XXX ESCUELA VENEZOLANA DE MATEMÁTICAS
EMALCA–VENEZUELA 2017

THE STOCHASTIC ALLEN-CAHN EQUATION WITH SMALL LÉVY PERTURBATIONS

Michael A. Högele

MÉRIDA, VENEZUELA, 03 al 08 de septiembre de 2017

XXX ESCUELA VENEZOLANA DE MATEMÁTICAS
EMALCA–VENEZUELA 2017

The stochastic Allen-Cahn equation with small Lévy Perturbations

Michael A. Högele
Universidad de los Andes, Bogotá, Colombia
ma.hoegele@uniandes.edu.co

MÉRIDA, 3 AL 8 DE SEPTIEMBRE DE 2017

XXX ESCUELA VENEZOLANA DE MATEMÁTICAS

La Escuela Venezolana de Matemáticas es una actividad de los postgrados en matemáticas de las instituciones siguientes: Centro de Estudios Avanzados del Instituto Venezolano de Investigaciones Científicas, Facultad de Ciencias de la Universidad Central de Venezuela, Facultad de Ciencias de la Universidad de Los Andes, Universidad Simón Bolívar, Universidad Centroccidental Lisandro Alvarado y Universidad de Oriente, y se realiza bajo el auspicio de la Asociación Matemática Venezolana.

La XXX Escuela Venezolana de Matemáticas recibió financiamiento de la Academia de Ciencias Físicas, Matemáticas y Naturales de Venezuela, el Fondo Nacional de Ciencia, Tecnología e Innovación (FONACIT), el Instituto Venezolano de Investigaciones Científicas (Centro de Estudios Avanzados, Departamento de Matemáticas y Ediciones IVIC), la Universidad de los Andes (CEP, CDCHT, Departamento de Matemáticas de la Facultad de Ciencias, Decanato de Ciencias y Vicerrectorado Administrativo), FUNDACITE–Mérida, Unión Matemática de América Latina y el Caribe (UMALCA) y Centre International de Mathématiques Pures et Appliquées (CIMPA).

2010 Mathematics Subject Classification:
60H15; 60G51; 60G52; 60G55; 35K05; 35K91; 35K57; 35K55; 37D15; 37L55.

Keywords:
first exit times; first exit loci; nonlinear reaction-diffusion equations; Morse-Smale property; small noise asymptotic; α-stable Lévy process; additive Lévy noise; regularly varying noise; stochastic heat equation with stable noise; stochastic Chafee-Infante equation with stable noise; stochastic Allen-Cahn equation

Instituto Venezolano de Investigaciones Científicas
Rif: G-20004206-0

The stochastic Allen-Cahn equation
with small Lévy Perturbations
Michael A. Högele

Diseño y edición: Escuela Venezolana de Matemáticas
Depósito legal DC2017001938
ISBN 978-980-261-181-2
Caracas, Venezuela
2017

für Cristina und Antonio

Contents

Preface

The mechanism of many physical phenomena is well-described by ordinary or partial differential equations giving rise to a dynamical system in finite or infinite dimensions. In many cases, however, on the time scales of interest these dynamical systems are subject to a (small) random residual perturbation inherited from the microdynamics of unresolved smaller time scales, see for instance [49].
In the case of finite dimensional dynamical systems perturbed by small Brownian motion there is a large literature on the qualitative behavior of the perturbed system in terms of large deviations theory going back to the pioneering work by Cramér in the 1940ies. Since then the field has grown enormously and is usually referred to as Freidlin-Wentzell theory. Classical texts comprise [3, 7, 17, 18, 21, 23, 27, 28, 50] and references therein. The asymptotic exit times of a dynamical system from the neighborhood of a stable state grows *exponentially* as a function of the inverse noise intensity with the precise exponents given as "minimal energy barriers" to cross and the exit paths evolve with overwhelming probability close to an energy minimizing trajectory.
In infinite dimensions the so-called Allen-Cahn equation is a particular case, where this is studied in detail is sometimes also referred to as Chafee-Infante equation, a dissipative non-linear parabolic partial differential equation (see for instance [34, 32]) subject to Gaussian noise and is often associated to the notion of *tunneling*. In the literature this phenomenon was treated in [5, 9, 10, 25, 26, 66]. In 2012 Stella Brassesco offered an introductory class at the XXV. Escuela Venezola de Matemáticas on the subject.
However, the Gaussian white noise structure is not the only possible probabilistic perturbation of interest. In general the class of limiting distributions and resulting processes is much larger and given by so-called Lévy processes, typically non-Gaussian processes with stationary and independent increments see for instance in [1, 61]. In the non-Gaussian case these processes exhibit jump discontinuities, which in particular cases, such as α-stable processes effectively change the dynamics. The literature on dynamical systems perturbed by this kind of discontinuous and heavy-tailed processes is much younger,

[40, 39, 42, 19, 20, 56, 36] and contains mainly perturbations of infinite dimensional dynamical systems. In [19, 20] the authors study the first exit times of the Allen-Cahn equation (referred to as Chafee-Infante equation in these publications) subject to additive α-stable noise.

The proof techniques rely on a carefully chosen flow decomposition of the driving noise splitting the large heavy-tailed jumps from the remaining noise, which occur only with finite intensity. Between (appropriately identified) large jumps of the random perturbation the process follows the deterministic trajectory with overwhelming (exponential) probability. In order to prove this result we combine the path-by-path version of the Burkholder-Davis-Gundy inequality by Siorpaes and Beiglböck [62], an estimate of the stochastic convolution by Salavati and Zanegeh [64] with Campbell's formula together with a method used in [36]. This result is much finer than the methods applied in [20]. In addition, the deterministic relaxation is sufficiently fast in order to occur on average before the next large jump, such that the main dynamics is described by large jump increments starting from neighborhoods of a stable state. Summing up this dynamics it turns out that the exit times behave *polynomially* as a function of the inverse noise intensity and the exit occurs due to one of the large jumps and hence yields a distribution (as a limiting object of the large jump distribution). Such a work is carried out for the Allen-Cahn equation in [20] perturbed by α-stable Lévy noise ($\alpha \in (0, 2)$) with additive noise. In [35], the results are generalized to multiplicative noise and general regularly varying noise and a generic class of dissipative scalar partial differential equations. This lecture course aims at presenting the results of [35] for the simpler case of additive noise.

The text is designed as a self-contained course taking seriously the prerequisites of the audience assumed as probability theory (including Kolmogorov's law of large numbers, the standard central limit theorem, characteristic functions), ordinary differential equations and dynamical systems and linear partial differential equations. As a consequence the course is divided in three parts:

The first part provides an introduction to general Lévy processes, the class of stochastic perturbations of interest. In Chapter 1 we introduce the laws of Lévy processes starting with central and local limit theorems, and the identification of the limiting distributions as infinitely divisible distributions and their characterization by the Lévy-Khinchin theorem. Chapter 2 is followed by the construction of general Lévy processes starting with the explicit construction of Q-Brownian motion and compound Poisson processes in separable Hilbert spaces. After the construction we establish the existence of càdlàg paths, the strong Markov property and the existence of moments for Lévy processes with bounded jumps. This paves the way to the derivation of the Lévy-Itô theorem, which characterizes Lévy processes as essentially a linear drift plus a Q-Brownian motion plus two Poisson random integrals. Chapter 2 finishes with the presentation of Burkholder-Davis-Gundy inequality, the essential estimate for pure jump processes.

The second part starts with a detailed discussion of the state space followed by an introduction to the dynamical system generated by the deterministic Allen-Cahn equation and a careful and detailed discussion of its dynamics. In Chapter 4 we establish the mild solution of the stochastic Allen-Cahn equation with additive Lévy noise and in particular the crucial strong Markov property.

The third part is entirely dedicated to the first exit problem, that is the asymptotic behavior of the first exit times and locus of the stochastic Allen-Cahn equation with small additive regularly varying Lévy noise. In Chapter 5 we present the general setting and a discussion of a "model" for each of the quantities. Chapter 6 is dedicated to the control of the "small" jumps stochastic Allen-Cahn equation between two consecutive "large" jumps, where we use all the machinery acquired in part 1. The final part contains the final book keeping of the large jumps dynamics.

The material in Part 1 and 2 is not original and can be found in many textbooks. Important sources of the presentation are among others [45], [57], [61], [22], [58], [20] and many more. Part 3 is a simplified version for the additive case of the article [35], where the first exit problem is treated for the stochastic Allen-Cahn equation with multiplicative regularly varying Lévy noise. On the channel of the Mathematics department of Universidad de los Andes, Bogotá, Colombia you can find the videos of this course.

`https://www.youtube.com/playlist?list=PLUH3ZFcFJmW4i7fSNSbujKVhP8tThYmwj`

Acknowledgements. I would like to thank S. Brassesco and C. DiPrisco for the kind invitation to the XXX. Escuela Venezolana de Matemáticas and the participants for the efforts to follow this video course. It is my pleasure to contribute to this distinguished institution of academic collaboration.

I also would like to thank the International Center for Theoretical Sciences, Bangalore, India, the new founded member institution of the Tata Institute of Fundamental Research Network, for the invitation to the first Summer Research Program on the Dynamics of Complex Systems in summer 2016, where I held a class on stochastic climate models with Lévy noise based on some of the material I present here. In particular I would like to thank A. Apte and his group at ICTS for their warm hospitality in Bangalore.

Michael A. Högele
Bogotá, December 2017

Part I

A short introduction to Lévy processes

Chapter 1

Limit theorems and infinite divisibility

1.1 Motivation: The central and the Poisson limit theorem

The sums of Bernoullis. Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ consider an i.i.d. family $(X_n)_{n\in\mathbb{N}}$ of Bernoulli random variables $X_n : \Omega \to \{0,1\}$, that is $\mathbb{P}_{X_n} = p\delta_0 + (1-p)\delta_1$ for some $p \in [0,1]$, where for all Borel sets $A \in \mathcal{B}(\mathbb{R})$ and $x \in \mathbb{R}$

$$\delta_x(A) := \begin{cases} 1 & \text{if } A \ni x \\ 0 & \text{if } A \not\ni x \end{cases}.$$

Clearly, $\mathbb{E}[X_n] = 0(1-p) + 1p = p$ and

$$\mathbb{V}(X_n) = \mathbb{E}[X_n^2] - \mathbb{E}[X_n]^2 = \mathbb{E}[X_n] - \mathbb{E}[X_n]^2 = p - p^2 = p(1-p).$$

Then for any $n \in \mathbb{N}$ the sum S_n of the first n-Bernoullis takes values in $\{0, \ldots, n\}$ and is **Binomially distributed** with parameters n and $p \in [0,1]$

$$S_n := \sum_{k=1}^{n} X_k \sim \mathcal{B}_{n,p},$$

in other words for all $k \in \{0, \ldots, n\}$

$$\mathbb{P}(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

This entails the well-known formulas

$$\mathbb{E}[S_n] = \mathbb{E}\Big[\sum_{k=1}^{n} X_k\Big] = \sum_{k=1}^{n} \mathbb{E}\Big[X_k\Big] = np,$$

and due to the well-known Bienaymé identity, the independence of the $(X_n)_{n\in\mathbb{N}}$ and their identical distribution

$$\mathbb{V}(S_n) = \mathbb{V}(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} \mathbb{V}(X_k) = np(1-p).$$

Remark 1.1.1.
- *$S_n - np$ is a centered random variable, that is $\mathbb{E}[S_n - np] = 0$.*
- *$S_n^* := \frac{S_n - np}{\sqrt{np(1-p)}}$ is a normalized random variable that is centered with variance $\mathbb{V}(S_n^*) = 1$.*

EXERCISE 1.1.2. *Check the preceding remark.*

Remark 1.1.3. *For $X \sim \mathcal{B}_{n_1,p}$ and $Y \sim \mathcal{B}_{n_2,p}$ and $X \perp Y$ we have*

$$X + Y \sim \mathcal{B}_{n_1+n_2,p}.$$

EXERCISE 1.1.4. *Check the preceding remark.*

There are two fundamental limit theorems for the behavior of S_n^* as n tends to infinity, the Central and the Poisson limit theorem.

The Central limit theorem. Given $(\Omega, \mathcal{A}, \mathbb{P})$ consider an i.i.d. sequence $(X_n)_{n\in\mathbb{N}}$ of Bernoulli random variables $X_n : \Omega \to \{0,1\}$ with $\mathbb{P}_{X_n} = (1-p)\delta_0 + p\delta_1$. In the *symmetric* case $p = \frac{1}{2}$ we have that for any $n \in \mathbb{N}$

$$S_n^* = \frac{S_n - n/2}{\sqrt{n\frac{1}{2}(1-\frac{1}{2})}} = \frac{2S_n - n}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \underbrace{(2X_k - 1)}_{\in\{-1,1\}}$$

has a centered, symmetric "Binomial" distribution with values in the set

$$\{-\frac{n}{\sqrt{n}}, -\frac{n-1}{\sqrt{n}}, \ldots, -\frac{1}{\sqrt{n}}, 0, \frac{1}{\sqrt{n}}, \ldots, \frac{n-1}{\sqrt{n}}, \frac{n}{\sqrt{n}}\},$$

which looks more and more "bell-shaped".

Consult the figures in the video.

Interestingly, the convergence remains true also for asymmetric Bernoullis with $p \in (0,1) \setminus \{\frac{1}{2}\}$ and yields the following well-known result.

Theorem 1.1.5 (Central limit theorem (De Moivre-Laplace))**.** *For an i.i.d. sequence $(X_n)_{n\in\mathbb{N}}$ of Bernoulli random variables $X_n : \Omega \to \{0,1\}$ on a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{P}_{X_n} = (1-p)\delta_0 + p\delta_1$, $p \in (0,1)$ we have that*

$$\mathcal{L}(S_n^*) = \mathcal{L}(\sum_{k=1}^{n} \frac{X_k - p}{\sqrt{np(1-p)}}) \xrightarrow{d} \mathcal{N}(0,1), \qquad \textit{as } n \to \infty,$$

that is for any finite $a < b$

$$\lim_{n\to\infty} \mathbb{P}(S_n^* \in [a,b]) = \int_a^b f(x)dx, \qquad \textit{for } f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Interestingly in the light of Remark 1.1.3 the limiting (normal) distribution satisfies.

Remark 1.1.6. *If $X \sim \mathcal{N}(m_1, \sigma_1^2)$ and $Y \sim \mathcal{N}(m_2, \sigma_2^2)$ and $X \perp Y$ we have*

$$X + Y \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2).$$

EXERCISE 1.1.7. *Show the preceding remark.*

More generally we have the following result .

Theorem 1.1.8 (The central limit theorem)**.** *For a sequence $(X_n)_{n\in\mathbb{N}}$ of i.i.d. random variables $X_n : \Omega \to \mathbb{R}$ with $0 < \sigma^2 = \mathbb{V}(X_n) < \infty$ and $m = \mathbb{E}[X_n]$ we have*

$$S_n^* = \sum_{k=1}^{n} \frac{X_k - m}{\sqrt{n}\,\sigma} \xrightarrow{d} \mathcal{N}(0,1) \qquad \textit{as } n \to \infty.$$

A short proof is given in Kallenberg [43] using the Taylor expansion of the characteristic function of S_n^*.

Non-central limit theorems: A second class of limit theorems considers sequences of "i.i.d. sequences" of Bernoullis $(X_k^n)_{k=0,\ldots,n}$, where the parameter $p = p_n$ and tends to 0, that is, we have more end more random variables, but the exit probability for each of them decreases, that is some kind of "thinning". Therefore it is also called the limit theorem of "rare events".
The easiest version is the following classical result.

Theorem 1.1.9 (The Poisson limit theorem).
For sequences $(m_n)_{n\in\mathbb{N}} \subset \mathbb{N}$ *with* $m_n \to \infty$ *and* $(p_n)_{n\in\mathbb{N}} \in (0,1)$, $n \in \mathbb{N}$ *satisfying* $\lim_{n\to\infty} m_n p_n = \lambda \in [0,\infty)$ *we have for any fixed* $\ell \in \mathbb{N}$

$$\lim_{n\to\infty} \mathcal{B}_{m_n,p_n}(\{\ell\}) = \lim_{n\to\infty} \binom{m_n}{k} p_n^\ell (1-p_n)^{m_n-\ell} = \frac{\lambda^\ell}{\ell!} e^{-\lambda} = \mathrm{Poi}_\lambda(\{\ell\}).$$

As explained above we can reformulate this as follows: For $n \in \mathbb{N}$ consider the finite i.i.d. sequence

$$(X_k^n)_{k=1,\dots,n}$$

with $X_k^n \sim \mathcal{B}_{p_n}$. Then we have that

$$\mathcal{L}(\sum_{k=1}^n X_k^n) \longrightarrow \mathrm{Poi}_\lambda\,.$$

Proof. The proof is elementary:

$$\begin{aligned}\frac{m_n!}{(m_n-\ell)!} p_n^\ell &= (m_n \cdot \dots \cdot (m_n-\ell+1)) p_n^\ell \\ &= \frac{m_n \cdot \dots \cdot (m_n-\ell+1)}{m_n^\ell} (m_n p_n)^\ell \longrightarrow \lambda^\ell, \quad \text{as } n\to\infty,\end{aligned}$$

by assumption and Euler's limit of the exponential function

$$(1-p_n)^{m_n-\ell} = \frac{(1-\frac{m_n p_n}{m_n})^{m_n}}{(1-p_n)^\ell} \longrightarrow \frac{e^\lambda}{1}, \quad \text{as } n\to\infty.$$

Putting together these partials results yields the statement. □

In the light of Remark 1.1.3 and 1.1.6 we have the following interesting property.

Remark 1.1.10. *If* $X \sim \mathrm{Poi}_{\lambda_1}$ *and* $Y \sim \mathrm{Poi}_{\lambda_2}$ *and* $X \perp Y$, *then*

$$\mathbb{P}_{X+Y} = \mathrm{Poi}_{\lambda_1+\lambda_2}\,.$$

EXERCISE 1.1.11. *Check the preceding remark.*

1.2 Arrays and infinitely divisible distributions

1.2.1 Arrays and the central limit theorem

A word about separable Hilbert spaces: For our applications in the second and third part of the course we will formulate all our results for separable Hilbert space $(H, \langle\cdot,\cdot\rangle, |\cdot|)$. We give some basic results about separable Hilbert spaces without proof. We refer for instance to Werner [67], Kapitel V or any other book on functional analysis.

- Separable Hilbert spaces are the natural generalization of the Euclidean space $\mathbb{R}^d$, equipped with the norm

$$|x| = \sqrt{\langle x, x\rangle}, \qquad \text{and inner product } \langle x, y\rangle = \sum_{i=1}^{d} x_i y_i, \qquad x, y \in \mathbb{R}^d.$$

- A Hilbert space H over $\mathbb{R}$ is a $\mathbb{R}$-vector space equipped with an inner product $\langle \cdot, \cdot \rangle$ and norm $|x| = \sqrt{\langle x, x\rangle}$, which is complete w.r.t. this norm, that is all Cauchy sequences converge.

- A separable Hilbert space H is a Hilbert space, with a dense countable subset, w.r.t. the topology of the norm $|\cdot|$. In particular, separable Hilbert spaces are exactly those Hilbert spaces which have a countable orthonormal basis $(e_n)_{n\in\mathbb{N}}$, such that any $x \in H$ has the norm expansion

$$x = \sum_{n\in\mathbb{N}} \langle x, e_n\rangle e_n$$

Note that this expression is a limit, which can be justified by the so-called Parseval's equality

$$|x| = \sum_{n\in\mathbb{N}} \langle x, e_n\rangle^2,$$

and

$$\langle x, y\rangle = \sum_{n\in\mathbb{N}} \langle x, e_n\rangle \langle y, e_n\rangle.$$

- The smallest sigma algebra containing all open balls in H is called the Borel-sigma algebra $\mathcal{B}(H)$ in H.

For most practical purposes until declared otherwise you can think of H being the $\mathbb{R}^d$.

Motivation of arrays: We can rewrite the convergence of the Theorem 1.1.5 and Theorem 1.1.9 in the following simple unified form.

1. We can rewrite the hypotheses of the De Moivre-Laplace Theorem 1.1.5 as

$$S_n^* = \frac{S_n - np}{\sqrt{np(1-p)}} = \sum_{k=1}^{n} \underbrace{\frac{1}{\sqrt{n}} \Big(\frac{X_k - p}{\sqrt{p(1-p)}} \Big)}_{=:X_k^n} = \sum_{k=1}^{n} X_k^n,$$

where $X_k^n : \Omega \to \{-\frac{p}{\sqrt{np(1-p)}}, \frac{1-p}{\sqrt{np(1-p)}}\}$ with $\mathbb{E}[X_k^n] = 0$ and $\mathbb{V}(X_k^n) = \frac{1}{n}$ such that

$$1 = \mathbb{V}(S_n^*) = \sum_{k=1}^{n} \underbrace{\mathbb{V}(X_k^n)}_{=\frac{1}{n}}.$$

Note that in this case the distribution of X_k^n does not depend on k.

2. Consider the case of the Poisson limit Theorem 1.1.9, where

$$S_n := \sum_{k=1}^{n} X_k^n$$

for the i.i.d. $(X_k^n)_{\substack{n\in\mathbb{N},\\ k=1,\dots,n}}$ and the law $X_k^n \sim \mathcal{B}_{p_n}$ being independent of k.

The concept of arrays. We generalize this concept to sequences of "sequences of random variables", which are not necessarily identically distributed.

Definition 1.2.1. *Let $(k_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence $k_\cdot : \mathbb{N} \to \mathbb{N}$ of natural numbers and H be a separable Hilbert space. Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$.*

- *A double sequence $(X_k^n)_{\substack{n\in\mathbb{N}\\ k\in\{1,\dots,k_n\}}}$ of random vectors $X_k^n : \Omega \to H$ is called an **array**.*
- *The array is called **independent** if for any $n \in \mathbb{N}$ the family $(X_k^n)_{k=1,\dots,k_n}$ is independent.*
- *The array is called **centered** if $\mathbb{E}[|X_k^n|] < \infty$ and $\mathbb{E}[X_k^n] = 0$ for any $n \in \mathbb{N}$ and $k \in \{1, \dots, k_n\}$.*
- *We define for $n \in \mathbb{N}$ as the **row sum***

$$S_n := \sum_{k=1}^{k_n} X_k^n.$$

- *The array is called **normalized** if $\mathbb{E}[|X_k^n|^2] < \infty$ and $\mathbb{V}(S_n) = 1$.*
- *The array is called **asymptotically negligible** or a **null array** if for any $\varepsilon > 0$*

$$\lim_{n\to\infty} \max_{k\in\{1,\dots,k_n\}} \mathbb{P}(|X_k^n| > \varepsilon) = 0,$$

that is uniform convergence to 0 in probability.

Remark 1.2.2. *For an independent normalized array of random variables $X_k^n : \Omega \to \mathbb{R}$, we have*

$$1 = \mathbb{V}(S_n) = \mathbb{V}(\sum_{k=1}^{k_n} X_k^n) = \sum_{k=1}^{k_n} \mathbb{V}(X_k^n).$$

The central limit theorem of Lindeberg-Feller

Theorem 1.2.3 (The central limit theorem of Lindeberg-Feller)**.** *Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ be $(X_k^n)_{\substack{n\in\mathbb{N} \\ k\in\{1,\dots,k_n\}}}$ an independent, centered, normalized array of random variables $X_k^n : \Omega \to \mathbb{R}$. Then the following statements are equivalent:*

1. *For any $\varepsilon > 0$ we have*

$$L_n(\varepsilon) := \sum_{k=1}^{k_n} \mathbb{E}[(X_k^n)^2 \mathbf{1}\{|X_k^n| > \varepsilon\}] \overset{n\to\infty}{\longrightarrow} 0. \tag{1.1}$$

2. *$(X_k^n)_{\substack{n\in\mathbb{N} \\ k\in\{1,\dots,k_n\}}}$ is asymptotically negligible and*

$$\mathbb{P}_{S_n} \overset{n\to\infty}{\longrightarrow} \mathcal{N}(0,1).$$

This result is found for instance in Klenke [45], Thm 15.43.

