathbb{P}$ a.s. on Ω_T, one has

$$S(s) = \sigma(u(s)), \;\; E(s) = E(u(s)) \text{ and } \; G(s) = G(u(s)).$$

To establish these relations we use the same idea as in Ref. [84]. Let

$$\begin{aligned}\mathcal{X} := \{&\psi \in L^4(\Omega_T, L^4(D)) \cap L^4\big(\Omega, L^\infty([0,T],H)\big) \cap L^2(\Omega_T, V),\\ &\partial_z\psi \in L^4\big(\Omega, L^\infty([0,T],H)\big) \cap L^2(\Omega_T, V)\}.\end{aligned}$$

Then (i)–(iii) yield $u \in \mathcal{X}$, let $\psi \in L^\infty(\Omega_T, H_m) \subset \mathcal{X}$. For every $t \in [0,T]$, set

$$r(t) = \int_0^t \Big[C_1(1 + |\partial_z\psi(s)|^4) + C_2\|\psi(s)\|^2 + \varepsilon L\Big]\,ds. \tag{5.5.64}$$

Then $r(t) < \infty$ for all $t \in [0,T]$.

Applying the Itô lemma to the function $e^{-r(t)|u_n(t)|^2}$, one obtains

$$\begin{aligned}d\big[e^{-r(t)}|u_n(t)|^2\big] = {}& e^{-r(t)}\langle 2E(u_n) - \dot r(t)u_n, u_n\rangle dt\\ &+ \varepsilon e^{-r(t)}|\sigma_n(u_n)|^2_{L_Q}dt + 2\sqrt{\varepsilon}e^{-r(t)}\langle\sigma_n(u_n), u_n\rangle dW\\ &+ e^{-r(t)}\varepsilon\int_Z |g_n(u_n(t-), z')|^2 N(dt, dz')\\ &+ 2e^{-r(t)}\varepsilon\int_Z \langle u_n(t-), g_n(u_n(t-), z')\rangle \widetilde{N}(dt, dz').\end{aligned}$$

Integrating and taking expectation, we get

$$\begin{aligned}\mathbf{E}\big[e^{-r(T)}|u_n(T)|^2 - |u_n(0)|^2\big] = {}& \mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(u_n) - \dot r(t)u_n, u_n\rangle dt\Big]\\ &+ \mathbf{E}\Big[\int_0^T \varepsilon e^{-r(t)}|\sigma_n(u_n)|^2_{L_Q}dt\Big]\\ &+ \mathbf{E}\int_0^T e^{-r(t)}\varepsilon\int_Z |g_n(u_n(t), z')|^2\lambda(dz')dt,\end{aligned}$$

where the term with $dW(t)$ and $\widetilde{N}(dt, dz)$ are martingales and having zero average.

By lower semicontinuity property of weak convergence,

$$\begin{aligned}\liminf_n \mathbf{E}\Big[&\int_0^T e^{-r(t)}\langle 2E(u_n) - \dot r(t)u_n, u_n\rangle dt\\ &+ \int_0^T \varepsilon e^{-r(t)}|\sigma_n(u_n)|^2_{L_Q}dt + \int_0^T e^{-r(t)}\varepsilon\int_Z |g_n(u_n(t), z')|^2\lambda(dz')dt\Big]\\ &= \liminf_n \mathbf{E}\big[e^{-r(T)}|u_n(T)|^2 - |u_n(0)|^2\big]\\ &\ge \mathbf{E}\big[e^{-r(T)}|u(T)|^2 - |u(0)|^2\big]\\ &= \mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E - \dot r(t)u, u\rangle dt + \int_0^T \varepsilon e^{-r(t)}|S|^2_{L_Q}dt\\ &\quad + \int_0^T e^{-r(t)}\varepsilon\int_Z |G|^2\lambda(dz')dt\Big].\end{aligned} \tag{5.5.65}$$

Then by monotonicity property (5.5.21), condition (**A.2**) and eq. (5.5.64), we obtain

$$
\begin{aligned}
&2\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle E(u_n)-E(\psi), u_n-\psi(t)\rangle dt\Big]-\mathbf{E}\Big[\int_0^T e^{-r(t)}\dot r(t)|u_n-\psi|^2 dt\Big]\\
&+\mathbf{E}\Big[\int_0^T \varepsilon e^{-r(t)}|\sigma_n(u_n)-\sigma_n(\psi)|^2_{L_Q}dt\\
&+\int_0^T e^{-r(t)}\varepsilon\int_Z |g_n(u_n(t),z')-g_n(\psi(t),z')|^2\lambda(dz')dt\Big]\\
&\le 0.
\end{aligned}\tag{5.5.66}
$$

Furthermore, we obtain

$$
\begin{aligned}
&\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(u_n)-\dot r(t)u_n, u_n\rangle dt\\
&\quad+\int_0^T \varepsilon e^{-r(t)}|\sigma_n(u_n)|^2_{L_Q}dt+\int_0^T e^{-r(t)}\varepsilon\int_Z |g_n(u_n(t),z')|^2\lambda(dz')dt\Big]\\
&\le \mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(u_n)-\dot r(t)(2u_n-\psi),\psi\rangle dt\Big]\\
&+\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(\psi),u_n-\psi\rangle dt\Big]\\
&\quad+\varepsilon\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2\sigma_n(u_n)-\sigma_n(v),\sigma_n(v)\rangle_{L_Q}\Big]\\
&\quad+\varepsilon\mathbf{E}\Big[\int_0^T e^{-r(t)}\int_Z\langle 2g_n(u_n,z')-g_n(v,z'),g_n(v,z')\rangle\lambda(dz')dt\Big].
\end{aligned}
$$

Letting $n\to\infty$, by eq. (5.5.65), we have

$$
\begin{aligned}
&\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(t)-\dot r(t)u, u\rangle dt\\
&\quad+\int_0^T \varepsilon e^{-r(t)}|S|^2_{L_Q}dt+\int_0^T e^{-r(t)}\varepsilon\int_Z |G|^2\lambda(dz')dt\Big]\\
&\le \mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(t)-\dot r(t)(2u-\psi),\psi\rangle dt\Big]+\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(\psi),u-\psi\rangle dt\Big]\\
&\quad+\varepsilon\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2S-\sigma(\psi),\sigma(\psi)\rangle_{L_Q}dt\Big]\\
&\quad+\varepsilon\mathbf{E}\Big[\int_0^T e^{-r(t)}\int_Z\langle 2G-g(\psi,z'),g(\psi,z')\rangle\lambda(dz')dt\Big].
\end{aligned}
$$

Rearranging the terms, we obtain

$$
\begin{aligned}
&\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(t)-2E(\psi),u-\psi\rangle dt\Big]+\mathbf{E}\Big[\int_0^T e^{-r(t)}\dot r(t)|u(t)-\psi(t)|^2 dt\Big]\\
&\quad+\mathbf{E}\Big[\int_0^T \varepsilon e^{-r(t)}|S-\sigma(\psi)|^2_{L_Q}dt\Big]+\mathbf{E}\Big[\int_0^T e^{-r(t)}\varepsilon\int_Z |G-g(\psi)|^2\lambda(dz')dt\Big]\\
&\le 0.
\end{aligned}
$$

Taking $\psi(t) = u(t)$, we obtain $S(t) = \sigma(u(t))$ and $G(t) = g(u(t))$. If we note $\psi(t) = u(t) - \mu\eta(t)$, $\mu > 0$, we have

$$\mu\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(t) - 2E(u - \mu\eta), \eta\rangle dt + \mu\int_0^T e^{-r(t)}\dot{r}(t)|\eta(t)|^2 dt\Big] \le 0.$$

Dividing by μ on both sides of above inequality and letting $\mu \to 0$, we have

$$\mathbf{E}\Big[\int_0^T e^{-r(t)}\langle 2E(t) - 2E(u - \mu\eta), \eta\rangle dt\Big] \le 0.$$

Since $\eta(t)$ is arbitrary, we obtain $E(t) = E(u(t))$. Thus the existence has been proved.

To complete the proof of Theorem 5.5.1, we should show that $u(t)$ is unique in $L^2(\Omega, \mathcal{D}(0, T, H) \cap L^2(0, T, V))$. Suppose $v(t)$ be another solution and let $\varrho(t) = u(t) - v(t)$, then $\varrho(t)$ satisfies the following equation:

$$\begin{aligned} d\varrho(t) &= (E(u) - E(v))dt + \sqrt{\varepsilon}(\sigma(u) - \sigma(v))dW \\ &\quad + \int_Z (g(u(t-), z') - g(v(t-), z'))\widetilde{N}(dt, dz'). \end{aligned} \tag{5.5.67}$$

Applying Itô lemma to $e^{-r(t)}|\varrho(t)|^2$, we get

$$\begin{aligned} d(e^{-r(t)}|\varrho(t)|^2) = \Big[&- e^{-r(t)}\dot{r}(t)|\varrho(t)|^2 + 2e^{-r(t)}\langle E(u) - E(v), \varrho(t)\rangle \\ &+ \varepsilon e^{-r(t)}|\sigma(u) - \sigma(v)|^2_{L_Q}\Big]dt \\ &+ 2\sqrt{\varepsilon}e^{-r(t)}\langle \sigma(u) - \sigma(v), \varrho\rangle dW(t) \\ &+ e^{-r(t)}\varepsilon\int_Z |g(u(t), z') - g(v(t)|^2 N(dt, dz') \\ &+ 2e^{-r(t)}\varepsilon\int_Z \langle g(u(t-), z') - g(v(t-), z'), \varrho(t)\rangle\widetilde{N}(dt, dz'). \end{aligned} \tag{5.5.68}$$

Taking expectation on both sides and recalling that the martingales have zero averages and

$$r(t) = \int_0^t \Big[C_1(1 + |\partial_z v(s)|^4) + C_2\|v(s)\|^2 + \varepsilon L\Big] ds, \tag{5.5.69}$$

noting eq. (5.5.66), we have

$$\mathbf{E}\Big[e^{-r(t)}|\varrho(t)|^2 + C\int_0^T e^{-r(t)}\|\varrho(t)\|^2 dt\Big] \le \mathbf{E}|\varrho(0)|^2 = 0, \tag{5.5.70}$$

which leads to the uniqueness of the solution. ∎

5.5.2 Large deviation for stochastic primitive equations

In this section, we first obtain the well-posedness and general a priori estimates for 2D stochastic primitive equations with small and more general multiplicative noise, which couple the temperature equation by another method which is different from that in Ref. [113]. Our second result is a Wentzell–Freidlin-type large deviation principle (LDP) for 2D stochastic primitive equations. There are already several interesting and important papers on LDP and its applications [53, 62, 84, 99, 141, 201, 234]. Especially Ref. [62] dealts with a class of abstract nonlinear stochastic models, which covers many 2D hydrodynamical models including 2D Navier–Stokes equations, 2D magnetohydrodynamic models and 2D magnetic Bénard problem and also some shell models of turbulence, but does not include our problem. It is because of the mapping B in this section does not satisfy the condition (C1) of Ref. [62]. The idea of our proof by weak convergence approach [42, 43] is similar to the one in Ref. [84]; therefore, in this section, we only give the outline of the proof.

5.5.2.1 Mathematical formulation

We study the following stochastic 2D primitive equations:

$$\frac{du^{\varepsilon}}{dt} = \nu_1 \Delta u^{\varepsilon} - u^{\varepsilon}\partial_x u^{\varepsilon} - w^{\varepsilon}\partial_z u^{\varepsilon} - \partial_x p + f + \sqrt{\varepsilon}\sigma_1(t, \phi^{\varepsilon})\dot{W}_1, \tag{5.5.71}$$

$$\partial_z p = -\theta, \tag{5.5.72}$$

$$\partial_x u^{\varepsilon} = -\partial_z w^{\varepsilon}, \tag{5.5.73}$$

$$\frac{d\theta^{\varepsilon}}{dt} = \nu_2 \Delta \theta^{\varepsilon} - u^{\varepsilon}\partial_x \theta^{\varepsilon} - w^{\varepsilon}\partial_z \theta^{\varepsilon} + q + \sqrt{\varepsilon}\sigma_2(t, \phi^{\varepsilon})\dot{W}_2, \tag{5.5.74}$$

with velocity $u^{\varepsilon} = u^{\varepsilon}(t, x, z) \in \mathbb{R}$, temperature $\theta^{\varepsilon} = \theta^{\varepsilon}(t, x, z) \in \mathbb{R}$, $\phi^{\varepsilon} = (u^{\varepsilon}, \theta^{\varepsilon})$, pressure p and f is an external forcing term, q is a given heat source, $(x, z) \in \mathcal{M} = [0, l] \times [-h, 0]$, $t > 0$ and $\dot{W}_1$ and $\dot{W}_2$ are the white noise processes. Here Δ is the Laplacian operator and without loss of generality in this section, we take ν_1, ν_2 to be 1.

We partition the boundary into the top $\Gamma_u = \{z = 0\}$, the bottom $\Gamma_b = \{z = -h\}$ and the sides $\Gamma_s = \{x = 0\} \cup \{x = l\}$. In this section, we consider the following boundary conditions:

$$\begin{aligned}
&\text{on } \Gamma_u: \ \partial_z u^{\varepsilon} = 0, \ w^{\varepsilon} = 0, \ \partial_z \theta^{\varepsilon} = 0,\\
&\text{on } \Gamma_b: \ \partial_z u^{\varepsilon} = 0, \ w^{\varepsilon} = 0, \ \partial_z \theta^{\varepsilon} = 0,\\
&\text{on } \Gamma_s: \ u^{\varepsilon} = 0, \ \ \partial_x \theta^{\varepsilon} = 0.
\end{aligned}$$

Due to eqs (5.5.72) and (5.5.73), we have that

$$w(x, z, t) = -\int_{-h}^{z} \partial_x u^{\varepsilon}(x, \xi, t)d\xi, \tag{5.5.75}$$

$$p(x, z, t) = p_s(x, t) - \int_{-h}^{z} \theta^{\varepsilon}(x, \xi, t)d\xi. \tag{5.5.76}$$

Note that p_s denotes the surface pressure (with only x-dependence). We define the function spaces H and V as follows:

$$H = H_1 \times H_2, \quad V = V_1 \times V_2, \tag{5.5.77}$$

$$H_1 = \left\{ v \in L^2(\mathcal{M}) \mid \int_{-h}^{0} v dz = 0 \right\}, \tag{5.5.78}$$

$$H_2 = \left\{ \theta \in L^2(\mathcal{M}) \right\}, \tag{5.5.79}$$

$$V_1 = \left\{ v \in H^1(\mathcal{M}) \mid \int_{-h}^{0} v dz = 0, \quad v|_{\Gamma_s} = 0 \right\}, \tag{5.5.80}$$

$$V_2 = \left\{ \theta \in H^1(\mathcal{M}) \right\}. \tag{5.5.81}$$

These spaces are endowed with the L^2 and H^1 norms, which we, respectively, denote by $|\cdot|$ and $\|\cdot\|$. The inner products and norms on V, H are given by

$$(U, U_1) = (v, v_1) + (\theta, \theta_1) = \int_{\mathcal{M}} (v v_1 + \theta\theta_1) d\mathcal{M},$$

$$((U, U_1)) = ((v, v_1)) + ((\theta, \theta_1)) = \int_{\mathcal{M}} (\nabla v \nabla v_1 + \nabla\theta\nabla\theta_1 + \theta\theta_1) d\mathcal{M},$$

$$|U| = (U, U)^{\frac{1}{2}}, \|U\| = ((U, U))^{\frac{1}{2}},$$

where $U = (v, \theta), U_1 = (v_1, \theta_1) \in V$. We shall also need the intermediate space

$$Y = Y_1 \times Y_2, \tag{5.5.82}$$

$$Y_1 = \{v \in H_1, \partial_z v \in H_1\}, \quad Y_2 = \{\theta \in H_2, \partial_z \theta \in H_2\}, \tag{5.5.83}$$

Let V' be the dual space of V. We have the dense and continuous embeddings $V \hookrightarrow H = H' \hookrightarrow V'$ and denote by $\langle \phi, \psi \rangle$ the duality between $\phi \in V$ (resp. V_i) and $\phi \in V'$ (resp. V_i').

