Qualitative study of randomly forced partial differential equations

Armen Shirikyan

Département de Mathématiques, Université de Cergy–Pontoise
2, av. Adolphe Chauvin, 95302 Cergy–Pontoise Cedex, France
E-mail: Armen.Shirikyan@u-cergy.fr

Contents

Introduction

The aim of this course is to give a self-contained concise introduction to the qualitative theory of nonlinear PDE's with random perturbation. We consider the parabolic equation

$$\dot{u} - \Delta u + g(u) = f(t, x), \quad x \in D, \tag{0.1}$$

where $D \in \mathbb{R}^n$ is a bounded domain and $f(t, x)$ is a given function depending on a random parameter. After studying the initial-boundary value problem for (0.1), we show that its solutions form a family of Markov processes in the infinite-dimensional phase space $L^2(D)$. This enables one to define an evolution in the space of probability measures on the phase space and, in particular, to introduce the concept of stationary measure. We next establish some a priori estimates for the time and ensemble averages of solutions and combine them with the Prokhorov compactness criterion to construct a stationary measure for the equation in question. Finally, we discuss the problem of uniqueness and mixing for Markov chains in finite-dimensional spaces.

Notation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X be a Polish space with metric d_X. We shall use the following notation:

$\mathcal{B}_X$ is the Borel σ-algebra on X.
$\mathcal{P}(X)$ is the set of probability Borel measures on X.
$C_b(X)$ is the space of bounded continuous functions $f : X \to \mathbb{R}$ endowed with the norm

$$\|f\|_\infty := \sup_{u \in X} |f(u)|.$$

If $\mu \in \mathcal{P}(X)$ and $f \in C_b(X)$, then we write

$$(f, \mu) = \int_X f(u)\, \mu(du) = \int_X f(u)\, d\mu.$$

If $\mu \in \mathcal{P}(X)$, then $L^1(X, \mu)$ is the set of measurable functions $f : X \to \mathbb{R}$ with finite norm

$$\|f\|_\mu := \int_X |f(u)|\, \mu(du).$$

$\mathcal{L}(X)$ is the space of functions $f \in C_b(X)$ such that

$$\|f\|_{\mathcal{L}} := \|f\|_\infty + \sup_{u \neq v} \frac{|f(u) - f(v)|}{d_X(u, v)} < \infty.$$

We use the following three metrics on the space $\mathcal{P}(X)$:

$$\|\mu_1 - \mu_2\|_{\mathrm{var}} := \sup_{\Gamma \in \mathcal{B}_X} |\mu_1(\Gamma) - \mu_2(\Gamma)|,$$

$$\|\mu_1 - \mu_2\|_\infty^* := \sup_{\|f\|_\infty \leq 1} |(f, \mu_1) - (f, \mu_2)|,$$

$$\|\mu_1 - \mu_2\|_{\mathcal{L}}^* := \sup_{\|f\|_{\mathcal{L}} \leq 1} |(f, \mu_1) - (f, \mu_2)|.$$

If $a, b \in \mathbb{R}$, then $a \vee b$ ($a \wedge b$) denotes the maximum (minimum) of a and b.

1 Preliminaries

1.1 Probability spaces, random variables, distributions

Let Ω be a set with σ-algebra $\mathcal{F}$, i.e., a family of subsets of Ω that contains Ω and satisfies the following two properties:

- if $B_i \in \mathcal{F}$ for $i = 1, 2, \dots$, then $\bigcap_i B_i \in \mathcal{F}$;
- if $B \in \mathcal{F}$, then $B^c = \Omega \setminus B \in \mathcal{F}$.

Any pair $(\Omega, \mathcal{F})$ possessing the above properties will be called a *measurable space*.

Example 1.1. Let X be a Polish space and let $\mathcal{B}_X$ be the Borel σ-algebra on X, i.e., the minimal σ-algebra generated by the open subsets of X. Then $(X, \mathcal{B}_X)$ is a measurable space. In what follows, we assume that all Polish spaces are endowed with their Borel σ-algebra.

Let $\mathbb{P}$ be a probability measure on a measurable space $(\Omega, \mathcal{F})$, i.e., $\mathbb{P}$ is a countably additive function from $\mathcal{F}$ to $[0, 1]$ such that $\mathbb{P}(\Omega) = 1$. Any such triple $(\Omega, \mathcal{F}, \mathbb{P})$ will be called a *probability space*.

Example 1.2. Let us consider the interval $I = [0, 1]$ endowed with the Borel σ-algebra $\mathcal{B}_I$, and let ℓ be the Lebesgue measure on I. Then $(I, \mathcal{B}_I, \ell)$ is a probability space.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(X, \mathcal{B})$ be a measurable space. Given an X-valued random variable ξ (i.e., a map from Ω to X such that $\xi^{-1}(\Gamma) \in \mathcal{F}$ for any $\Gamma \in \mathcal{B}$), we define its distribution $\mathcal{D}(\xi)$ as the image of $\mathbb{P}$ under ξ:

$$\mathcal{D}(\xi)(\Gamma) = \mathbb{P}\big(\xi^{-1}(\Gamma)\big) = \mathbb{P}\big(\{\omega \in \Omega : \xi(\omega) \in \Gamma\}\big).$$

Thus, the distribution of ξ is a probability measure on $(X, \mathcal{B})$. The space of all probability measures on a measurable space $(X, \mathcal{B})$ will be denoted by $\mathcal{P}(X)$.

Exercise 1.3. Let $(X, \mathcal{B})$ be a measurable space and let $\mu \in \mathcal{P}(X)$. Show that there is an X-valued random variable whose distribution coincides with μ.

Solution. Let us set $\Omega = X$, $\mathcal{F} = \mathcal{B}$, $\mathbb{P} = \mu$ and consider a random variable defined by the formula $\xi(\omega) = \omega$. Then, for any $\Gamma \in \mathcal{B}$, we have

$$\mathbb{P}\{\xi \in \Gamma\} = \mathbb{P}\{\omega \in \Gamma\} = \mu(\Gamma),$$

and therefore $\mathcal{D}(\xi) = \mu$. □

A random variable $\xi : \Omega \to X$ is said to be *simple* if takes on finitely many values. If X is a separable Banach space, then ξ is said to be *integrable* if

$$\int_\Omega \|\xi(\omega)\|_X \, \mathbb{P}(d\omega) < \infty. \tag{1.1}$$

Exercise 1.4. Let $\xi : \Omega \to X$ be an integrable random variable.

(i)* Show that there is a sequence of simple random variables $\xi_k : \Omega \to X$ such that

$$\int_\Omega \|\xi(\omega) - \xi_k(\omega)\|_X \, \mathbb{P}(d\omega) \to 0 \quad \text{as } k \to \infty. \tag{1.2}$$

(ii) Show that if $\xi_k : \Omega \to X$ is a sequence of simple random variables satisfying (1.1), then the limit

$$\lim_{k\to\infty} \int_\Omega \xi_k(\omega) \, \mathbb{P}(d\omega) \tag{1.3}$$

exists and does not depend on $\{\xi_k\}$.

If $\xi : \Omega \to X$ is an integrable random variable, then the limit (1.3) is called the *integral of* ξ *over* Ω and is denoted by

$$\mathbb{E}\,\xi = \int_\Omega \xi(\omega) \, \mathbb{P}(d\omega).$$

1.2 Independence, product of probability spaces

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{A}$ be a set of indices, and let $\mathcal{F}_\alpha \subset \mathcal{F}$, $\alpha \in \mathcal{A}$, be a family of sub-σ-algebras in Ω.

Definition 1.5. The family $\{\mathcal{F}_\alpha, \alpha \in \mathcal{A}\}$ is said to be *independent* if for any finite set of indices $\alpha_1, \dots, \alpha_n \in \mathcal{A}$ and any $B_i \in \mathcal{F}_{\alpha_i}$, $i = 1, \dots, n$, we have

$$\mathbb{P}(B_1 \cdots B_n) = \mathbb{P}(B_1) \cdots \mathbb{P}(B_n).$$

Let $(X_\alpha, \mathcal{B}_\alpha)$, $\alpha \in \mathcal{A}$, be a family of measurable spaces and let ξ_α be some X_α-valued random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let us denote by $\mathcal{F}_\alpha = \sigma(\xi_\alpha)$ the σ-algebra generated by ξ_α, i.e., the family of sets $B \in \mathcal{F}$ that can be represented in the form $\xi_\alpha^{-1}(\Gamma)$ for some $\Gamma \in \mathcal{B}_\alpha$.

Definition 1.6. The family $\{\xi_\alpha, \alpha \in \mathcal{A}\}$ is said to be *independent* if the corresponding family of σ-algebras $\{\mathcal{F}_\alpha, \alpha \in \mathcal{A}\}$ is independent, i.e., for any finite set of indices $\alpha_1, \dots, \alpha_n \in \mathcal{A}$ and any $\Gamma_i \in \mathcal{B}_{\alpha_i}$, $i = 1, \dots, n$, we have

$$\mathbb{P}\{\xi_{\alpha_1} \in \Gamma_1, \dots, \xi_{\alpha_n} \in \Gamma_n\} = \prod_{i=1}^n \mathbb{P}\{\xi_{\alpha_i} \in \Gamma_i\}.$$

Exercise 1.7. Show that a family of random variables $\{\xi_\alpha, \alpha \in \mathcal{A}\}$ is independent iff for any finite set of indices $\alpha_1, \dots, \alpha_n \in \mathcal{A}$ and any bounded measurable functions $f_i : X_{\alpha_i} \to \mathbb{R}$, $i = 1, \dots, n$, we have

$$\mathbb{E}\Big\{\prod_{i=1}^n f_i(\xi_{\alpha_i})\Big\} = \prod_{i=1}^n \mathbb{E}\, f_i(\xi_{\alpha_i}). \tag{1.4}$$

Hint: Begin with the case of simple functions.

We now describe a simple way for constructing independent random variables. Let $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbb{P}_\alpha)$, $\alpha \in \mathcal{A}$, be a family of probability spaces. Define the product space

$$\Omega = \prod_{\alpha \in \mathcal{A}} \Omega_\alpha = \big\{\omega = (\omega_\alpha, \alpha \in \mathcal{A}) : \omega_\alpha \in \Omega_\alpha \text{ for any } \alpha\big\}$$

and denote by $\mathcal{F}$ the product σ-algebra, i.e., the minimal σ-algebra generated by the sets of the form

$$B_{\alpha_1,\dots,\alpha_n} = \big\{(\omega_\alpha, \alpha \in \mathcal{A}) : \omega_{\alpha_1} \in B_1, \dots, \omega_{\alpha_n} \in B_n\big\}, \tag{1.5}$$

where n is a finite integer and $B_i \in \mathcal{F}_{\alpha_i}$ for $i = 1, \dots, n$.

Theorem 1.8. *There is a unique probability measure on* $(\Omega, \mathcal{F})$ *such that*

$$\mathbb{P}(B_{\alpha_1,\dots,\alpha_n}) = \prod_{i=1}^{n} \mathbb{P}_{\alpha_i}(B_i) \quad \text{for any set } B_{\alpha_1,\dots,\alpha_n}. \tag{1.6}$$

Proof. The uniqueness is based on the technique of monotone classes, and we first recall the corresponding result, which will be useful in many other situations.

Step 1: Monotone classes. Let Ω be a set and let $\mathcal{M}$ be a family of subsets of Ω. We shall say that $\mathcal{M}$ is a monotone class if it contains Ω and possesses the following properties:

- if $A, B \in \mathcal{M}$ and $A \cap B = \varnothing$, then $A \cup B \in \mathcal{M}$;
- if $A, B \in \mathcal{M}$ and $A \subset B$, then $B \setminus A \in \mathcal{M}$;
- if $A_i \in \mathcal{M}$ for $i = 1, 2 \dots$, and $A_1 \subset A_2 \subset \cdots$, then $\bigcup_i A_i \in \mathcal{M}$.

It is clear that any σ-algebra is a monotone class, but not vice versa. The following lemma gives a sufficient condition ensuring that the minimal monotone class generated by a family of subsets coincides with the minimal σ-algebra.

Lemma 1.9. *Let $\mathcal{C}$ be a family of subsets of Ω such that for any $A, B \in \mathcal{C}$ we have $A \cap B \in \mathcal{C}$. Then the minimal monotone class containing $\mathcal{C}$ coincides with the minimal σ-algebra generated by $\mathcal{C}$.*

Proof. It suffices to show that the minimal monotone class $\mathcal{M}$ containing $\mathcal{C}$ is a σ-algebra. This will be established if we prove that the intersection of any two sets in $\mathcal{M}$ is also an element of $\mathcal{M}$.