Remark 1.2.4. *How can we understand this result? The variance for the aggregate of tails vanishes*

$$\begin{aligned} L_n(\varepsilon) &= \mathbb{E}[\sum_{k=1}^{k_n} (X_k^n)^2 \mathbf{1}^2\{(X_k^n)^2 > \varepsilon^2\}] \\ &= \mathbb{V}\Big(\sum_{k=1}^{k_n} \underbrace{X_k^n \mathbf{1}\{|X_k^n| > \varepsilon\}}_{=:\overline{X_k^n}^{\varepsilon}}\Big) = \mathbb{V}(\overline{S_n}^{\varepsilon}) \end{aligned}$$

if

$$\overline{S_n}^{\varepsilon} := \sum_{k=1}^{k_n} \overline{X_k^n}^{\varepsilon}.$$

In addition, since $\mathbb{V}(S_n) = 1$ we can understand the limit (1.1) as

$$L_n(\varepsilon) = \frac{\mathbb{V}(\overline{S_n}^{\varepsilon})}{\mathbb{V}(S_n)} \longrightarrow 0, \qquad \textit{as } n \to \infty.$$

In other words, for any $\varepsilon > 0$ the influence of the variances of the random variables with the values larger than ε tends to 0 uniformly. Put differently, the collective behavior is determined by the infinite aggregate of the arbitrarily values of the random variables.

This result has also versions in higher dimensions, for instance see [51].

1.2.2 Infinitely divisible distributions

A) Basic properties:

Definition 1.2.5. *A distribution μ on $(H, \mathcal{B}(H))$ is called **infinitely divisible** if for a random vector $X : \Omega \to H$ with $\mathbb{P}_X = \mu$ and any $n \in \mathbb{N}$ there exists a probability distribution $\bar{\mu}_n$ on $(H, \mathcal{B}(H))$ and an independent family $(X_k^n)_{k=1,\dots,n}$ of random vectors $X_k^n : \Omega \to H$ with $\mathbb{P}_{X_k^n} = \bar{\mu}_n$ such that*

$$X_1^n + \cdots + X_n^n \stackrel{d}{=} X.$$

In other words: *A distribution μ on $(H, \mathcal{B}(H))$ is called infinitely divisible if for any $n \in \mathbb{N}$ there is a distribution $\bar{\mu}_n$ such that*

$$\underbrace{\bar{\mu}_n * \cdots * \bar{\mu}_n}_{n \text{ times}} = \bar{\mu}_n^{*n} = \mu.$$

By this property we write

$$\bar{\mu}_n =: \mu^{*\frac{1}{n}}.$$

Example 1.2.6 (Infinite divisibility of the normal distribution)**.** *Remark 1.1.6 tells us that for any $m \in \mathbb{R}$ and $\sigma^2 > 0$ and $X \sim \mathcal{N}(m, \sigma^2)$ we have that for two independent random variables X_1^2, X_2^2 with $X_i^2 \sim \mathcal{N}(\frac{m}{2}, \frac{1}{2}\sigma^2)$ we get*

$$X_1^2 + X_2^2 \sim X.$$

The same is true for finitely many random variables: For any $n \in \mathbb{N}$ any independent family $(X_1^n, \dots, X_n^n)$ with $X_i^n \sim \mathcal{N}(\frac{m}{n}, \frac{1}{n}\sigma^2)$ we have

$$X_1^n + \cdots + X_n^n \sim X.$$

Example 1.2.7 (Infinite divisibility of the Poisson distribution:)**.** *By Remark 1.1.10 we have in the same spirit that for any $\lambda > 0$ and $X \sim \mathrm{Poi}_\lambda$ any two independent random variables X_1^2, X_2^2 with $X_i^2 \sim \mathrm{Poi}_{\frac{\lambda}{2}}$ we have*

$$X_1^2 + X_2^2 \sim X.$$

The analogous result is true for finitely many random variables, that is for all $n \in \mathbb{N}$ and any i.i.d. family $(X_1^n, \dots, X_n^n)$ with $X_i^n \sim \mathrm{Poi}_{\frac{\lambda}{n}}$ we have that

$$X_1^n + \cdots + X_n^n \sim X.$$

Example 1.2.8. *Trivially, since* $\underbrace{\frac{1}{n}+\cdots+\frac{1}{n}}_{n \text{ times}} = 1$ *we have that for any* $b \in H$ *and* $X \sim \delta_b$ *that for* $X_i^n \sim \delta_{\frac{b}{n}}$

$$X_1^n + \cdots + X_n^n = \frac{1}{n}\underbrace{(b+\cdots+b)}_{n \text{ times}} = n\frac{b}{n} = b = X.$$

We have already seen, that the normal distribution, the Poisson distribution and the Dirac distribution δ_b (the constant vectors) are infinitely divisible. What about a counter example?

Example 1.2.9 (Counter example)**.** *The uniform distribution* $\mathcal{U}_{[-1,1]}$ *is not infinitely divisible. For* $X \sim \mathcal{U}_{[-1,1]}$ *and* $u \neq 0$ *we have*

$$\begin{aligned}
\phi_{\mathcal{U}_{[-1,1]}}(u) &= \frac{1}{2}\int_{-1}^{1} e^{iux}dx \\
&= \frac{1}{2}\int_{-1}^{1} \cos(ux)dx + i\int_{-1}^{1} \sin(ux)dx \\
&= \frac{\sin(u)}{u} \\
&= \sqrt{\frac{\sin(u)}{u}}\sqrt{\frac{\sin(u)}{u}}.
\end{aligned}$$

If $Z \sim \mathcal{U}_{[-1,1]}$*, does there exist a distribution with* μ*, such that for two independent copies* X_1, X_2 *with* $X_i \sim \mu$ *we have*

$$X_1 + X_2 \stackrel{d}{=} Z \quad ?$$

First note since Z *is symmetric and the convolution of symmetric random variables preserves symmetry, such that* μ *is necessarily symmetric. This implies that* $\phi_{X_1}(u) \in \mathbb{R}$ *for all* $u \in \mathbb{R}$*. However*

$$\sqrt{\frac{\sin(u)}{u}} \notin \mathbb{R}$$

for $u \in (\pi, 2\pi)$*, which is a contradiction.*

So why bother with infinite divisibility at all? Well, it characterizes the limits of <u>any</u> convergent array.

Theorem 1.2.10. *For a random vector* $X \sim \mu$ *the following are equivalent:*

1. *X is infinitely divisible.*

2. *There exists an independent array* $(X_k^n)_{1\leqslant k\leqslant n, n\in\mathbb{N}}$ *and a family of vectors* $(b_n)_{n\in\mathbb{N}}$ *in* H *such that*

$$\lim_{n\to\infty}(S_n - b_n) \stackrel{d}{=} \mu,$$

This result is given in Kallenberg [43], Theorem 13.12, for $H = \mathbb{R}^d$ and carries over directly to separable Hilbert spaces H.

B) Infinite dimensional Gaussian distributions: In this paragraph we construct the infinite dimensional H-valued analogue of the normal distribution.

Definition 1.2.11. *For a separable Hilbert space* $(H, \mathcal{B}(H))$ *with orthonormal basis* $(e_i)_{i\in\mathbb{N}}$, $m \in H$ *and a linear operator* $Q : H \to H$ *which is*

- *positive:* $\langle Qx, x\rangle \geqslant 0$ *for all* $x \in H$ *and*
- *symmetric:* $\langle Qx, y\rangle = \langle x, Qy\rangle$ *for all* $x, y \in H$

and satisfies

$$\sum_{i=1}^{\infty} |\langle Qe_i, e_i\rangle| < \infty$$

the ***Gaussian distribution*** $\mathcal{N}(m, Q)$ ***in*** $(H, \mathcal{B}(H))$ *is defined via the characteristic function of* $X \sim \mathcal{N}(m, Q)$

$$\phi_X(u) := \exp(i\langle u, m\rangle - \frac{1}{2}\langle Qu, u\rangle, \qquad u \in H.$$

Definition 1.2.12. *For a separable Hilbert space* $(H, \mathcal{B}(H))$ *the set of operators* $Q : H \to H$ *linear, symmetric operator satisfying*

$$\sum_{i=1}^{\infty} |\langle Qe_i, e_i\rangle| < \infty$$

is called the class of ***symmetric, positive trace class operators*** *or* ***covariance operatores*** *and denoted by* $L_1^+(H)$.

Example 1.2.13. *For* H *with orthonormal basis* $(e_n)_n$ *and nonnegative sequence* $(\lambda_n)_{n\in\mathbb{N}}$ *such that*

$$\sum_{n=1}^{\infty} \lambda_n < \infty \tag{1.2}$$

the operator

$$Qx := \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n, \qquad x \in H,$$

is an element in $L_1^+(H)$. *Equality (1.2) justifies the name "trace class".*

EXERCISE 1.2.14. *Check that that $Q \in L_1^+(H)$.*

Remark 1.2.15. *Why do we need trace class covariance operators? Note that for a random vector $Z \sim \mathcal{N}(0, Q)$, for some $Q \in L_1^+(H)$ we expect that*

$$\mathbb{V}(Z) = \mathbb{E}\big[|Z|^2\big] < \infty.$$

This is not guaranteed, if Q does not satisfy (1.2). Later we construct a Q-Brownian motion explicitly, where the respective calculus is carried out. What does it mean? It says, contrary to $\mathbb{R}^d$ where $\mathcal{N}(0, id_d)$ is the standard model of a Normal distribution ($\mathrm{trace}(id_d) = d$*), in a separable Hilbert space H there is no such thing as "$\mathcal{N}(0, id_H)$" in a classical sense! The covariance operator necessarily needs a summable diagonal.*

EXERCISE 1.2.16. *Given a $\mathbb{R}$-valued i.i.d. sequence $(Z_n)_{n\in\mathbb{N}}$ with $Z_n \sim \mathcal{N}(0, 1)$ and $(e_n)_{n\in\mathbb{N}}$ an ONB in the separable Hilbert space H. We consider a sequence $(\lambda_n)_{n\in\mathbb{N}}$ with $\lambda_n > 0$ and $\sum_{n\in\mathbb{N}} \lambda_n < \infty$ and define the $L_1^+(H)$-operator*

$$Qx := \sum_{n\in\mathbb{N}} \lambda_n \langle x, e_n\rangle e_n, \qquad x \in H.$$

Show that

$$X := \sum_{n\in\mathbb{N}} \sqrt{\lambda_n} Z_n e_n$$

satisfies $X \sim \mathcal{N}(0, Q)$.

The most important result of this paragraph is finally the infinite divisibility of $X \sim \mathcal{N}(m, Q)$.

Lemma 1.2.17. *Any $X \sim \mathcal{N}(m, Q)$ for any $m \in H$ and $Q \in L_1^+(H)$ is infinitely divisibile.*

EXERCISE 1.2.18. *Show the preceding lemma.*

C) Compound Poisson distributions: We now get to know the most important class of infinitely divisible distributions.

Definition 1.2.19 (Compound Poisson random variables)**.** *Fix a distribution μ on $(H, \mathcal{B}(H))$ and $\lambda > 0$. On a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ consider an i.i.d. family $(Z_k)_{k\in\mathbb{N}}$ of random vectors $Z_k : \Omega \to H$ with $Z_k \sim \mu$ and a random variable $\pi \sim \mathrm{Poi}_\lambda$ with $(Z_k)_{k\in\mathbb{N}} \perp \pi$. Define the random variable*

$$C(\omega) := \sum_{k=1}^{\pi(\omega)} Z_k(\omega), \qquad \omega \in \Omega.$$

We shall denote the distribution of C *by the* ***compound Poisson distribution with parameters*** λ ***and*** μ, $\mathrm{Cpp}(\lambda,\mu)$.

Remark 1.2.20 (The Poisson distribution is a trivial Compound Poisson distribution)**.** *Note that in general we do not know the density of such distributions. Trivial exception:* $d=1$ *and* $\mu=\delta_1$ *yields* $C=\sum_{k=0}^{\pi}=\pi$*, the Poisson distribution of* π *itself.*

In general compound Poisson distributions do not have known closed form densities, apart from particular examples. A natural form to characterize infinitely distributions is via its characteristic function. In addition, it provides an extraordinary useful tool to calculate moments etc.

Lemma 1.2.21 (The characteristic function of a Compound Poisson distribution)**.** *For any* $C\sim\mathrm{Cpp}(\lambda,\mu)$ *and* $u\in H$ *we have*

$$\phi_C(u)=\exp\Big(\int_H(e^{i\langle u,z\rangle}-1)\lambda\mu(dz)\Big).$$

Proof. We calculate the characteristic function for $C\sim\mathrm{Cpp}(\lambda,\mu)$ and

$$\begin{aligned}
\phi_C(u)&=\mathbb{E}[e^{i\langle u,C\rangle}]=\mathbb{E}[e^{i\langle u,\sum_{k=1}^{\pi}Z_k\rangle}]=\sum_{k=0}^{\infty}\mathbb{E}[e^{i\langle u,\sum_{\ell=1}^{\pi}Z_\ell\rangle}\mid\pi=k]\mathbb{P}(\pi=k)\\
&=\sum_{k=0}^{\infty}\mathbb{E}[e^{i\langle u,\sum_{\ell=1}^{k}Z_\ell\rangle}]\frac{\lambda^k}{k!}e^{-\lambda}=\sum_{k=0}^{\infty}\prod_{\ell=1}^{k}\mathbb{E}[e^{i\langle u,Z_\ell\rangle}]\frac{\lambda^k}{k!}e^{-\lambda}\\
&=\sum_{k=0}^{\infty}\mathbb{E}[e^{i\langle u,Z_1\rangle}]^k\frac{\lambda^k}{k!}e^{-\lambda}=\sum_{k=0}^{\infty}\frac{\Big(\mathbb{E}[e^{i\langle u,Z_1\rangle}]\lambda\Big)^k}{k!}e^{-\lambda}\\
&=e^{-\lambda}e^{\mathbb{E}[e^{i\langle u,Z_1\rangle}]\lambda}=e^{(\mathbb{E}[e^{i\langle u,Z_1\rangle}]-1)\lambda}\\
&=e^{(\phi_{Z_1}(u)-1)\lambda}=\exp\Big(\int_H(e^{i\langle u,z\rangle}-1)\lambda\mu(dz)\Big).
\end{aligned}$$

□

Lemma 1.2.22 (Compound Poisson distributions are infinitely divisible)**.** *Any compound Poisson distribution* $\mathrm{Cpp}(\lambda,\mu)$ *is infinitely divisible.*

Proof. Now for $C_1 \perp C_2$ with $C_j \sim \mathrm{Cpp}(\lambda_j, \mu_j)$ we have

$$\begin{aligned}\phi_{C_1+C_2}(u) &= \phi_{C_1}(u)\phi_{C_2}(u) \\ &= \exp\Big(\int_H (e^{i\langle u,z\rangle} - 1)\lambda_1\mu_1(dz)\Big)\exp\Big(\int_H (e^{i\langle u,z\rangle} - 1)\lambda_2\mu_2(dz)\Big) \\ &= \exp\Big(\int_H (e^{i\langle u,z\rangle} - 1)(\lambda_1\mu_1 + \lambda_2\mu_2)(dz)\Big) \\ &= \exp\Big(\int_H (e^{i\langle u,z\rangle} - 1)(\lambda_1 + \lambda_2)(\frac{\lambda_1}{\lambda_1+\lambda_2}\mu_1 + \frac{\lambda_2}{\lambda_1+\lambda_2}\mu_2)(dz)\Big).\end{aligned}$$

Hence $C_1 + C_2 \sim \mathrm{Cpp}(\lambda_1 + \lambda_2, \frac{\lambda_1}{\lambda_1+\lambda_2}\mu_1 + \frac{\lambda_2}{\lambda_1+\lambda_2}\mu_2)$. Given $C \sim \mathrm{Cpp}(\lambda, \mu)$ for any $n \in \mathbb{N}$ any i.i.d. vector $(C_1, \ldots, C_n)$, $C_i \sim \mathrm{Cpp}(\frac{\lambda}{n}, \mu)$ satisfies that

$$\begin{aligned}\phi_{C_1+\cdots+C_n}(u) &= \phi_{C_1}(u)\ldots\phi_{C_n}(u) \\ &= \exp\Big(\int_H (e^{i\langle u,z\rangle} - 1)\frac{\lambda}{n}\mu(dz)\Big)^n \\ &= \exp\Big(\int_H (e^{i\langle u,z\rangle} - 1)\lambda\mu(dz)\Big) = \phi_C(u), \qquad u \in H,\end{aligned}$$

that is

$$\mathcal{L}(C_1 + \cdots + C_n) = \mathcal{L}(C)$$

and $\mathcal{L}(C)$ is infinitely divisible. □

D) Infinitely divisibility so far: What have we achieved with Example 1.2.8, Lemma 1.2.17 and Lemma 1.2.22?

Proposition 1.2.23. *For any random variable X with values in H, which is the sum $X = b + Y + C$ for some $b \in H$*

$$Y \sim \mathcal{N}(m, Q) \qquad \text{and} \qquad C \sim \mathrm{Cpp}(\lambda, \mu)$$

satisfying $Y \perp C$ that is

$$\phi_X(u) = \exp(\eta(u)), \qquad \text{with } \eta(u) = i\langle b+m, u\rangle - \frac{1}{2}\langle Qu, u\rangle + \int_H (e^{iuz} - 1)\lambda\mu(dz)$$

we have that $\mathcal{L}(X)$ is infinitely divisible.

Is the shape of Proposition 1.2.23 everything for infinitely divisible distributions? Almost! Without proof we state the full characterization of infinitely divisible distributions, which will give us new non-trivial examples.

1.2.3 The Lévy-Khinchin decomposition

A) General infinitely divisible distributions and a first non-trivial example

Theorem 1.2.24 (Lévy-Khinchin)**.** *Consider a probability distribution μ on $(H, \mathcal{B}(H))$. Then the following statements are equivalent:*

1. *μ is infinitely divisible.*
2. *There exists a unique canonical triple (b, Q, ν) consisting of a vector $b \in H$, $Q \in L_1^+(H)$ is a symmetric, positive, trace class operator $Q : H \to H$ and a σ-finite measure ν on $(H, \mathcal{B}(H))$ satisfying*

$$\nu(\{0\}) = 0, \qquad \nu(B_1^c(0)) < \infty \qquad \textit{and} \qquad \int_{B_1(0)} |z|^2 \nu(dz) < \infty$$

such that

$$\phi_\mu(u) = e^{\eta(u)}, \qquad u \in H,$$

with the exponent

$$\eta(u) = i\langle b, r\rangle - \frac{1}{2}\langle Qu, u\rangle\sigma^2 + \int_H (e^{i\langle u,z\rangle} - 1 - i\langle u, z\rangle \mathbf{1}\{|z| \leqslant 1\})\nu(dz). \tag{1.3}$$

The measure ν is called ***Lévy measure****.*

For an even more general result for separable Banach spaces see [2].

Example 1.2.25 (A first non-trivial example in $\mathbb{R}^d$: the rotationally symmetric α-stable measure)**.** *Note that the Lévy measure ν is not necessarily finite. For instance the measure in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$*

$$\nu(dz) := \frac{dz}{|z|^{\alpha+d}}\mathbf{1}\{z \neq 0\}, \qquad \textit{for some } \alpha \in (0, 2) \tag{1.4}$$

satisfies $\nu(\{0\}) = 0$,

$$\nu(B_1^c(0)) = C\int_1^\infty \frac{dr}{r^{1+\alpha}} = Cr^{-\alpha}\Big|_{r=1}^{\infty} < \infty$$

and

$$\int_{B_1(0)} |z|^2\nu(dz) = C\int_0^1 r^{2-\alpha-1}dr = \frac{C}{2-\alpha}r^{2-\alpha}\Big|_{r=0}^{r=1} < \infty.$$

However, we have

$$\nu(B_1(0)) = C\int_0^1 r^{-\alpha-1}dr = Cr^{-\alpha}\Big|_0^1 = \infty.$$

The distribution determined by the Lévy triplet $(0, 0, \nu)$ is called ***rotationally invariant stable distribution****.*

Remark 1.2.26. *The previous expression looks complicated. Is it really well-defined? Writing* $h(u,z) := e^{i\langle u,z\rangle} - 1 - i\langle u,z\rangle \mathbf{1}\{|z| \leqslant 1\}$ *we obtain the estimate*

$$|h(u,z)| \leqslant 2|u|^2(1 \wedge |z|^2), \qquad u \in H, z \in H \tag{1.5}$$

such that

$$\begin{aligned} &\Big|\int_H (e^{i\langle u,z\rangle} - 1 - i\langle u,z\rangle \mathbf{1}\{|z| \leqslant 1\})\nu(dz)\Big| \\ &\leqslant \int_H |e^{i\langle u,z\rangle} - 1 - i\langle u,z\rangle \mathbf{1}\{|z| \leqslant 1\}|\nu(dz) \\ &\leqslant 2|u|^2 \int_H (1 \wedge |z|^2)\nu(dz) < \infty. \end{aligned} \tag{1.6}$$

EXERCISE 1.2.27. *Verify the inequality (1.5). Hint: distinguish the cases* $|z| > 1$ *and* $|z| \leqslant 1$.

B) Finite Lévy measure yields a compound Poisson distribution

Remark 1.2.28 (Finite Lévy measure implies a compound Poisson law)**.** *Assume for a measure* μ *that* $b = 0$ *and* $Q = 0$ *and* $\nu(H) < \infty$*. Then we have*

$$\underbrace{\Big|\int_{B_1(0)} z\nu(dz)\Big|}_{=:\bar{b}} \leqslant \int_{B_1(0)} |z|\nu(dz) \leqslant \nu(B_1(0)) < \infty$$

and hence

$$\begin{aligned} &\int_H (e^{i\langle u,z\rangle} - 1 - i\langle u,z\rangle \mathbf{1}\{|z| \leqslant 1\})\nu(dz) \\ &= \int_H (e^{i\langle u,z\rangle} - 1)\nu(dz) - i\int_{B_1(0)} \langle u, z\nu(dz) \\ &= \int_H (e^{i\langle u,z\rangle} - 1)\nu(dz) - i\langle u, \int_{B_1(0)} z\nu(dz)\rangle \\ &= \nu(H)\int_H (e^{i\langle u,z\rangle} - 1)\frac{\nu(dz)}{\nu(H)} - i\langle u, \int_{B_1(0)} z\nu(dz)\rangle. \end{aligned}$$

That is for $X \sim \mu$ *we have* $X = \bar{b} + Y$*, where* $Y \sim \mathrm{Cpp}(\nu(H), \nu/\nu(H))$*. In particual if* ν *is symmetric we have* $\bar{b} = 0$*.*

Remark 1.2.29. *Note that the constant* 1 *can be replaced by any constant* $\rho > 0$ *(at the price of a slightly changed expression for* b *and the integrals, of course).*

1.2.4 The class of alpha-stable distributions

A) Basic properties and examples of α-stable distributions in $\mathbb{R}$

Definition 1.2.30. *A probability distribution in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called* ***strictly stable*** *if for any $n \in \mathbb{N}$ there is a constant $a(n) > 0$ such that for $n+1$ independent copies X_j^n, X with that distribution satisfy*

$$\frac{1}{a(n)}\Big(X_1^n + \cdots + X_n^n\Big) \stackrel{d}{=} X.$$

*If $a(n) = n^{\frac{1}{\alpha}}$ it is called α-**stable**.*

Example 1.2.31. *Recall for $Y \sim \mathcal{N}(m, \sigma^2)$ we have that $\frac{Y-m}{\sigma} \sim \mathcal{N}(0,1)$ and hence $cY \sim \mathcal{N}(cm, c\sigma^2)$ for any $c > 0$. Hence for $X \in \mathcal{N}(0, \sigma^2)$ and $n \in \mathbb{N}$ with an i.i.d. sequence $(X_i^n)_{i=1,\ldots,n}$ with $X_i^n \sim \mathcal{N}(0, \sigma^2)$ we have*

$$\frac{1}{\sqrt{n}}(X_1^n + \cdots + X_1^n) \stackrel{d}{=} X$$

That is, centered normal distributions are 2-stable.