Consider an unbounded linear operator $A = (A_1, A_2) = (\Delta, \Delta) : D(A) \to H$ with $D(A) = D(A_1) \times D(A_2)$, where

$$D(A_1) = \{u \in V_1 \cap H^2(\mathcal{M}) : \partial_z u|_{\Gamma_u} = \partial_z u|_{\Gamma_b} = 0\} \subset V_1 \cap H^2(\mathcal{M}),$$

$$D(A_2) = \{\theta \in V_2 \cap H^2(\mathcal{M}) : \partial_x \theta|_{\Gamma_s} = 0, \partial_z \theta|_{\Gamma_b} = \partial_z \theta|_{\Gamma_u} = 0\} \subset V_2 \cap H^2(\mathcal{M}),$$

and define

$$\langle A_1 u, v \rangle = ((u, v)), \ \langle A_2 \theta, \eta \rangle = ((\theta, \eta)) \ \forall u, v \in D(A_1), \ \forall \theta, \eta \in D(A_2).$$

The Laplace operators A_1 and A_2 are self-adjoint, positive, with compact self-adjoint inverses. They map V_i to V_i', $i = 1, 2$.

In accordance with eq. (5.5.75) we take

$$\mathcal{W}(v) := -\int_{-h}^{z} \partial_x v(x,\tilde{z})d\tilde{z} \tag{5.5.84}$$

and let

$$B_1(u,v) := u\partial_x v + \mathcal{W}(u)\partial_z v, \tag{5.5.85}$$

where $u \in V_1$, and $v \in V_1$ or V_2.

One would like to establish that B_1 is a well-defined and continuous mapping from $V_1 \times V_1 \to V_1'$ or $V_1 \times V_2 \to V_2'$ according to

$$\langle B_1(u,v), w\rangle = b_1(u,v,w), \tag{5.5.86}$$

where the associated trilinear form is given by

$$b_1(u,v,w) = \int_{\mathcal{M}} (u\partial_x vw + \mathcal{W}(u)\partial_z vw)d\mathcal{M}.$$

In the sequel, when no confusion arises, we denote by C a constant which may change from one line to the next one.

Lemma 5.5.5 (Estimates for b_1 and B_1). *The trilinear forms b_1 and B_1 have the following properties. There exists a constant $C > 0$ such that*

$$\begin{aligned}
|b_1(u,v,w)| &\le C\Big(|u|^{\frac{1}{2}}\|u\|^{\frac{1}{2}}\|v\||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}} + |\partial_x u||\partial_z v||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}}\Big),\\
&\qquad u \in V_1, v \in V_1(or\ V_2), w \in V_1(or\ V_2), \qquad (5.5.87)\\
b_1(u,v,v) &= 0, \qquad u \in V_1, v \in V_1(or\ V_2), \qquad (5.5.88)\\
\langle B_1(u,u), \partial_{zz}u\rangle &= 0, \qquad u \in V_1. \qquad (5.5.89)
\end{aligned}$$

Proof. Following the similar estimates in Ref. [113], we easily obtain these results. In this section, we only give the proof of eqs (5.5.87) and (5.5.89). By Hölder's inequality and Ladyzhenskaya's inequality, we have

$$\begin{aligned}
|b_1(u,v,w)| &\le \int_{\mathcal{M}} \Big(|u\partial_x vw| + |\mathcal{W}(u)\partial_z vw|\Big)d\mathcal{M}\\
&\le C|u|^{\frac{1}{2}}\|u\|^{\frac{1}{2}}\|v\||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}}\\
&\quad + \int_0^l \Big(\sup_{z\in[-h,0]}\Big\{\int_{-h}^{z}\partial_x u dz\Big\}\int_0^l |\partial_z vw|dz\Big)dx\\
&\le C|u|^{\frac{1}{2}}\|u\|^{\frac{1}{2}}\|v\||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}}\\
&\quad + C\int_0^l \Big(\int_{-h}^{0}|\partial_x u|^2 dz \cdot \int_{-h}^{0}|\partial_z v|^2 dz \cdot \int_{-h}^{0}|w|^2 dz\Big)^{1/2} dx\\
&\le C|u|^{\frac{1}{2}}\|u\|^{\frac{1}{2}}\|v\||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}} + C\sup_{x\in[0,l]}\Big(\int_{-h}^{0}|w|^2 dz\Big)^{1/2}\times
\end{aligned}$$

$$\int_0^l \Big(\int_{-h}^0 |\partial_x u|^2 dz \cdot \int_{-h}^0 |\partial_z v|^2 dz\Big)^{1/2} dx$$
$$\leq C|u|^{\frac{1}{2}}\|u\|^{\frac{1}{2}}\|v\||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}} + C|\partial_x u||\partial_z v||w|^{\frac{1}{2}}\|w\|^{\frac{1}{2}}.$$

Noting eqs (5.5.85) and (5.5.86) and considering boundary conditions, we have

$$\begin{aligned}
\langle B_1(u,u), \partial_{zz}u\rangle &= \int_{\mathcal{M}} (u\partial_x u\partial_{zz}u + \mathcal{W}(u)\partial_z u\partial_{zz}u)d\mathcal{M} \\
&= -\int_{\mathcal{M}} [\partial_z(u\partial_x u)\partial_z u + \partial_z(\mathcal{W}(u)\partial_z u)\partial_z u]d\mathcal{M} \\
&= \int_{\mathcal{M}} [-\partial_x u(\partial_z u)^2 - u\partial_{xz}u\partial_z u + \frac{1}{2}\partial_x u(\partial_z u)^2]d\mathcal{M} \\
&= 0.
\end{aligned}$$

■

Remark 5.5.4. *Glatt-Holtz and Temam [112] considered $\partial_z u + \alpha_u u = 0, \partial_z\theta + \alpha_\theta\theta = 0$ on the top boundary, and defined different function spaces from this section. Due to the boundary conditions which they considered on the top boundary, in order to deal with the pressure term, they introduced a projection operator $\mathcal{Q}$ onto H from $L^2(\mathcal{M})^2$ such that $\mathcal{Q}\partial_x p_s = 0$ and $\langle B^1(u,u), \partial_{zz}u\rangle \neq 0$. In this section, since we consider free boundary condition on top boundary, we do not introduce the projection operator and thus the pressure term remains in equations. To obtain the estimates for $|u|^{2p}$ and $|\partial_z u|^{2p}$, noting the boundary conditions considered in this section and that p_s is only x-dependent, we have the inner product*

$$\begin{aligned}
\langle u, \partial_x p_s\rangle &= \int_0^l \int_{-h}^0 u\partial_x p_s dxdz \\
&= -\int_0^l p_s\Big(\partial_x \int_{-h}^0 udz\Big)dx \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
\langle \partial_{zz}u, \partial_x p_s\rangle &= \int_0^l \int_{-h}^0 \partial_{zz}u\partial_x p_s dxdz \\
&= -\int_0^l p_s\Big(\partial_x \int_{-h}^0 \partial_{zz}udz\Big)dx \\
&= -\int_0^l p_s\Big(\partial_x\partial_z u|_{z=-h}^{z=0}\Big)dx \\
&= 0.
\end{aligned}$$

But if we consider the same boundary conditions as in Ref. [112], we should introduce the projection operator $\mathcal{Q}$ and add estimates for $\langle B^1(u,u), \partial_{zz}u\rangle$ as the author did in Ref. [112].

In the present section, we assume that $W_1(t) \in H_1(\mathcal{M}), W_2(t) \in H_2(\mathcal{M})$ are independent Wiener processes defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, with linear symmetric positive covariant operators Q_1 and Q_2, respectively. We denote $Q = (Q_1, Q_2)$. It is a linear symmetric positive covariant operator in the Hilbert space H. We assume that Q_1, Q_2 and thus Q are trace class (and hence compact [75]), i.e., $\mathrm{tr}(Q) < \infty$.

We still need to introduce some definitions (see [84, 235]) in the following:

– Hilbert space $H_0 = Q^{\frac{1}{2}}H$ with the scalar product

$$(\phi, \psi)_0 = (Q^{-\frac{1}{2}}\phi, Q^{-\frac{1}{2}}\psi)\ \forall \phi, \psi \in H_0,$$

together with the induced norm $|\cdot|_0 = \sqrt{(\cdot,\cdot)_0}$.

– L_Q is the space of linear operators S such that $SQ^{\frac{1}{2}}$ is a Hilbert–Schmidt operator (and thus a compact operator [75]) from H to H. The norm in the space L_Q is defined by $|S|^2_{L_Q} = \mathrm{tr}\,(SQS^*)$, where S^* is the adjoint operator of S.

– $\mathcal{A}$ is the set of H_0 -valued $(\mathcal{F}_t)$-predictable stochastic processes ϕ with the property $\int_0^T |\phi(s)|_0^2 ds < \infty$, a.s. Define

$$\mathcal{A}_M = \{\phi \in \mathcal{A} : \phi(\omega) \in S_M,\ \text{a.s.}\}. \tag{5.5.90}$$

With these notations, the above primitive equations become

$$\begin{aligned} \frac{du^\varepsilon}{dt} &= \Delta u^\varepsilon - u^\varepsilon \partial_x u^\varepsilon - \mathcal{W}(u^\varepsilon)\partial_z u^\varepsilon - \partial_x p_s \\ &\quad + \int_{-h}^{z} \partial_x \theta^\varepsilon + f + \sqrt{\varepsilon}\sigma_1(t, \phi^\varepsilon)\dot{W}_1, \end{aligned} \tag{5.5.91}$$

$$\frac{d\theta^\varepsilon}{dt} = \Delta\theta^\varepsilon - u^\varepsilon \partial_x \theta^\varepsilon - \mathcal{W}(u^\varepsilon)\partial_z \theta^\varepsilon + q + \sqrt{\varepsilon}\sigma_2(t, \phi^\varepsilon)\dot{W}_2. \tag{5.5.92}$$

Thus, we rewrite this system for $\phi^\varepsilon = (u^\varepsilon, \theta^\varepsilon)$ as

$$d\phi^\varepsilon + [A\phi^\varepsilon + B(\phi^\varepsilon, \phi^\varepsilon) + F(\phi^\varepsilon)]dt = Rdt + \sqrt{\varepsilon}\sigma(t, \phi^\varepsilon)dW(t), \tag{5.5.93}$$

where $W(t) = (W_1(t), W_2(t))$ and

$$A\phi^\varepsilon = (A_1 u^\varepsilon,\quad A_2\theta^\varepsilon), \tag{5.5.94}$$

$$B(\phi^\varepsilon) = (B_1(u^\varepsilon, u^\varepsilon),\quad B_1(u^\varepsilon, \theta^\varepsilon)), \tag{5.5.95}$$

$$F\phi^\varepsilon = (\partial_x p - \int_{-h}^{z} \partial_x \theta^\varepsilon,\quad 0), \tag{5.5.96}$$

$$R = (f,\quad q), \tag{5.5.97}$$

$$\sigma(t, \phi^\varepsilon) = (\sigma_1(t, \phi^\varepsilon), \sigma_2(t, \phi^\varepsilon)). \tag{5.5.98}$$

The noise intensity $\sigma : [0,T] \times V \to L_Q(H_0, H)$ is assumed to satisfy the following conditions:

Assumption A. There exist positive constants K and L such that

(A.1) $\sigma \in C\big([0,T] \times H; L_Q(H_0,H)\big)$,

(A.2) $|\sigma(t,\phi)|^2_{L_Q} \le K(1 + \|\phi\|^2) \quad \forall t \in [0,T]\,, \forall \phi \in V$,

(A.3) $|\sigma(t,\phi) - \sigma(t,\psi)|^2_{L_Q} \le L\|\phi - \psi\|^2 \quad \forall t \in [0,T]\ \forall \phi, \psi \in V$.

In order to obtain large deviation, we should introduce the stochastic control equation, let $h \in \mathcal{A}$, $\varepsilon \ge 0$ and consider the following generalized primitive equations with initial condition $\phi^\varepsilon_h(0) = \xi$:

$$\begin{aligned} d\phi^\varepsilon_h(t) + \big[A\phi^\varepsilon_h(t) + B(\phi^\varepsilon_h(t)) + F(\phi^\varepsilon_h)\big]dt \\ = Rdt + \sqrt{\varepsilon}\sigma(\phi^\varepsilon_h(t))dW(t) + \tilde\sigma(\phi^\varepsilon_h(t))h(t)dt. \end{aligned} \tag{5.5.99}$$

Then, we introduce another intensity coefficient $\tilde\sigma \in C([0,T] \times V; L_Q(H_0,H))$ such that

Assumption Ã. There exist positive constants $\tilde K$ and $\tilde L$ such that

($\tilde A$.1) $|\tilde\sigma(t,\phi)|^2_{L_Q} \le \tilde K(1 + |\phi|^2_{L^4}) \quad \forall t \in [0,T], \forall \phi \in L^4(D)$,

($\tilde A$.2) $|\tilde\sigma(t,\phi) - \tilde\sigma(t,\psi)|^2_{L_Q} \le \tilde L|\phi - \psi|^2_{L^4} \quad \forall t \in [0,T], \forall \phi, \psi \in L^4(D)$.

Remark 5.5.5. *Continuity condition* **(A.1)** *and Lipschitz condition* **(A.3)** *imply the growth condition* **(A.2)**. *Meanwhile,* ($\tilde A$.2) *together with the assumption* $\tilde\sigma \in C([0,T] \times V; L_Q(H_0,H))$ *imply* ($\tilde A$.1). *We list* **(A.2)** *and* ($\tilde A$.1) *here only for convenience.*

To obtain Theorem 5.5.1, we have to give the additional assumptions on σ and $\widetilde\sigma$.

Assumption B. There exist positive constant $\overline{K}$ such that

(B.1) $|\partial_z\sigma(t,\phi)|^2_{L_Q} \le \overline{K}(1 + \|\partial_z\phi\|^2) \quad \forall t \in [0,T]\ \forall \partial_z\phi \in V$,

(B.2) $|\partial_z\tilde\sigma(t,\phi)|^2_{L_Q} \le \overline{K}(1 + \|\partial_z\phi\|^2) \quad \forall t \in [0,T]\ \forall \partial_z\phi \in V$.