Let us fix an arbitrary $A \in \mathcal{C}$ and set

$$\mathcal{M}_A = \{B \in \mathcal{M} : A \cap B \in \mathcal{M}\}.$$

It is clear that $\mathcal{M}_A$ is a monotone class and $\mathcal{M}_A \supset \mathcal{C}$. Therefore, by the definition of $\mathcal{M}$, we must have $\mathcal{M}_A \supset \mathcal{M}$. We have thus shown that, for any $A \in \mathcal{C}$ and $B \in \mathcal{M}$, the intersection $A \cap B$ belongs to $\mathcal{M}$.

We now fix $A \in \mathcal{M}$ and consider the family $\mathcal{M}_A$. Repeating literally the above argument, we can show that $\mathcal{M}_A \supset \mathcal{M}$. This completes the proof of the lemma. □

Step 2: Uniqueness of product measure. To prove the uniqueness, we suppose that there are two probability measures $\mathbb{P}$ and $\mathbb{P}'$ satisfying (1.6) and introduce the family

$$\mathcal{M} = \{A \in \mathcal{F} : \mathbb{P}(A) = \mathbb{P}'(A)\}.$$

It is easy to verify that $\mathcal{M}$ is a monotone class. Moreover, by assumption, it contains the family $\mathcal{C}$ of sets of the form (1.5). Since $\mathcal{C}$ satisfies the condition of Lemma 1.9, we conclude that $\mathcal{M}$ coincide with the σ-algebra $\mathcal{F}$. Thus, the measures $\mathbb{P}$ and $\mathbb{P}'$ are equal.

To prove the existence, we first show that the function $\mathbb{P}$ defined by formula (1.6) on the family $\mathcal{C}$ is continuous from above at $\varnothing$ and then use the Lebesgue extension theorem to extend it to the σ-algebra $\mathcal{F}$.

Step 3: Continuity of $\mathbb{P}$. We claim that if $\{\Gamma_k\} \subset \mathcal{C}$ is a decaying sequence of sets such that $\cap_k \Gamma_k = \varnothing$, then $\mathbb{P}(\Gamma_k) \to 0$ as $k \to \infty$. Without loss of generality, we can assume that there is a sequence of indices $\{\alpha_k\} \subset \mathcal{A}$ and measurable sets $B^k_{\alpha_i} \in \mathcal{F}_{\alpha_i}$, $i = 1, \dots, k$, such that

$$\Gamma_k = \{\omega = (\omega_\alpha, \alpha \in \mathcal{A}) : \omega_{\alpha_i} \in B^k_{\alpha_i} \text{ for } i = 1, \dots, k\}.$$

We shall show that if $\mathbb{P}(\Gamma_k) \geq \varepsilon$ for some $\varepsilon > 0$ and all $k \geq 1$, then there is a sequence $b_m \in \Omega_{\alpha_m}$, $m \geq 1$, such that any element $(\omega_\alpha, \alpha \in \mathcal{A})$ with $\omega_{\alpha_m} = b_m$ for $m \geq 1$ belongs to the intersection of Γ_k.

Let us set

$$g_m(\omega) = \lim_{n\to\infty} \int_{\Omega_{\alpha_n}} \cdots \int_{\Omega_{\alpha_{m+1}}} I_{\Gamma_n}(\omega)\, d\mathbb{P}_{\alpha_{m+1}} \cdots d\mathbb{P}_{\alpha_n}.$$

It is a matter of direct verification to show that the functions g_m are well defined and possess the following properties:

- The function g_m depends only on $\omega_{\alpha_1}, \dots, \omega_{\alpha_m}$, and $\operatorname{supp} g_m \subset \Gamma_m$.
- For any $m \geq 1$, we have
$$\int_{\Omega_{\alpha_{m+1}}} g_{m+1}(\omega)\, d\mathbb{P}_{\alpha_{m+1}} = g_m(\omega). \tag{1.7}$$
- If $\mathbb{P}(\Gamma_k) \geq \varepsilon$ for all $k \geq 1$, then
$$\int_{\Omega_{\alpha_1}} g_1(\omega)\, d\mathbb{P}_{\alpha_1} \geq \varepsilon. \tag{1.8}$$

We now assume that $\mathbb{P}(\Gamma_k) \geq \varepsilon$ for all $k \geq 1$ and construct the sequence b_m recursively. We shall write $g_m(\omega_{\alpha_1}, \dots, \omega_{\alpha_m})$ instead of $g_m(\omega)$. It follows from (1.8) that there is $b_1 \in \Omega_{\alpha_1}$ such that $g_1(b_1) \geq \varepsilon$. If $b_1, \dots . b_m$ are already constructed, then relation (1.7) implies that

$$\int_{\Omega_{\alpha_{m+1}}} g_{m+1}(b_1, \dots, b_m, \omega_{\alpha_{m+1}})\, d\mathbb{P}_{\alpha_{m+1}} = g_m(b_1, \dots, b_m) \geq \varepsilon.$$

Therefore there is $b_{m+1} \in \Omega_{\alpha_{m+1}}$ such that $g_{m+1}(b_1, \dots, b_{m+1}) \geq \varepsilon$. Since $\operatorname{supp} g_m \subset \Gamma_m$, we see that any point $(\omega_\alpha, \alpha \in \mathcal{A}) \in \Omega$ with $\omega_{\alpha_i} = b_i$ for $i = 1, \dots, m$ belongs to Γ_m. This shows that the sequence b_m possesses the required property.

Step 4: Extension to the σ-algebra. We now define a measure on $\mathcal{F}$ by the formula

$$\mathbb{P}(\Gamma) = \inf\biggl\{\sum_{i=1}^{\infty} \mathbb{P}(C_i) : \Gamma \subset \bigcup_{i=1}^{\infty} C_i,\ C_i \in \mathcal{C}\biggr\},$$

where $\Gamma \in \mathcal{F}$, and the infimum is taken over all countable covering $\{C_i\}$ of Γ by sets belonging to $\mathcal{C}$. Using the continuity property established in Step 3, it can be shown that $\mathbb{P}$ is a σ-additive measure on $\mathcal{F}$ (see Section 3.1 in [Dud02] for details). □

The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called the *product space* of $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbb{P}_\alpha)$, $\alpha \in \mathcal{A}$. It follows from (1.6) that if ξ_α are some random variables defined on Ω_α, then their natural extensions[1] to Ω are independent.

*Exercise** 1.10. Let $\xi_1, \dots, \xi_n$ be independent X-valued random variables and let $f : X \times \cdots \times X \to \mathbb{R}$ be a bounded measurable function of n variables. Then

$$\mathbb{E}\, f(\xi_1, \dots, \xi_n) = \mathbb{E}_{\omega_1} \dots \mathbb{E}_{\omega_n}\, f(\xi_1(\omega_1), \dots, \xi_n(\omega_n)). \tag{1.9}$$

Hint: Use the technique of monotone classes (cf. Steps 1 and 2 of the proof of Theorem 1.8).

1.3 Conditional expectation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let ξ be a real-valued integrable random variable.

Proposition 1.11. *For any sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$ there is a $\mathcal{G}$-measurable random variable η such that*

$$\int_B \xi(\omega)\mathbb{P}(d\omega) = \int_B \eta(\omega)\mathbb{P}(d\omega) \quad \text{for any } B \in \mathcal{G}. \tag{1.10}$$

If $\tilde{\eta}$ is another $\mathcal{G}$-measurable random variable satisfying (1.10)*, then $\eta(\omega) = \tilde{\eta}(\omega)$ for a.e. ω.*

Proof. Let us consider a signed measure on $(\Omega, \mathcal{G})$ defined by the formula

$$\mu(B) = \int_B \xi(\omega)\,\mathbb{P}(d\omega), \quad B \in \mathcal{G}. \tag{1.11}$$

[1] By the natural extension of ξ_α to Ω we mean the random variable defined by the relation $\tilde{\xi}_\alpha(\omega) = \xi_\alpha(\omega_\alpha)$.

The measure μ is absolutely continuous with respect to $\mathbb{P}$. Hence, by the Radon–Nikodym theorem (see Theorem 5.5.4 in [Dud02]), there is a $\mathcal{G}$-measurable function $\eta(\omega)$ such that

$$\mu(B) = \int_B \eta(\omega)\,\mathbb{P}(d\omega) \quad \text{for any } B \in \mathcal{G}.$$

Comparing this relation with (1.11), we arrive at (1.10).

If $\tilde{\eta}$ is another $\mathcal{G}$-measurable random variable satisfying (1.10), then

$$\int_B \big(\eta(\omega) - \tilde{\eta}(\omega)\big)\,\mathbb{P}(d\omega) = 0 \quad \text{for any } B \in \mathcal{G},$$

whence it follows that $\eta = \tilde{\eta}$ almost surely. □

Definition 1.12. The random variable η constructed in Proposition 1.11 is called the *conditional expectation of* ξ *given* $\mathcal{G}$ and is denoted by $\mathbb{E}\,(\xi\,|\,\mathcal{G})$.

Exercise 1.13. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Suppose that Ω is represented as a countable union of disjoint subsets Ω_i, $i \geq 1$, and let $\mathcal{G}$ be the sub-σ-algebra generated by $\{\Omega_i,\, i \geq 1\}$. Construct the conditional expectation of a real-valued random variable ξ given $\mathcal{G}$.

Exercise 1.14. Show that, if ξ is $\mathcal{G}$-measurable, then $\mathbb{E}\,(\xi\,|\,\mathcal{G}) = \xi$, and if $\sigma(\xi)$ and $\mathcal{G}$ are independent, then $\mathbb{E}\,(\xi\,|\,\mathcal{G}) = \mathbb{E}\,\xi$. Furthermore, if $\mathcal{G} \subset \mathcal{G}'$, then

$$\mathbb{E}\,\big(\mathbb{E}\,(\xi\,|\,\mathcal{G}')\,|\,\mathcal{G}\big) = \mathbb{E}\,(\xi\,|\,\mathcal{G}). \tag{1.12}$$

*Exercise** 1.15. Let ξ and η be random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and valued in a Polish space X and let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra such that ξ is $\mathcal{G}$-measurable and η is independent of $\mathcal{G}$. Show that for any bounded measurable function $f\colon X \times X \to \mathbb{R}$ we have

$$\mathbb{E}\,\big(f(\xi, \eta)\,|\,\mathcal{G}\big) = \big(\mathbb{E}\,f(x, \eta)\big)\big|_{x=\xi}.$$

Hint: Use the technique of monotone classes.

1.4 Metrics on the space of probability measures

Let X be a Polish space endowed with its Borel σ-algebra $\mathcal{B}_X$. We denote by $C_b(X)$ the space of continuous functions $f\colon X \to \mathbb{R}$ with finite norm

$$\|f\|_\infty := \sup_{u \in X} |f(u)|.$$

Since the family $\mathcal{P}(X)$ of probability measures on $(X, \mathcal{B}_X)$ is a subset in the dual space of $C_b(X)$, we can endow it with the dual metric

$$\|\mu_1 - \mu_2\|_\infty^* := \sup\big\{|(f, \mu_1) - (f, \mu_2)| : f \in C_b(X), \|f\|_\infty \leq 1\big\},$$

where, for any $f \in C_b(X)$ and $\mu \in \mathcal{P}(X)$, we set

$$(f,\mu) := \int_X f(u)\mu(du) = \int_X f(u)\,d\mu.$$

Let us introduce another metric on $\mathcal{P}(X)$.

Definition 1.16. The *total variation distance* between two probability measures μ_1 and μ_2 is defined by the formula

$$\|\mu_1 - \mu_2\|_{\mathrm{var}} := \sup\bigl\{|\mu_1(\Gamma) - \mu_2(\Gamma)| : \Gamma \in \mathcal{B}_X\bigr\}.$$

Theorem 1.17. *For any $\mu_1, \mu_2 \in \mathcal{P}(X)$, we have*

$$\|\mu_1 - \mu_2\|_\infty^* = 2\,\|\mu_1 - \mu_2\|_{\mathrm{var}}. \tag{1.13}$$

Proof. We shall need the following auxiliary assertion, which is of independent interest.

Proposition 1.18. *Let m be a positive Borel measure on X. Suppose that $\mu_1, \mu_2 \in \mathcal{P}(X)$ are absolutely continuous with respect to m. Then*

$$\|\mu_1 - \mu_2\|_{\mathrm{var}} = \frac{1}{2}\int_X \bigl|\rho_1(u) - \rho_2(u)\bigr|\,dm = 1 - \int_X (\rho_1 \wedge \rho_2)(u)\,dm, \tag{1.14}$$

where $\rho_i(u)$, $i = 1, 2$, is the density of μ_i with respect to m.