Example 1.2.32. *The* ***Cauchy distribution*** $\mathrm{Cau}(b, \gamma)$ *in $\mathbb{R}$ with location parameter b and scale parameter $\gamma > 0$ is given by the density*

$$f(x) := \frac{1}{\pi\gamma\Big(1 + \big(\frac{x-b}{\gamma}\big)^2\Big)} = \frac{\gamma}{\pi\gamma\Big(\gamma^2 + (x-b)^2\Big)}, \qquad x \in \mathbb{R}.$$

Check in the exercise below that for

$$Z \sim \mathrm{Cau}(b, \gamma) \text{ we have } \frac{Z-b}{\gamma} \sim \mathrm{Cau}(0,1). \tag{1.7}$$

Hence it is enough to work on $Z \sim \mathrm{Cau}(0,1)$, for which we have

$$\mathbb{E}[|Z|] = \int_{\mathbb{R}} \frac{x dx}{\pi(1+x^2)} \geqslant C_1 + C_2 \int_{|x|>R} \frac{dx}{|x|} = \infty.$$

that is no finite first moment. However, we have $\mathbb{E}[|Z|^\alpha] < \infty$ for any $\alpha < 1$. (Check that!)

EXERCISE 1.2.33. *Check for $Z \sim \mathrm{Cau}(b, \gamma)$ that $\frac{Z-b}{\gamma} \sim \mathrm{Cau}(0,1)$.*

EXERCISE 1.2.34.
- *Show that the characteristic function of a random variable X whose distribution is the symmetric exponential distribution with density*

$$f(x) = \frac{1}{2}e^{-|x|}, \qquad x \in \mathbb{R}.$$

is given by

$$\phi_X(u) = \frac{1}{1+u^2}, \qquad u \in \mathbb{R}.$$

Hint: Use integration by part and deduce a recursion formula.

- *Show that the characteristic function of a* Cau(0, 1) *distributed random variable is given by* $\phi_X(u) = e^{-|u|}, u \in \mathbb{R}$. *Use the inversion formula of the characteristic function given in the appendix for the previous example.*

Example 1.2.32 continued: In EXERCISE 1.2.34 you calculated the characteristic function of $Z \sim \text{Cau}(0,1)$ and by the scaling (1.7) we know that for $Z \sim \text{Cau}(b,\gamma)$ which is given as

$$\phi_Z(u) = \exp(iub - \gamma|u|), \qquad u \in \mathbb{R}.$$

This yields that Cau(0, 1) is infinitely divisible, since for $Y \sim \text{Cau}(0,1)$ and $n \in \mathbb{N}$ a family $(Y_1,\ldots,Y_n)$ of i.i.d. random variables $Y_n : \Omega \to \mathbb{R}$ with $Y_n \sim \text{Cau}(0, \frac{1}{n})$ satisfies

$$\phi_{Y_1+\cdots+Y_n}(u) = \phi_{Y_1}(u)\ldots\phi_{Y_n}(u) = \phi_{Y_1}(u)^n = e^{-n\frac{1}{n}|u|} = e^{-|u|} = \phi_Y(u).$$

With a similar calculus as in the α-stable case below it can be shown the representation

$$|u| = \int_{\mathbb{R}} (e^{iuz} - 1 - iuz\mathbf{1}\{|z| \leqslant 1\}\frac{dz}{z^2}, \tag{1.8}$$

which yields that

$$\eta(u) = |u|, \qquad u \in \mathbb{R}$$

and corresponds to a Lévy triplet $(0,0,\nu)$ for $\nu(dz) = \frac{dz}{z^2}$.

Stability: The previous scaling for $X \sim \text{Cau}(0,1)$ yields $\frac{1}{n}X \sim \text{Cau}(0,\frac{1}{n})$. Hence for a sequence of i.i.d. Cau(0, 1) distributed random variables $(X_n)_{n\in\mathbb{N}}$ we have

$$\frac{1}{n}(X_1 + \cdots + X_n) \stackrel{d}{=} X, \tag{1.9}$$

and hence the Cauchy distributions are 1-stable distributions.

EXERCISE 1.2.35. *Why is (1.9) not a contradiction to Kolmogorov's strong law of large number? Note how different they are! How do you interpret this?*

B) Some instructive calculations for α-stable distributions in $\mathbb{R}$ We calculate the characteristic functions and the Lévy measure of α-stable distributions in $d = 1$.

Remark 1.2.36. *For any $\alpha \in (0,1)$ and $u > 0$*

$$u^\alpha = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (e^{iuz} - 1) \frac{dz}{z^{1+\alpha}}.$$

We verify directly:

$$\begin{aligned}
\int_0^\infty (e^{iuz} - 1) z^{-(1+\alpha)} dz &= \int_0^\infty \left(\int_{-\infty}^z iue^{iuv} dv \right) z^{-(1+\alpha)} dz \\
&= \int_0^\infty \left(\int_{-\infty}^\infty \mathbf{1}\{v \le z\} iue^{iuv} dv \right) z^{-(1+\alpha)} dz \\
&= \int_0^\infty \int_{-\infty}^\infty \mathbf{1}\{v \le z\} iue^{iuv} z^{-(1+\alpha)} dv dz \\
&= \int_0^\infty \int_{-\infty}^\infty \mathbf{1}\{v \le z\} iue^{iuv} z^{-(1+\alpha)} dz dv \\
&= \int_0^\infty (\int_v^\infty z^{-(1+\alpha)} dz) iue^{iuv} dv \\
&= iu \int_0^\infty (\frac{1}{\alpha} v^{-\alpha}) e^{iuv} dv \\
&= \frac{iu}{\alpha} \int_0^\infty v^{-\alpha} e^{iuv} dv \\
&= \frac{iu^{1+\alpha-1}}{i\alpha} \int_0^\infty x^{-\alpha} e^{-x} dx = \frac{u^\alpha}{\alpha} \Gamma(1-\alpha)
\end{aligned}$$

for $x = iuv$ and $iudv = dx$. With the same arguments we obtain for $u < 0$ that

$$-|u|^\alpha = -\frac{\alpha}{\Gamma(1-\alpha)} \int_{-\infty}^0 (e^{iuz} - 1) \frac{dz}{z^{1+\alpha}}.$$

Hence for any $\alpha \in (0,1)$ there is an infinitely divisible distribution μ satisfying

$$\phi_\mu(u) = e^{\eta(u)}, \qquad u \in \mathbb{R},$$

such that for some $\min\{c^+, c^-\} > 0$ *there is* $c(\alpha) = \frac{2\alpha}{\Gamma(1-\alpha)} \frac{c^+ + c^-}{2}$ *and*

$$\begin{aligned}
\eta(u) &= \int_{\mathbb{R}} (e^{iuz} - 1)(c^- \mathbf{1}\{z < 0\} + c^+ \mathbf{1}\{z > 0\}) \frac{dz}{|z|^{1+\alpha}} = |u|^\alpha \\
&= iu \underbrace{\left(c^- \int_{-1}^{0} \frac{z\ dz}{|z|^{\alpha+1}} + c^+ \int_0^1 \frac{z\ dz}{|z|^{\alpha+1}} \right)}_{=: b(\alpha)} \\
&\quad + c(\alpha) \int_{\mathbb{R}} (e^{iuz} - 1 - iuz \mathbf{1}\{|z| \leqslant 1\})(c^- \mathbf{1}\{z < 0\} + c^+ \mathbf{1}\{z > 0\}) \frac{dz}{|z|^{1+\alpha}},
\end{aligned}$$

that is, we have a Lévy triplet $(b(\alpha), 0, \nu)$ *with*

$$\nu(dz) = (c^+ \mathbf{1}\{z > 0\} + c^- \mathbf{1}\{z < 0\}) \frac{dz}{|z|^{1+\alpha}}. \tag{1.10}$$

In the case of a symmetric distribution $c^+ = c^-$ *we have* $b(\alpha) = 0$.

Remark 1.2.37. *For* $\alpha \in (1, 2)$ *we have to admit an additional term in the Taylor expansion*

$$u^\alpha = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (e^{iuz} - 1 - iuz \mathbf{1}\{z \leqslant 1\}) \frac{dz}{z^{1+\alpha}}, \qquad u \in \mathbb{R}.$$

$$\eta(u) = \int_{\mathbb{R}} (e^{iuz} - 1 - iuz \mathbf{1}\{|z| \leqslant 1\}) \frac{(c^- \mathbf{1}\{z < 0\} + c^+ \mathbf{1}\{z > 0\}) dz}{|z|^{1+\alpha}} = c(\alpha)|u|^\alpha,$$

that is, we have a Lévy triplet $(0, 0, \nu)$ *with*

$$\nu(dz) = (c^- \mathbf{1}\{z < 0\} + c^+ \mathbf{1}\{z > 0\}) \frac{dz}{|z|^{1+\alpha}}. \tag{1.11}$$

Remark 1.2.38. *It can be shown that for an infinitely divisible distribution* μ *there is* $c > 0$ *such that*

$$\phi_\mu(u) \geqslant \exp(-c(\alpha)|u|^2), \qquad u \in \mathbb{R},$$

which is why there is no infinitely divisible (nor α*-stable) distribution for* $\alpha > 2$.

Remark 1.2.39. *The case* $\alpha = 1$ *is more complicated and implies necessarily that* $c^+ = c^-$*, that is the* 1*-stable distributions (the Cauchy distributions) are always symmetric.*

C) α-stable distributions in H:

Definition 1.2.40. *α-stable Lévy measure in a separable Hilbert space $(H, \mathcal{B}(H))$ are defined via the polar decomposition $r := |z|$ and $s := \frac{z}{|z|}$, $0 \neq z \in H$.*
An α-stable Lévy measure ν is given by a finite Radon measure σ on $\partial B_1(0)$ and index $\alpha \in (0,1)$

$$\nu(dz) := \sigma(ds)\frac{dr}{r^{1+\alpha}},$$

Remark 1.2.41. *For $H = \mathbb{R}$ we have seen these formulas in equation (1.10) for $\alpha \in (0,1)$, in equation (1.11) for $\alpha \in (1,2)$ and in (1.8) for $\alpha = 1$.*

Remark 1.2.42. *Recall that in Example 1.2.25 we have introduced in the space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ for $\sigma = U_{\partial B_1(0))}$ the rotationally invariant α-stable distribution.*

Remark 1.2.43. *A note on the infinite dimensional space $(H, \mathcal{B}(H))$. Note that in "radial" direction $r = |z| \geqslant 0$ we have a polynomial measure, which is absolutely continuous w.r.t. the Lebesgue measure in $\mathbb{R}$ or $\mathbb{R}^d$. The situation in "spherical" direction $s \in \partial B_1(0)$ is quite different. Recall that Radon measures (see for instance Elstrodt [24], Kapitel VIII) satisfy that they can be approximated from below by compact sets from above by open sets (always up to an arbitrary $\varepsilon > 0$). However, the unit sphere in infinite dimensions is never compact. Therefore most of the mass is concentrated on a (topologically) small set of directions of $\partial B_1(0)$. In addition, it is clearly possible to define a symmetric measure, however, there are no rotationally invariant α-stable processes, which due to simplicity represent the "usual" case in finite dimensions. This discussion is the equivalent phenomenon to the fact that there is no $\mathcal{N}(0, id_H)$ distribution in infinite dimensions.*

Some properties of α-stable distributions μ^α

- For $\alpha = 2$ we have $\mu^\alpha = N(0, Q)$, for some $Q \in L_1^+(H)$.
- For $\alpha \in (0,2)$ we have that μ^α is a non-Gaussian probability distribution.
- There is no α-stable distribution for $\alpha > 2$ (nor $\alpha < 0$).
- All α-stable distributions admit densities w.r.t. the Lebesgue measure, in general the densities are not know to have a closed form, except for $\mathrm{Cau}(0, b)$. The distributions are heavy-tailed, that is for large values R they behave as

$$\mathbb{P}(|Z| > R) \approx_{R\to\infty} \frac{1}{R^\alpha},$$

and therefore we see by the formula

$$\mathbb{E}[|Z|^p] = \int_0^\infty t^{p-1}\mathbb{P}(|Z| > t)dt$$

that

$$\mathbb{E}[|Z|^\alpha] = \begin{cases} < \infty & \text{for } \gamma < \alpha \\ = \infty & \text{for } \gamma \geqslant \alpha. \end{cases}$$

1.2.5 The class of regularly varying distributions

We introduce generalize α-stable distributions, to more "robust" versions, which also allow for -at least asymptotically- stability indices $\alpha > 2$.

A) Slowly varying functions

Example 1.2.44. *Consider a function $\ell : (0,\infty) \to (0,\infty)$ with $\lim_{r\to\infty} \ell(r) = c > 0$. Then for any $\lambda > 0$ we have*

$$\lim_{r\to\infty} \frac{\ell(\lambda r)}{\ell(r)} = \frac{\lim_{r\to\infty} \ell(\lambda r)}{\lim_{r\to\infty} \ell(r)} = \frac{c}{c} = 1.$$

The following definition generalizes this concept to functions, which do not necessarily converge.

Definition 1.2.45. *A* ***slowly varying function*** *ℓ (at infinity) is a measurable function $\ell : (0,\infty) \to (0,\infty)$ satisfying that for all $\lambda > 0$ we have*

$$\lim_{r\to\infty} \frac{\ell(\lambda r)}{\ell(r)} = 1.$$

EXERCISE 1.2.46.

1. *Check that any constant function is slowly varying at infinity.*
2. *Check that $\ell(r) := \ln(r)$ is a (diverging) slowly varying function at infinity.*
3. *Check that $\ell(r) := \ln(\ln(e + r))$ is a (diverging) slowly varying function at infinity.*
4. *Check that $\ell(r) := \exp(\ln(r)^\alpha)$ for $\alpha \in (0, 1)$ is a slowly varying function at infinity.*

Example 1.2.47. *The function* $\ell(r) := \exp(\ln(r)^{\frac{1}{3}} \cos(\ln(r)^{\frac{1}{3}}))$ *is a slowly varying function at infinity with* $\limsup_{r\to\infty} \ell(r) = \infty$ *and* $\liminf_{r\to\infty} \ell(r) = 0$, *see [12].*

EXERCISE 1.2.48.
- *Check that for any slowly varying function* ℓ *we have*

$$\lim_{r\to\infty} \ell(r) r^a = \begin{cases} \infty & a > 0 \\ 0 & a < 0 \end{cases}.$$

B) Regularly varying functions and distributions with index $-\alpha$ **for** $\alpha > 0$

Definition 1.2.49. *A* ***regularly varying function*** h ***with index*** $-\alpha$ *(at infinity) is a measurable function* $h : (0,\infty) \to (0,\infty)$ *satisfying that for all* $\lambda > 0$ *we have*

$$\lim_{r\to\infty} \frac{h(\lambda r)}{h(r)} = \lambda^{-\alpha}.$$

EXERCISE 1.2.50. *Check that the function* $h(r) = r^{-\alpha}$ *is regularly varying with index* $-\alpha$ *at infinity*

EXERCISE 1.2.51. *Check that a regularly varying function* h *with index* $-\alpha$ *at infinity satisfies that*

$$\ell(r) := \frac{h(r)}{r^{-\alpha}}$$

is a slowly varying function.

Definition 1.2.52. *Given a separable Hilbert space* $(H, \mathcal{B}(H))$. *Denote by* $\mathcal{M}_0(H)$ *the set of Radon measures* $\mu : \mathcal{B}(H) \to [0,\infty]$ *such that for all* $A \in \mathcal{B}(H)$ *with* $0 \notin \bar{A}$ *such that* $\mu(A) < \infty$.

Example 1.2.53. *For all* $\alpha \in (0,2)$ *any* α*-stable Lévy measure* ν *satisfies* $\nu \in \mathcal{M}_0(H)$.

Definition 1.2.54. *A regularly varying Lévy measure* $\nu \in \mathcal{M}_0(H)$ *with tail index* $-\alpha$ *satisfies that the function*

$$r \mapsto \nu(rB_1^c(0))$$

is a regularly function at infinity with index $-\alpha$.

EXERCISE 1.2.55. *Check that any* α*-stable Lévy measure* ν, $\alpha \in (0,2)$ *satisfies*

$$\nu(rB_1^c(0)) = r^{-\alpha}\nu(B_1^c(0)), \qquad r \geqslant 0,$$

and hence it is a regularly varying Lévy measure with tail index $-\alpha$.

Remark 1.2.56. *Any regularly varying Lévy measure* ν *with tail index* $-\alpha$, $\alpha > 0$ *satisfies that there exists an associated measure* $\mu \in \mathcal{M}_0(H)$ *(which is not necessarily a Lévy measure!) such that that*

$$\lim_{r\to\infty} \frac{\nu(r(B_1^c(0) \cap A))}{\nu(rB_1^c(0))} = \frac{\mu(A \cap B_1^c(0))}{\mu(B_1^c(0))}.$$

EXERCISE 1.2.57. *Check that in the case of* α*-stable* ν, $\alpha \in (0,2)$, *we have that* $\nu = \mu$.

1.2.6 Non-central local limit theorems

In Theorem 1.2.10 we have seen that the limit distributions of independent (centered) null arrays are necessarily infinitely divisibile distributions. Are there further sufficient conditions?
These are stated in so-called local limit theorems. We start with the easiest version of them, according to which the renormalized sum of i.d.d. random variables, whose distribution tails decay essentially as the tails of α-stable distributions converge to α-stable distributions.

Theorem 1.2.58 (Local limit theorem 1 in $\mathbb{R}$)**.** *Consider an i.i.d. sequence* $(X_n)_{n\in\mathbb{N}}$ *of random variables* $X_n : \Omega \to \mathbb{R}$ *with* $\mathbb{P}_{X_1} \neq \delta_c$ *for any* $c \in \mathbb{R}$ *and set* $S_n = \sum_{i=1}^n X_i$. *Then the following are equivalent:*

1. *There are sequences* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *of nonnegative numbers and a probability distribution* μ *in* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *such that*

$$\frac{S_n - b_n}{a_n} \xrightarrow{d} X, \qquad X \sim \mu.$$

2. *There is* $\alpha \in (0,2]$ *such that the function of the second moments of the values* $\leqslant x$

$$R(x) := \mathbb{E}[X_1^2 \mathbf{1}\{|X_1| \leqslant x\}]$$

is regularly varying with index $-(2-\alpha)$.
In case of $\alpha \in (0,2)$ *it holds additionally that the balance*

$$p := \lim_{x\to\infty} \frac{\mathbb{P}(X_1 > x)}{\mathbb{P}(|X_1| > x)} \in (0,\infty).$$

Remark 1.2.59. *Note that* $R(x) \in [0, x^2]$ *for any* $x > 0$. *Then Theorem 1.2.58 states that the divergence* $R(x) \sim x^{2-\alpha} \to \infty$ *(of* $\mathbb{P}_{X_1}$*) is essentially equivalent to the convergence of* S_n *(renormalized) to a distribution* μ, *which (by Theorem 1.2.10) is necessarily infinitely divisible.*

Remark 1.2.60. *More generally, the distribution is necessarily a* ***stable distribution in the wide sense with index*** $\alpha \in (0,2]$, *for* $X \sim \mu$ *and any i.i.d. family* $(X_n)_{n\in\mathbb{N}}$ *with* $X_n \sim \mu$ *we have sequences* $(a_n)_{n\in\mathbb{N}}$ *with* $a_n > 0$ *and* $(b_n)_{n\in\mathbb{N}}$, $b_n \in \mathbb{R}$ *such that for all* $n \in \mathbb{N}$

$$\frac{1}{n^{\frac{1}{\alpha}}}(X_1 + \cdots + X_n) \stackrel{d}{=} X + b_n.$$

In the following we give precise conditions.

Theorem 1.2.61 (Local limit theorem 2 in $\mathbb{R}$)**.** *Consider an i.i.d. sequence* $(X_n)_{n\in\mathbb{N}}$ *of random variables* $X_n : \Omega \to \mathbb{R}$ *with* $\mathbb{P}_{X_1} \neq \delta_c$ *for any* $c \in \mathbb{R}$ *and set* $S_n = \sum_{i=1}^n X_i$.

1. *Assume that the asymptotic balance* $p := \lim_{x\to\infty} \frac{\mathbb{P}(X_1 > x)}{\mathbb{P}(|X_1| > x)} \in (0,\infty)$ *is finite.*
2. *Further assume that the function of the second moments of the values* $\leqslant x$
$$R(x) := \mathbb{E}[X_1^2 \mathbf{1}\{|X_1| \leqslant x\}]$$
there is $\alpha \in (0,2]$ *such that the function* $R(x)$ *is regularly varying at infinity with index* $-(2-\alpha)$.
3. *We may assume that the sequence* $(a_n)_{n\in\mathbb{N}}$ *with* $a_n \to \infty$ *satisfies*
$$\lim_{n\to\infty} \frac{nR(a_n)}{a_n^2} = C \in (0,\infty).$$

Then there is a sequence $(b_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ *such that*

$$\frac{S_n - b_n}{a_n} \xrightarrow{d} \mu^\alpha, \qquad n \to \infty,$$

where μ^α *is the* α*-stable distribution with Lévy measure*

$$\nu(dz) = (c^+ \mathbf{1}\{z > 0\} + c^- \mathbf{1}\{z < 0\}) \frac{dz}{|z|^{1+\alpha}},$$

with $c^+ = Cp$, $c^- = C(1-p)$ *and*

1. *for* $\alpha \in (0,1)$ *with* $b_n \equiv 0$,
2. *for* $\alpha \in (1,2)$ *with* $b_n = n\mathbb{E}[X_1]$
3. *for* $\alpha = 1$ *and* $b_n = na_n\mathbb{E}[\sin(\frac{X_1}{n})]$.

Consult the excellent monograph Klenke [45] for further conditions and references.

Example 1.2.62. *For instance if* $a_n = n$ *and*

$$\frac{R(n)}{n} \longrightarrow c \in (0, \infty)$$

this implies $\alpha = 1$ *and*

$$\frac{S_n - n^2 \mathbb{E}[\sin(\frac{X_1}{n})]}{n} \longrightarrow \mu^1 = \text{Cau}(0, 1).$$

A version for separable Hilbert spaces is the following, implied by a more general version given in [2].