5.5.2.2 Well-posedness

Let $X := C\big([0,T]; H\big) \cap L^2\big((0,T); V\big)$ denote the Banach space endowed with the norm

$$\|\phi\|_X = \left\{ \sup_{0\le s\le T} |\phi(s)|^2 + \int_0^T \|\phi(s)\|^2 ds \right\}^{\frac{1}{2}}. \tag{5.5.100}$$

Recall that an $\{\mathcal{F}_t\}$-predictable stochastic process $\phi^\varepsilon_h(t,\omega)$ is called the weak solution for the generalized stochastic primitive problem (5.5.99) on $[0,T]$ with initial condition $\xi \in X$, if ϕ^ε_h is in $C([0,T]; H) \cap L^2((0,T); V)$, a.s., and satisfies

$$(\phi_h^\varepsilon(t),\psi)-(\xi,\psi)+\int_0^t[\langle\phi_h^\varepsilon(s),A\psi\rangle+\langle B(\phi_h^\varepsilon(s)),\psi\rangle+(F(\phi_h^\varepsilon(s)),\psi)]ds=$$

$$\int_0^t(R,\psi)ds+\sqrt{\varepsilon}\int_0^t\left(\sigma(\phi_h^\varepsilon(s))dW(s),\psi\right)+\int_0^t\left(\tilde{\sigma}(\phi_h^\varepsilon(s))h(s)\,,\,\psi\right)ds,\quad \text{a.s.,}\tag{5.5.101}$$

for all $\psi\in D(A)$ and all $t\in[0,T]$. Note that this solution is a strong one in the probabilistic meaning, that is written in terms of stochastic integrals with respect to the given Brownian motion W. The main result of this section is the following theorem.

Theorem 5.5.1 (Well-posedness and a priori bounds). *Fix $M>0$, then there exists $\varepsilon_0:=\varepsilon_0(\overline{K},K,L,\tilde{K},\tilde{L},T,M)>0$, such that the following existence and uniqueness result is true for $0\le\varepsilon\le\varepsilon_0$. Let the initial datum $\xi\in Y$ and satisfy $\mathbf{E}|\xi|^4<\infty$, $\mathbf{E}|\partial_z\xi|^4<\infty$, and let $h\in\mathcal{A}_M$, $f,\partial_z f\in L^4(\Omega;L^2(0,T;H))$, $q,\partial_z q\in L^4(\Omega;L^2(0,T;H))$ and $\varepsilon\in[0,\varepsilon_0]$, then there exists a unique weak solution ϕ_h^ε of the generalized stochastic primitive problem* (5.5.99) *with initial condition $\phi_h^\varepsilon(0)=\xi\in Y$ such that $\phi_h^\varepsilon\in X$ a.s. Furthermore, there exists a constant*

$$C:=C\Big(\overline{K},K,L,\tilde{K},\tilde{L},T,M,|f|_{L^4(\Omega;L^2(0,T;H))},|q|_{L^4(\Omega;L^2(0,T;H))},\\ |\partial_z f|_{L^4(\Omega;L^2(0,T;H))},|\partial_z q|_{L^4(\Omega;L^2(0,T;H))}\Big),$$

such that for $\varepsilon\in[0,\varepsilon_0]$ and $h\in\mathcal{A}_M$,

$$E\|\phi_h^\varepsilon\|_X^2\le 1+E\Big(\sup_{0\le t\le T}|\phi_h^\varepsilon(t)|^4+\int_0^T\|\phi_h^\varepsilon(t)\|^2\,dt\Big)\le C\left(1+E|\xi|^4\right),\tag{5.5.102}$$

and satisfy the additional regularity:

$$\partial_z\phi_h^\varepsilon(t)\in L^4(\Omega,L^\infty(0,T;H)\cap L^2(0,T;V)).\tag{5.5.103}$$

For $\phi=(u,\theta)\in V$, define

$$E(\phi)=-A\phi-B(\phi)-F(\phi)+R.\tag{5.5.104}$$

We first obtain monotonicity properties of E.

Lemma 5.5.6. *Assume that $\phi=(u,\theta)\in V$ and $\psi=(v,\eta)\in V$, we have*

$$\langle E(\phi)-E(\psi),\phi-\psi\rangle+\frac{1}{2}\|\phi-\psi\|^2\le C|\phi-\psi|\,\|\phi-\psi\|\,\|\psi\|+C(1+|\partial_z\psi|^4)|\phi-\psi|^2.\tag{5.5.105}$$

Proof. Setting $U = u - v$, $\Theta = \theta - \eta$ and $\Phi = \phi - \psi := (U, \Theta)$, we deduce

$$\begin{aligned}\langle E(\phi) - E(\psi), \Phi\rangle &= -\langle A(\phi) - A(\psi), \Phi\rangle - \langle B(\phi) - B(\psi), \Phi\rangle \\ &\quad -\langle F(\phi) - F(\psi), \Phi\rangle \\ &\equiv I_1 + I_2 + I_3.\end{aligned}$$

Integrating by parts and using Lemma 5.5.5, Hölder's inequality, Ladyzhenskaya's inequality for 2D domain and Young's inequality, we get

$$I_1 = -|\nabla(u - v)|^2 - |\nabla(\theta - \eta)|^2 = -\|\phi - \psi\|^2,$$

$$\begin{aligned}I_2 &= b_1(u - v, u - v, v) + b_1(u - v, \theta - \eta, \eta) \\ &\le |u - v|_{L^4}\|v\||u - v|_{L^4} + |u - v|_{L^4}\|\eta\||\theta - \eta|_{L^4} \\ &\quad +|\partial_x(u - v)||\partial_z v||u - v|^{\frac{1}{2}}\|u - v\|^{\frac{1}{2}} + |\partial_x(u - v)||\partial_z\eta||\theta - \eta|^{\frac{1}{2}}\|\theta - \eta\|^{\frac{1}{2}} \\ &\le C|u - v|\|u - v\|\|v\| + C|u - v|^{\frac{1}{2}}\|u - v\|^{\frac{1}{2}}\|\eta\||\theta - \eta|^{\frac{1}{2}}\|\theta - \eta\|^{\frac{1}{2}} \\ &\quad +\|u - v\|^{\frac{3}{2}}|\partial_z v||u - v|^{\frac{1}{2}} + \|u - v\|^{\frac{1}{2}}|\|\theta - \eta\|^{\frac{1}{2}}|\partial_z\eta||\theta - \eta|^{\frac{1}{2}} \\ &\le C|\phi - \psi|\|\phi - \psi\|(\|v\| + \|\eta\|) + C\|\phi - \psi\|^{\frac{3}{2}}|\phi - \psi|^{\frac{1}{2}}|\partial_z\psi| \\ &\le C|\phi - \psi|\|\phi - \psi\|\|\psi\| + \frac{1}{4}\|\phi - \psi\|^2 + C|\partial_z\psi|^4|\phi - \psi|^2,\end{aligned}$$

$$I_3 = \int_M\Big(\int_{-h}^{z}\partial_x(\theta - \eta)d\tilde{z}\cdot(u - v)\Big)dM \le \int_M\Big(\int_{-h}^{0}|\partial_x(\theta - \eta)|d\tilde{z}\cdot|u - v|\Big)dM$$

$$\le h|\partial_x(\theta - \eta)||u - v| \le \frac{1}{4}\|\phi - \psi\|^2 + C|\phi - \psi|^2.$$

Combining I_1, I_2 and I_3, we end the proof. ∎

We now introduce the Galerkin systems associated with the original equation and establish some uniform a priori estimates. For any $n \ge 1$, let $H_n = \mathrm{span}(e_1, \cdots, e_n) \subset \mathrm{Dom}(A)$ and $P_n : H \to H_n$ denote the orthogonal projection onto H_n. Note that P_n contracts the H and V norms. Suppose that the H-valued Wiener process W with covariance operator Q is such that

$$P_n Q^{\frac{1}{2}} = Q^{\frac{1}{2}} P_n, \quad n \ge 1,$$

which is true if $Qh = \sum_{n\ge1}\lambda_n e_n$ with trace $\sum_{n\ge1}\lambda_n < \infty$. Then for $H_0 = Q^{\frac{1}{2}}H$ and $(\phi, \psi)_0 = (Q^{-\frac{1}{2}}\phi, Q^{-\frac{1}{2}}\psi)$, for $\phi, \psi \in H_0$, we see that $P_n : H_0 \to H_0 \cap H_n$ is a contraction in both of the H and H_0 norms. Let $W_n = P_n W$, $\sigma_n = P_n\sigma$ and $\tilde{\sigma}_n = P_n\tilde{\sigma}$.

For $h \in \mathcal{A}_M$, we consider the following stochastic ordinary differential equation on the n-dimensional space H_n defined by

$$d(\phi^\varepsilon_{n,h}, \psi) = \left[\langle E(\phi^\varepsilon_{n,h}), \psi\rangle + (\tilde{\sigma}_n(\phi^\varepsilon_{n,h})h, \psi)\right]dt + \sqrt{\varepsilon}\,(\sigma_n(\phi^\varepsilon_{n,h})dW_n, \psi), \tag{5.5.106}$$

for $\psi = (v, \eta) \in H_n$ and $\phi^\varepsilon_{n,h}(0) = P_n\xi$.

We note that formulation (5.5.106) allows one to treat $\phi^\varepsilon_{n,h}$ as a process in R^n. Hence by a well-posedness result for stochastic ordinary differential equations, there exists a maximal solution to eq. (5.5.106), i.e., a stopping time $\tau^\varepsilon_{n,h} \le T$ such that eq. (5.5.106) holds for $t < \tau^\varepsilon_{n,h}$ and as $t \uparrow \tau^\varepsilon_{n,h} < T$, $|\phi^\varepsilon_{n,h}(t)| \to \infty$. As such one can apply the finite-dimensional Itô's calculus to the above Galerkin systems. We next establish some uniform estimates on $\phi^\varepsilon_{n,h}$ (independent of n). For every $N > 0$, set

$$\tau_N = \inf\{t :\ |\phi^\varepsilon_{n,h}(t)| \ge N\} \wedge \inf\{t :\ |\partial_z\phi^\varepsilon_{n,h}(t)| \ge N\} \wedge T. \tag{5.5.107}$$

The following proposition provides the (global) existence and uniqueness of approximate solutions and also their uniform (a priori) estimates. This is the main preliminary step in the proof of Theorem 5.5.1.

Proposition 5.5.4. *There exists $\varepsilon_{0,p} := \varepsilon_{0,p}(K, \tilde{K}, T, M)$ such that for $0 \le \varepsilon \le \varepsilon_{0,p}$ the following result holds for an integer $p \ge 1$ (with the convention $x^0 = 1$). Let $h \in \mathcal{A}_M$, $f, q \in L^{2p}(\Omega; L^2(0, T; H))$ and $\xi \in L^{2p}(\Omega, H)$. Then eq.* (5.5.106) *has a unique solution with a modification $\phi^\varepsilon_{n,h} \in C([0, T], H_n)$ satisfying*

$$\begin{aligned}\sup_n \mathbf{E}\Big(\sup_{0\le t\le T} |\phi^\varepsilon_{n,h}(t)|^{2p} + \int_0^T \|\phi^\varepsilon_{n,h}(s)\|^2\, |\phi^\varepsilon_{n,h}(s)|^{2(p-1)} ds\Big)&\\ \le C(p, K, \tilde{K}, T, M, |f|_{L^{2p}(\Omega;L^2(0,T;H))}, |q|_{L^{2p}(\Omega;L^2(0,T;H))})\big(\mathbf{E}|\xi|^{2p} + 1\big).&\end{aligned} \tag{5.5.108}$$

Proof. Itô's formula yields that for $t \in [0, T]$ and τ_N defined by eq. (5.5.107),

$$\begin{aligned}|\phi^\varepsilon_{n,h}(t \wedge \tau_N)|^2 = |P_n\xi|^2 &+ 2\sqrt{\varepsilon}\int_0^{t\wedge\tau_N} \big(\sigma_n(\phi^\varepsilon_{n,h}(s))dW_n(s), \phi^\varepsilon_{n,h}(s)\big)\\ &+ 2\int_0^{t\wedge\tau_N} \big\langle E(\phi^\varepsilon_{n,h}(s)), \phi^\varepsilon_{n,h}(s)\big\rangle ds\\ &+ 2\int_0^{t\wedge\tau_N} \big(\tilde{\sigma}_n(\phi^\varepsilon_{n,h}(s))h(s), \phi^\varepsilon_{n,h}(s)\big)\, ds\\ &+ \varepsilon\int_0^{t\wedge\tau_N} |\sigma_n(\phi^\varepsilon_{n,h}(s))\, P_n|^2_{L_Q}\, ds.\end{aligned} \tag{5.5.109}$$

Applying again Itô's formula to x^p for $p \ge 2$ and using Lemma 5.5.5 and with the convention $p(p-1)x^{p-2} = 0$ for $p = 1$ yields for $t \in [0, T]$,

$$
\begin{aligned}
|\phi^\varepsilon_{n,h}(t\wedge\tau_N)|^{2p} &+ 2p\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\left[\|u^\varepsilon_{n,h}(r)\|^2 + \|\theta^\varepsilon_{n,h}(r)\|^2\right] dr\\
&\le |P_n\xi|^{2p} + \sum_{1\le j\le 6} T_j(t),
\end{aligned}
\tag{5.5.110}
$$

where

$$
\begin{aligned}
T_1(t) &= 2p\int_0^{t\wedge\tau_N} |\langle\phi^\varepsilon_{n,h}, F(\phi^\varepsilon_{n,h})\rangle||\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
T_2(t) &= 2p\int_0^{t\wedge\tau_N} |\langle\phi^\varepsilon_{n,h}, R\rangle||\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
T_3(t) &= 2p\sqrt{\varepsilon}\left|\int_0^{t\wedge\tau_N}\left(\sigma_n(\phi^\varepsilon_{n,h}(r))\,dW_n(r), \phi^\varepsilon_{n,h}(r)\right)|\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\right|,\\
T_4(t) &= 2p\int_0^{t\wedge\tau_N} |(\tilde\sigma_n(\phi^\varepsilon_{n,h}(r))\,h(r), \phi^\varepsilon_{n,h}(r))|\,|\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
T_5(t) &= p\,\varepsilon\int_0^{t\wedge\tau_N} |\sigma_n(\phi^\varepsilon_{n,h}(r))\,P_n|^2_{L_Q}\,|\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
T_6(t) &= 2p(p-1)\varepsilon\int_0^{t\wedge\tau_N} |\Pi_n\,\sigma_n^*(\phi^\varepsilon_{n,h}(r))\,\phi^\varepsilon_{n,h}(r)|^2_{H_0}\,|\phi^\varepsilon_{n,h}(r)|^{2(p-2)}dr.
\end{aligned}
$$

The Hölder's inequality and Young's inequality imply that

$$
\begin{aligned}
T_1(t) &\le 2p\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\int_M\left(\left(\int_{-h}^0 |\partial_x\theta^\varepsilon_{n,h}(r)|dz\right)\cdot|u^\varepsilon_{n,h}(r)|\right)dM\,dr\\
&\le 2p\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}[h|\partial_x\theta^\varepsilon_{n,h}(r)||u^\varepsilon_{n,h}(r)|]dr\\
&\le \frac{1}{12}\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\|\phi^\varepsilon_{n,h}(r)\|^2 dr\\
&+ C_1\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2p}
\end{aligned}
\tag{5.5.111}
$$

and

$$
\begin{aligned}
T_2(t) &= 2p\int_0^{t\wedge\tau_N}[(u^\varepsilon_{n,h}(r), f) + (\theta^\varepsilon_{n,h}(r), q)]|\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr\\
&\le \frac{1}{12}\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\|\phi^\varepsilon_{n,h}(r)\|^2 dr\\
&\quad + C_2\sup_{0\le s\le t\wedge\tau_N}|\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\int_0^{t\wedge\tau_N}(|f(r)|^2 + |q(r)|^2)dr\\
&\le \frac{1}{12}\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\|\phi^\varepsilon_{n,h}(r)\|^2 dr + \frac{1}{2}\sup_{0\le s\le t\wedge\tau_N}|\phi^\varepsilon_{n,h}|^{2p}\\
&\quad + C_2\left(|f|^{2p}_{L^2(0,t\wedge\tau_N;H)} + |q|^{2p}_{L^2(0,t\wedge\tau_N;H)}\right).
\end{aligned}
\tag{5.5.112}
$$