Taking this proposition for granted, let us prove the theorem. Let m be a finite measure satisfying the conditions of Proposition 1.18. For instance, we can take $m = \mu_1 + \mu_2$. Using the first relation in (1.14), for any $f \in C_b(X)$ with $\|f\|_\infty \le 1$ we derive

$$\bigl|(f,\mu_1) - (f,\mu_2)\bigr| \le \int_X \bigl|f(u)\bigl(\rho_1(u) - \rho_2(u)\bigr)\bigr|\,dm \le 2\,\|\mu_1 - \mu_2\|_{\mathrm{var}},$$

which implies that

$$\|\mu_1 - \mu_2\|_\infty^* \le 2\,\|\mu_1 - \mu_2\|_{\mathrm{var}}.$$

To establish the converse inequality, we set

$$Y = \{u \in X : \rho_1(u) > \rho_2(u)\}. \tag{1.15}$$

Let us consider a function $f(u)$ that is equal to 1 on Y and to -1 on the complement of Γ. We have

$$\begin{aligned}(f,\mu_1) - (f,\mu_2) &= \int_X f(u)\bigl(\rho_1(u) - \rho_2(u)\bigr)\,dm \\ &= \int_X \bigl|\rho_1(u) - \rho_2(u)\bigr|\,dm = 2\,\|\mu_1 - \mu_2\|_{\mathrm{var}},\end{aligned} \tag{1.16}$$

where we used the first relation in (1.14). To complete the proof of (1.13), it suffices to choose a sequence $f_n \in C_b(X)$ such that $\|f_n\|_\infty \le 1$ for all n and $f_n(u) \to f(u)$ for m-a.e. $u \in X$ and note that $(f_n,\mu_1) - (f_n,\mu_2)$ tends to the left-hand side of (1.16) as $n \to \infty$. $\square$

*Exercise** 1.19. Let X be a Polish space endowed with its Borel σ-algebra and let m be a positive finite measure on X. Show that for any bounded measurable function $f\colon X\to\mathbb{R}$ there is a sequence of continuous functions uniformly bounded by $\sup|f|$ that converges to f almost surely. *Hint:* It suffices to prove that f can be approximated by continuous functions in the space $L^1(X,m)$; any bounded measurable function can be approximated uniformly by bounded simple functions; the indicator function of any measurable set can be approximated by bounded continuous functions.

Proof of Proposition 1.18. A direct verification shows that

$$\tfrac12\,|\rho_1-\rho_2| = \tfrac12\,(\rho_1+\rho_2)-\rho,$$

where $\rho=\rho_1\wedge\rho_2$. Integrating the above relation over X with respect to m, we obtain the second equality in (1.14).

We now show that

$$\|\mu_1-\mu_2\|_{\mathrm{var}} \le 1-\int_X \rho(u)\,dm. \tag{1.17}$$

Let Y be the set defined by (1.15). Then, for any $\Gamma\in\mathcal{B}_X$, we have

$$\begin{aligned}\mu_1(\Gamma)-\mu_2(\Gamma) &= \int_\Gamma(\rho_1-\rho_2)\,dm \le \int_{\Gamma\cap Y}(\rho_1-\rho_2)\,dm\\ &= \int_{\Gamma\cap Y}(\rho_1-\rho)\,dm \le \int_X(\rho_1-\rho)\,dm = 1-\int_X\rho(u)\,dm.\end{aligned}$$

In view of the symmetry, this inequality implies (1.17).

To prove the converse inequality, we denote by Y^c the complement of Y and note that $\rho=\rho_1$ on Y^c and $\rho=\rho_2$ on Y. It follows that

$$\begin{aligned}\mu_1(Y)-\mu_2(Y) &= \int_Y(\rho_1-\rho_2)\,dm\\ &= \left(\int_Y\rho_1 dm+\int_{Y^c}\rho\,dm\right)-\left(\int_Y\rho_2 dm+\int_{Y^c}\rho\,dm\right)\\ &= \left(\int_Y\rho_1 dm+\int_{Y^c}\rho_1 dm\right)-\left(\int_Y\rho\,dm+\int_{Y^c}\rho\,dm\right)\\ &= 1-\int_X\rho\,dm.\end{aligned}$$

This completes the proof of the proposition. □

In what follows, we shall need a weaker topology on $\mathcal{P}(X)$. Let $\mathcal{L}(X)$ be the space of functions $f\in C_b(X)$ such that

$$\|f\|_{\mathcal{L}} := \|f\|_\infty + \sup_{u\ne v}\frac{|f(u)-f(v)|}{d_X(u,v)} < \infty,$$

where d_X is the metric on X. For any $\mu_1,\mu_2\in\mathcal{P}(X)$, we set

$$\|\mu_1-\mu_2\|_{\mathcal{L}}^* := \sup\bigl\{|(f,\mu_1)-(f,\mu_2)| : f\in\mathcal{L}(X), \|f\|_{\mathcal{L}}\le 1\bigr\}. \tag{1.18}$$

*Exercise** 1.20. Show that $\|\mu_1 - \mu_2\|_{\mathcal{L}}^*$ defines a metric on $\mathcal{P}(X)$. *Hint:* The triangle inequality is obvious; to prove that $\mu_1 = \mu_2$ if $\|\mu_1 - \mu_2\|_{\mathcal{L}}^* = 0$, it suffices to show that $\mu_1(F) = \mu_2(F)$ for any closed set $F \subset X$; to this end, find a sequence $f_k \in \mathcal{L}(X)$ converging to the indicator function of F.

The following theorem is of fundamental importance. Its proof can be found in [Dud02] (see Theorem 11.3.3 and Corollary 11.5.5).

Theorem 1.21. **(i)** *The set $\mathcal{P}(X)$ is a complete metric space with respect to $\|\cdot\|_{\mathcal{L}}^*$.*

(ii) *A sequence $\{\mu_n\} \subset \mathcal{P}(X)$ converges to $\mu \in \mathcal{P}(X)$ in the metric $\|\cdot\|_{\mathcal{L}}^*$ if and only if*

$$(f, \mu_n) \to (f, \mu) \quad \text{as } n \to \infty \text{ for any } f \in C_b(X).$$

2 Randomly forced parabolic equations

In this section, we consider the problem

$$\dot{u} - \Delta u + g(u) = h(x) + \eta(t, x), \quad x \in D, \tag{2.1}$$

$$u\big|_{\partial D} = 0. \tag{2.2}$$

Here $D \subset \mathbb{R}^n$ is a bounded domain with boundary $\partial D \in C^2$, Δ is the Laplace operator, $g \in C^1(\mathbb{R})$ is a real-valued function, $h \in L^2(D)$, and η is a random process of the form

$$\eta(t, x) = \sum_{k=1}^{\infty} \eta_k(x)\delta(t - k), \tag{2.3}$$

where $\delta(t)$ is the Dirac measure concentrated at zero and $\{\eta_k\}$ is a sequence of independent identically distributed (i.i.d.) random variables in $L^2(D)$. We shall show that the initial value problem for (2.1) – (2.3) is well posed and that the restrictions of its solutions to integer times form a homogeneous family of Markov chains.

Here is the plan of this section. In § 2.1, we have compiled some properties of Sobolev spaces. Section 2.2 is devoted to studying the problem (2.1), (2.2) with $\eta \equiv 0$. In particular, we establish global existence of solutions and derive some a priori estimates. In § 2.3, we turn to the problem (2.1) – (2.3) and prove the Markov property for its solutions. We also introduce the corresponding Markov semigroups and establish their basic properties. Finally, in § 2.4, we prove the strong Markov property.

2.1 Some properties of Sobolev spaces

Let $D \subset \mathbb{R}^n$ be a bounded domain with smooth boundary ∂D and let $s \geq 0$ be an integer. Recall that the Sobolev space $H^s(D)$ is defined as the set of functions $u \in L^2(D)$ such that $\partial^\alpha u \in L^2(D)$ for any multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$

of length $|\alpha| = \alpha_1 + \cdots + \alpha_n \le s$. The space $H^s(D)$ is endowed with the norm

$$\|u\|_{H^s(D)} = \Big(\sum_{|\alpha| \le s} \|\partial^\alpha u\|^2 \Big)^{1/2},$$

where $\|\cdot\|$ stands for the norm in $L^2(D)$.

Let $C^s(\overline{D})$ be the space of s times continuously differentiable functions on D whose derivatives up to the order s admit a continuous extension to $\overline{D}$. The aim of the following exercise is to show that, for the elements of $H^s(D)$, one can define their restriction to the boundary ∂D.

Exercise 2.1. **(i)** Show that if $\partial D \in C^s$, then $C^s(\overline{D})$ is dense in $H^s(D)$.

(ii) Let $\gamma : C^1(\overline{D}) \to C(\partial D)$ be the operator that takes each function $u(x)$ to its restriction to the boundary ∂D. Show that there is a constant $C > 0$ such that

$$\|\gamma u\|_{L^2(\partial D)} \le C\, \|u\|_{H^1(D)} \quad \text{for all } u \in H^1(D). \tag{2.4}$$

Combine inequality (2.4) with (i) to prove that γ can be extended by continuity to an operator from $H^1(D)$ to $L^2(\partial D)$.

We now set

$$H_0^1(D) = \big\{ u \in H^1(D) : u\big|_{\partial D} = 0 \big\}.$$

The following exercise establishes the Friedrichs inequality, which shows that the L^2 norm of any function $u \in H_0^1(D)$ is bounded by the L^2 norm of its gradient and therefore the latter defines an equivalent norm on $H_0^1(D)$.

Exercise 2.2. Show that the constant

$$\alpha := \inf\big\{ \|\nabla u\|^2 : u \in H_0^1(D), \|u\| = 1 \big\} \tag{2.5}$$

is positive. Use this fact to prove the Friedrichs inequality:

$$\|u\|^2_{H^1(D)} \le \big(1 + \alpha^{-1}\big)\, \|\nabla u\|^2 \quad \text{for any } u \in H_0^1(D). \tag{2.6}$$

Finally, let us define $H^{-1}(D)$ as the dual space of $H_0^1(D)$, i.e., the space of continuous linear functionals $\ell : H_0^1(D) \to \mathbb{R}$ endowed with the norm

$$\|\ell\|_{H^{-1}(D)} = \sup_u |\ell(u)|,$$

where the supremum is taken over all $u \in H_0^1(D)$ such that $\|u\|_{H^1(D)} \le 1$.

Exercise 2.3. **(i)** Show that $H^{-1}(D)$ is included in the space of distributions on D and that a distribution ℓ belongs to $H^{-1}(D)$ iff there is a constant $C > 0$ such that

$$|\ell(\varphi)| \le C\, \|\varphi\|_{H^1(D)} \quad \text{for all } \varphi \in C_0^\infty(D).$$

(ii) Show that the operator Δ is continuous from $H_0^1(D)$ to $H^{-1}(D)$ and has a continuous inverse Δ^{-1}.

(iii)* Show that if $\partial D \in C^2$, then the operator Δ^{-1} is continuous from $L^2(D)$ to $H^2(D)$. *Hint:* see Theorem 8.12 in [GT01].

2.2 Well-posedness and a priori estimates for parabolic equations

Let us consider the following problem in a bounded domain $D \subset \mathbb{R}^n$ with boundary $\partial D \in C^2$:

$$\dot{u} - \Delta u + g(u) = h(x), \tag{2.7}$$

$$u\big|_{\partial D} = 0, \tag{2.8}$$

$$u(0,x) = u_0(x). \tag{2.9}$$

We assume that the function $g \in C^1(\mathbb{R})$ satisfies the condition

$$K := \sup_{u\in\mathbb{R}} |g'(u)| < \infty. \tag{2.10}$$

For any $T > 0$, let us set

$$\mathcal{X}_T := C(0,T; L^2(D)) \cap L^2(0,T; H_0^1(D)),$$
$$\mathcal{Y}_T := C(0,T; H_0^1(D)) \cap L^2(0,T; H^2(D)),$$

and denote by $\mathcal{X}$ the space of functions $u(t,x)$ on $\mathbb{R}_+ \times D$ whose restriction to $[0,T] \times D$ belongs to $\mathcal{X}_T$ for any $T > 0$.

Theorem 2.4. *Suppose that* (2.10) *is satisfied. Then for any* $u_0, h \in L^2(D)$ *the following assertions hold.*

Existence and uniqueness: *There is a unique function* $u \in \mathcal{X}$ *that satisfies Eq.* (2.7) *for* $t > 0$, $x \in D$ *in the sense of distributions and the initial condition* (2.9). *Moreover, the operator* S_t *that takes the pair* (u_0, h) *to* $u(t,\cdot)$ *is uniformly Lipschitz from* $L^2(D) \times L^2(D)$ *to* $L^2(D)$ *for any fixed* $t \geq 0$.