Theorem 1.2.63 (Araujo/Giné)**.** *Let* μ^α *be an* α*-stable law in a separable Hilbert space* $(H, \mathcal{B}(H))$ *and* $X \sim \mu$ *be a* H*-valued random variable. Then the following are equivalent:*

1. *There is a sequence of vectors* $(b_n)_{n\in\mathbb{N}}$ *such that for any i.i.d. sequence* $(X_n)_{n\in\mathbb{N}}$ *with* $X_n \sim \mu$ *we have*

$$\frac{1}{n^{\frac{1}{\alpha}}} \sum_{k=1}^{n} X_k - b_n \xrightarrow{d} \mu^\alpha, \qquad \text{as } n \to \infty.$$

2. *The following three conditions are satisfied:*

   (a) *For any* $x \in H$ *we have*

$$\frac{1}{n^{\frac{1}{\alpha}}} \sum_{k=1}^{n} \langle X_k, x\rangle - \langle b_n, x\rangle \xrightarrow{d} \mu^\alpha \circ \langle \cdot, x\rangle^{-1}$$

   (b) *For any* $\delta > 0$ *we have*

$$\sup_{n\in\mathbb{N}} n\mathbb{P}(|X| > \delta n^{\frac{1}{\alpha}}) < \infty.$$

   (c) *There exists an increasing sequence of finite dimensional subspaces* $F_m \subset H$ *and a constant* $\delta_0 > 0$ *such that*

$$\lim_{m\to\infty} \limsup_{n\to\infty} n\mathbb{P}(\inf_{y\in F_n} |X - y| > \delta_0 n^{\frac{1}{\alpha}}) = 0.$$

Chapter 2

Lévy processes

2.1 Examples, definition and basic properties of Lévy processes

In the following subsections we shall construct examples of different Lévy processes over a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

2.1.1 Construction of Q-Brownian motion

We start with the classical explicit construction of Brownian motion, going back to Lévy, see for instance [39].

A) The general definition of a Q-Brownian motion. We start with the general definition, which we shall recover later.

Definition 2.1.1. *Consider a separable Hilbert space $(H, \mathcal{B}(H))$ and given an operator $Q \in L_1^+(H)$. On a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a **Q-Brownian motion** is a stochastic process $(B_t)_{t\geqslant 0}$ with values in H satisfying the following.*

1. $B_0 = 0$, *$\mathbb{P}$-a.s.*
2. *For any $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_n$ the vector of increments*

$$(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}})$$

   is an independent family of random vectors.

3. *For any $0 \leqslant s < t$ we have the stationarity of increments*

$$B_t - B_s \sim B_{t-s} - B_0 = B_{t-s} \sim \mathcal{N}(0, Q(t-s)).$$

4. *For $\mathbb{P}$-almost any $\omega \in \Omega$ we have that $t \mapsto B_t(\omega)$ is continuous.*

B) The construction of a $\mathbb{R}$-valued Brownian motion on $[0,1]$:

Remark 2.1.2. *Note that for $H = \mathbb{R}$ and $Q = 1$ item 2) and 3) can be considered jointly in the sense that the random vector*

$$(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}})$$

is a Gaussian vector with distribution $\mathcal{N}(0, D)$, where
$D = diag(t_1, t_2 - t_1, \ldots, t_n - t_{n-1})$.

Definition 2.1.3. *A $\mathbb{R}$-valued stochastic process $(X_t)_{t\in[0,1]}$ on a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is called a Gaussian process if for all $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_n \leqslant 1$ the vector*

$$(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$$

is a Gaussian random vector, i.e. with a distribution $\mathcal{N}(m_n, Q)$ for some $m_n \in \mathbb{R}^n$ and $Q_n \in L_1^+(\mathbb{R}^n)$, that is $Q \in \mathbb{R}^{n\otimes n}$ being a symmetric, nonnegative definite matrix.

Lemma 2.1.4. *The (finite-dimensional) distribution(s) of a Gaussian process is uniquely defined via the $\mathbb{R}$-valued functions*

$$\mu(t) := \mathbb{E}[X_t] \qquad C(t,s) := \mathbb{E}\big[X_t X_s\big], \qquad 0 \leqslant s, t \leqslant 1.$$

EXERCISE 2.1.5. *Show the preceding lemma.*

EXERCISE 2.1.6. *Show that for a $\mathbb{R}$-valued Brownian motion $(B_t)_{t\in[0,1]}$ and $Q = 1$ we have $\mu(t) = 0$ and $C(t,s) = t \wedge s$.*

Lemma 2.1.7. *Let $(X_t)_{t\in[0,1]}$ be a Gaussian process with values in $\mathbb{R}$ satisfying $\mu(t) = 0$ for all $t \in [0,1]$ and $C(t,s) = s \wedge t$ for $s,t \in [0,1]$. Then $(X_t)_{t\in[0,1]}$ is a process with independent increments. If $t \mapsto X_t$ is continuous $\mathbb{P}$-a.s., then the process $(X_t)_{t\in[0,1]}$ is a Brownian motion in $\mathbb{R}$.*

EXERCISE 2.1.8. *Show the preceding lemma.*

The skeleton: the Haar basis and its integrals

Consult the figures in the video. Denote by

$$H(t) := \begin{cases} -1 & t \in [0, \frac{1}{2}) \\ 1 & t \in [\frac{1}{2}, 1] \\ 0 & \text{else} \end{cases}, \qquad t \in \mathbb{R},$$

the so-called **mother wavelet** and the **Haar functions**

$$
\begin{aligned}
&H_0(t) := \mathbf{1}_{[0,1]}(t)\\
&H_1(t) := H(t)\\
&H_2(t) := 2^{\frac{1}{2}} H(2t)\\
&H_3(t) := 2^{\frac{1}{2}} H(2t-1)\\
&H_4(t) := 2^{\frac{2}{2}} H(2^2 t)\\
&H_5(t) := 2^{\frac{2}{2}} H(2^2 t-1)\\
&H_6(t) := 2^{\frac{2}{2}} H(2^2 t-2)\\
&\vdots\\
&H_n(t) = 2^{\frac{j}{2}} H(2^j t-k), \qquad \text{for } n = 2^j+k, k = 0, \dots, 2^j-1
\end{aligned}
$$

Consult the figures in the video.

EXERCISE 2.1.9.
- *Sketch the first* 16 *functions* H_n *in the same (sufficiently large) plot.*
- *Check that the sequence of Haar functions* $(H_n)_{n\in\mathbb{N}}$ *is an orthonormal basis of* $L^2((0,1), dx)$*.*

By construction, the integrals over the elements of H_n are translated hat functions G_n on disjoint dyadic intervals.

Consult the figures in the video. We renormalize these hat functions uniformly to size 1 and by the recursion for any $t \in \mathbb{R}$

$$
G(t) := 2\int_0^t H(s)ds = \begin{cases} 2t & t \in [0, \frac{1}{2}) \\ 2-2t & t \in [\frac{1}{2}, 1] \\ 0 & \text{else.} \end{cases}, \qquad \text{note: } G([0,1]) \in [0,1].
$$

A simple calculation yields for any $t \in [0,1]$ and $\mathbb{N} \ni n = 2^j + k$ (this representation is unique!)

$$
\int_0^t H_n(r)dr = \lambda_n G_n(t), \qquad \text{where } \lambda_n := 2^{-\frac{j}{2}-1}, \tag{2.1}
$$

and set $G_0(t) = t$, $G_1(t) = G(t)$ and for $n \geqslant 2$

$$
G_n(t) = G(2^j t - k), \qquad t \in [0,1], n = 2^j + k.
$$

EXERCISE 2.1.10. • *Calculate $G_0, \ldots, G_{16}$ and sketch the first 16 functions G_n in the same (sufficiently large) plot.*

- *Verify equation (2.1).*

The actual construction: Consider an i.i.d. family $(Z_n)_{n\in\mathbb{N}}$ of $\mathcal{N}(0,1)$-distributed random variables $Z_n : \Omega \to \mathbb{R}$. Define the process $(B_t)_{t\in[0,1]}$ defined as the random function

$$B_t(\omega) := \sum_{n=0}^{\infty} \lambda_n G_n(t) Z_n(\omega), \qquad t \in [0,1],\ \omega \in \Omega. \tag{2.2}$$

The following lemma controls the random fluctuations in (2.2) and is not complicate to show with the help of the Borel-Cantelli lemma.

Lemma 2.1.11. *Consider an i.i.d. sequence $(Z_n)_{n\in\mathbb{N}}$ of $\mathbb{R}$-valued standard normal random variables $Z_n \sim \mathcal{N}(0,1)$. Then there exists a random variable $C : \Omega \to (0,\infty)$ such that for all $n \geqslant 2$ we have*

$$|Z_n(\omega)| \leqslant C(\omega)\sqrt{\ln(n)}, \qquad \text{for } \mathbb{P}\text{-a.a. } \omega \in \Omega.$$

With the help of Lemma 2.1.11 we show that (2.6) yields a $\mathbb{R}$-valued Brownian motion in $[0,1]$.

$\mathbb{P}$-a.s. convergence and continuous limit: First we show that the sequence $\sum_{n=0}^{m} \lambda_n G_n(t) Z_n$ converges $\mathbb{P}$-a.s. as $m \to \infty$ uniformly in $t \in [0,1]$ and that the limit $t \mapsto B_t$ is $\mathbb{P}$-a.s. continuous.
Consider for $t \in [0,1]$ the tail of the sequence, for $m \geqslant 2^j$ for some $j \geqslant 1$. As always $n = 2^j + k$ for $k = 0, \ldots, 2^{j-1}$, that is $j = j(n) = \lfloor \mathrm{lb}(n) \rfloor$. Then using that for $k = 0, \ldots 2^j - 1$ the functions G_{2^j+k} have disjoint support and $0 \leqslant G_n(t) \leqslant 1$ we obtain

$$\begin{aligned}
\sum_{n=m}^{\infty} \lambda_n |Z_n| G_n(t) &\leqslant C \sum_{n=m}^{\infty} \lambda_n \sqrt{\ln(n)} G_n(t) \\
&\leqslant C \sum_{j=\lfloor \mathrm{lb}(m) \rfloor}^{\infty} 2^{-\frac{j}{2}-1} \sqrt{j+1} \sum_{k=0}^{2^j-1} G_{2^j+k}(t) \\
&\leqslant C \sum_{j=\lfloor \mathrm{lb}(m) \rfloor}^{\infty} 2^{-\frac{j}{2}-1} \sqrt{j+1} \longrightarrow 0, \qquad \mathbb{P}\text{-a.s.}, m \to \infty.
\end{aligned}$$

Hence $(B_t)_{t\in[0,1]}$ is the limit of a family of continuous functions which converges $\mathbb{P}$-a.s. uniformly.

Convergence in L^2 and first and second moments: $(B_t)_{t\in[0,1]}$ is also limit in L^2 since the $(Z_n)_{n\in\mathbb{N}}$ are i.i.d.

$$\mathbb{E}\Big[\Big(\sum_{n=m}^{\infty}\lambda_n Z_n G_n(t)\Big)^2\Big] = \sum_{n=m}^{\infty}\lambda_n^2 G_n(t)^2 \leqslant \sum_{j=j_0}^{\infty} 2^{-j-2} \longrightarrow 0, n\to\infty$$

such that

$$\mathbb{E}[B_t^2] = \sum_{n=0}^{\infty}\lambda_n^2 G_n(t) < \infty. \tag{2.3}$$

Hence $\mathbb{E}[|B_t|]$ is finite and $\mathbb{E}[B_t]=0$ by construction.

The correct covariance structure: In order to check the covariance structure we now use the structure of $(H_n)_{n\in\mathbb{N}}$. In addition, we calculate

$$\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]}\rangle = \int_0^1 \mathbf{1}_{[0,t]}(r)\mathbf{1}_{[0,s]}(r)dr = \int_0^1 \mathbf{1}_{[0,t\wedge s]}(r)dr = t\wedge s.$$

Using the polarization identity and (2.3) we obtain

$$\begin{aligned}
C(t,s) = \mathbb{E}[B_tB_s] &= \frac{1}{4}\big(\mathbb{E}[(B_t+B_s)^2] + \mathbb{E}[(B_t-B_s)^2]\big)\\
&= \sum_{n=0}^{\infty}\lambda_n^2 G_n(t)G_n(s)\\
&= \sum_{n=0}^{\infty}\lambda_n^2 \int_0^t H_n(r)dr\int_0^s H_n(r)dr\\
&= \sum_{n=0}^{\infty}\lambda_n^2 \langle \mathbf{1}_{[0,t]}, H_n\rangle\langle \mathbf{1}_{[0,s]} H_n\rangle\\
&= \langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]}\rangle\\
&= \int_0^1 \mathbf{1}_{[0,t\wedge s]}(r)dr\\
&= t\wedge s.
\end{aligned}\tag{2.4}$$

Hence $(B_t)_{t\in[0,1]}$ is a continuous, centered Gaussian process with $C(t,s)=t\wedge s$ and by Lemma 2.1.7 $(B_t)_{t\in[0,1]}$ is a Brownian motion.

EXERCISE 2.1.12. *Verify step by step equation (2.4).*

C) Standard Brownian Motion in $\mathbb{R}$: Brownian motion $(B_t)_{t\geqslant 0}$ is defined via a sequence of i.i.d. Brownian motions $\big((B_t^n)_{t\in[0,1]}\big)_{n\in\mathbb{N}}$ and constructed via

$$B_t(\omega) := \sum_{n=1}^{\lfloor t\rfloor} B_1^n(\omega) + B_{t-\lfloor t\rfloor}^{\lfloor t\rfloor+1}(\omega).$$

D) Standard Brownian Motion in $\mathbb{R}^d$: Standard Brownian motion in $\mathbb{R}^d$ with canonical basis $(e_n)_{n\in\mathbb{N}}$ is defined via a sequence of i.i.d. Brownian motions $\big((B_t^i)_{t\geqslant 0}\big)_{i=1,\dots,d}$

$$B_t(\omega) := B_t^1(\omega)e_1 + \cdots + B_t^d(\omega)e_d \tag{2.5}$$

EXERCISE 2.1.13. *Show that (2.5) defines an Q-Brownian motion for $Q = id_{\mathbb{R}^d}$.*

E) Q-Brownian motion in a separable Hilbert space H: For an infinite dimensional separable Hilbert $(H, \mathcal{B}(H))$ with orthonormal basis $(e_n)_{n\in\mathbb{N}}$, a square summable sequence $(\lambda_n)_{n\in\mathbb{N}}$ with $0 < \lambda_{n+1} \leqslant \lambda_n$ and $\sum_{n=1}^\infty \lambda_n^2 < \infty$ and an i.i.d sequence of i.i.d. Brownian motions $\big((B_t^n)_{t\geqslant 0}\big)_{n\in\mathbb{N}}$ we define

$$B_t(\omega) := \sum_{n=1}^{\infty} \lambda_n B_t^n(\omega)e_n. \tag{2.6}$$

Note that

$$\mathbb{E}\Big[|B_t|^2\Big] = \sum_{n=1}^{\infty} \mathbb{E}[|\lambda_n B_t^n(\omega)|^2] = t\sum_{n=1}^{\infty} \lambda_n^2 < \infty.$$

Hence $\mathbb{E}[|B_t|] < \infty$ and $\mathbb{E}[B_t] = 0$.
Recall Exercise 1.2.16

EXERCISE 2.1.14. *Check that (2.6) defines a Q-Brownian motion for*

$$Qx := \sum_{n=1}^{\infty} \lambda_n^2 \langle e_n, x\rangle e_n, \qquad x \in H.$$

2.1.2 Construction of a Poisson and a compound Poisson process

A) The Poisson process: Consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, $\lambda > 0$ and an i.i.d. family $(\tau_k)_{k\in\mathbb{N}}$ of random variables $\tau_k : \Omega \to [0,\infty)$ with $\tau_k \sim$

Exp_λ, that is $\mathbb{P}(\tau_k > t) = e^{-\lambda t}$, $t > 0$. This family is called a family of (memoryless) **waiting times**.
We define the family $(T_k)_{k\in\mathbb{N}}$ of **arrival times** $T_k := \sum_{j=1}^k \tau_j$. Then a **Poisson process** $(\pi_t)_{t\geqslant 0}$ **on** $(\Omega, \mathcal{A}, \mathbb{P})$ **with intensity** $\lambda > 0$ is defined by

$$\pi_t(\omega) := \sum_{k=1}^{\infty} \mathbf{1}\{T_k(\omega) < t\}, \qquad t \geqslant 0,\ \omega \in \Omega.$$

Consult the figures in the video.

Lemma 2.1.15 (Poisson processes are Lévy processes with Poisson marginals)**.**
Let π be a Poisson process with intensity $\lambda > 0$. Then we have

a) $\mathcal{L}(\pi_t) = \mathrm{Poi}_{\lambda t}$ *for all* $t \geqslant 0$.

b) $\mathcal{L}(\pi_t - \pi_s) = \mathcal{L}(\pi_{t-s})$ *for all* $s \leqslant t$.

c) For all $0 = t_0 < t_1 < \ldots < t_n < \infty$, $n \in \mathbb{N}$ *the family of random variables*

$$(\pi_{t_n} - \pi_{t_{n-1}}, \ldots, \pi_{t_1} - \pi_{t_0})$$

is independent.

The core of the proof is the following sub simplex volume.

Lemma 2.1.16 (Subsimplex volume)**.**
For all $n \in \mathbb{N}$ und $t > 0$ we have

$$\int_0^\infty \cdots \int_0^\infty \mathbf{1}\{z_1 + \cdots + z_n \leqslant t\} dz_1 \ldots dz_n = \frac{t^n}{n!}.$$

EXERCISE 2.1.17. *Show the preceding lemma and illustrate it graphically in low dimensions.*

Proof. of Lemma 2.1.15: The proof is elementary and given for instance in Georgii [30]. However, it is crucial for the general understanding of Poisson processes, which is why we provide it here. It is sufficient to show, that for all $n, k_1, \ldots, k_n \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_n < \infty$ we have the identity

$$\mathbb{P}((\pi_{t_n} - \pi_{t_{n-1}}, .., \pi_{t_1} - \pi_{t_0}) = (k_n, \ldots, k_1)) = \\ \frac{(\lambda(t_n - t_{n-1}))^{k_n} e^{-\lambda(t_n - t_{n-1})}}{k_n!} \cdots \frac{(\lambda(t_1 - t_0))^{k_1} e^{-\lambda(t_1 - t_0)}}{k_1!}.$$

We show the case $n = 2$ setting $k_2 = \ell$ and $k_1 = k$. The case of general $n+1$ follows then by an induction step with induction assumption n analogous to the case $n = 2$.

The basic idea consists in an appropriate change of coordinates since the vector of waiting times $(\tau_1, \dots, \tau_{k+\ell+1})$ is independent by assumptions and has the $k+\ell+1$-fold product density of the family $f_{\tau_i}(x) = \lambda e^{-\lambda x}$ for any i, which we denote by $f_{(\tau_1,\dots,\tau_{k+\ell+1})} : [0,\infty)^{k+\ell+1} \to [0,\infty)$

$$\begin{aligned} f_{(\tau_1,\dots,\tau_{k+\ell+1})}(x_1,\dots,x_{k+\ell+1}) &:= \prod_{i=1}^{k+\ell+1} f_{\tau_i}(x_i) = \lambda e^{-\lambda x_1} \dots \lambda e^{-\lambda x_{k+\ell+1}} \\ &= \lambda^{k+\ell+1} e^{-\lambda(x_1+\dots+x_{k+\ell+1})}. \end{aligned}$$

Let $\ell \geqslant 1$ and $0 < s < t$

$$\begin{aligned} &\mathbb{P}(\pi_s = k, \pi_t - \pi_s = \ell) = \mathbb{P}(T_k \leqslant s < T_{k+1},\ T_{k+\ell} \leqslant t < T_{k+\ell+1}) \\ &= \mathbb{E}\left[\mathbf{1}\{\sum_{i=1}^{k} \tau_i \leqslant s < \sum_{i=1}^{k+1} \tau_i,\ \sum_{i=1}^{k+\ell} \tau_i \leqslant t < \sum_{i=1}^{k+\ell+1} \tau_i\}\right] \\ &= \int_0^\infty \Big(\int_0^\infty \cdots \int_0^\infty \mathbf{1}\{x_1 + \dots + x_k \leqslant s\}\mathbf{1}\{s < x_1 + \dots + x_k + x_{k+1}\} \\ &\qquad \mathbf{1}\{x_1 + \dots + x_{k+\ell} \leqslant t\}\mathbf{1}\{t < \underbrace{x_1 + \dots + x_k + x_{k+\ell+1}}_{=y}\} \\ &\qquad \lambda^{k+\ell+1} \exp\Big(\lambda(\underbrace{x_1 + \dots + x_{k+\ell+1}}_{=y})\Big) dx_1 \dots d_{k+\ell}\Big) dx_{k+\ell+1} =: J. \end{aligned}$$

We first integrate w.r.t. $x_{k+\ell+1}$ with all other nonnegative variables $x_1, \dots, x_{k+\ell}$ being fixed parameters and substitute

$$y := \sum_{i=1}^{k+\ell+1} x_i, \qquad \frac{dy}{dx_{k+\ell+1}} = 1,$$

$$\text{with the borders:} \quad x_{k+\ell+1} = 0 \quad \Rightarrow \quad y = \sum_{i=1}^{k+\ell} x_i = t,$$

$$\text{and} \quad x_{k+\ell+1} = \infty \quad \Rightarrow \quad y = \infty,$$

and obtain

$$J = \int_t^\infty \Big(\int_0^\infty \cdots \int_0^\infty \mathbf{1}\{x_1 + \cdots + x_k \leqslant s\}\mathbf{1}\{s < x_1 + \cdots + x_k + x_{k+1}\}$$

$$\mathbf{1}\{x_1 + \cdots + x_{k+\ell} \leqslant t\}\lambda^{k+\ell} dx_1 \dots dx_{k+\ell}\Big) \lambda e^{\lambda y} dy$$

$$= e^{-\lambda t} \int_0^\infty \cdots \int_0^\infty \Big(\int_0^\infty \cdots \int_0^\infty \mathbf{1}\{x_1 + \cdots + x_k \leqslant s\}\lambda^k$$

$$\mathbf{1}\{0 < \underbrace{x_1 + \cdots + x_k + x_{k+1} - s}_{=z_1}\}$$

$$\mathbf{1}\{\underbrace{x_1 + \cdots + x_{k+1} - s}_{=z_1} + x_{k+2} + \cdots + x_{k+\ell} \leqslant t - s\}$$

$$\lambda^\ell dx_1 \dots dx_k \Big) dx_{k+1} \dots dx_{k+\ell} =: K.$$

Now we integrate $x_{k+1}, \dots, x_{k+\ell}$ and substitute

$$z_1 = x_1 + \cdots + x_{k+1} - s, \qquad \text{and} \qquad z_i = x_{k+i} \quad \text{for } i = 2, \dots, \ell$$

and obtain with the help of Lemma 2.1.16

$$K = e^{-\lambda t} \int_0^\infty \cdots \int_0^\infty \Big(\int_0^\infty \cdots \int_0^\infty \mathbf{1}\{x_1 + \cdots + x_k \leqslant s\}\lambda^k dx_1 \dots dx_k \Big)$$

$$\mathbf{1}\{z_1 + \cdots + z_\ell \leqslant (t - s)\}\lambda^\ell dz_1 \dots dz_\ell$$

$$= e^{-\lambda t} \frac{(\lambda(t-s))^\ell}{\ell!} \int_0^\infty \cdots \int_0^\infty \mathbf{1}\{x_1 + \cdots + x_k \leqslant s\}\lambda^k dx_1 \dots dx_k$$

$$= e^{-\lambda(t-s)} \frac{(\lambda(t-s))^\ell}{\ell!}\; e^{-\lambda s} \frac{(\lambda s)^k}{k!}.$$

□

B) The compound Poisson processes: On a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a compound Poisson process $(C_t)_{t\geqslant 0}$ consists of a Poisson process $(\pi_t)_{t\geqslant 0}$ of intensity $\lambda > 0$ and an i.i.d. family $(Z_k)_{k\in\mathbb{N}}$ of random vectors $Z_k : \Omega \to H$ in a separable Hilbert space $(H, \mathcal{B}(H))$ with distribution $Z_k \sim \mu$ such that

$$C_t(\omega) := \sum_{k=1}^{\pi_t(\omega)} Z_k(\omega).$$

We have already calculated in Lemma 1.2.21 that

$$\phi_{C_t}(u) = \exp(t\lambda \int_H (e^{i\langle u,z\rangle} - 1)\mu(dz)), \qquad u \in H.$$

Lemma 2.1.18. *1. If $\int_H |z|\mu(dz) < \infty$, then*

$$\mathbb{E}[C_t] = t\lambda \int_H z\mu(dz).$$

2. If $\int_H |z|^2\mu(dz) < \infty$, then

$$\mathbb{V}(C_t) = \mathbb{E}[|C_t|^2] - |\mathbb{E}[C_t]|^2 = t\lambda \int_H |z|^2\mu(dz).$$

EXERCISE 2.1.19. *Proof the preceding result for $H = \mathbb{R}^d$ by differentiating the characteristic function. Consult the appendix.*

Properties of a compound Poisson process:

Lemma 2.1.20. *1. $C_0 = 0$ $\mathbb{P}$-a.s.*

2. For any $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_n$ the vector of increments

$$(C_{t_1}, C_{t_2} - C_{t_1}, \ldots, C_{t_n} - C_{t_{n-1}})$$

forms an independent family of random variables.