Using the Cauchy–Schwarz inequality and ($\widetilde{\mathrm{A}}$.1), we get

$$\begin{aligned} T_4(t) \le 2p \int_0^{t\wedge\tau_N} \Big[\tilde{K}(1+ \|\phi^\varepsilon_{n,h}(r)\|^2)\Big]^{\frac{1}{2}} |h(r)|_0 \, |\phi^\varepsilon_{n,h}(r)|^{2p-1} dr \\ \le \frac{1}{12}\int_0^{t\wedge\tau_N} \|\phi^\varepsilon_{n,h}(r)\|^2 \, |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\, dr \\ + C_4 \int_0^{t\wedge\tau_N} |h(r)|_0^2 \, |\phi^\varepsilon_{n,h}(r)|^{2p}\, dr \\ + \frac{1}{12}\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)} dr. \end{aligned} \tag{5.5.113}$$

Using (A.2), we deduce that

$$\begin{aligned} T_5(t) + T_6(t) \le 2p^2 K\varepsilon \int_0^{t\wedge\tau_N} \|\phi^\varepsilon_{n,h}(r)\|^2 \, |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\, dr \\ + 2p^2 K\varepsilon \int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\, dr. \end{aligned} \tag{5.5.114}$$

Finally, the BDG inequality (see Ref. [75]), (A.2) and Schwarz's inequality yield that for $t \in [0, T]$ and $\delta_3 > 0$,

$$\begin{aligned} \mathbf{E}\Big(\sup_{0\le s\le t} |T_2(s)|\Big) &\le 6p\sqrt{\varepsilon}\,\mathbf{E}\Big\{\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(2p-1)} \, |\sigma_{n,h}(\phi^\varepsilon_{n,h}(r))\, P_n|^2_{L_Q}\, dr\Big\}^{\frac{1}{2}} \\ &\le \delta_3 \mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N} |\phi^\varepsilon_{n,h}(s)|^{2p}\Big) \\ &\quad + \frac{9p^2K\varepsilon}{\delta_3}\,\mathbf{E}\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(p-1)} dr \\ &\quad + \frac{9p^2K\varepsilon}{\delta_3}\,\mathbf{E}\int_0^{t\wedge\tau_N} \|\phi^\varepsilon_{n,h}(r)\|^2 \, |\phi^\varepsilon_{n,h}(r)|^{2(p-1)} dr. \end{aligned} \tag{5.5.115}$$

Consider the following property $I(i)$ for an integer $i \ge 0$:
I(i) There exists $\varepsilon_{0,i} := \varepsilon_{0,i}(K, \tilde{K}, T, M) > 0$ such that for $0 \le \varepsilon \le \varepsilon_{0,i}$

$$\sup_n \mathbf{E}\int_0^{t\wedge\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2i} dr \le C(i) := C(i, K, \tilde{K}, T, M) < +\infty.$$

The property $I(0)$ obviously holds with $\varepsilon_{0,0} = 1$ and $C(0) = T$. Assume that for some integer i with $1 \le i \le p$, the property $I(i-1)$ holds; we prove that $I(i)$ holds. Here we mainly use a version of Gronwall's lemma [62, 84].

By setting

$$\begin{aligned} \varphi_i(r) &= 2(C_1 + C_4|h(r)|_0^2), \\ Z &= 2\Big(\frac{1}{12} + 2i^2K\varepsilon\Big)\int_0^{\tau_N} |\phi^\varepsilon_{n,h}(r)|^{2(i-1)} dr + 2|\xi|^{2i} \end{aligned}$$

$$
\begin{aligned}
&\quad + 2C_2(|f|^{2i}_{L^2(0,T;H)} + |q|^{2i}_{L^2(0,T;H)}),\\
X(t) &= \sup_{0\le s\le t} |\phi^\varepsilon_{n,h}(s\wedge\tau_N)|^{2i},\\
Y(t) &= \int_0^{t\wedge\tau_N} \|\phi^\varepsilon_{n,h}(s)\|^2\, |\phi^\varepsilon_{n,h}(s)|^{2(i-1)}\, ds,\\
I(t) &= \sup_{0\le s\le t} |2i\sqrt{\varepsilon}\, \Big| \int_0^{t\wedge\tau_N} \big(\sigma_n(\phi^\varepsilon_{n,h}(r))\, dW_n(r), \phi^\varepsilon_{n,h}(r)\big)\, |\phi^\varepsilon_{n,h}(r)|^{2(i-1)} \Big|.
\end{aligned}
$$

Then $\int_0^T \varphi_i(s)ds \le C_i(M) := 2C_1T + C_4M$. Let $\alpha = 2(2p - \frac{1}{4} - 2p^2K\varepsilon)$, $\beta = \delta_3 = \frac{1}{2\left[1+C_i(M)e^{C_i(M)}\right]}$ and $\tilde{C} = \frac{9i^2K}{\delta_3}\mathbf{E}\int_0^{\tau_N} |\phi^\varepsilon_{n,h}(s)|^{2(i-1)}ds$. If we choose ε small enough and satisfying $\varepsilon \le \frac{2\delta_3^2(2i-\frac{1}{4})}{9i^2K+4i^2K\delta_3^2}$, then we have $y = \frac{9p^2K\varepsilon}{\delta_3} \le \alpha\beta$. Finally, letting $\varepsilon_{0,i} = \frac{2\delta_3^2(2i-\frac{1}{4})}{9i^2K+4i^2K\delta_3^2} \wedge \varepsilon_{0,i-1}$ and using Gronwall's lemma [62, 84], we yield **I(i)** is valid.

An induction argument shows that $I(p-1)$ holds, and hence the previous computations with $i = p$ yield that for $t = T$ and $0 \le \varepsilon \le \varepsilon_{0,p}$,

$$
\sup_n \mathbf{E}\Big(\sup_{0\le s\le \tau_N} |\phi^\varepsilon_{n,h}(s)|^{2p} + \int_0^{\tau_N} \|\phi^\varepsilon_{n,h}(s)\|^2\, \phi^\varepsilon_{n,h}(s)|^{2(p-1)}\, ds\Big) \le C(p, K, \tilde{K}, T, M).
$$

By the definition of $\tau_{n,h}$ and eq. (5.5.107), we know $\sup_{0\le s\le t\wedge\tau_N} |\phi_{n,h}(s)| \to \infty$, and $\tau_N \uparrow \tau_{n,h}$ on $\{\tau_{n,h} < T\}$, as $N \to \infty$. Hence by the above estimate, we have $\mathbb{P}(\tau_{n,h} < T) = 0$ for almost all ω, for $N(\omega)$ large enough and $\tau_{N(\omega)}(\omega) = T$. Thus we complete the proof of the proposition. ∎

Proposition 5.5.5. *There exists $\varepsilon_{0,p} := \varepsilon_{0,p}(K, \tilde{K}, \overline{K}, T, M)$ such that for $0 \le \varepsilon \le \varepsilon_{0,p}$ the following result holds. Let*

$$
h \in \mathcal{A}_M, \quad \partial_z f, \partial_z q \in L^{2p}(\Omega; L^2(0, T; H))
$$

and $\partial_z\xi \in L^{2p}(\Omega, H)$. Then we have

$$
\begin{aligned}
\sup_n \mathbf{E}\Big(&\sup_{0\le t\le T} |\partial_z\phi^\varepsilon_{n,h}(t)|^{2p} + \int_0^T \|\partial_z\phi^\varepsilon_{n,h}(s)\|^2\, |\partial_z\phi^\varepsilon_{n,h}(s)|^{2(p-1)}ds\Big)\\
&\le C(K, \tilde{K}, \overline{K}, T, M, |\partial_z f|_{L^{2p}(\Omega;L^2(0,T;H))},\\
&\quad |\partial_z q|_{L^{2p}(\Omega;L^2(0,T;H))})\big(\mathbf{E}|\partial_z\xi|^{2p} + 1\big).
\end{aligned}
\tag{5.5.116}
$$

Proof. Applying Itô's formula to $|\partial_z\phi^\varepsilon_{n,h}|^{2p}$ for $t \in [0, T]$ and τ_N defined by eq. (5.5.107), we get

$$
\begin{aligned}
|\partial_z\phi^\varepsilon_{n,h}(t\wedge\tau_N)|^{2p} &+ 2p\int_0^{t\wedge\tau_N} |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\left[\|\partial_z u^\varepsilon_{n,h}(r)\|^2 + \|\partial_z\theta^\varepsilon_{n,h}(r)\|^2\right] dr\\
&\le |P_n\partial_z\xi|^{2p} + \sum_{1\le j\le 6} J_j(t),
\end{aligned}
\tag{5.5.117}
$$

where

$$
\begin{aligned}
J_1(t) &= 2p \int_0^{t\wedge\tau_N} |\langle \partial_{zz}\phi^\varepsilon_{n,h}, B(\phi^\varepsilon_{n,h})\rangle||\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
J_2(t) &= 2p \int_0^{t\wedge\tau_N} |\langle \partial_{zz}\phi^\varepsilon_{n,h}, F(\phi^\varepsilon_{n,h})\rangle||\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
J_3(t) &= 2p \int_0^{t\wedge\tau_N} |\langle \partial_{zz}\phi^\varepsilon_{n,h}, R\rangle||\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
J_4(t) &= 2p\sqrt{\varepsilon}\left| \int_0^{t\wedge\tau_N} \big(\sigma_n(\phi^\varepsilon_{n,h}(r))\, dW_n(r), \partial_{zz}\phi^\varepsilon_{n,h}(r)\big)\ |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\right|,\\
J_5(t) &= 2p \int_0^{t\wedge\tau_N} |(\tilde{\sigma}_n(\phi^\varepsilon_{n,h}(r))\, h(r), \partial_{zz}\phi^\varepsilon_{n,h}(r))|\ |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
J_6(t) &= p\,\varepsilon \int_0^{t\wedge\tau_N} |\partial_z\sigma_n(\phi^\varepsilon_{n,h}(r))\, P_n|^2_{L_Q}\ |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr,\\
J_7(t) &= 2p(p-1)\varepsilon \times\\
&\quad \int_0^{t\wedge\tau_N} |\Pi_n\,(\partial_z\sigma_n)^*(\phi^\varepsilon_{n,h}(r))\ \partial_z\phi^\varepsilon_{n,h}(r)|^2_{H_0}\ |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-2)}dr.
\end{aligned}
$$

Note that by eq. (5.5.89), we have

$$
J_1(t) = 2p\int_0^{t\wedge\tau_N} |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\left|b_1\Big(u^\varepsilon_{n,h}(r), \theta^\varepsilon_{n,h}(r), \partial_{zz}\theta^\varepsilon_{n,h}(r)\Big)\right|dr.
$$

Integrating by parts and using Hölder's inequality, Ladyzhenskaya's inequality and Young's inequality, one obtains that

$$
\begin{aligned}
J_1(t) &\le 2p\int_0^{t\wedge\tau_N} |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\int_{\mathcal{M}} |\partial_z\theta^\varepsilon_{n,h}(r)||\partial_z u^\varepsilon_{n,h}(r)||\partial_x\theta^\varepsilon_{n,h}(r)|d\mathcal{M}dr\\
&\quad + 2p\int_0^{t\wedge\tau_N} |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\int_{\mathcal{M}} |\partial_z\theta^\varepsilon_{n,h}(r)|^2|\partial_x u^\varepsilon_{n,h}(r)|d\mathcal{M}dr\\
&\le 4p\int_0^{t\wedge\tau_N} |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}|\partial_z\phi^\varepsilon_{n,h}(r)|^2_{L^4}|\partial_x\phi^\varepsilon_{n,h}(r)|dr\\
&\le \frac{1}{12}\int_0^{t\wedge\tau_N} |\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}\|\partial_z\phi^\varepsilon_{n,h}(r)\|^2dr\\
&\quad + C_1\int_0^{t\wedge\tau_N} |\partial_z\phi^\varepsilon_{n,h}(r)|^{2p}\|\phi^\varepsilon_{n,h}(r)\|^2dr
\end{aligned}
\tag{5.5.118}
$$

and

$$
\begin{aligned}
J_2(t) &= 2p \int_0^{t\wedge\tau_N} |\langle \partial_{zz}u^\varepsilon_{n,h}, \int_{-h}^z \partial_x\theta^\varepsilon_{n,h}dz\rangle||\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr\\
&= 2p \int_0^{t\wedge\tau_N} |\langle \partial_z u^\varepsilon_{n,h}, \partial_x\theta^\varepsilon_{n,h}\rangle||\partial_z\phi^\varepsilon_{n,h}(r)|^{2(p-1)}dr
\end{aligned}
\tag{5.5.119}
$$

$$
\begin{aligned}
&\le 2p \int_0^{t\wedge\tau_N} \|\theta^\varepsilon_{n,h}\| |\partial_z \phi^\varepsilon_{n,h}(r)|^{2p-1} dr \\
&\le \frac{1}{4} \sup_{0\le s\le t\wedge\tau_N} |\partial_z \phi^\varepsilon_{n,h}(s)|^{2p} + C_2 \Big(\int_0^{t\wedge\tau_N} \|\theta^\varepsilon_{n,h}(r)\|^2 dr \Big)^{2p}.
\end{aligned}
$$

As we obtain estimate (5.5.112), we have

$$
\begin{aligned}
J_3 \le{}& \frac{1}{12} \int_0^{t\wedge\tau_N} |\partial_z \phi^\varepsilon_{n,h}(r)|^{2(p-1)} \|\partial_z \phi^\varepsilon_{n,h}(r)\|^2 dr + \frac{1}{4} \sup_{0\le s\le t\wedge\tau_N} |\partial_z \phi^\varepsilon_{n,h}|^{2p} \\
&+ C_3 \Big(|\partial_z f|^{2p}_{L^2(0,t\wedge\tau_N;H)} + |\partial_z q|^{2p}_{L^2(0,t\wedge\tau_N;H)} \Big).
\end{aligned} \tag{5.5.120}
$$

As in eq. (5.5.113), using the Cauchy–Schwarz inequality and (B.2), we get

$$
\begin{aligned}
J_5 \le{}& \frac{1}{12} \int_0^{t\wedge\tau_N} \|\partial_z \phi^\varepsilon_{n,h}(r)\|^2 \, |\partial_z \phi^\varepsilon_{n,h}(r)|^{2(p-1)} \, dr \\
&+ C_5 \int_0^{t\wedge\tau_N} |h(r)|_0^2 \, |\partial_z \phi^\varepsilon_{n,h}(r)|^{2p} \, dr \\
&+ \frac{1}{12} \int_0^{t\wedge\tau_N} |\partial_z \phi^\varepsilon_{n,h}(r)|^{2(p-1)} dr.
\end{aligned} \tag{5.5.121}
$$