Regularity: *For any* $T > 0$, *the function* $u(t,x)$ *satisfies the inclusions*

$$\dot{u} \in L^2(0,T; H^{-1}(D)), \quad t^{\frac{1}{2}} u \in \mathcal{Y}_T. \tag{2.11}$$

Moreover, the operator S_t *is uniformly Lipschitz from* $L^2(D) \times L^2(D)$ *to* $H_0^1(D)$ *for any fixed* $t > 0$.

Boundedness: *Suppose that*

$$L := \sup_{u\in\mathbb{R}} (-g(u)u) < \infty. \tag{2.12}$$

Then

$$\|u(t,\cdot)\|^2 + 2\int_0^t \|\nabla u(s,\cdot)\|^2 ds = 2L\,\mathrm{vol}(D)\,t + 2\int_0^t (h, u(s,\cdot))\,ds, \tag{2.13}$$

where $t \geq 0$, *and* $(\cdot,\cdot)$ *stands for the scalar product in* $L^2(D)$. *In particular, we have the estimate*

$$\|u(t,\cdot)\|^2 \leq e^{-\alpha t}\|u_0\|^2 + C^2, \quad t \geq 0, \tag{2.14}$$

where $C^2 = \alpha^{-2}(2L\alpha\,\mathrm{vol}(D) + \|h\|^2)$.

The proof of this theorem is based on the following lemma, which is of independent interest. Let us consider the linear equation

$$\dot{u} - \Delta u = f(t,x), \quad x \in D. \tag{2.15}$$

Lemma 2.5. *Let $T > 0$. Then for any $u_0 \in L^2(D)$ and $f \in L^1(0,T;L^2(D))$ there is a unique function $u \in \mathcal{X}_T$ that satisfies (2.15) for $t \in (0,T)$, $x \in D$ in the sense of distributions and the initial condition (2.9). Moreover, the assertions below hold.*

Duhamel representation: *Let Δ_D be the Laplace operator in $L^2(D)$ regarded as an unbounded operator with the domain $\mathcal{D} = H^2(D) \cap H_0^1(D)$. Then*

$$u(t,x) = e^{t\Delta_D} u_0 + \int_0^t e^{(t-s)\Delta_D} f(s,\cdot)\, ds, \quad 0 \le t \le T, \tag{2.16}$$

where $\{e^{t\Delta_D}, t \ge 0\}$ is the semigroup generated by Δ_D.

Energy identity: *For $0 \le t \le T$, we have*

$$\|u(t)\|^2 + 2\int_0^t \|\nabla u(s)\|^2 ds = \|u_0\|^2 + 2\int_0^t (f(s), u(s))\, ds. \tag{2.17}$$

A priori estimates: *The following estimate holds for $0 \le t \le T$:*

$$\|u(t)\|^2 + 2\int_0^t \|\nabla u(s)\|^2 ds \le 2\|u_0\|^2 + \left(2\int_0^t \|f(s)\|\, ds\right)^2. \tag{2.18}$$

Furthermore, if $h \in L^2(0,T,L^2(D))$, then

$$t\|\nabla u(t)\|^2 + \int_0^t s\|\Delta u(s)\|^2 ds \le \|u_0\|^2 + 2\left(\int_0^t \|f\|\, ds\right)^2 + \int_0^t s\|f\|^2 ds, \tag{2.19}$$

and the solution $u(t,x)$ satisfies inclusions (2.11).

Proof. We only outline the proof, which is divided into several steps. We shall need the concept of an analytic semigroup and a sectorial operator (see Section 1.3 in [Hen81]).

Step 1: Semigroup $e^{t\Delta_D}$. Let us show that the operator Δ_D generates an analytic semigroup in $L^2(D)$. To this end, it suffices to prove that Δ_D is a sectorial operator (see Theorem 1.3.4 in [Hen81]). This fact will be established if we show that it is a negative self-adjoint operator.

It is a matter of direct verification to show that Δ_D is symmetric and that

$$(\Delta_D u, u) = -\int_D |\nabla u|^2 dx \le 0 \quad \text{for any } u \in \mathcal{D}.$$

Therefore, by the corollary of Theorem 1 in [Yos95, Section VII.3], to prove that the operator Δ_D is self-adjoint, it suffices to show that its range coincides with $L^2(D)$. This is a straightforward consequence of Exercise 2.3 (iii).

Step 2: Existence. To prove existence of a solution, we first assume that $u_0 \in \mathcal{D}$ and $f \in C(0,T;\mathcal{D})$. In this case, well-known properties of the semigroup $e^{t\Delta_D}$ (see [Hen81, Theorem 1.3.4]) imply that formula (2.16) defines a function

$$u \in C(0,T;\mathcal{D}) \cap C^1(0,T;L^2(D)) \tag{2.20}$$

that satisfies Eq. (2.15) and the initial condition (2.9). To construct a solution in the general case, we consider arbitrary sequences $u_{0k} \in \mathcal{D}$ and $f_k \in C(0,T;\mathcal{D})$ that converge to u_0 and f in the spaces $L^2(D)$ and $L^1(0,T;L^2(D))$, respectively. Using inequality (2.18) for solutions satisfying (2.20) (its proof is given below), we can show that the corresponding functions $u_k(t,x)$ converges in $\mathcal{X}_T$, and the limiting function is a solution of problem (2.15), (2.9). The above argument also implies the Duhamel representation (2.16).

Step 3: Uniqueness. To prove uniqueness, we assume that $u \in \mathcal{X}_T$ is a solution of problem (2.15), (2.9) with $u_0 = 0$ and $f = 0$. Let us fix an arbitrary $\tau \in (0,T]$ and consider the dual problem

$$\dot{v} + \Delta v = 0, \quad 0 \le t \le \tau, \tag{2.21}$$

$$v(\tau,x) = \varphi(x), \tag{2.22}$$

where $\varphi \in C_0^\infty(D)$ is an arbitrary function. Repeating the argument in Step 2, it can be shown that problem (2.21), (2.22) has a solution $v(t,x)$ satisfying (2.20) with $T = \tau$. We have

$$\frac{d}{dt}(u(t),v(t)) = (\dot{u},v) + (u,\dot{v}) = 0,$$

and since $u(0) \equiv 0$, we conclude that $(u(\tau),\varphi) = 0$ for any $\varphi \in C_0^\infty(D)$ and $\tau \in (0,T]$. Thus, we see that $u \equiv 0$.

Step 4: Estimates and regularity. We first assume that $f \in C(0,T,\mathcal{D})$ and $u_0 \in \mathcal{D}$. In this case, the solution satisfies (2.20). Multiplying Eq. (2.15) by $2u$ and integrating in $x \in D$, we obtain

$$\frac{d}{dt}\|u(t)\|^2 + 2\|\nabla u(t)\|^2 = 2(f(t),u(t)).$$

Integrating this equation in time, we derive the energy identity (2.17).

To prove (2.18), we replace t by τ in (2.17) and take the supremum of both sides with respect to $\tau \in]0,t]$. This results in

$$\sup_{0\le\tau\le t}\left(\|u(\tau)\|^2 + 2\int_0^\tau \|\nabla u(s)\|^2 ds\right) \le \|u_0\|^2 + 2\Big(\sup_{0\le\tau\le t}\|u(\tau)\|\Big)\int_0^t \|f(s)\|\, ds.$$

This implies inequality (2.18).

Finally, to prove (2.19), we multiply Eq. (2.15) by $-2t\Delta u$ and integrate in $x \in D$. After some simple transformations, we obtain

$$\frac{d}{dt}\big(t\,\|\nabla u(t)\|^2\big) + 2t\,\|\Delta u(t)\|^2 = \|\nabla u(t)\|^2 - 2t(f(t),\Delta u(t)).$$

Noting that $|2t(f,\Delta u)| \le t\|f\|^2 + t\|\Delta u\|^2$, integrating in time and using (2.18), we arrive at (2.19).

To prove the required assertions in the general case, we use a standard limiting argument. For instance, let us establish inequality (2.19) and the second inclusion in (2.11). We choose sequences $u_{0k} \in \mathcal{D}$ and $f_k \in C(0,T,\mathcal{D})$ that converge to u_0 and f in the spaces $L^2(D)$ and $L^2(0,T,L^2(D))$, respectively. Let $u_k(t,x)$ be the corresponding solutions. Applying inequalities (2.18) and (2.19) to $u_k - u_m$ and using the continuity of $\Delta_D^{-1} : L^2(D) \to H^2(D)$, we see that $\{u_k\}$ and $\{t^{\frac{1}{2}}u_k\}$ are Cauchy sequences in the spaces $\mathcal{X}_T$ and $\mathcal{Y}_T$, respectively. It follows that the limiting function satisfies the second inclusion in (2.11) and that we can pass to the limit as $k \to \infty$ in inequality (2.19) with $u = u_k$. The proof of the lemma is complete. □

Proof of Theorem 2.4. The proof is divided into four steps.

Step 1: Existence. We shall use the contraction mapping principle to construct a solution. Namely, let us fix an arbitrary $T > 0$ and endow the space $\mathcal{X}_T$ with the norm

$$\|u\|_{\mathcal{X}_T} := \sup_{0\le t\le T} e^{-\lambda t}\Big(\|u(t)\|^2 + 2\int_0^t \|\nabla u(s)\|^2 ds\Big)^{1/2},$$

where $\lambda > 0$ is a large parameter and will be chosen later. Consider a mapping $F\colon \mathcal{X}_T \to \mathcal{X}_T$ that takes each function $v \in \mathcal{X}_T$ to the solution of the problem

$$\dot u - \Delta u = h(x) - g(v), \quad u(0,x) = u_0(x).$$

It is clear that $u \in \mathcal{X}_T$ is a solution of the problem (2.7) – (2.9) if and only if it is a fixed point of F. We claim that F is a contraction for sufficiently large $\lambda > 0$. Indeed, if $v_i \in \mathcal{X}_T$, $i = 1,2$, and $u_i = F(v_i)$, then the difference $u = u_1 - u_2 \in \mathcal{X}_T$ satisfies Eq. (2.15) with $f(t,x) = g(v_2) - g(v_1)$ and vanishes at $t = 0$. Therefore, by (2.18), we have

$$\begin{aligned}\|F(v_1) - F(v_2)\|_{\mathcal{X}_T} &\le 2 \sup_{0\le t\le T}\Big(e^{-\lambda t}\int_0^t \|g(v_1(s)) - g(v_2(s))\|\, ds\Big)\\ &\le 2 \sup_{0\le t\le T}\Big(e^{-\lambda t}\int_0^t e^{\lambda s}\,\|v_1 - v_2\|_{\mathcal{X}_T}\, ds\Big)\\ &\le 2K\lambda^{-1}\|v_1 - v_2\|_{\mathcal{X}_T}. \end{aligned} \tag{2.23}$$

Thus, if $\lambda > 2K$, then the mapping F is a contraction in $\mathcal{X}_T$ and has a unique fixed point. Since $T > 0$ is an arbitrary constant, we conclude that there is a solution $u \in \mathcal{X}$.

Step 2: Uniqueness and Lipschitz property. Let $u_{0i}, h_i \in L^2(D)$ for $i = 1,2$ and let $u_i = S_t(u_{0i}, h_i)$. Then, for any $T > 0$, the difference $u = u_1 - u_2 \in \mathcal{X}_T$ is a solution of the problem

$$\dot u - \Delta u = h(x) - (g(u_1) - g(u_2)), \quad u(0,x) = u_0(x), \tag{2.24}$$

where $h = h_1 - h_2$ and $u_0 = u_{01} - u_{02}$. Repeating the argument used in the proof of (2.23), we can show that

$$\begin{aligned}\|u_1 - u_2\|_{\mathcal{X}_T} &\leq \sqrt{2}\|u_0\| + 2 \sup_{0\leq t\leq T}\Big(e^{-\lambda t}\int_0^t \big(\|h\| + e^{\lambda s}K\,\|u_1 - u_2\|_{\mathcal{X}_T} ds\Big)\\ &\leq \sqrt{2}\|u_0\| + 2\big(\|h\| + K\lambda^{-1}\|u_1 - u_2\|_{\mathcal{X}_T}\big).\end{aligned}$$

Choosing $\lambda = 4K$, we conclude that

$$\|u_1(t) - u_2(t)\| \leq 2\sqrt{2}\, e^{\lambda t}\big(\|u_{01} - u_{02}\| + T\,\|h_1 - h_2\|\big), \quad 0 \leq t \leq T. \tag{2.25}$$

This inequality implies the uniqueness of solution and the Lipschitz property of the operator S_t.