3. For any $0 \leqslant s < t$ we have the stationarity of increments

$$C_t - C_s \sim C_{t-s} - C_0 = C_{t-s} \sim \mathrm{Cpp}(t\lambda, \mu).$$

4. For $\mathbb{P}$-a.a. ω we have $t \mapsto C_t(\omega)$ is càdlàg.

Proof. 1) By construction.

4) By construction, the trajectories $t \mapsto C_t(\omega)$ are right-continuous and have left limits.

2) Stationarity of the increments. We play everything back to the same properties of the Poisson process and the independence of π and the

family $(Z_k)_{k\in\mathbb{N}}$.

$$
\begin{aligned}
&\phi_{C_t-C_s}(u)\\
&= \mathbb{E}[e^{i\langle u, C_t-C_s\rangle}] = \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=\pi_s+1}^{\pi_t} Z_\ell\rangle} \mid \pi_t-\pi_s=k]\mathbb{P}(\pi_t-\pi_s=k)\\
&= \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=\pi_s+1}^{\pi_s+k} Z_\ell\rangle} \mid \pi_t-\pi_s=k]\mathbb{P}(\pi_t-\pi_s=k)\\
&= \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=\pi_s+1}^{\pi_s+k} Z_\ell\rangle}]\mathbb{P}(\pi_t-\pi_s=k)\\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=\pi_s+1}^{\pi_s+k} Z_\ell\rangle} \mid \pi_s=j]\mathbb{P}(\pi_s=j)\mathbb{P}(\pi_t-\pi_s=k)\\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=j+1}^{j+k} Z_\ell\rangle}]\mathbb{P}(\pi_s=j)\mathbb{P}(\pi_t-\pi_s=k)\\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=j+1}^{j+k} Z_\ell\rangle}]\mathbb{P}(\pi_s=j)\mathbb{P}(\pi_t-\pi_s=k)\\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=j+1}^{j+k} Z_\ell\rangle}]\mathbb{P}(\pi_t-j=k, \pi_s=j) =: J_1.
\end{aligned}
$$

We continue

$$
\begin{aligned}
J_1 &= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=1}^{k} Z_\ell\rangle}]\mathbb{P}(\pi_t=k+j, \pi_s=j)\\
&= \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=1}^{k} Z_\ell\rangle}]\sum_{j=0}^{\infty}\mathbb{P}(\pi_t-\pi_s=k, \pi_s=j)\\
&= \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=1}^{k} Z_\ell\rangle}]\sum_{j=0}^{\infty}\mathbb{P}(\pi_t-\pi_s=k, \pi_s=j)\\
&= \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=1}^{k} Z_\ell\rangle}]\mathbb{P}(\pi_t-\pi_s=k)\\
&= \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=1}^{k} Z_\ell\rangle}]\mathbb{P}(\pi_{t-s}=k) =: J_2.
\end{aligned}
$$

$$J_2 = \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=1}^{k} Z_\ell \rangle} \mid \pi_{t-s} = k]\mathbb{P}(\pi_{t-s} = k)$$
$$= \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{\ell=1}^{\pi_{t-s}} Z_\ell \rangle} \mid \pi_{t-s} = k]\mathbb{P}(\pi_{t-s} = k)$$
$$= \mathbb{E}[e^{i\langle u, \sum_{\ell=1}^{\pi_{t-s}} Z_\ell \rangle}] = \phi_{C_{t-s}}(u).$$

3) Independence of the increments. We show the case of two increments. The general case follows by induction. Let $0 < s < t$ then for $P_{t,s,k,\ell} = \mathbb{P}(\pi_t - \pi_s = k, \pi_s = \ell)$ we have

$$\phi_{(C_s, C_t - C_s)}(u, v) = \mathbb{E}[e^{i\langle (u,v),(C_s, C_{t-C_s})\rangle}] = \mathbb{E}[e^{i,\langle u, C_s\rangle} e^{i\langle v, C_t - C_s\rangle}]$$
$$= \sum_{k,\ell=0}^{\infty} \mathbb{E}[e^{i,\langle u, \sum_{j=1}^{\pi_s+1} Z_j\rangle} e^{i\langle v, \sum_{m=\pi_s}^{\pi_t} Z_m\rangle} \mid \pi_s = \ell, \pi_t - \pi_s = k]P_{t,s,k,\ell}$$
$$= \sum_{k,\ell=0}^{\infty} \mathbb{E}[e^{i,\langle u, \sum_{j=1}^{\pi_s+1} Z_j\rangle} e^{i\langle v, \sum_{m=\pi_s}^{\pi_t} Z_m\rangle} \mid \pi_s = \ell, \pi_t - \pi_s = k]P_{t,s,k,\ell}$$
$$= \sum_{k,\ell=0}^{\infty} \mathbb{E}[e^{i,\langle u, \sum_{j=1}^{\pi_s+1} Z_j\rangle} e^{i\langle v, \sum_{m=\pi_s}^{\pi_t} Z_m\rangle} \mid \pi_s = \ell, \pi_t - \pi_s = k]P_{t,s,k,\ell}$$
$$= \sum_{k,\ell=0}^{\infty} \mathbb{E}[e^{i,\langle u, \sum_{j=1}^{\ell+1} Z_j\rangle} e^{i\langle v, \sum_{m=\ell}^{\ell+k} Z_m\rangle} \mid \pi_s = \ell, \pi_t - \pi_s = k]P_{t,s,k,\ell} =: J_3.$$

We continue

$$J_3 = \sum_{k,\ell=0}^{\infty} \mathbb{E}[e^{i,\langle u, \sum_{j=1}^{\ell} Z_j\rangle} e^{i\langle v, \sum_{m=\ell+1}^{\ell+k} Z_m\rangle}]\mathbb{P}(\pi_t - \pi_s = k, \pi_s = \ell)$$
$$= \sum_{k,\ell=0}^{\infty} \mathbb{E}[e^{i,\langle u, \sum_{j=1}^{\ell} Z_j\rangle}]\mathbb{E}[e^{i\langle v, \sum_{m=\ell+1}^{\ell+k} Z_m\rangle}]\mathbb{P}(\pi_t - \pi_s = k)\mathbb{P}(\pi_s = \ell)$$
$$= \sum_{\ell=0}^{\infty} \mathbb{E}[e^{i,\langle u, \sum_{j=1}^{\ell} Z_j\rangle}]\mathbb{P}(\pi_s = \ell) \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle v, \sum_{m=0}^{k} Z_m\rangle}]\mathbb{P}(\pi_t - \pi_s = k) = J_4$$

and hence

$$J_4 = \sum_{\ell=0}^{\infty} \mathbb{E}[e^{i,\langle u, \sum_{j=1}^{\ell} Z_j\rangle} \mid \pi_s = \ell]\mathbb{P}(\pi_s = \ell)$$
$$\cdot \sum_{k=0}^{\infty} \mathbb{E}[e^{i\langle v, \sum_{m=1}^{k} Z_m\rangle} \mid \pi_t - \pi_s = k]\mathbb{P}(\pi_t - \pi_s = k)$$
$$= \mathbb{E}[e^{i\langle u, C_s\rangle}]\mathbb{E}[e^{i\langle v, C_{t-s}\rangle}]$$
$$= \mathbb{E}[e^{i\langle u, C_s\rangle}]\mathbb{E}[e^{i\langle v, C_t - C_s\rangle}]$$
$$= \phi_{C_s}(u)\phi_{C_t - C_s}(v).$$

Hence a compound Poisson process has independent increments.

□

EXERCISE 2.1.21. *Go through the proof of Lemma 2.1.20 and clarify for yourself each single identity.*

2.1.3 Lévy processes

Looking at the Definition 2.6 of a Q-Brownian motion, Lemma 2.1.15 for Poisson process and Lemma 2.1.20 for the compound Poisson process, we abstract what a Lévy process should be:

- a process **starting in** 0,
- having **stationary, infinitely divisible increments**
- and **independent non-overlapping increments**,
- with some kind of **(weaker than path-by-path, think of the Poisson process) continuity property**.

A) Definition and examples:

Definition 2.1.22. *On a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a **Lévy process** $(L_t)_{t\geqslant 0}$ **with values in a separable Hilbert space** $(H, \mathcal{B}(H))$ is a family of random vectors $L_t : \Omega \to H$ such that*

1. $L_0 = 0$ $\mathbb{P}$-*a.s.*
2. *For any $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_n$ the vector of non-overlapping increments*
$$(L_{t_1}, L_{t_2} - L_{t_1}, \ldots, L_{t_n} - L_{t_{n-1}})$$
forms an independent family of random vectors.

3. *For any* $0 \leqslant s < t$ *we have the* ***stationarity of increments***

   $L_t - L_s \sim L_{t-s} - L_0 = L_{t-s}$ *has an infinitely divisible distribution in* H.

4. $(L_t)_{t\geqslant 0}$ *is* ***continuous in probability****, that is for any* $t > 0$ *and* $\varepsilon > 0$ *we have*

$$\lim_{s\to 0} \mathbb{P}(|L_{t+s} - L_t| > \varepsilon) = 0.$$

Remark 2.1.23. *Continuity in probability can be formulated as*

$$\mathbb{P}(|\Delta_{t_0} L| > \varepsilon) = 0 \qquad \forall t_0, \varepsilon > 0.$$

for the jump increment of $\Delta_t L := L_t - L_{t-}$. *That is the probability of a discontinuity for any deterministic point in time* t_0 *is* 0.
This continuity is obviously wrong if t_0 *is a random point in time.*

Remark 2.1.24. *Note that the last property implies that for any* $t > 0$ *we have*

$$\lim_{s\to t} \mathcal{L}(L_s) = \mathcal{L}(L_t) \qquad \text{in the weak sense}$$

and in particular implies that

$$t \mapsto \phi_{L_t}(u), \qquad u \in H$$

is continuous.

Example 2.1.25.

1. *For a fixed vector* $b \in H$ *any linear function* $t \mapsto L_t := tb$ *is a Lévy process.*

2. *Q-Brownian motion in* H *is a Lévy process by definition.*

3. *A Poisson process is a Lévy process with values in* $\mathbb{N}_0$ *by Lemma 2.1.15.*

4. *A compound Poisson process with values in* H *is a Lévy process in* H *by Lemma 2.1.20.*

EXERCISE 2.1.26. *Check all missing steps in the previous examples.*

B) The connection between Lévy processes and infinitely divisible distributions:

Theorem 2.1.27. *Given a Lévy process* $(L_t)_{t\geqslant 0}$ *on a given probability space* $(\Omega, \mathcal{A}, \mathbb{P})$. *Then for any* $t_0 > 0$ *the marginal distribution* $\mathcal{L}(L_{t_0})$ *determines the distribution (in the sense of the finite-dimensional distributions) of the process* $(L_t)_{t\geqslant 0}$ *and it is infinitely divisible.*

Proof. For any $n \in \mathbb{N}$ we write $t_i = t\frac{i}{n}$, $i \in \{0, \dots, n\}$.

$$L_t = L_t - L_0 = \sum_{i=1}^{n} (L_{t_i} - L_{t_{i-1}}) = \sum_{i=1}^{n} L^i.$$

Now, for each $i \in \{1, \dots, n\}$ we have that $L^i = L_{t_i} - L_{t_{i-1}} \sim L_{t_i - t_{i-1}} - L_0 = L_{t_1} = L_{\frac{1}{n}}$. Due to the independence of increments we have that the family $(L^i)_{i=1,\dots,n}$ is i.i.d. Hence L_t has an infinitely divisible distribution. This proves the second statement.
For the first statement we have to show that $\mathcal{L}(L_{t_0})$ determines for any $n \in \mathbb{N}$ and $0 = s_0 < s_1 < \dots s_n$ the distribution of

$$(L_{s_1}, L_{s_2}, \dots, L_{s_n}).$$

Note that

$$(L_{s_1}, L_{s_2}, \dots, L_{s_n}) = D^{-1}(L_{s_1}, L_{s_2} - L_{s_1}, \dots, L_{s_n} - L_{s_{n-1}})$$

for the matrix of operators

$$D = \begin{pmatrix} id_H & 0 & & & 0 \\ -id_H & id_H & 0 & & \\ 0 & -id_H & id_H & 0 & \\ & 0 & -id_H & id_H & \ddots \\ 0 & & 0 & \ddots & \ddots \end{pmatrix}$$

Since the time increments of $(L_{s_1}, L_{s_2} - L_{s_1}, \dots, L_{s_n} - L_{s_{n-1}})$ are non-overlapping the entries form an independent family of random vectors and the product measure is uniquely determined by the laws of the entries $L_{s_i} - L_{s_{i-1}}$. Since $L_{s_i} - L_{s_{i-1}} \sim L_{s_i - s_{i-1}}$ it is sufficient to show that for any $s > 0$ the law of L_{t_0} determines the law of L_s.

Case 1: Assume first $s = \frac{m}{n} t_0$ for some $m, n \in \mathbb{N}$. Since $\mathcal{L}(L_{t_0})$ is infinitely divisible for each $n \in \mathbb{N}$ exists a unique distribution $\mathcal{L}(L_{\frac{t_0}{n}})$. Now, $\mathcal{L}(L_{\frac{t_0}{n}}) = \mathcal{L}(L_{\frac{t}{n}} - L_0) = \mathcal{L}(\underbrace{L_{\frac{it}{n}} - L_{\frac{(i-1)t}{n}}}_{L^i})$ and

$$\mathcal{L}(L_s) = \mathcal{L}(L_{\frac{m}{n} t_0}) = \mathcal{L}(\sum_{i=1}^{m} L^i) = \underbrace{\mathcal{L}(L_{\frac{t_0}{n}}) * \dots * \mathcal{L}(L_{\frac{t_0}{n}})}_{m \text{ times}} =: \mathcal{L}(L_{\frac{t_0}{n}})^{*m}$$

is uniquely determined.

Case 2: For $\frac{t_0}{s} \in \mathbb{R} \setminus \mathbb{Q}$ there exits a sequence m_n such that $\frac{m_n}{n} \longrightarrow \frac{s}{t_0}$. For any $n \in \mathbb{N}$ we then have by the same arguments that

$$\mathcal{L}(L_{\frac{m_n t_0}{n}}) = \mathcal{L}(L_{\frac{t_0}{n}})^{*m_n}$$

is uniquely defined and by the continuity in probability we may pass to the limit w.r.t. the weak convergence

$$\lim_{n\to\infty} \mathcal{L}(L_{\frac{m_n t_0}{n}}) = \mathcal{L}(L_{\lim_{n\to\infty} \frac{m_n t_0}{n}}) = \mathcal{L}(L_s).$$

□

2.1.4 Càdlàg version

In the sequel we establish the $\mathbb{P}$-a.s. regularity of the paths $t \mapsto L_t$ for any given Lévy process $(L_t)_{t\geqslant 0}$.

Example 2.1.28. *1. Linear functions $t \mapsto bt$, $b \in H$ are obviously continuous surely.*

2. A Q-Brownian motion $t \mapsto B_t$ is $\mathbb{P}$-almost surely continuous.

*3. A Poisson process $t \mapsto \pi_t$ for any intensity $\lambda > 0$ is <u>not</u> $\mathbb{P}$-almost surely continuous, but it is $\mathbb{P}$-almost surely right continuous $\lim_{s\searrow t} L_s = L_t$ and the left limits $\lim_{s\nearrow t} L_s \in \mathbb{R}$ do exist. This is the generalization of continuity is called **càdlàg** (French acronym of continu à droit, avec limite à gauche).*

4. Any compound Poisson process $t \mapsto C_t$ has $\mathbb{P}$-a.s. càdlàg trajectories.

EXERCISE 2.1.29. *Check the previous examples in detail.*

EXERCISE 2.1.30. *Check that the Poisson process is continuous in probability.*

Remark 2.1.31. *So, why bother with càdlàg versions? Couldn't we just consider paths in $L^\infty_{loc}([0,\infty), H)$ or $L^1_{loc}([0,\infty), H)$? Yes, we certainly could, however, trajectories over a bounded interval, will not be uniquely determined by a dense set any more, in fact they are null-sets. This property is know as "separability".*

A) Càdlàg versions are indistinguishable We show in this subsection that all Lévy processes have càdlàg trajectories $\mathbb{P}$-a.s. In order to fully appreciate this fact we have to distinguish different forms of equality for stochastic process.
The weakest form for our purposes is the identity of its finite-dimensional distributions.

Definition 2.1.32. *Let $(X_t)_{t\geqslant 0}$ be a stochastic processes with values in H over $(\Omega, \mathcal{A}, \mathbb{P})$ and $(Y_t)_{t\geqslant 0}$ a* ***stochastic process*** *over the probability space $(\Omega', \mathcal{A}', \mathbb{P}')$. X and Y have the* ***same finite dimensional distributions*** *or* ***law****, if for any $n \in \mathbb{N}$, $0 < t_1 < \ldots < t_n$ and $A \in \mathcal{B}(\mathbb{R}^n)$ we have*

$$\mathbb{P}((X_{t_1}, \ldots, X_{t_n}) \in A) = \mathbb{P}((Y_{t_1}, \ldots, Y_{t_n}) \in A).$$

in other words

$$\mathcal{L}((X_t)_{t\geqslant 0}) = \mathcal{L}((Y_t)_{t\geqslant 0}).$$

EXERCISE 2.1.33. *Construct an example of two stochastic processes X and Y on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with the same finite dimensional distributions, but satisfying $\mathbb{P}(X_t = Y_t) = 0$ for any $t \geqslant 0$.*

Definition 2.1.34. *Let X and Y be two stochastic processes with values in H over the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Y is called a* ***version*** *of X, if*

$$\mathbb{P}(X_t = Y_t) = 1, \qquad t \geqslant 0.$$

Remark 2.1.35. *Beware: There are in general uncountably many t and for each of it there may be a different null set $N_t \in \mathcal{B}(H)$, but its intersection is not necessarily a null set again.*

The strongest version of identity is having $\mathbb{P}$-a.s. for all $t \geqslant 0$ identical paths.

Definition 2.1.36. *Let X and Y be two processes $\mathbb{R}$ over the same probability space. Y is called* ***indistinguishable*** *from X, if*

$$\mathbb{P}(X_t = Y_t \ \forall\ 0 \leqslant t < \infty) = \mathbb{P}(\bigcap_{t\in[0,\infty)} \{X_t = Y_t\}) = 1.$$

EXERCISE 2.1.37. *Construct an example of two processes X and Y on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ which are versions of each other but which are not indistinguishable.*

Remark 2.1.38.

1. *Y, X indistinguishable $\Longrightarrow$ Y is a version of X.*
2. *Y is a version of $X \Longrightarrow \mathcal{L}(X) = \mathcal{L}(Y)$.*

Any version of a càdlàg process is indistinguishable from it.

Lemma 2.1.39. *Let Y be a version of X and for both processes we have $\mathbb{P}$-a.s. for all $t \geqslant 0$ the right limits*

$$\lim_{s\searrow t} X_s = X_{t+} \qquad \text{and} \qquad \lim_{s\searrow t} Y_s = Y_{t+}$$

Then they are indistinguishable.

EXERCISE 2.1.40. *Show the previous lemma.*
Note that the right limits of right-continuous functions are uniquely defined on countable, dense subsets.

Lemma 2.1.41. *Consider a Lévy process X and Y a version of X. Then Y is also a Lévy process.*

EXERCISE 2.1.42. *Show the previous lemma.*

B) Càdlàg versions of a $\mathbb{R}$-valued Lévy process: In order to understand the existence of right limits we need to understand when this limit does not exist.
For any function $f : [0, \infty) \to \mathbb{R}$ we have the following. The limit from the right

$$\lim_{s \to t+} f(s) = f(t+)$$

does not exist if and only if, the value

$$\varepsilon := \limsup_{s \to t+} f(s) - \liminf_{s \to t+} f(s) > 0.$$

This is equivalent to the existence of a sequence $(t_j)_{j \in \mathbb{N}}$, $t_j \searrow t$ as $j \to \infty$ such that

$$|f(t_j) - f(t_{j-1})| > \frac{\varepsilon}{2},$$

that is f has "infinitely many $\varepsilon/2$ oscillations" right to t_0.

Definition 2.1.43. *A function $f : [0, \infty) \to \mathbb{R}$ has at least $m \in \mathbb{N}$ ε-oscillations, $\varepsilon > 0$, in a set $I \subset [0, \infty)$, if there are values $t_0 < \ldots < t_m$ in I satisfying*

$$|f(t_j) - f(t_{j-1})| > \varepsilon \qquad j = 1, 2, \ldots, m.$$

We define the maximal number of ε-oscillations of f in an interval $I \subset [0, \infty)$

$$\mathrm{Osc}_\varepsilon(f; I) := \sup\{m \in \mathbb{N}_0 \mid \exists\, t_0, \ldots, t_n \in I, t_{j-1} < t_j : |f(t_j) - f(t_{j-1})| > \varepsilon\}.$$

Set $\mathbb{Q}_+ := \mathbb{Q} \cap [0, \infty)$.

The following lemma connects oscillations and the existence of limits.

Lemma 2.1.44. *For any function $f : \mathbb{Q}_+ \to \mathbb{R}$ the following statements are equivalent:*

1. *The limit from the right and from the left*

$$f(t+) := \lim_{\mathbb{Q}_+ \ni s \to t+} f(s) \qquad \textit{resp.} \qquad f(t-) := \lim_{\mathbb{Q}_+ \ni s \to t-} f(s)$$

   along values in $\mathbb{Q}_+$ exist for any $t \in [0, \infty)$.