Using again (B.1), we deduce that

$$
\begin{aligned}
J_6(t) + J_7(t) \le{}& 2p^2 \overline{K} \varepsilon \int_0^{t\wedge\tau_N} \|\partial_z \phi^\varepsilon_{n,h}(r)\|^2 \, |\partial_z \phi^\varepsilon_{n,h}(r)|^{2(p-1)} \, dr \\
&+ 2p^2 \overline{K} \varepsilon \int_0^{t\wedge\tau_N} |\partial_z \phi^\varepsilon_{n,h}(r)|^{2(p-1)} \, dr.
\end{aligned} \tag{5.5.122}
$$

Finally, the BDG inequality, (B.1) and Schwarz's inequality yield that for $t \in [0, T]$ and $\delta_4 > 0$,

$$
\begin{aligned}
&\mathbf{E}\Big(\sup_{0\le s\le t} |J_4(s)| \Big) \\
&\le 6p\sqrt{\varepsilon}\, \mathbf{E}\Big\{ \int_0^{t\wedge\tau_N} |\partial_z \phi^\varepsilon_{n,h}(r)|^{2(2p-1)} \, |\partial_z \sigma_{n,h}(\phi^\varepsilon_{n,h}(r)) \, P_n|^2_{L_Q} \, dr \Big\}^{\frac{1}{2}} \\
&\le \delta_4 \mathbf{E}\Big(\sup_{0\le s\le t\wedge\tau_N} |\partial_z \phi^\varepsilon_{n,h}(s)|^{2p} \Big) + \frac{9p^2\overline{K}\varepsilon}{\delta_4} \, \mathbf{E} \int_0^{t\wedge\tau_N} |\partial_z \phi^\varepsilon_{n,h}(r)|^{2(p-1)} dr \\
&\quad + \frac{9p^2\overline{K}\varepsilon}{\delta_4} \, \mathbf{E} \int_0^{t\wedge\tau_N} \|\partial_z \phi^\varepsilon_{n,h}(r)\|^2 \, |\partial_z \phi^\varepsilon_{n,h}(r)|^{2(p-1)} dr.
\end{aligned} \tag{5.5.123}
$$

Using the similar steps as those in Proposition 5.5.4, due to eqs (5.5.118)–(5.5.123), we get eq. (5.5.116). ∎

Due to Ladyzhenskaya's inequality for 2D domain, we now have the following bound in $L^4(\mathcal{M})$.

Proposition 5.5.6. *Let $h \in \mathcal{A}_M$, $\xi \in L^4(\Omega, H)$ and $\varepsilon_{2,0}$ be defined as in Proposition 5.5.4 with $p = 2$. Then there exists a constant*

$$C_2 := C_2(K, \tilde{K}, T, M, |f|_{L^4(\Omega;L^2(0,T;H))}, |q|_{L^4(\Omega;L^2(0,T;H))})$$

such that

$$\sup_n \mathbf{E} \int_0^T |\phi^\varepsilon_{n,h}(s)|^4_{L^4} ds \le C_2(1 + \mathbf{E}|\xi|^4). \tag{5.5.124}$$

The following result is a consequence of Itô's formula.

Lemma 5.5.7. *Let $\rho' : [0, T] \times \Omega \to [0 + \infty)$ be adapted such that for almost every ω $t \to \rho'(t, \omega) \in L^1([0, T])$ and for $t \in [0, T]$, set $\rho(t) = \int_0^t \rho'(s)\, ds$. For $i = 1, 2$, let σ_i satisfy assumptions (A) and (B), $\bar{\sigma}_i \in C([0, T] \times H, L^2_Q)$ and let $\bar{\sigma}$ satisfy assumptions $\tilde{A}$ and B. Let E satisfy condition* (5.5.105), *$h_\varepsilon \in \mathcal{A}_M$ and $\phi_i \in L^2([0, T], V) \cap L^\infty([0, T], H)$ a.s. with $\phi_i(0) = \xi \in L^4(\Omega, H)$, for ξ $\mathcal{F}_0$ measurable and satisfy the equation*

$$d\phi_i(t) = R(\phi_i(t))dt + \sqrt{\varepsilon}\sigma_i(t, \phi_i(t))\, dW(t) + \big(\bar{\sigma}(t, \phi_i(t)) + \bar{\sigma}_i(t, \phi_i(t))\big)\, h_\varepsilon(t)\, dt. \tag{5.5.125}$$

Let $\Phi = \phi_1 - \phi_2$, then for every $t \in [0, T]$,

$$\begin{aligned} e^{-\rho(t)}\, |\Phi(t)|^2 \le \int_0^t e^{-\rho(s)} \Big\{ &-\frac{1}{2}\, \|\Phi(s)\|^2 + \varepsilon\big|\sigma_1(s, \phi_1(s)) - \sigma_2(s, \phi_2(s))\big|^2_{L^2_Q} \\ &+ |\Phi(s)|^2 \Big[-\rho'(s) \\ &\quad + 2C(1 + |\partial_z \phi_2(s)|^4) + C\, \|\phi_2(s)\|^2 + C\, |h_\varepsilon(s)|^2_0 \Big]\Big\}\, ds \\ &+ 2\int_0^t e^{-\rho(s)} \big(\bar{\sigma}_1(s) - \bar{\sigma}_2(s)\, ,\ \Phi(s)\big)\, ds + I(t), \end{aligned} \tag{5.5.126}$$

where $I(t) = 2\sqrt{\varepsilon}\, \int_0^t e^{-\rho(s)}\Big(\big[\sigma_1(s, \phi_1(s)) - \sigma_2(s, \phi_2(s))\big]\, dW(s)\, ,\ \Phi(s)\Big)$.

Proof. Itô's formula, eq. (5.5.105) and conditions ($\tilde{A}$.2) imply that for $t \in [0, T]$,

$$\begin{aligned} e^{-\rho(t)}\, |\Phi(t)|^2 = \int_0^t e^{-\rho(s)} \Big\{ &-\rho'(s)|\Phi(s)|^2 + \varepsilon\big|\sigma_1(s, \phi_1(s)) - \sigma_2(s, \phi_2(s))\big|^2_{L_Q} \\ &+ 2\big\langle E(\phi_1(s)) - E(\phi_2(s))\, ,\ \Phi(s)\big\rangle \\ &+ 2\big(\bar{\sigma}(s, \phi_1(s))h_\varepsilon(s) - \bar{\sigma}(s, \phi_2(s))h_\varepsilon(s)\, ,\ \Phi(s)\big)\Big\}\, ds \end{aligned}$$

$$
\begin{aligned}
&+ \int_0^t e^{-\rho(s)} 2\big(\big[\bar{\sigma}_1(s) - \bar{\sigma}_2(s)\big] h_\varepsilon(s)\,,\ \Phi(s)\big)\, ds + I(t) \\
\le& \int_0^t e^{-\rho(s)} \Big\{ -\rho'(s)\, |\Phi(s)|^2 + \varepsilon \big|\sigma_1(s,\phi_1(s)) - \sigma_2(s,\phi_2(s))\big|_{L_Q}^2 - \ \|\Phi(s)\|^2 \\
&+ 2C|\Phi(s)|\ \|\Phi(s)\|\ \|\phi_2(s)\| + 2C(1 + |\partial_z \phi_2(s)|^4)|\Phi(s)|^2 \\
&+ 2C\sqrt{\tilde{L}}\|\Phi(s)\|\ |h_\varepsilon(s)|_0\ |\Phi(s)| \Big\} ds \\
&+ \int_0^t e^{-\rho(s)} 2\big(\bar{\sigma}_1(s) - \bar{\sigma}_2(s)\,,\ \Phi(s)\big)\, ds + I(t).
\end{aligned}
$$

The inequalities

$$
2C|\Phi(s)|\ \|\Phi(s)\|\ \|\phi_2(s)\| \le \frac{1}{4}\|\Phi(s)\|^2 + C\|\phi_2(s)\|^2|\Phi(s)|^2
$$

and

$$
2C\sqrt{\tilde{L}}\|\Phi(s)\|\ |h_\varepsilon(s)|_0\ |\Phi(s)| \le \frac{1}{4}\|\Phi(s)\|^2 + C|h_\varepsilon(s)|_0^2\ |\Phi(s)|^2
$$

conclude the proof of eq. (5.5.126).

Proof of Theorem 5.5.2. As in Ref. [84], due to monotonicity property (5.5.105), priori estimates (5.5.108), (5.5.116) and Lemma 5.5.7, we obtain Theorem 5.5.1. ■

5.5.2.3 Large deviations for stochastic primitive equation

Since primitive equation is a large-scale model, one may neglect the effect of small scale and intermediate scale in its modeling. One may consider this effect by adding small noise in the equations. Large deviations theory concerns with the study of precise asymptotics governing the decay rate of probabilities of rare events. A classical area of large deviations is the Wentzell–Freidlin theory that deals with path probability asymptotics for small noise stochastic dynamical systems. More exactly, we will put a bound on the probability that the random perturbed trajectory goes very far from the unperturbed trajectory, and see the rate at which this probability goes to zero as the noise shrinks ($\varepsilon \to 0$). We consider large deviations via a weak convergence approach (originated with Budhiraja and Dupuis [42, 43], Sritharan and Sundar [235], Duan and Millet [84], among others). In this section, the idea of the proof for large deviations is the same as in Ref. [84], here we only give the outline of the proof. First, we recall some classical definitions for large deviations.

Definition 5.5.4. *The random family $\{\phi^\varepsilon\}$ is said to satisfy an LDP on X with the good rate function I if the following conditions hold:*

I is a good rate function. The function $I : X \to [0,\infty]$ is such that for each $M \in [0,\infty[$ the level set $\{\phi \in X : I(\phi) \le M\}$ is a compact subset of X.
For $A \in \mathcal{B}(X)$, set $I(A) = \inf_{\phi \in A} I(\phi)$.

Large deviation upper bound. For each closed subset F of X:

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(\phi^\varepsilon \in F) \le -I(F).$$

Large deviation lower bound. For each open subset G of X:

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(\phi^\varepsilon \in G) \ge -I(G).$$

To establish the LDP, we need to strengthen the hypothesis on the growth condition and Lipschitz property of σ (and $\tilde{\sigma}$) as follows:

Assumption A. There exist positive constants K and L such that

(A.4) $|\sigma(t,\phi)|^2_{L_Q} \le K(1+|\phi|^2) \quad \forall t \in [0,T], \forall \phi \in V,$

(A.5) $|\sigma(t,\phi) - \sigma(t,\psi)|^2_{L_Q} \le L\,|\phi-\psi|^2 \quad \forall t \in [0,T], \forall \phi,\psi \in V.$

The following theorem is the main result of this section.

Theorem 5.5.2. *Let σ do not depend on time and satisfy (A.1), (B), (A.4) and (A.5), ϕ^ε be the solution of the stochastic primitive eq* (5.5.93). *Then $\{\phi^\varepsilon\}$ satisfies the LDP in $C([0,T];H) \cap L^2((0,T);V)$, with the good rate function*

$$I_\xi(\psi) = \inf_{\{h\in L^2(0,T;H_0):\ \psi=\mathcal{G}^0(\int_0^\cdot h(s)ds)\}} \left\{\frac{1}{2}\int_0^T |h(s)|_0^2\, ds\right\}. \tag{5.5.127}$$

Here the infimum of an empty set is taken as infinity.

The proof of the LDP will use the following technical lemma, which studies time increments of the solution to the stochastic control equation. For any integer $k = 0,\cdots,2^n-1$, and $s \in [kT2^{-n},(k+1)T2^{-n}[$, set $\underline{s}_n = kT2^{-n}$ and $\bar{s}_n = (k+1)T2^n$. Given $N > 0$, $h \in \mathcal{A}_M$, $\varepsilon \ge 0$ small enough, let ϕ_h^ε denote the solution to eq. (5.5.99) given by Theorem 5.5.1, and for $t \in [0,T]$, let

$$G_N(t) = \Big\{\omega : \Big(\sup_{0\le s\le t} |\phi_h^\varepsilon(s)(\omega)|^2\Big) \vee \Big(\int_0^t \|\phi_h^\varepsilon(s)(\omega)\|^2 ds\Big) \\ \vee \Big(\sup_{0\le s\le t} |\partial_z \phi_h^\varepsilon(s)(\omega)|^2 \le N\Big)\Big\}.$$

Lemma 5.5.8. *Let $M,N > 0$, σ and $\tilde{\sigma}$ satisfy Assumptions (A.1), (B), (A.4) and (A.5), $\partial_z\xi,\ \xi \in L^4(H)$. Then there exists a positive constant*

$$C := C\Big(K, L, |f|_{L^4(\Omega;L^2(0,T;H))}, |q|_{L^4(\Omega;L^2(0,T;H))}, \\ |\partial_z f|_{L^4(\Omega;L^2(0,T;H))}, |\partial_z q|_{L^4(\Omega;L^2(0,T;H))}, T, M, N, \varepsilon_0\Big)$$

such that for any $h \in \mathcal{A}_M$, $\varepsilon \in [0, \varepsilon_0]$,

$$I_n(h,\varepsilon) := \boldsymbol{E}\Big[1_{G_N(T)} \int_0^T |\phi_h^\varepsilon(s) - \phi_h^\varepsilon(\bar{s}_n)|^2\, ds\Big] \le C\, 2^{-\frac{n}{2}}. \tag{5.5.128}$$

Proof. The proof is close to that of Lemma 4.2 in Ref. [84]. ■

Now we return to the setting of Theorem 5.5.2. Let ε_0 be defined as in Theorem 5.5.1 and $(h_\varepsilon), (0 < \varepsilon \le \varepsilon_0)$, be a family of random elements taking values in $\mathcal{A}_M$. Let $\phi_{h_\varepsilon}^\varepsilon$ be the solution of the corresponding stochastic control equation with initial condition $\phi_{h_\varepsilon}^\varepsilon(0) = \xi \in H$:

$$\begin{aligned} d\phi_{h_\varepsilon}^\varepsilon &+ [A\phi_{h_\varepsilon}^\varepsilon + B(\phi_{h_\varepsilon}^\varepsilon) + F(\phi_{h_\varepsilon}^\varepsilon)]dt \\ &= Rdt + \sigma(\phi_{h_\varepsilon}^\varepsilon)h_\varepsilon dt + \sqrt{\varepsilon}\, \sigma(\phi_{h_\varepsilon}^\varepsilon)dW(t). \end{aligned} \tag{5.5.129}$$

Note that $\phi_{h_\varepsilon}^\varepsilon = \mathcal{G}^\varepsilon\Big(\sqrt{\varepsilon}\big(W_\cdot + \frac{1}{\sqrt{\varepsilon}}\int_0^\cdot h_\varepsilon(s)ds\big)\Big)$ due to the uniqueness of the solution.

For all ω and $h \in L^2([0,T], H_0)$, let ϕ_h be the solution of the corresponding control equation with initial condition $\phi_h(0) = \xi(\omega)$:

$$d\phi_h + [A\phi_h + B(\phi_h) + F(\phi_h)]dt = Rdt + \sigma(\phi_h)hdt. \tag{5.5.130}$$

Note that here we may assume that h and ξ are random, but ϕ_h may be defined pointwise by eq. (5.5.130).