Step 3: Regularity. To prove (2.11), we note that $u(t,x)$ is a solution of the problem (2.15), (2.9), where $f = h - g(u) \in L^2(0,T;L^2(D))$. Hence, by Lemma 2.5, inclusions (2.11) hold.

We now use inequality (2.19) to establish the Lipschitz property of S_t. Retaining the notation of Step 2, we recall that the difference $u = u_1 - u_2$ is the solution of the problem (2.24). Therefore, have inequality (2.19) with $f = h - (g(u_1) - g(u_2))$. Applying the Cauchy inequality to estimate the second term on the right-hand side of (2.19), we derive

$$\begin{aligned}\sqrt{t}\,\|\nabla u(t)\| &\leq \|u_0\| + \left(3t\int_0^t \|f(s,\cdot)\|^2 ds\right)^{1/2}\\ &\leq \|u_0\| + \sqrt{3t}\,\|h\| + \left(3t\int_0^t \|g(u_1) - g(u_2)\|^2 ds\right)^{1/2}\\ &\leq \|u_0\| + \sqrt{3t}\,\|h\| + \sqrt{3}Kt \sup_{0\leq s\leq t}\|u_1(s) - u_2(s)\|.\end{aligned}$$

Combining this with (2.25), we obtain the inequality

$$\|\nabla(u_1(t) - u_2(t))\| \leq C_T\big(t^{-\frac{1}{2}}\|u_{01} - u_{02}\| + \|h_1 - h_2\|\big), \quad 0 < t \leq T,$$

which implies that $S_t\colon L^2(D) \times L^2(D) \to H_0^1(D)$ is uniformly Lipschitz continuous.

Step 4: Boundedness. Recall that $u(t,x)$ is a solution of the problem (2.15), (2.9) with where $f = h - g(u) \in L^2(0,T;L^2(D))$. Applying the energy identity (2.17), we derive

$$\|u(t)\|^2 + 2\int_0^t \|\nabla u(s)\|^2 ds = \|u_0\|^2 + 2\int_0^t (h - g(u), u)\, ds.$$

Taking into account condition (2.12), we obtain (2.13).

Using the Cauchy and Friedrichs inequalities (see Exercise 2.2) to estimate the second term on the right-hand side of (2.13), we obtain

$$\|u(t,\cdot)\|^2 + 2\alpha \int_0^t \|u(s,\cdot)\|^2 ds \le 2L\,\mathrm{vol}(D)\,t + 2\int_0^t \|h\|\,\|u(s,\cdot)\|\,ds$$
$$\le 2L\,\mathrm{vol}(D)\,t + \alpha^{-1}t\,\|h\|^2 + \alpha\int_0^t \|u(s,\cdot)\|^2 ds,$$

whence it follows that

$$\|u(t,\cdot)\|^2 + \alpha \int_0^t \|u(s,\cdot)\|^2 ds \le \left(2L\,\mathrm{vol}(D) + \alpha^{-1}\|h\|^2\right) t.$$

Application of the Gronwall inequality now results in (2.14). The proof of Theorem 2.4 is complete. □

We now consider the problem (2.1) – (2.3), where $\{\eta_k\}$ is a sequence of functions in $L^2(D)$. As before, we assume that $D \subset \mathbb{R}^n$ is a bounded domain with C^2 boundary, $g(u)$ is a continuously differentiable function with bounded derivative, and $h \in L^2(D)$.

Definition 2.6. A function $u(t,x)$ defined for $t \ge 0$, $x \in D$ is called a *solution of* (2.1) – (2.3) if the following two properties hold for any integer $k \ge 1$.

(i) The restriction of $u(t,x)$ to the interval $I_k := [k-1,k)$ belongs to the space $C(I_k, L^2(D)) \cap L^2(I_k, H_0^1(D))$ and satisfies Eq. (2.7).

(ii) There is a limit $\lim_{t\to k^-} u(t,\cdot) = u_k^-$, and $u_k = u_k^- + \eta_k$, where $u_k = u(k,\cdot)$ (see Figure 1).

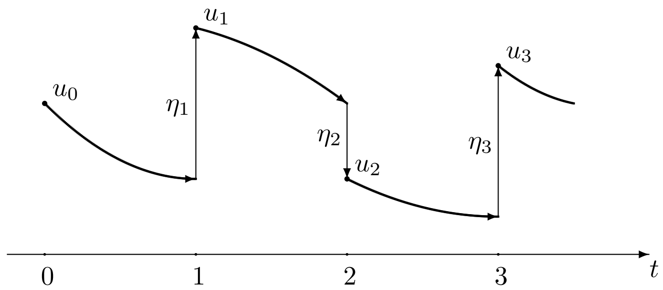

Figure 1: Evolution defined by Equations (2.1) – (2.2)

Let us fix an arbitrary function $u_0 \in L^2(D)$ and consider the initial-boundary value problem (2.1) – (2.3), (2.9). Recall that we denote by $S_t(u_0, h)$ the resolving operator for (2.7) – (2.9) and set $S(u_0) = S_1(u_0, h)$. Thus, S is a continuous operator in $L^2(D)$.

Theorem 2.7. *Suppose that the conditions of Theorem 2.4 are satisfied. Then, for any $u_0 \in L^2(D)$, the problem* (2.1) – (2.3), (2.9) *has a unique solution $u(t,x)$ in the sense of Definition 2.6. Moreover, for any integer $k \geq 1$, we have*

$$u_k = S(u_{k-1}) + \eta_k. \tag{2.26}$$

Proof. For integer values of t, we define $u(t,x)$ inductively by relation (2.26), and, for $t \in [k, k+1)$, we set $u(t,\cdot) = S_{t-k}(u_k)$. The resulting function is the unique solution of the problem in question. □

2.3 Markov chain associated with the problem (2.1) − (2.3)

From now on, we shall study the discrete-time random dynamical system (RDS)

$$u_k = S(u_{k-1}) + \eta_k, \tag{2.27}$$

$$u_0 = u, \tag{2.28}$$

where $\{\eta_k\}$ is a sequence of independent identically distributed (i.i.d.) random variables valued in the space $H = L^2(D)$ and $u = u(x)$ is an initial (random) function. We shall sometimes write $u_k(u)$ to indicate the dependence of the trajectory on the initial function u.

Theorem 2.8. *Let $\{u_k\}$ be a sequence defined by* (2.27), (2.28), *where u is an H-valued random variable independent of $\{\eta_k, k \geq 1\}$. Then it satisfies the Markov property. Namely, for any integers $k, n \geq 0$ and any bounded measurable function $f\colon H \to \mathbb{R}$, we have*

$$\mathbb{E}\left(f(u_{k+n}) \,|\, \mathcal{F}_k\right) = \left(\mathbb{E}\, f(u_n(v))\right)\big|_{v=u_k}, \tag{2.29}$$

where $\mathcal{F}_k$ is the σ-algebra generated[2] *by $\eta_1, \ldots, \eta_k$ and u, and the equality holds almost surely.*

Proof. We shall write $u_k(u) = u_k(u; \eta_1, \ldots, \eta_k)$ to indicate the dependence of the trajectory of (2.27), (2.28) on the random variables $\{\eta_m\}$. We have

$$u_{k+n}(u; \eta_1, \ldots, \eta_{k+n}) = u_n(u_k(u); \eta_{k+1}, \ldots, \eta_{k+n}).$$

Since $u_k(u)$ is $\mathcal{F}_k$-measurable and $\{\eta_i,\, i \geq k+1\}$ is independent of $\mathcal{F}_k$, it follows from the above relation and Exercise 1.15 that

$$\mathbb{E}\left(f(u_{k+n}(u)) \,|\, \mathcal{F}_k\right) = \left(\mathbb{E}\, f(u_n(v; \eta_{k+1}, \ldots, \eta_{k+n})\right)\big|_{v=u_k(u)}. \tag{2.30}$$

Now note that the distributions of the vectors $(\eta_{k+1}, \ldots, \eta_{k+n})$ and $(\eta_1, \ldots, \eta_n)$ coincide. Therefore,

$$\mathbb{E}\, f(u_n(v; \eta_{k+1}, \ldots, \eta_{k+n}) = \mathbb{E}\, f(u_n(v; \eta_1, \ldots, \eta_n)),$$

where $v \in H$ is an arbitrary deterministic function. Substitution of the above relation into the right-hand side of (2.30) completes the proof of (2.29). □

[2] We denote by $\mathcal{F}_0$ the σ-algebra generated by u.

Exercise 2.9. In the notation of Theorem 2.8, show that, if $f\colon H\times\cdots\times H\to\mathbb{R}$ is a bounded measurable function of $n+1$ arguments, then

$$\mathbb{E}\left(f(u_k,u_{k+1},\dots,u_{k+n})\,|\,\mathcal{F}_k\right)=\left(\mathbb{E}\,f(v,u_1(v),\dots,u_n(v))\right)\big|_{v=u_k}.$$

The Markov property implies two important corollaries. To formulate them, we introduce the *transition function* for the RDS (2.27). Namely, for any deterministic function $v\in H$ and any integer $k\geq 0$, we denote by $P_k(v,\cdot)$ the distribution of $u_k(v)$:

$$P_k(v,\Gamma)=\mathbb{P}\{u_k(v)\in\Gamma\},\quad \Gamma\in\mathcal{B}_H. \tag{2.31}$$

Corollary 2.10. *Let $u(x)$ be an H-valued random variable independent of $\{\eta_k\}$ and let μ be the distribution of u. Then the distribution of $u_k=u_k(u)$ is given by the formula*

$$\mathcal{D}(u_k)(\Gamma)=\int_H P_k(v,\Gamma)\mu(dv). \tag{2.32}$$

In particular, the measure $\mathcal{D}(u_k)$ depends only on μ (but not on the random variable u).

Proof. Let us fix an arbitrary Borel set $\Gamma\subset H$. In view of relation (2.29) with $f(z)=I_\Gamma(z)$, we have

$$\mathbb{E}\,I_\Gamma(u_k)=\mathbb{E}\left\{\mathbb{E}\left(I_\Gamma(u_k)\,|\,\mathcal{F}_0\right)\right\}=\mathbb{E}\left\{\left(\mathbb{E}\,I_\Gamma(u_k(v))\right)\big|_{v=u_0}\right\}.$$

It remains to note that $\mathbb{E}\,I_\Gamma(u_k(v))=\mathbb{P}\{u_k(v)\in\Gamma\}=P_k(u,\Gamma)$. □

Corollary 2.11. *The transition function $P_k(v,\Gamma)$ satisfies the Chapman–Kolmogorov relation. Namely, for any $k,n\geq 0$, $v\in H$, and $\Gamma\in\mathcal{B}_H$, we have*

$$P_{k+n}(v,\Gamma)=\int_H P_k(v,dz)P_n(z,\Gamma). \tag{2.33}$$

Proof. In view of (2.29), we have

$$\begin{aligned}P_{k+n}(v,\Gamma)&=\mathbb{E}\,I_\Gamma(u_{k+n}(v))=\mathbb{E}\left\{\mathbb{E}\left(I_\Gamma(u_{k+n}(v))\,|\,\mathcal{F}_k\right)\right\}\\&=\mathbb{E}\left\{\mathbb{E}\left(I_\Gamma(u_n(z))\right)\big|_{z=u_k(v)}\right\}=\mathbb{E}\left\{P_n(u_k(v),\Gamma)\right\}.\end{aligned}$$

This expression coincides with the integral on the right-hand side of (2.33). □

We now introduce the *Markov semigroups* corresponding to the transition function (2.31):

$$\mathfrak{P}_k\colon C_b(H)\to C_b(H),\qquad \mathfrak{P}_kf(v)=\int_H P_k(v,dz)f(z),$$
$$\mathfrak{P}_k^*\colon \mathcal{P}(H)\to\mathcal{P}(H),\qquad \mathfrak{P}_k^*\mu(\Gamma)=\int_H P_k(v,\Gamma)\mu(dv).$$

Exercise 2.12. Show that the operators $\mathfrak{P}_k$ and $\mathfrak{P}_k^*$ are well defined. Show also that they form semigroups, that is, $\mathfrak{P}_0 = \mathrm{Id}$ and $\mathfrak{P}_{k+n} = \mathfrak{P}_n \circ \mathfrak{P}_k$, and similarly for $\mathfrak{P}_k^*$. *Hint:* Use the Chapman–Kolmogorov relation (2.33).