2. *For any $t \in [0,\infty)$, $\varepsilon > 0$ and any bounded and closed set $K \subset [0,\infty)$ we have*
$$\mathrm{Osc}_\varepsilon(f,K) < \infty.$$

Lemma 2.1.45. *For a Lévy process X with values in $\mathbb{R}$ and $m \in \mathbb{N}$. Then for any $\varepsilon > 0$ there exists a constant $\delta > 0$ such that for any repartition $0 < t_1 < \ldots < t_m$ satisfying $t_m - t_1 < \delta$ we have*
$$\max_{k=1\ldots n} \mathbb{P}(|X_{t_k} - X_{t_1}| \geqslant \varepsilon) < \frac{1}{6}.$$
and as a consequence we may infer
$$\mathbb{E}[\mathrm{Osc}_{6\varepsilon}(X,\{t_1,\ldots,t_m\})] \leqslant 2,$$
which implies
$$\mathrm{Osc}_{6\varepsilon}(X,\{t_1,\ldots,t_m\}) < \infty, \qquad \mathbb{P}-a.s.$$

This result is shown by deriving the recursion for $\ell \in \mathbb{N}$
$$\mathbb{P}(\mathrm{Osc}_{6\varepsilon}(X,\{t_1,\ldots,t_m\}) > \ell) \leqslant \frac{1}{2}\mathbb{P}(\mathrm{Osc}_{6\varepsilon}(X,\{t_1,\ldots,t_m\} > \ell - 1)$$
using Etemadi's maximal inequality and the formula
$$\mathbb{E}[\mathrm{Osc}_{6\varepsilon}(X,\{t_1,\ldots,t_m\})] = \sum_{\ell=1}^{\infty} \mathbb{P}(\mathrm{Osc}_{6\varepsilon}(X,\{t_1,\ldots,t_m\} > \ell).$$

Theorem 2.1.46 (Etemadi)**.** *For an independent family of random variables $X_1,\ldots,X_n$ with values in $\mathbb{R}$ denote its partial sum*
$$S_k := \sum_{i=1}^{k} X_i, \qquad k = 1,\ldots,n.$$
Then we have for all $r \geqslant 0$ that
$$\mathbb{P}(\max_{k=1,\ldots,n} |S_k| \geqslant 3r) \leqslant 3 \max_{k=1,\ldots,n} \mathbb{P}(|S_k| \geqslant r).$$

A proof is given for instance in [6], Theorem M19.

Theorem 2.1.47. *Consider a Lévy process X with values in $\mathbb{R}$. Then there exists a version Y of X with cádlág paths and hence indistinguishable.*

Proof. Fix $\varepsilon > 0$. By the previous lemma we have for any sufficiently small interval $[s,s+h]$ and any finite repartition $\{t_1,\ldots,t_m\} \subset [s,s+h]$ that
$$\mathbb{E}[\mathrm{Osc}_{6\varepsilon}(X,\{t_1,\ldots,t_m\})] \leqslant 2.$$

The monotone convergence theorem implies

$$\mathbb{E}[\mathrm{Osc}_{6\varepsilon}(X, \mathbb{Q} \cap [s, s+h])) \leqslant 2.$$

Since any interval $[0, T]$ can be covered by a finite number of intervals of the shape $[s, s+h]$ we obtain

$$\mathbb{E}[\mathrm{Osc}_{6\varepsilon}(X, \mathbb{Q} \cap [0, T])] \leqslant \mathbb{E}[\sum_{k=1}^{\lceil \frac{T}{h} \rceil} \mathrm{Osc}_{6\varepsilon}(X, \mathbb{Q} \cap [k, k+h])) \leqslant 2\lceil \frac{T}{h} \rceil < \infty.$$

Therefore the set

$$\widetilde{\Omega} := \bigcap_{T\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} \{\mathrm{Osc}_{\frac{6}{n}}(X, \mathbb{Q} \cap [0, T]) < \infty\}$$

is the countable intersections of events of probability 1 and $\mathbb{P}(\Omega_1) = 1$. Hence for all $\omega \in \widetilde{\Omega}$ the limits

$$\lim_{\mathbb{Q}\ni r \nearrow t} X_r(\omega) \qquad \text{und} \qquad \lim_{\mathbb{Q}\ni r \searrow t} X_r(\omega)$$

do exist and Lemma implies that

$$Y_t(\omega) := \begin{cases} \lim_{\mathbb{Q}\ni q \searrow t} X_t(\omega) & \omega \in \widetilde{\Omega} \\ 0 & \text{sonst.} \end{cases}$$

is càdlàg for any $\omega \in \Omega$. By definition of Y we have a sequence $(r_j)_j$ mit $r_j \searrow t$ such that

$$X_{r_j} \longrightarrow Y_t$$

$\mathbb{P}$-a.s. and hence also this convergence in probability. The continuity in probability implies

$$X_{r_j} \longrightarrow X_t.$$

Since the limits of the convergence in probability is $\mathbb{P}$-a.s. uniquely determined we have that $\mathbb{P}(X_t = Y_t) = 1$, in other words Y is a modification of X with cádlág paths. □

We have seen that any Lévy process with values in $\mathbb{R}$ can be considered as a random vector in the space $\mathbb{D}([a, b], \mathbb{R})$ of càdlàg functions.
The analogue remains true for all Lévy proces with values in $\mathbb{R}^d$ and any separable Hilbert space H.

C) The space of càdlàg functions:

Lemma 2.1.48 (Properties of the Skorohod space)**.** *For all any separable Hilbert space H we have the following.*

1. *The space $\mathbb{D}([a,b],H)$ of all cádlag functions $f:[a,b]\to H$ is a vectors space over $\mathbb{R}$. It is called* ***Skorohod space****.*

2. *For $f,g\in\mathbb{D}([a,b],\mathbb{R})$ we have $fg\in\mathbb{D}([a,b],\mathbb{R})$.*

3. *For $f\in\mathbb{D}([a,b],\mathbb{R})$ with $f(x)\neq 0 \qquad x\in[a,b]$ we have $\frac{1}{f}\in\mathbb{D}([a,b],\mathbb{R})$*

4. *$\mathbb{D}([a,b],H)\subset L^\infty_{loc}([a,b],H)$, that is for any $\bar{I}\subset[a,b]$ we have that $f\mathbf{1}_I$ is an essentially bounded function.*

5. *For any $f\in\mathbb{D}([a,b],H)$ and $\kappa>0$ the number $\#\{s\in[a,b]\mid|\Delta_s f|>\kappa\}<\infty$ and hence the number of discontinuities of f in the interval $[a,b]$ is at most countable.*

6. *On any finite interval càdlàg functions are uniformly right-continuous.*

7. *For any $(f_n)_{n\in\mathbb{N}}\subset\mathbb{D}([a,b],H)$ and $f:[a,b]\to H$ satisfying*

$$\lim_{n\to\infty}\sup_{s\in[a,b]}|f_n(s)-f(s)|=0,$$

we have that $f\in\mathbb{D}([a,b],H)$.

8. *Any càdlàg is Borel measurable and hence for any càdlàg $f\in\mathbb{D}([a,b],H)$ there exists a sequence of step functions $s_n\in\mathbb{D}([a,b],H)$, satisfying*

$$\lim_{n\to\infty}\sup_{s\in[a,b]}|f(s)-s_n(s)|=0.$$

9. *For $f,g\in\mathbb{D}([a,b],H)$ and $\widetilde{\mathbb{Q}}$ a dense countable set in $\mathbb{R}$ and $f(q)=g(q)$ for all $q\in[a,b]\cap\widetilde{\mathbb{Q}}$. Then $f=g$.*

EXERCISE 2.1.49. *Show the previous lemma.*

Remark 2.1.50. 1. *The last item in Lemma 2.1.48 is the reason why stochastic processes, which do not have continuous paths, are considered as random vector in $\mathbb{D}([a,b],H)$ instead of other spaces such as $L^2([a,b],H)$ etc.*

2. *The space $\mathbb{D}([a,b],H)$ equipped with the uniform norm $|f|_\infty:=\sup_{s\in[a,b]}|f(s)|$ is easily seen to be a Banach space. However, it is not a separable! However, it is possible to introduce another metrics in $\mathbb{D}([a,b],H)$, the so-called J_1 metrics, see for instance [6].*

2.2 The strong Markov property and the moments of Lévy processes with bounded jumps

In the sequel we shall establish the property of independent increments, the so-called strong Markov property, also for a class of random times, so-called stopping times, which are "adapted" to the process X.

2.2.1 Stopping times and hitting times

A) Filtrations

Definition 2.2.1. *On* $(\Omega, \mathcal{A}, \mathbb{P})$ *let* $(X_t)_{t\geqslant 0}$ *be a Lévy process with values in* H *and* $(\mathcal{F}_t)_{t\geqslant 0}$ *be a* ***filtration in*** $\mathcal{A}$, *that is a family of sub-sigma algebras* $\mathcal{F}_t \subset \mathcal{F}$ *satisfying*

$$0 \leqslant s \leqslant t \quad \Rightarrow \quad \mathcal{F}_s \subset \mathcal{F}_t.$$

Definition 2.2.2. *We say that* $(X_t)_{t\geqslant 0}$ *is* $(\mathcal{F}_t)$*-adapted if* $X_t \sim \mathcal{F}_t$, *that is the set of preimages of* $\mathcal{B}(H)$ *via* X_t *satisfies* $X_t^{-1}(\mathcal{B}(H)) \subset \mathcal{F}_t$ *a sub-sigma algebra.*

Remark 2.2.3.

1. *Note for any random vector* $Y : \Omega \to H$ *the collection of* $Y^{-1}(\mathcal{B}(H)) := \{Y^{-1}(B) \mid B \in \mathcal{B}(H)\}$ *is a sigma algebra, in particular* $Y^{-1}(\mathcal{B}(H))$ *is a sub-sigma-algebra of* $\mathcal{A}$. *It is often denoted by* $\sigma(X)$.

2. *Note that the union of sigma algebras is not a sigma algebra! For any collection* $\widetilde{\mathcal{B}} \subset 2^\Omega$ *there is a unique smallest sigma algebra containing* $\widetilde{\mathcal{B}}$.

$$\sigma(\widetilde{\mathcal{B}}) := \bigcap_{\substack{\mathcal{H}\supseteq\widetilde{\mathcal{B}} \\ \mathcal{H} \text{ sigma algebra}}} \mathcal{H}.$$

3. *For a stochastic process* $X = (X_t)_{t\geqslant 0}$ *on* $(\Omega, \mathcal{A}, \mathbb{P})$ *the family of sigma algebras* $\mathcal{F}_t := \sigma\Big(\bigcup_{0\leqslant s\leqslant t} X_s^{-1}(\mathcal{B}(H))\Big)$ *is a filtration in* $(\Omega, \mathcal{A}, \mathbb{P})$. *It is called the* ***natural filtration of*** X. *It is the smallest filtration such that* X *is* $(\mathcal{F}_t)_{t\geqslant 0}$*-adapted.*

For later use we define the following property of filtrations.

Definition 2.2.4.

1. *For a given filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *in* $(\Omega, \mathcal{A}, \mathbb{P})$ *the filtration* $(\mathcal{F}_{t+})_{t\geqslant 0}$ *of*

$$\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$$

is called the right-continuous filtration of $(\mathcal{F}_t)_{t\geqslant 0}$.

2. *A filtration* $(\mathcal{F}_{t\geqslant 0})_{t\geqslant 0}$ *in* $(\Omega, \mathcal{A}, \mathbb{P})$ *is called* ***right continuous****, if for all* $t \geqslant 0$
$$\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s.$$

Remark 2.2.5. *Obviously* $\mathcal{F}_t \subset \mathcal{F}_{t+}$ *for all* $t \geqslant 0$*. Note that* $(\mathcal{F}_{t+})_{t\geqslant 0}$ *is right continuous.*

Definition 2.2.6. *A filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *in* $(\Omega, \mathcal{A}, \mathbb{P})$ *is called* $\mathbb{P}$***-complete*** *if*
$$\mathcal{N}_{\mathbb{P}} \subset \mathcal{F}_0, \qquad \text{where} \quad \mathcal{N}_{\mathbb{P}} := \{A \in \mathcal{A} \mid \mathbb{P}(A) = 0\}.$$

Lemma 2.2.7. *Let* $(\mathcal{F}_t)_{t\geqslant 0}$ *be a filtration in* $(\Omega, \mathcal{A}, \mathbb{P})$*. Then there exists a minimal right continuous and* $\mathbb{P}$*-complete filtration* $(\mathcal{F}_t^+)_{t\geqslant 0}$ *of* $(\mathcal{F}_t)_{t\geqslant 0}$ *w.r.t. inclusion, which is given by*
$$\mathcal{F}_t^+ := \bigcap_{s>t} \sigma(\mathcal{N}_{\mathbb{P}} \cup \mathcal{F}_s). \tag{2.7}$$

It is called the ***enhanced canonical filtration****.*

EXERCISE 2.2.8. *Verify the preceding lemma.*

B) Stopping times

Definition 2.2.9. *Let* $(\mathcal{F}_t)_{t\geqslant 0}$ *be a filtration in* $(\Omega, \mathcal{A}, \mathbb{P})$*. Then a* $(\mathcal{F}, \mathcal{B}([0,\infty])$*-measurable random variable*
$$T : \Omega \to [0, \infty]$$
is called $(\mathcal{F}_t)_{t\geqslant 0}$***-stopping time****, if for all* $t \in [0,\infty)$ *we have*
$$\{T \leqslant t\} \in \mathcal{F}_t.$$

Remark 2.2.10. *In other words, being a* $(\mathcal{F}_t)_{t\geqslant 0}$*-time* T *means that* $T^{-1}([0,t]) \in \mathcal{F}_t$*. Imagine the natural filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *of a stochastic process* $X = (X_t)_{t\geqslant 0}$*. That is* $\mathcal{F}_t$ *represents the sigma algebra of all 'known events' of* X *on the time interval* $[0,t]$*. Hence for any point in time* $t \geqslant 0$ *with the knowledge of* $\mathcal{F}_t$ *it is 'known' whether* T *has occurred so far or not.*
If $(\mathcal{F}_t)_{t\geqslant 0}$ *is larger than the natural filtration, then at any time* t *there is more "information" available to determine whether* T *has occurred so far or not.*

Stopping times satisfy the following easy properties.

Lemma 2.2.11 (Properties of stopping times)**.** *For a filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *in* $(\Omega, \mathcal{A}, \mathbb{P})$ *and* T, S *two* $(\mathcal{F}_t)_{t\geqslant 0}$ *stopping times. Then the random variables* $S+T$*,* $S \wedge T$*,* $S \vee T$*,* aT*,* $a \geqslant 1$ *are also* $(\mathcal{F}_t)_{t\geqslant 0}$*-stopping times.*

EXERCISE 2.2.12. *Show Lemma 2.2.11.*

Sequences of stopping times also define stopping times.

Lemma 2.2.13 (Properties of stopping times 2)**.** *For a filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *in* $(\Omega, \mathcal{A}, \mathbb{P})$ *and* $(T_n)_{n\in\mathbb{N}}$ *a family of* $(\mathcal{F}_t)_{t\geqslant 0}$*-stopping times. Then the random variables* $\sup_{n\in\mathbb{N}} T_n$, $\inf_{n\in\mathbb{N}} T_n$, $\limsup_{n\in\mathbb{N}} T_n$ *and* $\liminf_{n\in\mathbb{N}} T_n$ *are also* $(\mathcal{F}_t)$*-stopping times.*

EXERCISE 2.2.14. *Show Lemma 2.2.13.*

Example 2.2.15. *Any constant* $T(\omega) = a$ *is a stopping time w.r.t. any filtration.*

However, the typical stopping times are the "hitting times" stochastic processes, that is the "first" times a given stochastic process hits a given set. In order to make this examples rigorous, we have to

Definition 2.2.16. *Let* $(F_t)_{t\geqslant 0}$ *be a filtration in* $(\Omega, \mathcal{A}, \mathbb{P})$*. We say that* $(F_t)_{t\geqslant 0}$ *satisfies the* ***usual conditions*** *in the sense of Protter [57] if it is a right-continuous, complete filtration.*

Proposition 2.2.17. *Let* $(F_t)_{t\geqslant 0}$ *be a filtration in* $(\Omega, \mathcal{A}, \mathbb{P})$ *satisfying the usual conditions and* $T : \Omega \to [0, \infty]$ *a* $(\mathcal{F}_t)$*-random variable. Then the following are equivalent:*

1) T is a $(\mathcal{F}_t)_{t\geqslant 0}$*-stopping time.*

2) We have $\{T < t\} \in \mathcal{F}_t \qquad \forall\, t \geqslant 0.$

EXERCISE 2.2.18.
- *Prove that 2) implies 1) using 2), the previous lemma and the right continuity.*
- *Prove that 1) implies 2), this is the easy case.*

We can now define the main class of examples of stopping times.

Definition 2.2.19. *Let* $(F_t)_{t\geqslant 0}$ *be a filtration in* $(\Omega, \mathcal{A}, \mathbb{P})$ *satisfying the usual conditions and* $(X_t)_{t\geqslant 0}$ *be* $(\mathcal{F}_t)$*-adapted càdlàg process with values in a separable Hilbert space* H*.*

1. For $B \in \mathcal{B}(H)$ ***open*** *we call*

$$T_B(\omega) := \inf\{t > 0 \mid X_t(\omega) \in B\}$$

the ***hitting time*** *of* B *by* X*.*

2. *For $B \in \mathcal{B}(H)$ **closed**, we call*

$$T_B(\omega) := \inf\{t > 0 \mid X_t(\omega) \in B \text{ or } X_{t-}(\omega) \in B\}$$

*the **hitting time** of B by X.*

Lemma 2.2.20. *Let $(F_t)_{t\geqslant 0}$ be a filtration in $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying the usual conditions and $(X_t)_{t\geqslant 0}$ be $(\mathcal{F}_t)$-adapted càdlàg process with values in a separable Hilbert space H. Then we have the following*

1. *For any open set $B \in \mathcal{B}(\mathbb{R}^d)$ the hitting time T_B is a $(\mathcal{F}_t)$-stopping time.*
2. *For any closed set $B \in \mathcal{B}(\mathbb{R}^d)$ the hitting time T_B is a $(\mathcal{F}_t)$-stopping time.*

Proof. 1. Let B be open. Since X has càdlàg trajectories, which are determined by a countable dense set, such as $\mathbb{Q}$ we obtain the countable union

$$\{T_B < t\} = \bigcup_{s\in[0,t)} \{X_s \in B\} = \bigcup_{s\in[0,t)\cap\mathbb{Q}} \underbrace{\{X_s \in B\}}_{\in\mathcal{F}_s} \in \mathcal{F}_t.$$

2. Let B be closed. Then the open neighborhoods $B_n := \bigcup_{x\in B} B_{\frac{1}{n}}(x)$ of B are open and hence

$$\begin{aligned}\{T_B \leqslant t\} &= \{T_B = t\} \cup \{T_B < t\} \\ &= \underbrace{\{X_t \in B\}}_{\in\mathcal{F}_t} \cup \underbrace{\{X_{t-} \in B\}}_{\in\mathcal{F}_t} \cup \underbrace{\bigcap_{n\in\mathbb{N}} \bigcup_{s\in[0,t)\cap\mathbb{Q}} \{X_s \in B_n\}}_{\in\mathcal{F}_t \text{ by (1)}} \in \mathcal{F}_t.\end{aligned}$$

□

Remark 2.2.21. *The so-called Début theorem shows more generally that the preceding result is true for any Borel-measurable set. In addition the requirement of càdlàg paths and adaptedness can be relaxed to the notion of $(\mathcal{F}_t)_{t\geqslant 0}$-progressive measurability. There is also an inverse of the theorem, saying that any stopping time can be represented as a stopping time w.r.t. the appropriate filtration and set.*

Example 2.2.22. *Consider a Lévy process $(X_t)_{t\geqslant 0}$ with this enhanced canonical filtration $(\mathcal{F}_t^+)_{t\geqslant 0}$ given in (2.7). Then the hitting time*

$$T_n := \inf\{t > 0 \mid |X_t| > n\}$$

is a $(\mathcal{F}_t^+)_{t\geqslant 0}$-stopping time.

C) Stopping time sigma algebras

Definition 2.2.23. *Let $(F_t)_{t\geqslant 0}$ be a filtration in $(\Omega, \mathcal{A}, \mathbb{P})$ and T a $(\mathcal{F}_t)_{t\geqslant 0}$-stopping time. Then*

$$\mathcal{F}_T := \{A \in \mathcal{A} \mid \{T \leqslant t\} \cap A \in \mathcal{F}_t \text{ for all } t \geqslant 0\}$$

*is called the T**-stopped sigma algebra**.*

Remark 2.2.24. *This definition tells us, that all events $A \in \mathcal{F}_t$ known (or available) up to time t which allow for the decision on A if T has occurred or not on so far are gathered in $\mathcal{F}_T$.*

Lemma 2.2.25. *Let $(F_t)_{t\geqslant 0}$ be a filtration in $(\Omega, \mathcal{A}, \mathbb{P})$ and S, T two $(\mathcal{F}_t)_{t\geqslant 0}$-stopping times. Then we have*

$$\mathcal{F}_{S\wedge T} = \mathcal{F}_T \cap \mathcal{F}_S.$$

This result is standard in stochastic analysis, a result can be found in [43].

2.2.2 Conditional expectation and a glimpse of martingale theory

In order to prove the strong Markov property of a Lévy process we need some basics on martingales. We follow the lines of Protter [57] and [45].

A) Discrete conditional expectation: We recall the elementary conditional expectation. Let $X : \Omega \to H$ be an H-valued random variable on $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{E}[|X|] < \infty$ and $A \in \mathcal{A}$. Then the **conditional expectation of X w.r.t A** is defined

$$\mathbb{E}[X \mid A] := \int_\Omega X(\omega) d\mathbb{P}(\omega \mid A) = \frac{1}{\mathbb{P}(A)} \int_\Omega X(\omega)\mathbf{1}_A(\omega) d\mathbb{P}(\omega) = \frac{\mathbb{E}[X\mathbf{1}_A]}{\mathbb{P}(A)}.$$

We now extend this notion from *events* to sigma algebras. First we consider the case that of a sub sigma algebra $\mathcal{F} \subset \mathcal{A}$ which is countably generated.

Definition 2.2.26. *Let $X : \Omega \to H$ be an $\mathbb{R}$-valued random variable on $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{E}[|X|] < \infty$ and $\mathcal{F} = \sigma((B_n)_{n\in\mathbb{N}})$ for a disjoint measurable decomposition $\Omega = \dot{\bigcup}_{n=1}^{\infty} B_n$. We define the **conditional expectation of** X **under** $\mathcal{F}$ by*

$$\mathbb{E}[X \mid \mathcal{F}](\omega) := \sum_{n\in\mathbb{N}} \mathbb{E}[X \mid B_n]\mathbf{1}_{B_n}(\omega)$$

Remark 2.2.27. *Note since the sets B_n are disjoint, theres is always exactly one of the terms non zero.*

Lemma 2.2.28. *Let $X : \Omega \to H$ be an random vector on $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{E}[|X|] < \infty$ and $\mathcal{F} = \sigma((B_n)_{n\in\mathbb{N}})$ for a disjoint measurable decomposition $\Omega = \dot{\bigcup}_{n=1}^{\infty} B_n$. Then $\mathbb{E}[X \mid \mathcal{F}]$ has the following properties:*

1. *$\mathbb{E}[X \mid \mathcal{F}]$ is $\mathcal{F}$-measurable.*
2. *$\mathbb{E}[|\mathbb{E}[X \mid \mathcal{F}]|] < \infty$ and*

$$\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{F}]\mathbf{1}_A] \qquad \forall A \in \mathcal{F}.$$

EXERCISE 2.2.29. *Prove the preceding lemma. Use the fact that $\mathbb{E}[X \mid \mathcal{F}] = f \circ g(\omega)$ for some discrete function g.*

EXERCISE 2.2.30. *Consider $X : [0,1] \to \mathbb{R}, X(\omega) = \omega^2$, $B_1 = [0, 1/2)$, $B_2 = [1/2, 3/4)$, $B_3 = [3/4, 1]$, $\mathbb{P} = d\omega$. Calculate $\mathbb{E}[X \mid \mathcal{F}]$ for $\mathcal{F} = \sigma(B_1, B_2, B_3)$ and draw a sketch of the function $\omega \mapsto \mathbb{E}[X \mid \mathcal{F}]$.*

B) General conditional expectations: For the more general case we need the Theorem of Radon-Nikodym, which is a consequence of the **Lebesgue decomposition theorem** for measures.