Let $\mathcal{C}_0 = \{\int_0^\cdot h(s)ds \;:\; h \in L^2([0,T], H_0)\} \subset C([0,T], H_0)$. For every $\omega \in \Omega$, define $\mathcal{G}^0 \;:\; C([0,T], H_0) \to X$ by $\mathcal{G}^0(g)(\omega) = \phi_h(\omega)$ for $g = \int_0^\cdot h(s)ds \in \mathcal{C}_0$ and $\mathcal{G}^0(g) = 0$ otherwise.

Proposition 5.5.7 (Weak convergence). *Suppose that σ does not depend on time and satisfies Assumptions (A.1), (B), (A.4) and (A.5). Let $\xi \in Y$, be $\mathcal{F}_0$ measurable such that $E|\xi|_H^4 < +\infty$, $E|\partial_z\xi|_H^4 < +\infty$ and let h_ε converge to h in distribution as random elements taking values in $\mathcal{A}_M$. (Note that here $\mathcal{A}_M$ is endowed with the weak topology induced by the norm* (5.5.100)*). Then as $\varepsilon \to 0$, $\phi_{h_\varepsilon}^\varepsilon$ converges in distribution to ϕ_h in $X = C([0,T];H) \cap L^2((0,T);V)$ endowed with the norm (5.5.100). That is, $\mathcal{G}^\varepsilon\Big(\sqrt{\varepsilon}\big(W_\cdot + \frac{1}{\sqrt{\varepsilon}}\int_0^\cdot h_\varepsilon(s)ds\big)\Big)$ converges in distribution to $\mathcal{G}^0\big(\int_0^\cdot h(s)ds\big)$ in X, as $\varepsilon \to 0$.*

The proof can be obtained by delicate estimates and by the method in Ref. [84].

The following compactness result is the second ingredient that allows us to transfer the LDP from $\sqrt{\varepsilon}W$ to u^ε. Its proof is similar to that of Proposition 5.5.7 and easier and hence will be sketched (see also Ref. [84], Proposition 4.4).

Proposition 5.5.8 (Compactness). *Let M be any fixed finite positive number and let $\xi \in Y$ be deterministic. Define*

$$K_M = \{\phi_h \in C([0,T];H) \cap L^2((0,T);V) : h \in S_M\},$$

where ϕ_h is the unique solution of the deterministic control equation:

$$\begin{aligned} &d\phi_h(t) + \big[A\phi_h(t) + B(\phi_h(t)) + R\phi_h(t)\big]dt = F\phi_h(t)dt + \sigma(\phi_h(t))h(t)dt, \\ &\phi_h(0) = \xi, \end{aligned} \tag{5.5.131}$$

and σ does not depend on time and satisfies (A.1), (B), (A.4) and (A.5). Then K_M is a compact subset of X.

Proof of Theorem 5.5.3. Propositions 5.5.8 and 5.5.7 imply that $\{\phi^\varepsilon\}$ satisfies the Laplace principle, which is equivalent to the LDP in $X = C([0,T],H) \cap L^2((0,T),V)$ with the above-mentioned rate function (see Theorem 4.4 in Ref. [42] or Theorem 5 in Ref. [43]). ∎

Bibliography

[1] R.A. Adams and J.J.F. Fournier, Sobolev spaces, 2nd edition, Elsevier/Academic Press, Amsterdam, 2003.

[2] S. Albeverio and A.B. Cruzeiro, Global flow and invariant (Gibbs) measures for Euler and Navier-Stokes two dimensional fluids, Commun. Math. Phys., **129**(1990), 431–444.

[3] S. Albeverio, J.L. Wu and T.S. Zhang, Parabolic SPDEs driven by Poisson white noise, Stochastic Process. Appl., **74**(1998), 21–36.

[4] F. Alouges and A. Soyeur, On global weak solutions for Landau-Lifshitz equations: existence and nonuniqueness, Nonlinear Anal. TAM, **18**(1992), 1071–1084.

[5] L. Arnold, Random dynamical system, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.

[6] M. Artola, Sur un tháorème d'interpolation, J. Math. Anal. Appl., **41**(1973), 148–163.

[7] D. Applebaum, Levy processes and stochastic calculus, Cambridge University Press, Cambridge, 2004.

[8] O. Bang, P.L. Christiansen, F. If and K.O. Rasmussen, Temperature effects in a nonlinear model of monolayer Scheibe aggregates, Phys. Rev. E, **49**(1994), 4627–4636.

[9] O. Bang, P.L. Christiansen, F. If and K.O. Rasmussen, White noise in the two-dimensional nonlinear Schrödinger equation, Appl. Anal., **57**(1995), 3–15.

[10] E.S. Benilov, On the surface waves in a shallow channel with an uneven bottom, Stud. Appl. Math., **87**(1992), 1–14.

[11] A.F. Bennett and P.E. Kloeden, The dissipative quasi-geostrophic equations: existence and uniqueness of strong solutions, Mathematika, **27**(1980), 287–311.

[12] A. Bensoussan and R. Temam, Equations Stochastiques du type Navier-Stokes, J. Funct. Anal., **13**(1973), 195–222.

[13] J. Bergh and J. Lofström, Interpolation spaces, Springer Verlag, Berlin, 1976.

[14] J. Bertoin, Levy Processes, Cambridge Tracts in Mathematics, Vol. **121**, Cambridge University Press, Cambridge, 1996.

[15] J.M. Bismut, An introductory approach to duality in optimal stochastic control, SIAM Rev., **20**(1978), 62–78.

[16] F. Black and M. Scholes, The pricing of options and corporate liabilities, J. Political Economy, **81**(1973), 637–659.

[17] O. Blasco, Interpolation between $H^1_{B_0}$ and $L^p_{B_1}$, Studia Math., **92**(1989), 205–210.

[18] J.L. Bona and R. Scott, Solutions of the Korteweg-de Vries equation in fractional order Sobolev spaces, Duke Math. J., **43**(1976), 87–99.

[19] J.L. Bona and R. Smith, The initial-value problem for the Korteweg-de Vries equation, Philos. Trans. Roy. Soc. London, Ser. A, **278**(1975), 555–604.

[20] J.L. Bona and B.Y. Zhang, The initial value problem for the forced KdV equation, Proc. Soc. Edinburgh A, **126**(1996), 571–598.

[21] A. de Bouard and A. Debussche, On the stochastic Korteweg-de Vries equation, J. Funct. Anal., **154**(1998), 215–251.

[22] A. de Bouard and A. Debussche, A stochastic nonlinear Schrödinger equation with multiplicative noise, Commun. Math. Phys., **205**(1999), 161–181.

[23] A. de Bouard and A. Debussche, On the effect of a noise on the solutions of the focusing supercritical nonlinear Schrödinger equation, Probab. Theory Related Fields, **123**(2002), 76–96.

[24] A. de Bouard and A. Debussche, The stochastic nonlinear Schrödinger equation in H^1, Stochastic Anal. Appl., **21**(2003), 97–126.

[25] A. de Bouard and A. Debussche, A semi-discrete scheme for the stochastic nonlinear Schrödinger equation, Numer. Math., **96**(2004), 733–770.

[26] A. de Bouard and A. Debussche, Blow-up for the stochastic nonlinear Schrodinger equation with multiplicative noise, Ann. Probab., **33**(2005), 1078–1110.

[27] A. de Bouard, A. Debussche and Y. Tsutsumi, White noise driven Korteweg de-Vries equation, J. Funct. Anal., **169**(1999), 532–558.

[28] A. de Bouard and A. Debussche, A Stochastic nonlinear Schrodinger equation with multiplicative noise. Commun. Math. Phys., **205**(1999), 161–181.

[29] J. Bourgain, Fourier restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, Part II, Geom. Funct. Anal., **3**(1993), 209–262.

[30] A. Debussche and L. Di Menza, Numerical simulation of focusing stochastic nonlinear Schrödinger equation, Phys. D, **162**(2002), 131–154.

[31] H.R. Brand and R.J. Deissler, Interaction of localized solutions for subcritical bifurcations, Phys, Rev. Lett., **63**(1992), 2801–2804.

[32] J. Brannan, J. Duan and T. Wanner, Dissipative quasi-geostrophic dynamics under random forcing, J. Math. Anal. Anal., **228**(1998), 221–233.

[33] W.F. Brown, Thermal fluctuations of a single-domain particle, Phys. Rev., **130**(1963), 1677–1686.

[34] Z. Brzezniak, On stochastic convolution in Banach spaces and applications, Stochastics Stochastic Rep., **61**(1997), 245–295.

[35] Z. Brzezniak and K.D. Elworthy, Stochastic differential equations on Banach manifolds, Meth. Funct. Anal. Top., **6**(2000), 43–84.

[36] Z. Brzezniak and Y. Li, Asymptotic compactness and absorbing sets for 2d stochastic Navier-Stokes equations on some unbounded domains, Trans. Amer. Math. Soc., **358**(2006), 5587–5629.

[37] Z. Brzezniak, B. Maslowski and J. Seidler, Stochastic nonlinear beam equations, Probab. Theory Related Fields, **132**(2005), 119–149.

[38] Z. Brzezniak and S. Peszat, Space-time continuous solutions to SPDE's driven by a homogeneous Wiener process, Studia Math., **137**(1999), 261–299.

[39] Z. Brzezniak, M. Capinski and F. Flandoli, Pathwise global attractors for stationary random dynamical systems, Probab. Theory Related Fields, **95**(1993), 87–102.

[40] Z. Brzezniak, On stochastic convolution in Banach spaces and applications, Stochastics Stochastic Rep., **61**(1997), 245–295.

[41] Z. Brzezniak and S. Peszat, Space-time continuous solutions to SPDE's derived by a homogeneous Wiener process. Studia Math., **137**(1999), 261–299.

[42] A. Budhiraja and P. Dupuis, A variational representation for positive functionals of infinite dimensional Brownian motion, Probab. Math. Stat., **20**(2000), 39–61.

[43] A. Budhiraja, P. Dupuis and V. Maroulas, Large deviations for infinite dimensional stochastic dynamical systems, Ann. Probab., **36**(2008), 1390–1420.

[44] C. Cao and E. S. Titi, Global well-posedness of the three-dimensional primitive equations of large scale ocean and atmosphere dynamics, Ann. Math., **166**(2007), 245–267.

[45] G. Carbou and P. Fabrie, Regular Solutions for Landau-Lifschitz equations in a bounded domain, Differ. Integr. Equ., **14**(2001), 213–229.

[46] T. Caraballo, H. Crauel, J.A. Langa and J.C. Robinson, The effect of noise on the Chafee-Infante equation: a nonlinear case study, Proc. Amer. Math. Soc., **135**(2007), 373–382.

[47] T. Caraballo and J.A. Langa, Stability and random attractors for a reaction-diffusion equation with multiplicative noise, Discrete Contin. Dyn. Syst., **6**(2000), 876–892.

[48] T. Caraballo, J.A. Langa and J.C. Robinson, Upper semicontinuity of attractors for random perturbations of dynamical systems, Commun. Partial Differ. Equ., **23**(1998), 1557–1581.

[49] T. Caraballo, J.A. Langa and J.C. Robinson, A stochastic pitchfork bifurcation in a reaction-diffusion, Proc. R. Soc. London A, **457**(2001), 2041–2061.

[50] R. Carmona and D. Nualart, Random non-linear wave equations: Smoothness of the solutions, Probab. Theory Related Fields, **79**(1988), 469–508.

[51] C. Castaing and M. Valadier, Convex analysis and measurable multifunctions, Lecture Notes in Mathematics **580**, Springer-Verlag, Berlin, 1977.

[52] T. Cazenave, An introduction to nonlinear Schrodinger equations, Texts de Métodos Matématicos, 26, Instituto de Matématica-UFRJ Rio de Janeiro, Brazil, 1993.

[53] S. Cerrai and M. Rockner, Large deviations for stochastic reaction-diffusion systems with multiplicative noise and non-Lipschitz reaction term, Ann. Probab., **32**(2004), 1100–1139.

[54] F. Chenal and A. Millet, Uniform large deviations for parabolic SPDEs and applications, Stochastic Process. Appl., **72**(1997), 161–186.

[55] H.Y. Chang, C. Lien and S. Sukarto, et al, Propagation of ion-acoustic solitons in a non-quiescent plasma. Plasma Phys. Controlled Fusion, **28**(1986), 675–681.

[56] J.G. Charney and N.A. Philips, Numerical integration of the quasi-geostrophic equations for barotropic simple baroclinic flows, J. Meteor., **10**(1953), 71–99.

[57] D. Cheban, P.E. Kloeden and B. Schmalfuss, The relationship between pullback, forwards and global attractors of nonautonomous dynamical systems, Nonlinear Dyn. Syst. Theory, **2**(2002), 9–28.

[58] Y. Chen, H.J. Gao and B.L. Guo, Well-posedness for stochastic Camassa-Holm equation, J. Differ. Equ., **253**(2012), 2353–2379.

[59] Y.M. Chen, The weak solutions to the evolution problems of harmonic maps, Math. Z., **201**(1989), 69–74.

[60] H. Choi, R. Temam, P. Moin and J. Kim, Feedback control for unsteady flow and its application to Burgers equation, J. Fluid Mech., **253**(1993), 509–543.

[61] P.L. Christiansen, K.O. Rasmussen, O. Bang and Yu.B. Gaididei, The temperature-dependent collapse regime in a nonlinear dynamical model of Scheibe aggregates, Phys. D, **87**(1995), 321–324.

[62] I. Chueshov and A. Millet, Stochastic 2D hydrodynamical type systems: well posedness and large deviations, Appl. Math. Optim., **61**(2010), 379–420.

[63] Z. Ciesielski, Holder condition for realizations of Gaussian processes, Trans. Amer. Math. Soc., **99**(1961), 403–413.

[64] J. Colliander, M. Keel, G. Staffilani, H. Takaoka and T. Tao, Global well-posedness for KdV in Sobolev spaces of negative index, Electron. J. Differ. Equ., **26**(2001), 1–7.

[65] J.C. Cox and M. Rubinstein, Options Markets, Prentice-Hall, Englewood Cliffs, N.J., 1985.

[66] H. Crauel, A. Debussche and F. Flandoli, Random attractors, J. Dynam. Differ. Equ., **9**(1997), 307–341.

[67] H. Crauel and F. Flandoli, Hausdorff dimension of invariant set for random dynamical systems, J. Dynam. Differ. Equ., **10**(1998), 449–474.

[68] H. Crauel and F. Flandoli, Attractors for random dynamical systems, Prob. Theory Related Fields, **100**(1994), 365–393.

[69] G. Da Prato and A. Debussche, Ergodivity for the 3D stochastic Navier-Stokes equations, J. Math. Pures Appl., **82**(2003), 877–947.

[70] G. Da Prato, A. Debusschde and R. Temam, Stochastic Burgers equation, Nonlin. Diff. Equ. Appl., **1**(1994), 389–402.

[71] G. Da Prato and D. Gatarek, Stochastic Burgers equation with correlated noise, Stochastics and Stochastic Reports, **52**(1995), 29–41.

[72] G. Da Prato, S. Kwapien and J. Zabczyk, Regularity of solutions of linear stochastic equations in Hilbert spaces, Stochastics, **23**(1987), 1–23.