Exercise 2.13. Show that $\mathfrak{P}_k$ and $\mathfrak{P}_k^*$ are dual semigroups in the sense that

$$(\mathfrak{P}_k f, \mu) = (f, \mathfrak{P}_k^* \mu) \quad \text{for any } f \in C_b(H),\ \mu \in \mathcal{P}(H).$$

2.4 Strong Markov property

Let u_k be a trajectory of (2.27), (2.28). Recall that $\mathcal{F}_k$ denotes the σ-algebra generated by $u, \eta_1, \dots, \eta_k$. We shall assume that u is independent of $\{\eta_k, k \ge 1\}$.

Definition 2.14. An integer-valued non-negative random variable τ is called a *stopping time* for the RDS (2.27), (2.28) if the event $\{\tau \le k\}$ belongs to $\mathcal{F}_k$ for any $k \ge 0$.

Exercise 2.15. **(i)** Show that if τ and σ are stopping times, then the random variables $\tau + \sigma$, $\tau \wedge \sigma$, and $\tau \vee \sigma$ are also stopping times.

(ii) Let $r > 0$ a constant. Show that the random variable

$$\tau_u(r) = \min\{k \ge 0 : \|u_k(u)\| \le r\}$$

is a stopping time.

For any stopping time τ, we shall denote by $\mathcal{F}_\tau$ the σ-algebra of the events $\Gamma \in \mathcal{F}$ such that

$$\Gamma \cap \{\tau \le k\} \in \mathcal{F}_k \quad \text{for any } k \ge 0.$$

Exercise 2.16. Show that if a random variable τ is a stopping time, then it is $\mathcal{F}_\tau$-measurable.

The following theorem establishes the strong Markov property for trajectories of (2.27) (cf. Theorem 2.8).

Theorem 2.17. *Let $f : H \to \mathbb{R}$ be a bounded measurable function and let τ be an almost surely finite stopping time. Then for any integer $k \ge 0$, we have*

$$\mathbb{E}\left(f(u_{\tau+k}(u)) \,|\, \mathcal{F}_\tau\right) = \left\{\mathbb{E} f(u_k(v))\right\}\big|_{v=u_\tau(u)} \quad \textit{almost surely.}$$

Outline of the proof. We need to show that, for any $\Gamma \in \mathcal{F}_\tau$,

$$\mathbb{E}\left(I_\Gamma f(u_{\tau+k}(u))\right) = \mathbb{E}\Big(I_\Gamma \left\{\mathbb{E} f(u_k(v))\right\}\big|_{v=u_\tau(u)}\Big), \tag{2.34}$$

where I_Γ is the indicator function of Γ. To prove (2.34), let us note that

$$I_\Gamma = \sum_{n=0}^{\infty} I_{\Gamma \cap \{\tau = n\}}. \tag{2.35}$$

Since $\Gamma_n := \Gamma \cap \{\tau = n\}$ belongs to $\mathcal{F}_n$, the Markov property implies that

$$\begin{aligned}\mathbb{E}\left(I_{\Gamma_n} f(u_{\tau+k}(u))\right) &= \mathbb{E}\left(I_{\Gamma_n} f(u_{n+k}(u))\right)\\ &= \mathbb{E}\left(I_{\Gamma_n} \mathbb{E}\{f(u_{n+k}(u)) \,|\, \mathcal{F}_n\}\right)\\ &= \mathbb{E}\left(I_{\Gamma_n} \{\mathbb{E}\, f(u_k(v))\}\big|_{v=u_n(u)}\right).\end{aligned}$$

Combining this relation with (2.35), we obtain (2.34). □

Exercise 2.18. **(i)** Fill the missing details of the proof.

(ii) Formulate and prove a generalisation of Theorem 2.16 in the spirit of Exercise 2.9.

3 Qualitative properties of solutions

In this section, we show that the time and ensemble averages of the L^2-norm of any trajectory for the RDS (2.27), (2.28) is bounded, for sufficiently large k, by a constant not depending on the initial condition. We also introduce the concept of a stationary measure and prove its existence.

3.1 Stabilisation of the time and ensemble averages

Recall that we set $H = L^2(D)$. In what follows, we shall need the following inequality for S, which is a straightforward consequence of (2.14):

$$\|S(u)\|^2 \leq q^2\|u\|^2 + C^2, \tag{3.1}$$

where $u \in H$ is an arbitrary function and $q = e^{-\frac{\alpha}{2}} < 1$.

Let us set $\mathfrak{m}_p = \mathbb{E}\,|\eta_k|^p$, $p \geq 1$. The following establish an upper bound for the ensemble average of the L^2-norm of trajectories.

Theorem 3.1. *Suppose that the conditions of Theorem 2.4 are satisfied and $\mathfrak{m}_1 < \infty$. Let $\{u_k\}$ be defined by (2.27), (2.28), where $u = u(x)$ is an H-valued random variable such that $\mathbb{E}\,\|u\| < \infty$. Then*

$$\mathbb{E}\,\|u_k\| \leq q^k \mathbb{E}\,\|u\| + \tfrac{\mathfrak{m}_1 + C}{1-q}, \quad k \geq 1. \tag{3.2}$$

Proof. We note that

$$\mathbb{E}\,\|u_k\| \leq \mathbb{E}\,\|S(u_{k-1})\| + \mathbb{E}\,\|\eta_k\| \leq q\mathbb{E}\,\|u_{k-1}\| + \mathfrak{m}_1 + C,$$

where we used (3.1). Iteration of this inequality results in (3.2). □

For any trajectory $\{u_k\}$ of the RDS (2.27) and any integer $k \geq 1$, we set

$$\langle \|u_l\|^2 \rangle_1^k := \frac{1}{k} \sum_{l=1}^{k} \|u_l\|^2.$$

Theorem 3.2. *Suppose the conditions of Theorem 2.4 are satisfied and* $\{u_k\}$ *is defined by* (2.27), (2.28), *where* η_k, $k \geq 1$, *are* H*-valued i.i.d. random variables and* $u = u(x)$ *is a random function in* H *independent of* $\{\eta_k, k \geq 1\}$*. Assume that*

$$\mathfrak{m}_p < \infty, \quad \mathbb{E}\,\|u\|^p < \infty \quad \text{for all } p \geq 1.$$

Then there is a constant $M > 0$ *not depending on* u *and an integer-valued random variable* $T_u \geq 1$ *such that*

$$\langle \|u_l\|^2 \rangle_1^k \leq M \quad \text{for } k \geq T_u, \tag{3.3}$$

$$\mathbb{E}\,T_u^q < \infty \quad \text{for all } q \geq 1. \tag{3.4}$$

Proof. Let us fix an arbitrary $\varepsilon > 0$. It follows from (3.1) and the Cauchy inequality that

$$\begin{aligned} \|u_l\|^2 &= \|S(u_{l-1})\|^2 + \|\eta_l\|^2 + 2(S(u_{l-1}), \eta_l) \\ &\leq (1+\varepsilon)\|S(u_{l-1})\|^2 + (1+\varepsilon^{-1})\|\eta_l\|^2 \\ &\leq (1+\varepsilon)q^2\|u_{l-1}\|^2 + (1+\varepsilon^{-1})\|\eta_l\|^2. \end{aligned} \tag{3.5}$$

Choosing $\varepsilon > 0$ so small that $\gamma := (1+\varepsilon)q^2 < 1$ and summing up inequalities (3.5) for $1 \leq l \leq k$, we obtain

$$\sum_{l=1}^{k} \|u_l\|^2 \leq \gamma \sum_{l=0}^{k-1} \|u_l\|^2 + (1+\varepsilon^{-1}) \sum_{l=1}^{k} \|\eta_l\|^2,$$

whence it follows that

$$\langle \|u_l\|^2 \rangle_1^k \leq \frac{a\|u\|^2}{k} + b\mathfrak{m}_2 + \frac{b}{k}\sum_{l=1}^{k} \xi_l, \tag{3.6}$$

where we set

$$a = \frac{\gamma}{1-\gamma}, \quad b = \frac{1+\varepsilon}{\varepsilon(1-\gamma)}, \quad \xi_l = \|\eta_l\|^2 - \mathfrak{m}_2.$$

We now need the following lemma, whose proof is given below.

Lemma 3.3. *Let* ξ_l, $l \geq 1$, *be i.i.d scalar random variables such that* $\mathbb{E}\,\xi_l = 0$ *and* $\mathbb{E}\,|\xi_l|^p < \infty$ *for any* $p \geq 1$*. Then there is a random integer* $T \geq 1$ *such that*

$$\frac{1}{k}\left|\sum_{l=1}^{k} \xi_l\right| \leq 1 \quad \text{for } k \geq T, \tag{3.7}$$

$$\mathbb{E}\,T^q < \infty \quad \text{for all } q \geq 1. \tag{3.8}$$

Let us set

$$M = b(\mathfrak{m}_2 + 1) + 1, \quad T_u = T \vee (a\|u\|^2).$$

Using (3.6) – (3.8) it is not difficult to verify that the required inequalities (3.3) and (3.4) hold. □

Proof of Lemma 3.3. Step 1. We first show that

$$M_k(p) := \mathbb{E}\left|\sum_{l=1}^{k}\xi_l\right|^{2p} \le C_p\,\mathbb{E}\,|\xi_1|^{2p}k^p \quad \text{for any } p\ge 1, \tag{3.9}$$

where $C_p>0$ is a constant not depending on k. Indeed, since the mean value of ξ_l is zero, using the Hölder inequality, we derive

$$\begin{aligned} M_k(p) &= \mathbb{E}\sum_{l_1,\dots,l_{2p}}\xi_{l_1}\cdots\xi_{l_{2p}} \le c_p\sum_{m_i,s_i}\mathbb{E}\,\bigl|\xi_{m_1}^{s_1}\cdots\xi_{m_p}^{s_p}\bigr| \\ &\le c_p\sum_{m_i,s_i}\Bigl(\mathbb{E}\,\bigl|\xi_{m_1}\bigr|^{2p}\Bigr)^{\frac{s_1}{2p}}\cdots\Bigl(\mathbb{E}\,\bigl|\xi_{m_p}\bigr|^{2p}\Bigr)^{\frac{s_p}{2p}} \le C_p\mathbb{E}\,\bigl|\xi_i\bigr|^{2p}k^p, \end{aligned} \tag{3.10}$$

where the second and third sums extend over all m_i and s_i such that $1\le m_i\le k$, $s_i\ge 0$, $s_1+\cdots+s_p=2p$.

Exercise 3.4. Justify the first and third inequalities in (3.10).

Step 2. We now set

$$T=\min\Bigl\{n\ge 1: \frac{1}{k}\Bigl|\sum_{l=1}^{k}\xi_l\Bigr|\le 1 \text{ for } k\ge n\Bigr\}.$$

The definition of T implies that, for any $N\ge 1$,

$$\mathbb{P}\{T=\infty\}\le\sum_{k=N}^{\infty}\mathbb{P}\Bigl\{\frac{1}{k}\Bigl|\sum_{l=1}^{k}\xi_l\Bigr|>1\Bigr\}\le\sum_{k=N}^{\infty}C_2\mathbb{E}\,|\xi_m|^4k^{-2}\le C\,N^{-1},$$

where we used (3.9) with $p=2$ and the Chebyshev inequality. Passing to the limit as $N\to+\infty$, we conclude that $\mathbb{P}\{T=\infty\}=0$.

Let us estimate the moments of T. Using inequality (3.9) with $p=q+2$, we derive

$$\begin{aligned} \mathbb{E}\,T^q &= \sum_{n=1}^{\infty}\mathbb{P}\{T=n\}\,n^q \le 1+\sum_{k=1}^{\infty}\mathbb{P}\Bigl\{\frac{1}{k}\Bigl|\sum_{l=1}^{k}\xi_l\Bigr|>1\Bigr\}\,(k+1)^q \\ &\le 1+C_p\mathbb{E}\,|\xi_1|^{2p}\sum_{k=1}^{\infty}k^{-p}(k+1)^q<\infty. \end{aligned}$$

This completes the proof of the lemma. □

3.2 Existence of stationary measures

Definition 3.5. A measure $\mu\in\mathcal{P}(H)$ is said to be *stationary* for the RDS (2.27) if $\mathfrak{P}_1^*\mu=\mu$.

Let us note that, if $\mu \in \mathcal{P}(H)$ is a stationary measure and $u(x)$ is a random function in H with distribution μ, then the distribution of the trajectory u_k for (2.27), (2.28) coincides with μ for any $k \geq 1$. This assertion is a straightforward consequence of relation (2.32) and Definition 3.5.

Theorem 3.6. *Suppose that $\mathfrak{m}_1 < \infty$. Then the RDS (2.27) has at least one stationary measure.*

Proof. We shall apply the classical Bogolyubov–Krylov argument.