Theorem 2.2.31 (Radon-Nikodym)**.** *Let $\mu, \nu : (\Omega, \mathcal{A}) \to [0, \infty)$ be two finite measures. Then the following statements are equivalent:*

1. *μ is absolutely continuous w.r.t. ν, that is*

$$\forall\, A \in \mathcal{A} : \qquad \nu(A) = 0 \quad \Longrightarrow \quad \mu(A) = 0.$$

2. *μ has a density with respect to ν, that is there exists a $(\mathcal{F}, \mathcal{B}([0, \infty))$-measurable map $f : \Omega \to [0, \infty)$ with*

$$\mu(A) = \int_A f(\omega)\nu(d\omega) \qquad \text{for all } A \in \mathcal{F}, \mathbb{P}(A) > 0.$$

A proof is given for instance in Klenke [45], Korollar 7.34.

Definition 2.2.32. *Given $(\Omega, \mathcal{A}, \mathbb{P})$ and a random vector $X : \Omega \to H$ with $\mathbb{E}[|X|] < \infty$ and $\mathcal{F} \subset \mathcal{A}$ a sub-sigma-algebra. A random vector $Y : \Omega \to \mathbb{R}^d$ is called **conditional expectation of** X **under** $\mathcal{F}$, if*

1) Y is $\mathcal{F}$-measurable and

2) for any $A \in \mathcal{F}$ we have the identity

$$\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A].$$

In this case we write $\mathbb{E}[X \mid \mathcal{F}] := Y$.

Proposition 2.2.33. *Given $(\Omega, \mathcal{A}, \mathbb{P})$ and a random vector $X : \Omega \to H$ with $\mathbb{E}[|X|] < \infty$ and $\mathcal{F} \subset \mathcal{A}$ a sub-sigma-algebra. Then the random vector $\mathbb{E}[X \mid \mathcal{F}] : \Omega \to H$ exists and is $\mathbb{P}$-a.s. unique.*

Proof. We sketch the proof for $X : \Omega \to \mathbb{R}$.

Uniqueness: Consider two random variables Y, Y' both satisfying 1) and 2). Setting $A = \{Y > Y'\}$ we see by 1) that $A \in \mathcal{F}$ and by 2) that

$$0 = \mathbb{E}[Y\mathbf{1}_A] - \mathbb{E}[Y'\mathbf{1}_A] = \mathbb{E}[(Y - Y')\mathbf{1}_A].$$

By construction $(Y - Y')\mathbf{1}_A \geqslant 0$ $\mathbb{P}$-a.s. and hence also $\mathbb{E}[(Y - Y')\mathbf{1}_A] \geqslant 0$. Now, the monotonicity of the expectation (it is an integral!) then implies that $\mathbb{P}(A) = 0$, which proves $Y \leqslant Y'$ $\mathbb{P}$-a.s. Analogously we can also prove $Y \leqslant Y'$ $\mathbb{P}$-a.s.

Existence: We set $X^+ := X \vee 0$ and $X^- = -(X - X^+)$ and define for either case

$$Q^{\pm}(A) := \mathbb{E}[X^{\pm}\mathbf{1}_A], \qquad A \in \mathcal{F}.$$

this defines two finite measures on $(\Omega, \mathcal{F})$, which are both absolutely continuous w.r.t. $\mathbb{P}$. In this situation the Radon-Nikodym theorem 2.2.31 guarantees the existence of a $\mathcal{F}$-measurable density $Y^{\pm} \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$Q^{\pm}(A) = \mathbb{E}[Y^{\pm}\mathbf{1}_A].$$

Finally set $Y = Y^+ - Y^-$. □

Definition 2.2.34. *For any two random vectors $X, Y : \Omega \to H$ we set*

$$\mathbb{E}[X \mid Y] := \mathbb{E}[X \mid \sigma(Y)].$$

Proposition 2.2.35 (Properties of the conditional expectation:)**.**
Given $(\Omega, \mathcal{A}, \mathbb{P})$, $X, Y : \Omega \to H$ random vectors satisfying $\mathbb{E}[|X|], \mathbb{E}[|Y|] < \infty$ and $\mathcal{G} \subset \mathcal{H} \subset \mathcal{F}$ sub sigma algebras. Then we have two groups of properties:

1. ***Properties of an integral:***

a) Linearity: *For all* $a, b \in \mathbb{R}$ *we have*

$$\mathbb{E}[aX + bY \mid \mathcal{H}] = a\mathbb{E}[X \mid \mathcal{H}] + b\mathbb{E}[Y \mid \mathcal{H}] \qquad \mathbb{P} - a.s.$$

b) Monotonicity: *If* X, Y *are* $\mathbb{R}$*-valued, and* $X \leqslant Y$ $\mathbb{P}$*-a.s., then*

$$\mathbb{E}[X \mid \mathcal{H}] \leqslant \mathbb{E}[Y \mid \mathcal{H}] \qquad \mathbb{P} - a.s.$$

c) Dominated convergence: *If* $\mathbb{E}[|Y|] < \infty$, $Y \geqslant 0$ *and* $(X_n)_{n\in\mathbb{N}}$ *a sequence of random vectors with* $|X_n| \leqslant Y$ $\mathbb{P}$*-a.s. for all* $n \in \mathbb{N}$ *and* $X_n \to X$ $\mathbb{P}$*-a.s.., then*

$$\lim_{n\to\infty} \mathbb{E}[X_n \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}] \qquad \mathbb{P} - a.s. \text{ and in } L^1(\mathbb{P}).$$

d) Conditional Jensen's inequality *If* $X : \Omega \to H$, $\Phi : \mathbb{R} \to \mathbb{R}$ *is a convex function and* $\mathbb{E}[|\Phi(X)|] < \infty$*. Then*

$$\Phi\Big(\mathbb{E}\Big[\Phi(X) \mid \mathcal{G}\Big]\Big) \leqslant \mathbb{E}\Big[\Phi(X) \mid \mathcal{G}\Big], \qquad \mathbb{P} - a.s.$$

e) In the same spirit: *conditional monotonic convergence, conditional Minkowski's inequality, conditional Hölder's inequality, conditional Young's inequality.*

2. ***Measurability properties:***

a') Triviality: *If* $\mathbb{P}(A) \in \{0, 1\}$ *for all* $A \in \mathcal{H}$*, then*

$$\mathbb{E}[X \mid \mathcal{H}] = \mathbb{E}[X] \qquad \mathbb{P} - a.s.$$

b') Global average: $\mathbb{E}[\mathbb{E}[X \mid \mathcal{H}]] = \mathbb{E}[X]$ $\mathbb{P}$*-a.s.*

c') Independence: *If* X *and* $\mathcal{H}$ *are independent, then we have*

$$\mathbb{E}[X \mid \mathcal{H}] = \mathbb{E}[X] \qquad \mathbb{P} - a.s.$$

d') Homogeneity for measurable factors: *If* X, Y *are* $\mathbb{R}$*-valued satisfying* $\mathbb{E}[|XY|] < \infty$ *and* $Y \sim \mathcal{H}$*, then*

$$\mathbb{E}[XY \mid \mathcal{H}] = Y\mathbb{E}[Y \mid \mathcal{H}] \qquad \mathbb{P} - a.s.,$$

in particular $\mathbb{E}[Y \mid \mathcal{H}] = Y$ $\mathbb{P}$*-a.s.*

e') Tower property: "The coarser one wins."

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{H}] \mid \mathcal{G}] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{G}] \qquad \mathbb{P} a.s.$$

Proofs are given for instance in Klenke [45], Kapitel 8, or Koshnevisan [48], Chapter 8.

Remark 2.2.36. *Conditional expectations can be understood in terms of optimal prediction or projection (on your preknowledge) in* $L^2(\Omega, \mathcal{A}, \mathbb{P}; H)$ *onto* $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$*. This is the topic for a proper course on stochastic processes.*

C) Martingales Martingales are a class of stochastic processes, which merit an entire class on itself. They are the natural objects with respect to which a stochastic intgral is defined. In this course we shall only introduce the mere definition and its connection to Lévy processes, since we shall use it in order to establish for us the much more essential strong Markov property.

Definition 2.2.37. *Let $(F_t)_{t\geqslant 0}$ be a filtration in $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying the usual conditions and $X = (X_t)_{t\geqslant 0}$ be $(\mathcal{F}_t)$-adapted càdlàg process with values in a separable Hilbert space H satisfying $\mathbb{E}[|X_t|] < \infty$ for all $t \geqslant 0$. If X satisfies*

$$\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s \qquad \mathbb{P}-a.s. \text{ for all } t \geqslant s.$$

*it is called a $(\mathcal{F}_t)_{t\geqslant 0}$-**martingale**.*

Remark 2.2.38. *All martingales have constant expectation*

$$\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[X_t \mid \mathcal{F}_0]] = \mathbb{E}[X_0] \qquad \text{for all } t \geqslant 0.$$

Example 2.2.39. 1. *Let X be a $(\mathcal{F}_t)_{t\geqslant 0}$-adapted Lévy process with $\mathbb{E}[|X_1|] < \infty$, then $\mathbb{E}[X_t] - \mathbb{E}[X_t]$ is a $(\mathcal{F}_t)$-martingale, since*

$$\begin{aligned}\mathbb{E}[X_t - \mathbb{E}[X_t] \mid \mathcal{F}_s] &= \mathbb{E}[X_t - X_s + X_s - t\mathbb{E}[X_1] \mid \mathcal{F}_s] \\ &= \mathbb{E}[X_t - X_s] + X_s - t\mathbb{E}[X_1] \\ &= \mathbb{E}[X_{t-s}] + X_s - t\mathbb{E}[X_1] = X_s - s\mathbb{E}[X_1] \\ &= X_s - \mathbb{E}[X_s], \qquad \mathbb{P}-a.s.\end{aligned}$$

2. *Let X be a $(\mathcal{F}_t)_{t\geqslant 0}$-adapted Lévy process. Then for all $u \in H$ we have*

$$\frac{e^{i\langle u, X_t\rangle}}{\mathbb{E}[e^{i\langle u, X_t\rangle}]} = e^{i\langle u, X_t - t\eta(u)\rangle}$$

a $\mathbb{C}$-valued $(\mathcal{F}_t)_{t\geqslant 0}$-martingale, since

$$\begin{aligned}\mathbb{E}[e^{i\langle u, X_t\rangle - t\eta(u)}] \mid \mathcal{F}_s]] &= \mathbb{E}[e^{i\langle u, X_t - X_s + X_s\rangle - t\eta(u)}] \mid \mathcal{F}_s]] \\ &= \mathbb{E}[e^{i\langle u, X_t - X_s\rangle} \mid \mathcal{F}_s]\mathbb{E}[e^{i\langle u, X_s}] \mid \mathcal{F}_s]e^{-t\eta(u)} \\ &= \mathbb{E}[e^{i\langle u, X_t - X_s\rangle}]e^{i\langle u, X_s\rangle}e^{-t\eta(u)} \\ &= \frac{e^{(t-s)\eta(u)}}{e^{t\eta(u)}}e^{i\langle u, X_s\rangle} \\ &= e^{i\langle u, X_s\rangle - s\eta(u)}, \qquad \mathbb{P}-a.s.\end{aligned}$$

In other words for any $t \geqslant s$

$$\mathbb{E}[e^{i\langle u, X_t - X_s\rangle - (t-s)\eta(u)}] \mid \mathcal{F}_s] = 1, \qquad \mathbb{P}-a.s..$$

Theorem 2.2.40 (Optional stopping theorem)**.** *Given a filtration $(\mathcal{F}_t)_{t\geqslant 0}$ in $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying the usual conditions. Consider two $(\mathcal{F}_t)_{t\geqslant 0}$-stopping times $S, T : \Omega \to [0, \infty]$ satisfying $S \vee T \leqslant K < \infty$ $\mathbb{P}$-a.s. and $X = (X_t)_{t\geqslant 0}$ a cádlág $(\mathcal{F}_t)_{t\geqslant 0}$-martingale. Let $S(\omega), T(\omega) \leqslant K < \infty$ two $(\mathcal{F}_t)$-stopping times and X a càdlàg $(\mathcal{F}_t)$-martingale. Then $\mathbb{E}[|X_S|] < \infty$ and $\mathbb{E}[|X_T|] < \infty$ and*

$$\mathbb{E}\Big[X_T \mid \mathcal{F}_S\Big] = X_{T\wedge S} \qquad \mathbb{P} - a.s.$$

Proof. We show the result in the case of discrete time $t \in \mathbb{N}$ following Klenke [45] for $0 \leqslant S \leqslant T \leqslant K$ $\mathbb{P}$-a.s.

$$\begin{aligned} X_S &= \mathbb{E}[X_S \mid \mathcal{F}_S] = \mathbb{E}[X_K \mid \mathcal{F}_S] \\ &= \mathbb{E}[\mathbb{E}[X_K \mid \mathcal{F}_T] \mid \mathcal{F}_S] = \mathbb{E}[\mathbb{E}[X_T \mid \mathcal{F}_T] \mid \mathcal{F}_S] = \mathbb{E}[X_T \mid \mathcal{F}_S], \quad \mathbb{P} - \text{a.s.} \end{aligned}$$

□

Example 2.2.41. *Let X be a $(\mathcal{F}_t)_{t\geqslant 0}$-adapted Lévy process and T be a uniformly bounded $(\mathcal{F}_t)$-stopping time. Then for any $t \geqslant s \geqslant 0$ we have that*

$$\mathbb{E}[e^{i\langle u, X_{T+t} - X_{T+s}\rangle - (t-s)\eta(u)}] \mid \mathcal{F}_{T+s}] = 1, \qquad \mathbb{P} - a.s.$$

In other words, the process $t \mapsto e^{i\langle u, X_{T+t} - (T+t)\eta(u)\rangle}$ is a $(\mathcal{F}_{T+t})_{t\geqslant 0}$-martingale.

2.2.3 The strong Markov property and the moments of Lévy processes with bounded jumps

A) The strong Markov property. We are now in the position to prove the strong Markov property with the help of Example 2.2.41.

Theorem 2.2.42. *Consider an $(\mathcal{F}_t)$-adapted Lévy process X and T a $(\mathcal{F}_t)$ stopping time. On the event $\{T < \infty\}$ the process $(Y_t)_{\geqslant 0}$ defined as*

$$Y_t := X_{T+t} - X_T$$

is a $(\mathcal{H}_t)_{t\geqslant 0}$-Lévy process with $\mathcal{H}_t := \mathcal{F}_{T+t}$. In addition Y is independent of $\mathcal{F}_T$ and X and Y have the same finite dimensional distributions.

Proof. This is a slight adaption of the proof given in [57].

- First assume that $T(\omega) \leqslant K$ for some constant $K > 0$ for all $\omega \in \Omega$.

- Let $A \in \mathcal{F}_T$ and $(u_0, \ldots, u_n)$, $(t_0, \ldots, t_n)$ for some $n \in \mathbb{N}$, where $(u_n)_{n\in\mathbb{N}}$ is the enumeration of a countable dense subset of H. Then

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_A \; e^{i\sum_{j=1}^n \langle u_j, X_{T+t_j} - X_{T+t_{j-1}}\rangle}\right] \\
&= \mathbb{E}\left[\mathbf{1}_A \prod_{j=1}^n \frac{e^{i\langle u_j, X_{T+t_j}\rangle - (T+t_j)\eta(u_j)}}{e^{i\langle u_j, X_{T+t_{j-1}} - (T+t_{j-1})\eta(u_j)\rangle}}\right] \prod_{j=1}^n \frac{\phi_{X_{t_j}}(u_j)}{\phi_{X_{t_{j-1}}}(u_j)}.
\end{aligned}$$

For the second factor we obtain

$$\begin{aligned}
\prod_{j=1}^n \frac{\phi_{X_{t_j}}(u_j)}{\phi_{X_{t_{j-1}}}(u_j)} &= \prod_{j=1}^n \left(e^{t_j\eta(u_j)} - e^{t_{j-1}\eta(u_j)}\right) \\
&= \prod_{j=1}^n e^{(t_j - t_{j-1})\eta(u_j)} = \prod_{j=1}^n \phi_{X_{t_j} - X_{t_{j-1}}}(u_j).
\end{aligned}$$

For the first factor we calculate

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_A \prod_{j=1}^n \frac{e^{i\langle u_j, X_{T+t_j}\rangle - (T+t_j)\eta(u_j)}}{e^{i\langle u_j, X_{T+t_{j-1}} - (T+t_{j-1})\eta(u_j)\rangle}}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_A \prod_{j=1}^n \frac{e^{i\langle u_j, X_{T+t_j}\rangle - (T+t_j)\eta(u_j)}}{e^{i\langle u_j, X_{T+t_{j-1}} - (T+t_{j-1})\eta(u_j)\rangle}} \mid \mathcal{F}_{T+t_{n-1}}\right]\right] \\
&= \mathbb{E}\Bigg[\mathbf{1}_A \prod_{j=1}^{n-1} \frac{e^{i\langle u_j, X_{T+t_j}\rangle - (T+t_j)\eta(u_j)}}{e^{i\langle u_j, X_{T+t_{j-1}} - (T+t_{j-1})\eta(u_j)\rangle}} \\
&\qquad\qquad \mathbb{E}\left[e^{i\langle u_j, X_{T+t_n} - X_{T+t_{n-1}}\rangle - (t_n - t_{n-1})\eta(u_j)} \mid \mathcal{F}_{T+t_{n-1}}\right]\Bigg] \\
&= \mathbb{E}\left[\mathbf{1}_A \prod_{j=1}^{n-1} \frac{e^{i\langle u_j, X_{T+t_j}\rangle - (T+t_j)\eta(u_j)}}{e^{i\langle u_j, X_{T+t_{j-1}} - (T+t_{j-1})\eta(u_j)\rangle}}\right] \\
&= \mathbb{P}(A).
\end{aligned}$$

Hence we have obtained

$$\mathbb{E}\left[\mathbf{1}_A \; e^{i\sum_{j=1}^n \langle u_j, X_{T+t_j} - X_{T+t_{j-1}}\rangle}\right] = \mathbb{P}(A) \prod_{j=1}^n \phi_{X_{t_j} - X_{t_{j-1}}}(u_j).$$

Therefore the process $(Y_t)_{t\geqslant 0}$ given by $Y_t = X_{T+t} - X_T$ is independent form $\mathcal{F}_T$ and has stationary and independent increments, with the same finite dimensional distributions as X.

- If T is unbounded, we may set $T^m := T \wedge m$ for any $m \in \mathbb{N}$. If then $A \in \mathcal{F}_T$, then the formula is valid on $A_m := A \cap \{T \leqslant m\}$, that is $A_m \in \mathcal{F}_{T\wedge m}$. Since on the right-hand side $\mathbb{P}(A_m) \leqslant \mathbb{P}(A)$ for all $m \in \mathbb{N}$ and

$$\lim_{m\to\infty} \mathbb{E}\left[\mathbf{1}_{A_m}\ e^{i\sum_{j=1}^n \langle u_j, X_{T^m+t_j} - X_{T^m+t_{j-1}}\rangle}\right] \leqslant \prod_{j=1}^n \phi_{X_{t_j-t_{j-1}}}(u_j)$$

the dominated convergence theorem yields

$$\mathbb{E}\left[\mathbf{1}_{A\cap\{T<\infty\}}\ e^{i\sum_{j=1}^n \langle u_j, X_{T+t_j} - X_{T+t_{j-1}}\rangle}\right]$$
$$= \underbrace{\mathbb{P}(A\cap\{T<\infty\})}_{=\mathbb{P}(A)} \prod_{j=1}^n \phi_{X_{t_j-t_{j-1}}}(u_j).$$

□

Remark 2.2.43. *The result of Theorem 2.2.42 tells us that the law of a Lévy process "restarted" at a (finite) stopping time has no memory of the past and evolves with the same law as if started at time zero* 0 *from the stopped position. In other words, the memory of this process is "thin" in the sense that the law (the finite dimensional distributions) of the process only depend on the current state.*

This strong Markov property can be reformulated as follows:
First note that on a separable Hilbert space $(H, \mathcal{B}(H))$ any probability distribution is uniquely determined by the collection of the values

$$\{\int_H f(x)\mu(dx) \mid f \in \mathcal{C}_b(H, \mathbb{R})\}$$

Note that $\int f(x)\mu(dx) = \mathbb{E}[f(X)]$ for <u>any</u> random vector with $X \sim \mu$.

Definition 2.2.44. *Given $(\mathcal{F}_t)_{t\geqslant 0}$ a filtration in $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying the usual conditions, and τ a $(\mathcal{F}_t)_{t\geqslant 0}$-stopping time satisfying $\tau < \infty$ $\mathbb{P}$-a.s. A $(\mathcal{F}_t)_{t\geqslant 0}$-adapted càdlàg process $X = (X_t)_{t\geqslant 0}$ satisfies the **strong Markov property** if for any bounded continuous function f on path space, $f \in \mathcal{C}_b(\mathbb{D}([0,\infty), H), \mathbb{R})$ we have*

$$\mathbb{E}\Big[f(X_{\tau+\cdot}) \mid \mathcal{F}_\tau\Big] = \mathbb{E}\Big[f(X_{\tau+\cdot}) \mid X_\tau\Big] \qquad \mathbb{P} - a.s.$$

We also call X a $(\mathcal{F}_t)_{t\geqslant 0}$-strong Markov process.

Corollary 2.2.45. *Given $(\mathcal{F}_t)_{t\geqslant 0}$ a filtration in $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying the usual conditions. Then any $(X_t)_{t\geqslant 0}$ an $(\mathcal{F}_t)_{t\geqslant 0}$-adapted Lévy process is a strong Markov process.*

We shall see in Part II that the solutions of stochastic (partial) differential equations driven by a Lévy process inherit this property from the Lévy process. This property turns out crucial in order to determine the first exit problem in Part III.

B) Moments of Lévy processes with bounded jumps. In this section we use the property that Lévy processes with bounded jumps have all moments.