[73] G. Da Prato, K.D. Elworthy and J. Zabczyk, Strong Feller property for stochastic semilinear equations, Stochastic Anal. Appl., **13**(1995), 35–46.

[74] G. Da Prato and M. Röchner, Weak solutions to stochastic porous media equations, J. Evol. Equ., **4**(2004), 249–271.
[75] G. Da Prato and J. Zabczyk, Stochastic equations in infinite dimensions, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1992.
[76] G. Da Prato and J. Zabczyk, Ergodicity for infinite dimensional systems, Cambridge University Press, Cambridge, 1996.
[77] A. Debussche, Hausdorff dimension of a random invariant set, J. Math. Pures Appl., **77**(1998), 967–988.
[78] A. Debussche, N. Glatt-Holtz and R. Temam, Local martingale and pathwise solutions for an abstract fluids model, Physica D, **240**(2011), 1123–1144.
[79] A. Debussche and C. Odasso, Markov solutions for the 3D stochastic Navier-Stokes equations with state dependent noise, J. Evol. Equ., **6**(2006), 305–324.
[80] A. Desimone, R.V. Kohn, S. Müller and F. Otto, A reduced theory for thin-film micromagnetics, Commun. Pure Appl. Math., **55**(2002), 1408–1460.
[81] S. Ding and B. Guo, Hausdorff measure of the singular set of Landau-Lifshitz equations with a nonlocal term, Commun. Math. Phys., **250**(2004), 95–117.
[82] Z. Dong, L. Xu and X. Zhang, Invariance measures of stochastic 2D Navier-Stokes equations driven by α-stable processes, Elect. Commun. in Probab., **16**(2011), 678–688.
[83] J. Duan, P.E. Kloeden and B. Schmalfuss, Exponential stability of the quasi-geostrophic equation under random perturbations, Prog. Probab., **49**(2001), 241–265.
[84] J. Duan and A. Millet, Large deviations for the Boussinesq equations under random influences, Stochastic Process. Appl., **119**(2009), 2052–2081.
[85] D. Duffie and L. Epstein, Asset pricing with stochastic differential utilities, Review of Financial Studies, **5**(1992), 411–436.
[86] W.N. E and J.C. Mattingly, Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation, Commun. Pure Appl. Math., **54**(2001), 1386–1402.
[87] N. El Karoui, S. Peng and M.C. Quenez, Backward stochastic differential equations in Finance, Math. Fin., **71**(1997), 1–71.
[88] N. El Karoui and M.C. Quenez, Dynamic programming and pricing of contingent claims in an incomplete market, SIAM J. Control Optim., **33**(1995), 29–66.
[89] L.C. Evans, An introduction to stochastic differential equations, Version 1.2, available online: http://math.berkeley.edu/ evans/SDE.course.pdf.
[90] B. Ewald, M. Petcu and R. Temam, Stochastic solutions of the two-dimensional primitive equations of the ocean and atmosphere with an additive noise, Anal. Appl., **5**(2007), 183–198.
[91] H.J. Engelbert, On the theorem of T.Yamada and S.Watanabe, Stochastics and Stochastics Rep., **36**(1991), 205–216.
[92] E.B. Fabes, B.F. Jones and N.M. Rivière, The initial value problem for the Navier-Stokes equations with data in L^p, Arch. Rational Mech. Anal., **45**(1972), 222–240.
[93] G.E. Falkovich, I. Kolokolov, V. Lebedev and S.K. Turitsyn, Statistics of soliton-bearing system with additive noise, Phys. Rev. E, **63**(2001), 025601.
[94] G. Fibich and G. Papanicolaou, Self-focusing in the perturbed and unperturbed nonlinear Schrödinger equation in critical dimension, SIAM J. Appl. Math., **60**(1999), 183–240.
[95] F. Flandoli, Dissipativity and invariant measures for stochastic Navier-Stokes equations, Nonlin. Differ. Equ. Appl., **1**(1994), 403–423.
[96] F. Flandoli and D. Gatarek, Martingale and Stationary solutions for Stochastic Navier-Stokes equations, Probab. Theory Related Fields, **102**(1995), 367–391.
[97] F. Flandoli and B. Maslowski, Ergodicity of the 2-D Navier-Stokes equation under random perturbations, Commun. Math. Phys., **172**(1995), 119–141.

[98] F. Flandoli and B. Schmalfuss, Random attractors for the 3D stochastic Navier-Stokes equation with multiplicative noise, Stochastics and Stochastics Rep., **59**(1996), 21–45.

[99] M.I. Freidlin and A.D. Wentzell, Reaction-diffusion equation with randomly perturbed boundary condition, Annals of Probab., **20**(1992), 963–986.

[100] A. Friedman, Partial differential equations of parabolic type, Prentice-Hall, Englewood Cliffs, N.J., 1964.

[101] A. Friedman, Stochastic differential equations and applications, Academic Press, New York, 1975.

[102] V.N. Galkin and Yu.A. Stepanyants, On the existence of stationary solitary waves in a rotating fluid, J. Appl. Math. Mech., **55**(1991), 939–943.

[103] H. Gao and C. Sun, Random attractor for the 3D viscous stochastic primitive equations with additive noise, Stochastics Dynamics **9**(2009), 293–313.

[104] H. Gao and C. Sun, Well-posedness and large deviations for the stochastic primitive equations in two space dimensions, Commun. Math. Sci., **10**(2012), 575–593.

[105] L. Garcia-Palacios and F.J. Lazaro, Langevin-dynamics study of the dynamical properties of small magnetic particles, Phys. Rev. B, **58**(1998), 14937–14958.

[106] C.S. Gardner, Korteweg-de Vries equation and generalizations IV: The Korteweg-de Vries equation as a Hamiltonian system, J. Math. Phys., **12**(1971), 1548–1551.

[107] D. Gątarek and B. Goldys, On weak solutions of stochastic equations in Hilbert spaces, Stochastics and Stochastics Rep., **46**(1994), 41–51.

[108] J.M. Ghidaglia, Finite dimensional behavior for weakly damped driven Schrödinger equations, Ann. Inst. H. Poincaré Anal. Nonlinéaire, **5**(1988), 365–405.

[109] O. A. Gilman, R. Grimshaw and Yu. A. Stepanyants, Approximate and numerical solutions of the stationary Ostrovsky equation, Stud. Appl. Math., **95**(1995), 115–126.

[110] J. Ginibre and G. Velo, The global Cauchy problem for the nonlinear Schrödinger equation revisited, Ann. Inst. H. Poincaré Anal. Nonlinéaire, **2**(1985), 309–327.

[111] J. Ginibre and Y. Tsutsumi, Uniqueness of solutions for the generalized Korteweg de-Vries equation, SIAM J. Math. Anal., **20**(1989), 1388–1425.

[112] N. Glatt-Holtz and R. Temam, Pathwise solutions of the 2-d stochastic primitive equations, Appl Math. Optim., **63**(2011), 401–433.

[113] N. Glatt-Holtz and M. Ziane, The stochastic primitive equations in two space dimensions with multiplicative noise, Discrete Contin. Dyn. Syst. Ser. B, **10**(2008), 801–822.

[114] G. Gong, Introduction to stochastic differential equations (In Chinese), Science Press Publication, Beijing, 1987.

[115] F. Guillén-Gonzáez, N. Masmoudi and M. A. Rodríguez-Bellido, Anisotropic estimates and strong solutions of the Primitive Equations, Differ. Integr. Equ., **14**(2001), 1381–1408.

[116] B. Guo and F. Chen, Finite dimensional behavior of global attractors for weakly damped and forced KdV equations coupling with nonlinear Schrödinger equations, Nonlinear Anal.: TMA, **29**(1997), 569–584.

[117] B. Guo and S. Ding, Landau-Lifshitz equations, Frontiers of Research with the Chinese Academy of Sciences, Vol. **1**, World Scientific, Singapore, 2008.

[118] B. Guo and M. Hong, The Landau-Lifshitz equation of the ferromagnetic spin chain and harmonic maps, Calc. Var., **1**(1993), 311–334.

[119] B. Guo and D. Huang, 3D Stochastic primitive equations of the large-scale ocean: Global well-posedness and attractors, Commun. Math. Phys., **286**(2009), 697–723.

[120] D. Huang, B. Guo and Y. Han, Random attractors for a dissipative quasi-geostrophic dynamics under stochastic forcing, Int. J. Dyn. Syst. Differ. Equ., **1**(2008), 147–154.

[121] B. Guo and Z. Huo, The well-posedness of the Korteweg de Vries-Benjamin-Ono equation, J. Math. Anal. Appl., **295**(2004), 444–458.

[122] B. Guo and Z. Huo, The global attractor of the damped forced Ostrovsky equation, J. Math. Anal. Appl., **329**(2007), 392–407.

[123] B. Guo and G. Wang, Well-posedness of stochastic Korteweg-de Vries-Benjamin-Ono equation, Front. Math. China, **5**(2010), 161–177.

[124] B. Guo and G. Wang, Attractors for nonlinear Schrödinger equation coupling with stochastic weakly damped, forced KdV equation, Front. Math. China, **3**(2008), 495–510.

[125] B. Guo, G. Wang and D. Li, The attractor of the stochastic generalized Ginzburg-Landau equation (Chinese), Sci. China, Ser. Z, **51**(2008), 955–964.

[126] M. Hairer and J.C. Mattingly, Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing, Ann. Math., **164**(2006), 993–1032.

[127] J.K. Hale, Asymptotic behavior of dissipative dynamical systems, Mathematical Surveys and Monographs, **25**, AMS, Providence, 1988.

[128] A. Haraux, Attractors of asymptotically compact processes and applications to nonlinear partial differential equations, Commun. PDE, **13**(1988), 1383–1414.

[129] R. Herman, The stochastic, damped Korteweg-de Vries equation, J. Phys. A, **23**(1990), 1063–1084.

[130] I. Hosokawa and K. Yamamoto, Turbulence in the randomly forced one dimensional Burgers flow, J. Stat. Phys., **13**(1975), 245–272.

[131] C. Hu, R. Temam and M. Ziane, The primitive equations on the large scale ocean under the small depth hypothesis, Discrete Contin. Dyn. Syst., **9**(2003), 97–131.

[132] Y. Hu, Heat equations with fractional white noise potentials, Appl. Math. Optim., **43**(2001), 221–243.

[133] Z. H. Huo and Y. L. Jia, Low regularity solutions for the Ostrovsky equation, Proc. Edinb. Math. Soc., **49**(2006), 87–100.

[134] N. Ikeda and S. Watanabe, Stochastic Differential Equations and diffusion process, North-Holland, Amsterdam, 1981.

[135] P. Isaza and J. Mejía, Cauchy problem for the Ostrovsky equation in spaces of low regularity, J. Diff. Equ., **230**(2006), 661–681.

[136] K. Itô, Lectures on stochastic processes, Tata Inst. Fundamental Research, Bombay, 1961.

[137] J. Jacod and A.N. Shiryaev, Limit theorems for stochastic processes, Springer, Berlin, 1987.

[138] A. Jakubowski, Towards a general Doob-Meyer decomposition theorem, Probab. Math. Stat., **26**(2006), 143–153.

[139] N. Ju, The global attractor for the solutions to the 3d viscous primitive equations, Discrete Contin. Dyn. Syst., **17**(2007), 159–179.

[140] O. Kallenberg, Foundations of modern probability, Springer, Berlin, 2002.

[141] G. Kallianpur and J. Xiong, Large deviations for a class of stochastic partial differential equations, Ann. Probab., **24**(1996), 320–345.

[142] G. Kallianpur, J. Xiong, G. Hardy and S. Ramasubramanian, The existence and uniqueness of solutions of nuclear space-valued stochastic differential equations driven by Poisson random measures, Stochastics and Stochastics Rep., **50**(1994), 85–122.

[143] N.G. Van Kampen, Stochastic process in physics and chemistry, North-Holland, Amsterdam, 1981.

[144] T. Kamppeter and F.G. Mertens et al., Stochastic vortex dynamics in two-dimensional easy-plane ferromagnets: Multiplicative versus additive noise, Phys. Rev. B, **59**(1999), 11349–11357.

[145] I. Karatzas and S. E. Shreve, Brownian motion and Stochastic calculus, Springer-Verlag, New York, 1988.

[146] T. Kato, Quasilinear equation of evolution with applications to partial differential equations, Lect. Notes Math., **448**(1975), 27–50.

[147] T. Kato, On nonlinear Schrodinger equation. Ann. Inst. H. Poincaré Phys. Theor., **46**(1987), 113–129.
[148] T. Kato, On the Cauchy problem for the (generalized) Korteweg-de Vries equation, Stud. Appl. Math., **8**(1983), 93–128.
[149] T. Kato and T. Ogawa, Analyticity and smoothing effect for the Korteweg de-Vries equation with a single point singularity, Math. Ann., **316**(2000), 577–608.
[150] C.E. Kenig, G.P. Ponce and L. Vega, Well-posedness of the initial value problem for the Korteweg-de Vries equation, J. Amer. Math. Soc., **4**(1991), 323–347.
[151] C.E. Kenig, G.P. Ponce and L. Vega, Oscillatory integrals and regularity of dispersive equations, Indiana Univ. Math. J., **40**(1991), 33–69.
[152] C.E. Kenig, G.P. Ponce and L. Vega, A bilinear estimate with application to the KdV equation, J. Amer. Math. Soc., **9**(1996), 573–603.
[153] C.E. Kenig, G.P. Ponce and L. Vega, The Cauchy problem for the Korteweg-de Vries equation in Sobolev spaces of negative indices, Duke Math. J., **71**(1993), 1–21.
[154] J.U. Kim, On the stochastic Benjamin-Ono equation, J. Differ. Equ., **228**(2006), 737–768.
[155] G. Kobelkov, Existence of a solution in “whole" for the large-scale ocean dynamics equations, C. R. Math. Acad. Sci. Paris, **343**(2006), 283–286.
[156] G. M. Kobelkov and V. B. Zalesny, Existence and uniqueness of a solution to primitive equations with stratification ‘in the large’, *Russian J. Numer. Anal. Math. Modelling*, **23**(2008), 39–61.
[157] C. Knoche, Mild solutions of SPDE’s driven by Poisson Noise in Infinite dimensions and their dependence on initial conditions, PhD thesis, University of Bielefeld, 2005.
[158] A. Kolmogorov, Wienersche spiralen und einige andere interessante kurven im Hilbertschen Raum S.R. (Doklady), **26**(1940), 115–118.
[159] V. Konotop and L. Vázquez, Nonlinear random waves, World Scientific Publishing Co., Inc. River Edge, NJ, 1994.
[160] N.V. Krylov and B.L. Rozovskii, Stochastic differential equations, J. Soviet Math., **14**(1983), 1233–1277,(Translated from the Russian edition published in 1979).
[161] I. Kukavica and M. Ziane, On the regularity of the primitive equations of the ocean, Nonlinearity, **20**(2007), 2739–2753.
[162] T.G. Kurtz, The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities, Electron. J. Probab., **12**(2007), 951–965.
[163] T.G. Kurtz, E. Pardoux and P. Protter, Stratonovitch stochastic differential equations driven by general semimartingales, Ann. Inst. H. Poincaré B, **31**(1985), 351–377.
[164] Ladyzenskaya, The boundary value problems of mathematical physics, Applied Mathematical Science, Vol. **49**, Springer-Verlag, Berlin, Heidelberg, New York, 1985.
[165] D. Lamberton and B. Lapeyre, Introduction to Stochastic calculus applied to finance, Chapman and Hall, London, 1996 (English translation by N. Rabeau and F. Mantion).
[166] L. Landau and E. M. Lifshitz, On the theory of the dispersion of magnetic permeability in ferromagnetic bodies, Phys. Z. Sowj, **8**(1935), 153 (Reproduced in Collected Papers of L.D.Landau, Pergamon Press, New York, 1965, 101–114)
[167] J.A. Langa, Finite-dimensional limiting dynamics of random dynamical systems, Dynam. Syst., **18**(2003), 57–68.
[168] J.A. Langa and J.C. Robinson, A finite number of point observations which determines a non-autonomous fluid flow, Nonlinearity, **14**(2001), 673–682.
[169] P. Levy, Processus Stochastiques et Mouvement Brownien, Gauthier-Villars, Paris, 1948.
[170] F. Lin and C. Wang, The analysis of Harmonic maps and their heat flows, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.
[171] Y.L. Lin, Applied Stochastic Process, Tsinghua University Press, Beijing, 2002 (in Chinese).
[172] F. Linares, L^2 global well-posedness of the initial value problem associated to the Benjamin equation, J. Differ. Equ., **152**(1999), 1425–1433.