Step 1. Let u_k be the trajectory of (2.27), (2.28) with $u \equiv 0$ and let μ_k be the distribution of u_k. We set

$$\bar{\mu}_k = \frac{1}{k} \sum_{l=0}^{k-1} \mu_l.$$

Suppose we have shown that the sequence $\{\mu_k\}$ is relatively compact in the space $\mathcal{P}(H)$ endowed with the metric $\|\cdot\|_{\mathcal{L}}^*$ (see (1.18)). Then there is a subsequence μ_{k_m} and a measure $\mu \in \mathcal{P}(H)$ such that $\mu_{k_m} \rightharpoonup \mu$ as $m \to +\infty$, where $\rightharpoonup$ stands for convergence with respect to the metric $\|\cdot\|_{\mathcal{L}}^*$. We claim that μ is a stationary measure. Indeed, for any $f \in \mathcal{L}(H)$, we have

$$\begin{aligned} (f, \mathfrak{P}_1^* \mu) &= \lim_{m \to \infty} \bigl(f, \mathfrak{P}_1^* \bar{\mu}_{k_m}\bigr) = \lim_{m \to \infty} \frac{1}{k_m} \sum_{l=0}^{k_m - 1} (f, \mathfrak{P}_1^* \mu_l) \\ &= \lim_{m \to \infty} \Bigl\{ \bigl(f, \bar{\mu}_{k_m}\bigr) - \frac{1}{k_m} \bigl(f, \mu_0\bigr) + \frac{1}{k_m} \bigl(f, \mu_{k_m}\bigr) \Bigr\} = (f, \mu). \end{aligned} \tag{3.11}$$

Since this relation is true for any $f \in \mathcal{L}(H)$, we conclude that $\mathfrak{P}_1^* \mu = \mu$.

Step 2. Let us show that $\{\mu_k\}$ is relatively compact. We resort to the following assertion due to Prokhorov (see Theorem 11.5.4 in [Dud02]).

Proposition 3.7. *A family $\{\mu_\alpha\}$ of probability Borel measures on a Polish space is relatively compact iff for any $\varepsilon > 0$ there is a compact subset K_ε such that $\mu_\alpha(K_\varepsilon) \geq 1 - \varepsilon$ for any α.*

We shall show that for any $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset H$ such that $\mu_k(K_\varepsilon) \geq 1 - \varepsilon$ for any $k \geq 1$. This will imply that $\{\bar{\mu}_k\}$ is relatively compact.

Since $u_k = S(u_{k-1}) + \eta_k$, the required assertion will be established if we prove that

$$\mathbb{P}\bigl\{S(u_{k-1}) \notin K_\varepsilon^1\bigr\} \leq \varepsilon/2, \quad \mathbb{P}\bigl\{\eta_k \notin K_\varepsilon^2\bigr\} \leq \varepsilon/2. \tag{3.12}$$

where K_ε^1 and K_ε^2 are compact sets in H. (We can take $K_\varepsilon = K_\varepsilon^1 + K_\varepsilon^2$.)

Step 3. It follows from (3.2) that $\mathbb{E}\,\|u_k\| \leq (\mathfrak{m}_1 + C)(1-q)^{-1}$ for all $k \geq 1$. Therefore we can choose $R_\varepsilon > 0$ so large that

$$\mathbb{P}\bigl\{\|u_{k-1}\| > R_\varepsilon\bigr\} \leq R_\varepsilon^{-1} \mathbb{E}\,\|u_{k-1}\| \leq \varepsilon/2. \tag{3.13}$$

Furthermore, since the embedding $H_0^1(D) \subset H$ is compact and S is Lipschitz continuous from H to $H_0^1(D)$, the image under S of any bounded set in H is relatively compact. Hence, setting $K_\varepsilon^1 = S\big(B_H(R_\varepsilon)\big)$, from (3.13) we derive

$$\mathbb{P}\big\{S(u_{k-1}) \notin K_\varepsilon^1\big\} \le \mathbb{P}\big\{\|u_{k-1}\| > R_\varepsilon\big\} \le \varepsilon/2.$$

Finally, recall that, by Ulam's theorem, any probability Borel measure on a Polish space is regular (see Theorem 7.1.4 in [Dud02]). Hence, if χ is the distribution of η_k, then there is a compact set $K_\varepsilon^2 \subset H$ such that $\chi(K_\varepsilon^2) \ge 1 - \varepsilon/2$. This is equivalent to the second inequality in (3.12). □

Exercise 3.8. Justify the first relation in (3.11). *Hint:* Use assertion (ii) of Theorem 1.21.

Exercise 3.9. Show that any stationary measure μ has a finite moment, that is,

$$\int_H \|v\|\mu(dv) < \infty.$$

Hint: Use inequality (3.2) and Fatou's lemma (see Lemma 4.3.3 in [Dud02]).

4 Ergodicity for finite-dimensional systems

In this section, we consider Markov chains with finite-dimensional phase space and discuss the uniqueness of stationary measure and its mixing properties. We first formulate the main result and reduce its proof to an inequality for the semigroup $\mathfrak{P}_k^*$. In Section 4.2, we introduce the concept of maximal coupling and establish its existence. The maximal coupling is a key ingredient of the proof of the above-mentioned inequality, which is given in Section 4.3.

4.1 Uniqueness of stationary measure and mixing

Let us consider the following RDS in $\mathbb{R}^n$:

$$u_k = S(u_{k-1}) + \eta_k, \tag{4.1}$$

$$u_0 = u. \tag{4.2}$$

Here $S \in C(\mathbb{R}^n, \mathbb{R}^n)$ is a given function and $\{\eta_k\}$ is a sequence of i.i.d. random variables in $\mathbb{R}^n$ such that $\mathfrak{m}_1 = \mathbb{E}\,\|\eta_k\| < \infty$. As was shown in Section 2.3, the RDS (4.1) generates a family of Markov chains in $\mathbb{R}^n$, and we denote by $P_k(u,\Gamma)$ its transition function and by $\mathfrak{P}_k$ and $\mathfrak{P}_k^*$ the corresponding Markov semigroups.

Let $\mathcal{P}_1(\mathbb{R}^n)$ be the set of measures $\lambda \in \mathcal{P}(\mathbb{R}^n)$ such that

$$\mathfrak{m}(\lambda) := \int_{\mathbb{R}^n} \|u\|\,\lambda(du) < \infty.$$

Definition 4.1. We shall say that the RDS (4.1) is mixing if it has a unique stationary measure $\mu \in \mathcal{P}_1(\mathbb{R}^n)$, and for any $\lambda \in \mathcal{P}_1(\mathbb{R}^n)$,

$$\|\mathfrak{P}_k^*\lambda - \mu\|_{\mathrm{var}} \to 0 \quad \text{as } k \to +\infty. \tag{4.3}$$

Let us assume that the above-mentioned conditions hold, and the operator S satisfies the inequality

$$\|S(u)\| \le q\,\|u\| + C \quad \text{for all } u \in \mathbb{R}^n, \tag{4.4}$$

where $q < 1$ and C are positive constants. The following theorem shows that if the distribution of η_k is non-degenerate, then the RDS in question in mixing.

Theorem 4.2. *Suppose that the distribution of the random variables η_k has a continuous density ρ with respect to the Lebesgue measure, and that $\rho(x) > 0$ for almost all $x \in \mathbb{R}^n$. Then the RDS (4.1) is mixing in the sense of Definition 4.1.*

Proof. Step 1. Existence of a stationary measure can be established by the Bogolyubov–Krylov argument (see the proof of Theorem 3.6). Let us show that any stationary measure $\mu \in \mathcal{P}(\mathbb{R}^n)$ belongs to $\mathcal{P}_1(\mathbb{R}^n)$ and satisfies the inequality

$$\mathfrak{m}(\mu) \le \frac{\mathfrak{m}_1 + C}{1-q}, \tag{4.5}$$

where C and q are the constants in (4.4). Indeed, let us fix an arbitrary $R > 0$ and consider the function

$$f_R(u) = \begin{cases} \|u\|, & \|u\| \le R, \\ R & \|u\| > R. \end{cases}$$

Since $\mathfrak{P}_k^*\mu = \mu$ for any $k \ge 1$, we have

$$(f_R, \mu) = (f_R, \mathfrak{P}_k^*\mu) = (\mathfrak{P}_k f_R, \mu) = \int_{\mathbb{R}^n} \mathfrak{P}_k f_R(u)\mu(du). \tag{4.6}$$

Let us define u_k by (4.1), (4.2), where $u \in \mathbb{R}^n$ is fixed. It follows from (4.4) that (see the proof of (3.2))

$$\mathbb{E}\,\|u_k\| \le q^k\|u\| + \tfrac{\mathfrak{m}_1 + C}{1-q},$$

and, hence,

$$\mathfrak{P}_k f_R(u) = \mathbb{E}\, f_R(u_k) \le \mathbb{E}\,\|u_k\| \le q^k r + \tfrac{\mathfrak{m}_1 + C}{1-q} \quad \text{for } \|u\| \le r. \tag{4.7}$$

On the other hand,

$$\mathfrak{P}_k f_R(u) \le R \quad \text{for all } u \in \mathbb{R}^n. \tag{4.8}$$

Combining (4.6) – (4.8), we derive

$$\begin{aligned} (f_R, \mu) &= \int_{B_r} \mathfrak{P}_k f_R(u)\mu(du) + \int_{B_r^c} \mathfrak{P}_k f_R(u)\mu(du) \\ &\le \left(q^k r + \tfrac{\mathfrak{m}_1 + C}{1-q}\right)\mu(B_r) + R\mu(B_r^c), \end{aligned} \tag{4.9}$$

where $B_r \subset \mathbb{R}^n$ is the ball of radius r centred at zero. Letting $k \to +\infty$ and then $r \to +\infty$ in (4.9), we obtain

$$(f_R, \mu) \le \tfrac{\mathfrak{m}_1 + C}{1-q}.$$

Application of Fatou's lemma now results in (4.5).

Step 2. To prove the uniqueness of stationary measure and convergence (4.3), it suffices to establish the following assertion:

- For any $M > 1$ there is a constant $\gamma \in (0,1)$ such that

$$\|\mathfrak{P}_1^*\lambda_1 - \mathfrak{P}_1^*\lambda_2\|_{\mathrm{var}} \le \gamma\,\|\lambda_1 - \lambda_2\|_{\mathrm{var}}, \tag{4.10}$$

  where $\lambda_1, \lambda_2 \in \mathcal{P}_1(\mathbb{R}^n)$ are arbitrary measures satisfying the conditions

$$\mathfrak{m}(\lambda_1) \vee \mathfrak{m}(\lambda_2) \le M, \quad \|\lambda_1 - \lambda_2\|_{\mathrm{var}} \ge M^{-1}. \tag{4.11}$$

Indeed, suppose that above assertion is established. If $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^n)$ are two different stationary measures, then $\mathfrak{m}(\mu_i) < \infty$ for $i = 1,2$, and therefore there is $M > 0$ such that (4.11) holds with $\lambda_i = \mu_i$. Applying (4.10), we arrive at contradiction.

To prove (4.3), we fix an arbitrary measure $\lambda \in \mathcal{P}_1(\mathbb{R}^n)$ different from μ and set $d_k := \|\mathfrak{P}_k^*\lambda - \mu\|_{\mathrm{var}}$. We claim that d_k is a non-increasing sequence. Indeed, since μ is a stationary measure, in view of Theorem 1.17, we have

$$\begin{aligned} d_k &= \tfrac{1}{2}\sup_f \bigl|(f, \mathfrak{P}_k^*\lambda) - (f,\mu)\bigr| = \tfrac{1}{2}\sup_f \bigl|(\mathfrak{P}_1 f, \mathfrak{P}_{k-1}^*\lambda) - (\mathfrak{P}_1 f, \mu)\bigr| \\ &\le \tfrac{1}{2}\sup_g \bigl|(g, \mathfrak{P}_{k-1}^*\lambda) - (g,\mu)\bigr| = d_{k-1}, \end{aligned} \tag{4.12}$$

where the supremums are taken over all functions $f, g \in C_b(H)$ whose norm does not exceed 1.

Exercise 4.3. Show that $\mathfrak{P}_k$ possesses the *Feller property*, that is, if $f \in C_b(H)$, then $\mathfrak{P}_k f \in C_b(H)$. Use this fact to justify the inequality in (4.12).

Hence, it suffices to show that for any $\varepsilon > 0$ there is an integer $k \ge 0$ for which $d_k \le \varepsilon$. To this end, we note that, if $\mathfrak{m}(\lambda) \le M$, then (see (3.2))

$$\mathfrak{m}(\mathfrak{P}_k^*\lambda) \le q^k \mathfrak{m}(\lambda) + \tfrac{\mathfrak{m}_1 + C}{1-q}.$$

Therefore, if $M > \varepsilon^{-1}$ is sufficiently large, then $\mathfrak{m}(\mathfrak{P}_k^*\lambda) \le M$ for all $k \ge 1$ and $\|\lambda - \mu\|_{\mathrm{var}} \ge M^{-1}$. Inequality (4.10) now implies that $d_k \le \gamma^k$ as long as $d_{k-1} \ge M^{-1}$. It follows that there is $k \ge 1$ such that $d_k \le M^{-1} < \varepsilon$.