[173] F. Linares and A. Milanés, Local and global well-posedness for the Ostrovsky equation, J. Differ. Equ., **222**(2006), 325–340.
[174] J.L. Lions, Quelques méthods de résolutions des problèmes aux limites nonlinéaires, Dunod, Paris, 1969.
[175] J.L. Lions, R. Temam and S. Wang, Models for the coupled atmosphere and ocean, Comput. Mech. Adv., **1**(1993), 1–120.
[176] J.L. Lions, R. Temam and S. Wang, New formulations of the primitive equations of atmosphere and applications, Nonlinearity, **5**(1992), 237–288.
[177] J.L. Lions, R. Temam and S. Wang, On the equations of the large-scale ocean, Nonlinearity, **5**(1992), 1007–1053.
[178] V.G. Makhankov, Dynamics of classical solitons, Phys. Rep. (Sect. C of Physics Letters), **35**(1978), 1–128.
[179] R. Manthey and B. Maslowski, Qualitative behaviour of solutions of stochastic reaction-diffusion equations, Stochastic Process. Appl., **43**(1992), 265–289.
[180] B. Maslowski, On probability distributions of solutions of semilinear stochastic evolution equations, Stochastics, **45**(1993), 17–44.
[181] B. Maslowski and J. Seidler, Invariant measures for Nonlinear SPDE's: uniqueness and stability, Arch. Math. (Brno), **34**(1998), 153–172.
[182] R. Merton, Theory of rational option pricing, Bell J. Econ. Manage. Sci., **4**(1973), 141–183.
[183] M. Métivier, Semimartingales: A course on stochastic process, De Gruyter, Berlin, 1982.
[184] R. Mikulevicius and B.L. Rozovskii, Global L^2-solution of stochastic Navier-Stokes equation, Ann. Probab., **33**(2005), 137–176.
[185] H. Morimoto, Attractors of probability measures for semilinear stochastic evolution equations, Stochastics Anal. Appl., **10**(1992), 205–212.
[186] R. Moser, Partial regularity for the Landau-Lifshitz equation in small dimensions, MPI Preprint **26**, 2002.
[187] C. Muller, The heat equation with Levy noise, Stochastic Process. Appl., **74**(1998), 67–82.
[188] C. Muller, L. Mytnik and A. Stan, The heat equation with time-independent multiplicative stable Levy noise, Stochastic Process. Appl., **116**(2006), 70–100.
[189] L. Mytnik, Stochastic partial differential equation driven by stable noise, Probab. Theory Related Fields, **123**(2002), 157–201.
[190] J. Neveu, Bases mathematiques du Calcul des probabilites, Masson and Cie, Editeurs, Paris, 1970.
[191] C. Odasso, Ergodicity for the stochastic complex Ginzburg-Landau equations, Annales de l'Institut Henri Poincare (B) Probab. Stat., **42**(2006), 417–454.
[192] B. Øksendal, Stochastic differential equations: An introduction with applications (6th edition), Springer-Verlag, Berlin, 2005.
[193] L.A. Ostrovsky, Nonlinear internal waves in a rotating ocean, Okeanologia, **18**(1978), 181–191.
[194] E. Pardoux, Equaytions aux dérivées partielles stochastiques non lineaires monotones: Etude de solutions fortes de typr Ito, Thése Doct. Sci. Math. Univ. Paris Sud., 1975.
[195] E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes, Stochastics, **3**(1980), 127–167.
[196] E. Pardoux, Stochastic partial differential equations, a review, Bull. Sci. Math., **117**(1993), 29–47.
[197] E. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation, Syst. Control Lett., **14**(1990), 55–61.
[198] E. Pardoux and T. Zhang, Absolute continuity of the law of the solution of a parabolic SPDE, J. Funct. Anal., **112**(1993), 447–458.
[199] A. Pazy, Semigroup of linear operators and applications to partial differential equations, Springer-Verlag, Berlin and New York, 1983.

[200] S. Peng, A nonlinear Feynman-Kac formula and applications, in Proceedings of Symposium of System Sciences and Control Theory, Chen & Yong eds., 173–184, World Scientific, Singapore, 1992.

[201] S. Peszat, Large deviation estimates for stochastic evolution equations, Probab. Theory Related Fields, **98**(1994), 113–136.

[202] S. Peszat and J. Zabczyk, Stochastic partial differential equations with Levy noise, Cambridge University Press, Cambridge, 1992.

[203] S. Peszat and J. Zabczyk, Strong Feller property and irreducibility for diffusions on Hilbert spaces, Ann. Probab., **23**(1995), 157–172.

[204] S. Peszat and J. Zabczyk, Stochastic partial differential equations with Lévy noise, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 2007.

[205] M. Petcu, Gevrey class regularity for the primitive equations in space dimension 2, Asymptot. Anal., **39**(2004), 1–13.

[206] M. Petcu, R. Temam, and D. Wirosoetisno, Existence and regularity results for the primitive equations in two space dimensions, Commun. Pure Appl. Anal., **3**(2004), 115–131.

[207] M. Petcu, R. Temam, and M. Ziane, Some mathematical problems in geophysical fluid dynamics. In *Special Volume on Computational Methods for the Atmosphere and the Oceans*, volume 14 of *Handbook of Numerical Analysis*, 577–750, Elsevier, 2008.

[208] C. Prévôt and M. Röckner, A concise course on stochastic partial differential equations, Lect. Notes Math., **1905**, Springer, Berlin, 2007.

[209] J. Printems, The stochastic Korteweg de-Vries equation in $L^2(\mathbb{R})$, J. Differ. Equ., **153**(1999), 338–373.

[210] P. Protter, Volterra equations driven by semimartingales, Ann. Probab., **13**(1985), 519–530.

[211] P. Protter, Stochastic Integration and Differential equations: A new approach, Springer-Verlag, Berlin, 1990.

[212] X. Pu and B. Guo, Regular solutions for stochastic Landau-Lifshitz-Gilbert equation and blow up phenomena, Sci. China, Math., **53**(2010), 3115–3130.

[213] X. Pu and B. Guo, Momentum estimates and ergodicity for the 3D stochastic cubic Ginzburg CLandau equation with degenerate noise, J. Differ. Equ., **251**(2011), 1747–1777.

[214] M. Qian and G. Gong, Stochastic process (2 edition), Peking University Press, Beijing, 1997 (in Chinese).

[215] K.O. Rasmussen, Yu.B. Gaididei, O. Bang and P.L. Christiansen, The influence of noise on critical collapse in the nonlinear Schrödinger equation, Phys. Lett. A, **204**(1995), 121–127.

[216] L.F. Richardson, Weather prediction by numerical process, Cambridge University Press, 1922 (Reprinted Dover, New York, 1965).

[217] J. C. Robinson, Infinite-dimensional dynamical systems, Cambridge University Press, Cambridge 2001.

[218] M. Rochner and X. Zhang, Stochastic tamed 3D Navier-Stokes equations: Existence, uniqueness and ergodicity, Probab. Theory Related Fields, **145**(2009), 211–267.

[219] L.C.G. Rogers and D. Williams, Diffusion, Markov processes and martingales, Vols. I.II, Cambridge University Press, Cambridge, 2000.

[220] M. Röchner, Introduction to stochastic partial differential equations, Lect. Notes in Purdue University, 2006.

[221] R. Rosa, The global attractor of a weakly damped, forced Korteweg-de Vries equation in $H^1(\mathbb{R})$, Math. Contemp., **19**(2000), 129–152.

[222] J. Rosinski and W.A. Woyczynski, On Ito stochastic integration with respect to p-stable motion: inner clock, integrability of sample paths, double and multiple integrals, Ann. Probab., **14**(1986), 271–286.

[223] C. Roth, Difference Methods for stochastic partial differential equations, Z. Angew. Math. Mech., **82**(2002), 821–830.

[224] B.L. Rozovskii, Stochastic evolution systems. Linear theory and applications to nonlinear filtering. Mathematics and its Applications (Soviet Series), **35**, Kluwer Academic Publishers, Dordrecht, 1990.

[225] B. Rüdiger, Stochastic integration with respect to compensated Poisson random measures on separable Banach spaces, Stochastics and Stochastics Rep., **76**(2004), 213–242.

[226] W. Rudin, Real and complex analysis, 2nd edition, McGraw-Hill, New York, 1974.

[227] J.L. Rubio de Francia, F.J. Ruiz and J.L. Torrea, Calderon-Zygmund theory for operator-valued kernels, Adv. Math., **62**(1986), 7–48.

[228] M. Scalerandi, A. Romano and C.A. Condat, Korteweg-de Vries solitons and additive stochastic perturbations, Phys. Rev. E, **58**(1998), 709–712.

[229] B. Schmalfuss, Measure Attractors of the stochastic Navier-Stokes equation, Report **258**, Institut für Dynamische System, Bermen, 1991.

[230] B. Schmalfuss, Backward cocycles and attractors of stochastic differential equation, in *international Seminar on Applied Mathematics–Nonlinear Dynamics: Attractor Approximation and Global Behaviour*, edited by V. Reitmann, T. Riedrich and N. Kokch (Technische Uniberität, Dresden, 1992), pp. 185–192.

[231] B. Schmalfuss, The stochastic attractor of the stochastic Lorenz system, Z. Angew. Math. Phys., **48**(1997), 951–975.

[232] J. Seidler, Ergodic behaviour of stochastic parabolic equations, Czechoslovak Math. J., **47**(1997), 277–316.

[233] A.N. Shiryaev, Essential of stochastic finance, World Scientific Publishing, INc., River Edge, NJ, 1999.

[234] R. Sowers, Large deviations for a reaction-diffusion system with non-Gaussian perturbations, Ann. Probab., **20**(1992), 504–537.

[235] S. S. Sritharan and P. Sundar, Large deviations for the two-dimensional Navier-Stokes equations with multiplicative noise, Stochastic Process. Appl., **116**(2006), 1636–1659.

[236] C.F. Sun and H.J. Gao, Well-posedness for the stochastic 2D primitive equation with Lévy noise, Sci. China Math., **56**(2013), 1629–1645.

[237] C. Sulem and P.L. Sulem, The nonlinear Schrödinger equation: Self-Focusing and Wave Collapse, Springer, New York (1999).

[238] P.L. Sulem, C. Sulem and C. Bardos, On the continuous limit for a system of classical spins, Commun. Math. Phys., **107**(1986), 431–454.

[239] E.M. Stein, Singular integrals and differentiability properties of functions, Princeton University Press, Princeton, New Jersey, 1970.

[240] R.L. Stratonovich, A new representation for stochastic integrals and equations, J. Soc. Ind. Appl. Math., **4**(1966), 362–371.

[241] T. Taniguchi, K. Liu and A. Truman, Existence, Uniqueness, and Asymptotic behavior of mild solutions to stochastic functional differential equations in Hilbert spaces, J. Differ. Equ., **181**(2002), 72–91.

[242] T. Tao, Multilinear weighted convolution of L^2 functions, and applications to nonlinear dispersive equation, Amer. J. Math., **123**(2001), 839–908.

[243] T. Tao, Nonlinear dispersive equations: local and global analysis, CBMS, American Mathematical Society, Providence, Rhode Island, 2006.

[244] R. Temam, Navier-Stokes equations and nonlinear functional analysis, CBMS-NSF Regional Conference Series in Applied Mathematics, **41**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1983.

[245] R. Temam, Infinite dimensional dynamical systems in mechanics and physics, Springer-Verlag, New York, 1998.

[246] R. Temam and M. Ziane, Some mathematical problems in geophysical fluid dynamics, Handbook of mathematical fluid dynamics, vol. **3**, 2004.

[247] Y. Tsutsumi, L^2-solitions for nonlinear Schrodinger equations and nonlinear groups, Funkcial. Ekvac., **30**(1987), 115–125.

[248] T. Ueda and W.L. Kath, Dynamics of optical pulses in randomly birefringent fibers, Physica D, **55**(1992), 166–181.

[249] V. Varlamov, Y. Liu, Cauchy problem for the Ostrovsky equation, Discrete Contin. Dyn. Syst., **10**(2004), 731–753.

[250] M. I. Vishik and A. V. Fursikov, Mathematical problems of statistical hydromechanics, Kluwer, Dordrecht, 1980.

[251] A. Visintin, On Landau-Lifshitz equation for ferromagnetism, Japan J. Appl. Math., **2**(1985), 69–84.

[252] M. Wadati, Stochastic Korteweg de-Vries equation, J. Phys. Soc. Japan, **52**(1983), 2642–2648.

[253] M. Wadati and Y. Akutsu, Stochastic Korteweg de-Vries equation with and without damping, J. Phys. Soc. Japan, **53**(1984), 3342.

[254] J.B. Walsh, An introduction to stochastic partial differential equations, Lecture Notes in Math., **1180**, Springer, Berlin, New York, 265–439.

[255] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations, J. Math. Kyoto Univ., **11**(1971), 155–167.

[256] J-A Yan, Shige Peng, Shizhan Fan and Liming Wu, Selected lectures on stochastic analysis, Science Press, Beijing, 1997 (in Chinese).

[257] D. Yang, The asymptotic behavior of the stochastic Ginzburg-Landau equation with multiplicative noise, J. Math. Phys., **45**(2004), 4064–4076.

[258] H.F. Yashima, Equations de Navier-Stokes Stochastiques Non-Homogenes et Applications, Tesi di Perfezionamento, Scuola Normale Superiore, Pisa, 1992.

[259] K. Yosida, Functional analysis (6th edition), Springer-Verlag, Berlin, 1980.

[260] K-C Chang, Lectures on functional analysis, Peking University Press, Beijing, 1990 (in Chinese).

[261] X.C. Zhang, Smooth solutions of non-linear stochastic partial differential equations driven by multiplicative noises, Sci. China Math., **53**(2010), 2949–2972.

[262] Y.L. Zhou, Applications of discrete functional analysis to the finite difference method, International Academic Publishers, Beijing, 1990.

[263] Y.L. Zhou, B.L. Guo and S.B. Tan, Existence and uniqueness of smooth solution for system of ferromagnetic chain, Sci. Sin., Ser. A, **34**(1991), 257–266.

Index