Step 3. To prove inequality (4.10), we first estimate the distance between the measures $\mathfrak{P}_1^*\delta_{u_1}$ and $\mathfrak{P}_1^*\delta_{u_2}$, where $u_1, u_2 \in \mathbb{R}^n$ are some points. It is easy to see that the function $\rho_u(x) = \rho(x - S(u))$ is the density of $\mathfrak{P}_1^*\delta_u$ with respect to the Lebesgue measure. Therefore, according to Proposition 1.18, we have

$$d(u_1, u_2) := \|\mathfrak{P}_1^*\delta_{u_1} - \mathfrak{P}_1^*\delta_{u_2}\|_{\mathrm{var}} = 1 - \int_{\mathbb{R}^n} \rho(u_1, u_2, x)\,dx, \tag{4.13}$$

where $\rho(u_1, u_2, x) = \rho_{u_1}(x) \wedge \rho_{u_2}(x)$. Since the density ρ is continuous and $\rho(x) > 0$ for almost all $x \in \mathbb{R}^n$, we see that

$$\min_{u_1, u_2 \in B_R} \int_{\mathbb{R}^n} \rho(u_1, u_2, x)\,dx > 0 \quad \text{for any } R > 0,$$

where $B_R \subset \mathbb{R}^n$ is the ball of radius R centred at zero. Combining this with (4.13), we conclude that

$$\max_{u_1,u_2\in B_R} d(u_1,u_2) \le 1-\delta_R, \tag{4.14}$$

where $\delta_R > 0$ is a constant.

Step 4. The proof of (4.10) is based on maximal coupling of measures. This concept is studied in the next subsection, and the proof of the theorem is concluded in Section 4.3. □

4.2 Maximal coupling of measures

Let X be a Polish space and let $\lambda_1, \lambda_2 \in \mathcal{P}(X)$.

Definition 4.4. A pair of random variables (ξ_1, ξ_2) defined on the same probability space is called a *coupling for* (λ_1, λ_2) if

$$\mathcal{D}(\xi_1) = \lambda_1, \quad \mathcal{D}(\xi_2) = \lambda_2.$$

Let (ξ_1, ξ_2) be a coupling for (λ_1, λ_2). Then for any $\Gamma \in \mathcal{B}_X$ we have

$$\begin{aligned}\lambda_1(\Gamma) - \lambda_2(\Gamma) &= \mathbb{P}\{\xi_1 \in \Gamma\} - \mathbb{P}\{\xi_2 \in \Gamma\} \\ &= \mathbb{E}\big\{I_{\{\xi_1\ne\xi_2\}}\big(I_\Gamma(\xi_1) - I_\Gamma(\xi_2)\big)\big\} \le \mathbb{P}\{\xi_1 \ne \xi_2\},\end{aligned}$$

whence it follows that

$$\mathbb{P}\{\xi_1 \ne \xi_2\} \ge \|\lambda_1 - \lambda_2\|_{\rm var}.$$

Definition 4.5. A coupling (ξ_1, ξ_2) for (λ_1, λ_2) is said to be *maximal* if

$$\mathbb{P}\{\xi_1 \ne \xi_2\} = \|\lambda_1 - \lambda_2\|_{\rm var}.$$

Theorem 4.6. *For any pair of measures $\lambda_1, \lambda_2 \in \mathcal{P}(X)$, there is a maximal coupling.*

Proof. If $\delta := \|\lambda_1 - \lambda_2\|_{\rm var} = 1$, then any pair (ξ_1, ξ_2) of independent random variables with $\mathcal{D}(\xi_i) = \lambda_i$, $i = 1, 2$, is a maximal coupling for (λ_1, λ_2). If $\delta = 0$, then $\lambda_1 = \lambda_2$, and for any random variable ξ with distribution λ_1 the pair (ξ, ξ) is a maximal coupling. Hence, we can assume that $0 < \delta < 1$.

Let $m(du)$ be a measure satisfying the conditions of Proposition 1.18 and let

$$\rho_i = \frac{d\lambda_i}{dm}, \quad \rho = \rho_1 \wedge \rho_2, \quad \hat\rho_i = \delta^{-1}(\rho_i - \rho).$$

Direct verification shows that $\hat\lambda_i = \hat\rho_i dm$ and $\mu = (1-\delta)^{-1}\rho\, dm$ are probability measures on X. Let ζ_1, ζ_2, ζ, and α be independent random variables defined on the same probability space such that

$$\mathcal{D}(\zeta_i) = \hat\lambda_i, \quad \mathcal{D}(\zeta) = \mu, \quad \mathbb{P}\{\alpha = 1\} = 1-\delta, \quad \mathbb{P}\{\alpha = 0\} = \delta.$$

We claim that the random variables

$$\xi_i = \alpha\zeta + (1-\alpha)\zeta_i, \quad i = 1, 2, \tag{4.15}$$

form a maximal coupling for (λ_1, λ_2). Indeed, for any $\Gamma \in \mathcal{B}_X$, we have

$$\begin{aligned}\mathbb{P}\{\xi_i \in \Gamma\} &= \mathbb{P}\{\xi_i \in \Gamma, \alpha = 0\} + \mathbb{P}\{\xi_i \in \Gamma, \alpha = 1\} \\ &= \mathbb{P}\{\alpha = 0\}\mathbb{P}\{\zeta_i \in \Gamma\} + \mathbb{P}\{\alpha = 1\}\mathbb{P}\{\zeta \in \Gamma\} \\ &= \delta \int_\Gamma \hat\rho_i(u)\, dm + \int_\Gamma \rho(u)\, dm = \lambda_i(\Gamma),\end{aligned}$$

where we used the independence of $(\zeta_1, \zeta_2, \zeta, \alpha)$, the relation $\rho_i = \rho + \delta\hat\rho_i$, and also the fact that $\xi_i = \zeta_i$, $i = 1, 2$, on the set $\{\alpha = 0\}$ and $\xi_1 = \xi_2 = \zeta$ on the set $\{\alpha = 1\}$. Furthermore,

$$\begin{aligned}\mathbb{P}\{\xi_1 \neq \xi_2\} &= \mathbb{P}\{\xi_1 \neq \xi_2, \alpha = 0\} + \mathbb{P}\{\xi_1 \neq \xi_2, \alpha = 1\} \\ &= \mathbb{P}\{\alpha = 0\}\mathbb{P}\{\zeta_1 \neq \zeta_2\} = \delta,\end{aligned}$$

where we used again the independence of $(\zeta_1, \zeta_2, \zeta, \alpha)$ and also the relation

$$\mathbb{P}\{\zeta_1 = \zeta_2\} = \delta^{-2} \iint\limits_{\{u_1 = u_2\}} \big(\rho_1(u_1) - \rho(u_1)\big)\big(\rho_2(u_2) - \rho(u_2)\big)\, m(du_1)m(du_2) = 0.$$

This completes the proof of Theorem 4.6. □

4.3 Conclusion of the proof of Theorem 4.2

Step 1. We need to prove inequality (4.10). Let us fix two measures $\lambda_1, \lambda_2 \in \mathcal{P}_1(\mathbb{R}^n)$ satisfying (4.11) and denote by (ξ_1, ξ_2) their maximal coupling constructed in Theorem 4.6 (see (4.15)). Note that, for any Borel subset $\Gamma \subset \mathbb{R}^n$, we have

$$\mathfrak{P}_1^*\lambda_i(\Gamma) = \int_{\mathbb{R}^n} P_1(u, \Gamma)\lambda_i(du) = \mathbb{E}\, P_1(\xi_i, \Gamma), \quad i = 1, 2.$$

It follows that

$$\|\mathfrak{P}_1^*\lambda_1 - \mathfrak{P}_1^*\lambda_2\|_{\mathrm{var}} \le \mathbb{E}\, d(\xi_1, \xi_2) = \mathbb{E}\,\big(I_{\{\xi_1 \neq \xi_2\}} d(\xi_1, \xi_2)\big),$$

where $d(u_1, u_2)$ is defined in (4.13). By virtue of (4.15), we have $\xi_i = \zeta_i$ on the set $\{\xi_1 \neq \xi_2\} = \{\alpha = 0\}$. This implies that

$$\begin{aligned}\|\mathfrak{P}_1^*\lambda_1 - \mathfrak{P}_1^*\lambda_2\|_{\mathrm{var}} &\le \mathbb{E}\,\big\{I_{\{\alpha=0\}}\, d(\zeta_1, \zeta_2)\big\} = \mathbb{P}\{\alpha = 0\}\, \mathbb{E}\, d(\zeta_1, \zeta_2) \\ &= \|\lambda_1 - \lambda_2\|_{\mathrm{var}}\, \mathbb{E}\, d(\zeta_1, \zeta_2),\end{aligned} \tag{4.16}$$

where we used the independence of $(\alpha, \zeta_1, \zeta_2)$ and also the fact that (ξ_1, ξ_2) is a maximal coupling for (λ_1, λ_2). Thus, it suffices to show that, for any two measures $\lambda_1, \lambda_2 \in \mathcal{P}_1(\mathbb{R}^n)$ satisfying (4.11) for some $M > 1$, we have

$$\mathbb{E}\, d(\zeta_1, \zeta_2) \le \gamma, \tag{4.17}$$

where $\gamma \in (0,1)$ is a constant depending on M.

Step 2. To prove (4.17), let us fix a constant $R > 0$ and set

$$\Gamma_R = \{\|\xi_1\| \vee \|\xi_2\| \le R\}. \tag{4.18}$$

We have

$$\begin{aligned}\mathbb{E}\, d(\zeta_1, \zeta_2) &= \mathbb{E}\big\{ I_{\{\alpha=1\}} d(\zeta_1, \zeta_2) + I_{\{\alpha=0\}} d(\zeta_1, \zeta_2) \big\} \\ &= \mathbb{P}\{\alpha = 1\}\, \mathbb{E}\, d(\zeta_1, \zeta_2) + \mathbb{E}\big\{ I_{\{\alpha=0\}} \big(I_{\Gamma_R} + I_{\Gamma_R^c} \big)\, d(\xi_1, \xi_2) \big\}. \end{aligned} \tag{4.19}$$

Recalling inequality (4.14), we see that

$$\mathbb{E}\big\{ I_{\{\alpha=0\}} I_{\Gamma_R} d(\xi_1, \xi_2) \big\} \le (1 - \delta_R) \mathbb{P}\big(\{\alpha = 0\} \cap \Gamma_R \big).$$

Substituting this into (4.19) and using the inequality $d(u_1, u_2) \le 1$, we derive

$$\begin{aligned}\mathbb{E}\, d(\zeta_1, \zeta_2) &\le \mathbb{P}\{\alpha = 1\} + \mathbb{P}\big(\{\alpha = 0\} \cap \Gamma_R^c \big) + (1 - \delta_R) \mathbb{P}\big(\{\alpha = 0\} \cap \Gamma_R \big) \\ &\le 1 - \delta_R \mathbb{P}\big(\{\alpha = 0\} \cap \Gamma_R \big). \end{aligned} \tag{4.20}$$

Furthermore, it follows from condition (4.11) and the Chebyshev inequality that, if $R = 4M^2$, then

$$\begin{aligned}\mathbb{P}(\{\alpha = 0\}^c) &= \mathbb{P}(\{\alpha = 1\}) = 1 - \|\lambda_1 - \lambda_2\|_{\mathrm{var}} \le 1 - M^{-1}, \\ \mathbb{P}(\Gamma_R^c) &\le \mathbb{P}\{\|\xi_1\| > R\} + \mathbb{P}\{\|\xi_2\| > R\} \le 2MR^{-1} = (2M)^{-1}. \end{aligned}$$

Thus, we have

$$\mathbb{P}\big(\{\alpha = 0\} \cap \Gamma_R \big) \ge 1 - \mathbb{P}(\{\alpha = 0\}^c) - \mathbb{P}(\Gamma_R^c) \ge (2M)^{-1}.$$

Comparing this with (4.20), we arrive at the required inequality (4.17). This completes the proof of Theorem 4.2.

References

[Dud02] R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2002.

[GT01] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 2001.

[Hen81] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Springer-Verlag, Berlin, 1981.

[Yos95] K. Yosida, *Functional Analysis*, Springer-Verlag, Berlin, 1995.