Proposition 2.2.46. *Given $(\mathcal{F}_t)_{t\geqslant 0}$ a filtration in $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying the usual conditions, and $X = (X_t)_{t\geqslant 0}$ the càdlàg version of a $(\mathcal{F}_t)$-Lévy proces with uniformly bounded jumps, that is there exists a constant $K > 0$ such that*

$$|\Delta_t X| \leqslant K \qquad \forall\, t \geqslant 0 \qquad \mathbb{P} - a.s.$$

Then we have for all $p > 0$, that

$$\mathbb{E}[|X_t|^p] < \infty.$$

Proof. We define the sequence of hitting times $(T_n)_{n\in\mathbb{N}}$

$$T_1 := \inf\{t > 0 \mid |X_t| \geqslant K\}$$

$$\vdots$$

$$T_{n+1} := \mathbf{1}\{t > T_n \mid |X_t - X_{T_n}| \geqslant K\}.$$

Since X has càdlàg paths and $(\mathcal{F}_t)_{t\geqslant 0}$ satisfies the usual conditions $(T_n)_{n\in\mathbb{N}}$ is a sequence of $(\mathcal{F}_t)_{t\geqslant 0}$-stopping times. Since the jumps $|\Delta_t X| \leqslant K$ and the paths are right-continuous, we have that the sequence (T_n) is strictly increasing, that is $T_n - T_{n-1} > 0$, $\mathbb{P}$-a.s. In addition we have $|\Delta_T L|$ for any $(\mathcal{F}_t)$-stopping time T. Therefore we obtain for the stopped process $X_{t\wedge T_n}$

$$\sup_{s\in[0,\infty)} |X_{s\wedge T_n}| \leqslant \sup_{s\in[0,\infty)} \sum_{k=1}^{n} |X_{s\wedge T_k} - X_{s\wedge T_{k-1}}| \leqslant 2nK.$$

In other words, we have for all $n \in \mathbb{N}$

$$\mathbb{P}(|X_t| > 2nK) \leqslant \mathbb{P}(t < T_n).$$

The $(\mathcal{F}_t)_{t\geqslant 0}$-strong Markov property of X tells us that $T_n - T_{n-1}$ is independent from $\mathcal{F}_{T_{n-1}}$ and additionally that $T_n - T_{n-1} \sim T_1$. The n-fold iteration of this argument yields

$$\mathbb{E}[e^{-T_n}] = \mathbb{E}[e^{-\sum_{k=1}^{n}(T_k - T_{k-1})}] = \mathbb{E}[e^{-T_1}]^n =: q^n \in [0, 1).$$

Now, Markov's inequality yields

$$\mathbb{P}(|X_t| > 2nK) \leqslant \mathbb{P}(T_n < t) = \mathbb{P}(e^{-T_n} \geqslant e^{-t}) \leqslant \frac{\mathbb{E}[e^{-T_n}]}{e^{-t}} \leqslant e^t q^n$$

for all $n \in \mathbb{N}$. We finally calculate

$$\begin{aligned}
\mathbb{E}[|X_t|^p] &= \mathbb{E}[\int_0^{|X_t|} ps^{p-1} ds] = \mathbb{E}[\int_0^\infty \mathbf{1}_{\{|X_t|>s\}} ps^{p-1} ds] \\
&= p\int_0^\infty \mathbb{P}(|X_t| > s) s^{p-1} ds \\
&\leqslant p\int_0^\infty \mathbb{P}(|X_t| > 2K\lfloor \frac{s}{2K} \rfloor)(2K\lceil \frac{s}{2K} \rceil)^{p-1} ds \\
&\leqslant p(2K)^{p-1} \sum_{n=1}^\infty \mathbb{P}(|X_t| > 2Kn)(n+1)^{p-1} \\
&\leqslant e^t (2K)^{p-1} \sum_{n=1}^\infty q^n (n+1)^{p-1} < \infty.
\end{aligned}$$

□

2.3 Representation in terms of Poisson random measures

2.3.1 Motivation

In order to construct a general Lévy process on a given probability space we introduce the concept of a Poisson random measure. What is Poisson random measure?

Example 2.3.1. *1. Recall the uniform distribution $\mathcal{U}_{[-1,2]}$ on the interval $([-1,2],\mathcal{B}([-1,2]))$, which is defined as for any $a<b$ by*

$$\mathcal{U}_{[-1,2]}([a,b)) = \int_a^b \frac{1}{3}\mathbf{1}_{\{[-1,2]\}}(x)dx = \int_a^b \frac{1}{\lambda([-1,2])}\mathbf{1}_{\{[-1,2]\}}(x)dx.$$

This construction is possible not only for $[-1,2]$ but for any Borel-set $A\in\mathcal{B}(\mathbb{R})$ such that $\lambda(A)<\infty$.

2. The obvious problem: there is no uniform distribution $\mathcal{U}_{\mathbb{R}}$ on the real numbers $\mathbb{R}$. Since for any uniform distribution since in this case

$$\mathcal{U}_{\mathbb{R}}([a,b)) = \int_a^b \frac{1}{\lambda(\mathbb{R})}dx = 0, \text{ since } \lambda(\mathbb{R}) = \infty.$$

3. In many situations, however, we need a good "model" for an i.i.d. $(X_n)_{n\in\mathbb{N}}$ of "uniformly" distributed random points in $\mathbb{R}$.

Idea: We change the perspective! Instead of looking at the "location" of each single point x we only look at any arbitrary interval $[a,b)$ and ask ourselves: How many points are in there?

We denote the number of points by $N([a,b))$. What should it satisfy reasonably?

1. First of all, since it represents the number of elements of something we have $N([a,b))\in\{0,1,2,\dots\}\cup\{\infty\}$.
2. We should exclude to find infinitely many particles in any finite interval $[a,b)$.
3. We should exclude to find two particles in precisely the same location in any finite interval $[a,b)$.
4. Since it is a (surprise!) random number, the map $\omega\to N([a,b))(\omega)$ is a random variable over a probability space $(\Omega,\mathcal{A},\mathbb{P})$.

5. For $[a, b)$ and $[c, d)$ with $b-a = d-c$ we should expect that the distribution of the number of particles found in there is equal

$$N([a, b)) = N([c, d)).$$

6. The number we should expect to find there should be proportional to the size of the set $[a, b)$:

$$\frac{\mathbb{E}[N([a, b))]}{\lambda([a, b))} = \text{ const. indep. from } [a, b) \qquad \text{for any } a < b.$$

7. If we have a set which is the union of two disjoint intervals: $A = [a, b) \cup [c, d)$ for $a < b < c < d$ we should simply sum up the (random) numbers:

$$N([a, b)\dot{\cup}[c, d))(\omega) = N([a, b))(\omega) + N([c, d))(\omega).$$

More generally, whenever there is a sequence $(A_n)_{n\in\mathbb{N}}$ of disjoint events, then

$$N(\dot{\bigcup}_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} N(A_n), \qquad \mathbb{P} - \text{a.s.}$$

8. Since we have a model of an i.i.d. sequence $(X_n)_{n\in\mathbb{N}}$ the number of particles in an interval $[a, b)$ should be independent from the number of any other interval $[c, d)$, which is disjoint from $[a, b)$ (since they cannot share particles). In general, whenever there is a sequence $(A_n)_{n\in\mathbb{N}}$ of disjoint events, then the random variables

$$(N(A_n))_{n\in\mathbb{N}}$$

are an independent family of random variables.

Remark 2.3.2. *Item 3) implies the following condition another should decay more than linearly*

$$\lim_{s\to 0+} \mathbb{P}(N([t, t+s)) \geqslant 2)/s = 0.$$

If $\varepsilon := \limsup_{s\to 0} \frac{\mathbb{P}(N(0,s])\geqslant 2)}{s} > 0$ *then*

$$\begin{aligned}
&\mathbb{P}(\text{“two particle on the same spot inside”}[0,1]) \\
&= \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=0}^{2^n-1} \{N(k2^{-n},(k+1)2^{-n}]) \geqslant 2\}) \\
&= 1 - \lim_{n\to\infty} \mathbb{P}(\bigcap_{k=0}^{2^n-1} \{N((k2^{-n},(k+1)2^{-n}]) \leqslant 1\}) \\
&= 1 - \lim_{n\to\infty} \prod_{k=0}^{2^n-1} \mathbb{P}(N((k2^{-n},(k+1)2^{-n}]) \leqslant 1) \\
&= 1 - \lim_{n\to\infty} \left(1 - \frac{2^n\mathbb{P}(N((0,2^{-n}]) \geqslant 2)}{2^n}\right)^{2^n} \\
&= 1 - e^{-\varepsilon} > 0,
\end{aligned}$$

by the exponential formula $\lim_{n\to\infty}(1-\frac{a_n}{n})^n = e^{-a}$ *for all sequences* $a_k \to a$.

Remark 2.3.3. 1. *Of course this is true not only for particles in the real line, for any space, where we have a “volume”, for instance given by the Lebesgue measure. For instance:* $(0,\infty)$ *or in* $\mathbb{R}^2$ *or in* $(0,\infty)\times\mathbb{R}^d$.

2. *This construction does not only work for the Lebesgue measures,*

$$\lambda([a,b)) = b-a,$$

but for any σ*-finite measure* ν*, which satisfies, that finite intervals (or balls) always have finite intensity.*

3. *Note that this construction already implies that* $N([a,b))$ *has a Poisson distribution.*

$$N([0,1)) = N(\bigcup_{k=1}^{n}[\frac{k-1}{n},\frac{k}{n})) = \sum_{k=1}^{n} N([\frac{k-1}{n},\frac{k}{n})) \approx \mathcal{B}_{n,\frac{1}{n}} \longrightarrow \mathrm{Poi}_1,$$

in distribution as $n\to\infty$.

2.3.2 Definition

Definition 2.3.4. *Given* σ*-finite measure on* ν *on* $(H,\mathcal{B}(H))$ *and a probability space* $(\Omega,\mathcal{A},\mathbb{P})$*. Then we call a measurable mapping* $N:\mathcal{B}(H)\times\Omega\to\mathbb{N}_0\cup\{\infty\}$ *a* ***Poisson random measure with intensity measure*** ν *if we have the following:*

1. *For each $A \in \mathcal{B}(H)$ with $\nu(A) < \infty$ there is a random variable $N(A) : \Omega \to \mathbb{N}_0$ such that $N(A) \sim \mathrm{Poi}_{\nu(A)}$.*

2. *For any family $(A_n)_{n\in\mathbb{N}}$ of pairwise disjoint events $A_n \in \mathcal{B}(H)$ the family $(N(A_n))_{n\in\mathbb{N}}$ is a family of independent random variables.*

3. *For any $\omega \in \Omega$ the map $A \mapsto N(A)(\omega)$ is a measure on $(H, \mathcal{B}(H))$, that is for all $\omega \in \Omega$ we have that $N(\emptyset)(\omega) = 0$ and for $(A_n)_{n\in\mathbb{N}}$ pairwise disjoint we have $N(\bigcup_{n\in\mathbb{N}} A_n)(\omega) = \sum_{n\in\mathbb{N}} N(A_n)(\omega)$.*

Does such a hybrid object always exist?

Construction of a Poisson random measure on probability space.

Lemma 2.3.5. *Given $(H, \mathcal{B}(H), \nu)$ a measure space with a finite measure ν. On $(\Omega, \mathcal{A}, \mathbb{P})$ carrying a family of an i.i.d. sequence $Z = (Z_n)_{n\in\mathbb{N}}$ of random vectors $Z_n : \Omega \to H$ with $Z_1 \sim \frac{\nu}{\nu(H)}$ and a Poisson random variable $\pi : \Omega \to \mathbb{N}_0$ with $\pi \perp Z$ there exists a Poisson random measure N for the intensity measure $\bar{\nu}$.*

Proof. This proof is given in Sato [61], we provide it for convenience. For $\nu = 0$ associate $N(A)(\omega) = 0$. For short we write $\nu_d = \nu(H)$. Define

$$N(A)(\omega) := \sum_{k=1}^{\pi(\omega)} \mathbf{1}_A(Z_k(\omega)) = \sum_{k=1}^{\pi(\omega)} \delta_{Z_k(\omega)}(A), \qquad \omega \in \mathbb{N},\ A \in \mathcal{B}(H). \tag{2.8}$$

For each $\omega \in \Omega$ this is a counting measure and hence satisfies 3). Now, for $j \geqslant 2$ and pairwise disjoint sets $A_1, \ldots, A_j \in \mathcal{B}(H)$ such that $\bigcup_{\ell=1}^{j} A_\ell = H$ and $k_1, \ldots, k_j \in \mathbb{N}_0$ with $\sum_{\ell=1}^{j} k_\ell = k$. Then we have

$$\begin{aligned}
&\mathbb{P}(N(A_1) = k_1, \ldots, N(A_j) = k_j) \\
&= \mathbb{P}(N(A_1) = k_1, \ldots, N(A_j) = k_j \mid N(H) = k)\mathbb{P}(N(H) = k) \\
&= \mathbb{P}\Big(\sum_{\ell=1}^{k} \mathbf{1}_{A_1}(Z_\ell) = k_1, \ldots, \sum_{\ell=1}^{k} \mathbf{1}_{A_j}(Z_\ell) = k_j\Big)\, e^{-\nu_d}\, \frac{\nu_d^k}{k!} \\
&= \binom{k_1, \ldots, k_j}{k} \Big(\frac{\nu(A_1)}{\nu_d}\Big)^{k_1} \cdots \Big(\frac{\nu(A_j)}{\nu_d}\Big)^{k_j}\, e^{-\nu_d}\, \frac{\nu_d^k}{k!} \\
&= \frac{k!}{k_1! \ldots k_j!} \Big(\frac{\nu(A_1)}{\nu_d}\Big)^{k_1} \cdots \Big(\frac{\nu(A_j)}{\nu_d}\Big)^{k_j}\, e^{-\nu_d}\, \frac{\nu_d^k}{k!} \\
&= \prod_{\ell=1}^{k} e^{-\nu(A_\ell)}\, \frac{\nu(A_\ell)^{k_\ell}}{k_\ell!}.
\end{aligned}$$

Summing over $k_2, \dots, k_j$ we obtain

$$\begin{aligned}\mathbb{P}(N(A_1) = k_1) &= \sum_{k_2,\dots,k_j \in \mathbb{N}_0} \mathbb{P}(N(A_1) = k_1, \dots, N(A_j) = k_j) \\ &= \sum_{k_2,\dots,k_j \in \mathbb{N}_0} \prod_{\ell=1}^{k} e^{-\nu(A_\ell)} \frac{\nu(A_\ell)^{k_\ell}}{k_\ell!} \\ &= e^{-\nu(A_1)} \frac{\nu(A_1)^{k_1}}{k_1!}.\end{aligned}$$

Therefore we also satisfy 1) and 2). □

The same can be carried out for a σ-finite intensity measure ν.

Lemma 2.3.6. *Given $(H, \mathcal{B}(H), \nu)$ a measure space with a σ-finite measure ν with repartition of $H = \bigcup_{k=0}^{\infty} B_k$ such that $\nu(B_k) < \infty$. On $(\Omega, \mathcal{A}, \mathbb{P})$ carrying a family of an i.i.d. sequence $Z = (Z_n^k)_{n\in\mathbb{N}}$ of random vectors $Z_n^k : \Omega \to H$ with $Z_n^k \sim \frac{\nu(\cdot \cap B_k)}{\nu(B_k)}$ and a Poisson random variable $\pi : \Omega \to \mathbb{N}_0$ with $\pi \perp Z$ there exists a Poisson random measure N for the intensity measure ν.*

EXERCISE 2.3.7. *Show the preceding lemma with the help of the Borel-Cantelli lemma.*

Example 2.3.8. *For λ being the Lebesgue measure we define $\nu([a, b)) := \beta\lambda([a, b))$ for some $\beta > 0$ and any $0 \leqslant a < b$ on $[0, \infty)$ we have the repartition of $B_i = [i, i + 1)$. Note that for any $0 \leqslant a < b$ we have $\nu([a, b)) < \infty$. Hence we have for $Z_n^k \sim \mathcal{U}([k, k + 1))$*

$$N([a, b)) \sim \mathrm{Poi}_{\beta(b-a)}$$

Now a Poisson process was defined via the i.i.d. sequence of waiting times $(\tau_k)_{k\in\mathbb{N}}$ with $\tau_k \sim \mathrm{Exp}_\beta$ and the arrival times $T_k = \tau_1 + \cdots + \tau_k$

$$\pi_t(\omega) = \sum_{k=1}^{\infty} \mathbf{1}\{T_k(\omega) \leqslant t\} = \sum_{k=1}^{\infty} \mathbf{1}\{T_k(\omega) \in [0, t]\} = N([0, t]).$$

This is a Poisson random measure for intensity β and intensity measure $\nu = \delta_1$.

Example 2.3.9.

Consider N being a Poisson random measure on $[0, \infty) \times H$ with intensity measure $\bar{\nu}(dt, dz) = dt \otimes \nu(dz)$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ for some jump increment distribution ν.

2.3.3 Integrating functions with Poisson random measures

A) The Poisson random integral: Consider a measurable function $f : [0,\infty)\times H \to \mathbb{R}$ and N being a Poisson random measure on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ for intensity measure $\bar{\nu}(dt, dz) = dt \otimes \nu(dz)$ on $([0,\infty)\times H, \mathcal{B}([0,\infty)\times H))$.

Rewrite the Poisson random measure: For $[a,b)\times A$ and $\omega \in \Omega$ we rewrite

$$N([a,b)\times A)(\omega) =: \int_{(0,\infty)\times H} \mathbf{1}_{[a,b)\times A}(s,z)\ N(dsdz)(\omega).$$

B) The Poisson random integral for step functions: For any simple function $f(s,z) := \sum_{\ell=1}^{k} c_\ell \mathbf{1}_{[a_\ell,b_\ell)\times A_\ell}(s,z)$ we obtain by linearity

$$\begin{aligned}
\int_{[0,\infty)\times H} f(s,z)\ N(dsdz)(\omega) &= \int_{[0,\infty)\times H} \sum_{\ell=1}^{k} c_\ell \mathbf{1}_{[a_\ell,b_\ell)\times A_\ell}(s,z)\ N(dsdz)(\omega) \\
&= \sum_{\ell=1}^{k} c_\ell \int_{[0,\infty)\times H} \mathbf{1}_{[a_\ell,b_\ell)\times A_\ell}(s,z)\ N(dsdz)(\omega) \\
&= \sum_{\ell=1}^{k} c_\ell N([a_\ell,b_\ell)\times A_\ell)(\omega)
\end{aligned}$$

and hence

$$\begin{aligned}
\mathbb{E}\big[\int_{(0,\infty)\times H} f(s,z)\ N(dsdz)\big] &= \mathbb{E}\big[\sum_{\ell=1}^{k} c_\ell N([a_\ell,b_\ell)\times A_\ell)\big] \\
&= \sum_{\ell=1}^{k} c_\ell \mathbb{E}\big[N([a_\ell,b_\ell)\times A_\ell)\big] \\
&= \sum_{\ell=1}^{k} c_\ell \bar{\nu}([a_\ell,b_\ell)\times A_\ell) \\
&= \int_{(0,\infty)\times H} f(s,z)\bar{\nu}(ds,dz).
\end{aligned}$$

C) General scalar measurable functions: Any measurable $f : H \to \mathbb{R}$ can be decomposed into the difference of nonnegative functions $f = f^+ - f^-$ for $f^+ = \max\{f, 0\}$ and $f^- = f - f^+$.

D) General scalar nonnegative functions: For any measurable, nonnegative function f there exists a sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ with $f_n(s,x) \nearrow f(s,x)$ for Lebesgue almost all $(s,x) \in [0,\infty) \otimes H$. In this case we may define for any $\omega \in \Omega$ by monotone convergence (Beppo-Levi)

$$\int_{(0,\infty)\times H} f(s,z)\ N(dsdz)(\omega) := \lim_{n\to\infty} \int_{(0,\infty)\times H} f_n(s,z)\ N(dsdz)(\omega).$$

In particular we obtain again by monotonic convergence

$$\begin{aligned}\mathbb{E}\big[\int_{(0,\infty)\times H} f(s,z)\ N(dsdz)\big] &= \mathbb{E}\big[\int_{(0,\infty)\times H} \lim_{n\to\infty} f_n(s,z)\ N(dsdz)\big] \\ &= \mathbb{E}\big[\lim_{n\to\infty} \int_{(0,\infty)\times H} f_n(s,z)\ N(dsdz)\big] \\ &= \lim_{n\to\infty} \mathbb{E}\big[\int_{(0,\infty)\times H} f_n(s,z)\ N(dsdz)\big] \\ &= \lim_{n\to\infty} \int_H f_n(s,z)\bar{\nu}(ds,dz) \\ &= \int_{(0,\infty)} \int_H \lim_{n\to\infty} f_n(s,z)\bar{\nu}(ds,dz) \\ &= \int_{(0,\infty)} \int_H f(s,z)\bar{\nu}(ds,dz).\end{aligned}$$

E) The special case of essentially time homogeneous functions: Now let $f(s,z) = \mathbf{1}_{[0,t]\times A}(s,z)f(z)$ for some measurable function $f : H \to \mathbb{R}$. We shall study the Poisson random integrals of type

$$N_t^f(A) := \int_0^\infty \int_H f(z)N(ds,dz) = \int_0^t \int_A f(z)N(ds,dz).$$

We use the approximation by simple functions $f_n(s,z) = \mathbf{1}_{[0,t]\times A}(s,z)f(z) = \mathbf{1}_{[0,t]\times A}(s)\sum_{i=1}^n c_i \mathbf{1}_{B_i\cap A}(z)$ to calculate its law via the characteristic function

$$\begin{aligned}
\mathbb{E}\left[\exp\left(i\langle u, \int_0^t \int_A f(z)N(ds,dz)\rangle\right)\right] &= \mathbb{E}\Big[\exp\Big(i\langle u, \sum_{j=1}^n c_j N([0,t], B_j\cap A)\rangle\Big)\Big] \\
&= \prod_{j=1}^n \mathbb{E}\left[\exp(i\langle u, c_j\rangle N([0,t], B_j\cap A))\right] \\
&= \prod_{j=1}^n \exp(t\nu(B_j\cap A)(e^{i\langle u,c_j\rangle -1})) \\
&= \exp(t\int_A (e^{i\langle u, f(z)\rangle} - 1)\nu(dz)) \\
&= \exp(t\int_{f(A)} (e^{i\langle u, z\rangle} - 1)(\nu\circ f^{-1})(dz)).
\end{aligned}$$

This is the characteristic function of a compound Poisson process with Lévy measure $\nu\circ f^{-1}$. In other words

$$\mathbb{E}\Big[\exp\Big(i\langle u, \underbrace{\int_B f(z)N([0,t],dz)\rangle}_{N_t^f(B)}\Big)\Big] = \exp(t\int_B (e^{i\langle u, f(z)\rangle} - 1)\nu(dz)).$$

Example 2.3.10. *For $f : H \to H$ and $f(z) = z$ we have that*

$$\begin{aligned}
\phi_{N^f(A)}(u) &= \mathbb{E}\left[\exp\left(i\langle u, \int_0^t \int_A f(z)N(ds,dz)\rangle\right)\right] \\
&= \exp(t\int_A (e^{i\langle u,z\rangle} - 1)\nu(dz)),
\end{aligned}$$

in other words, the process

$$N_t^f(A) = \int_0^t \int_A zN(ds,dz) = \int_A zN([0,t],dz)$$

is a compound Poisson process with intensity $\lambda(A)$ and jump measure $\frac{\nu(\cdot\cap A)}{\nu(A)}$.

Properties of $N_t^f(A)$:

1. $(N_t^f(B))_{t\geqslant 0}$ is a compound Poisson process with

   intensity $\lambda^f(B) := (\nu\circ f^{-1})(f(B))$ and jump distribution $M \mapsto \frac{\nu(M\cap f(B))}{\lambda^f(B)}$.

2. $(N_t^f(B_1))_{t\geqslant 0}, \ldots, (N_t^f(B_n))_{t\geqslant 0}$ are independent
   for disjoint $B_1, \ldots, B_n$ with $\lambda^f(B_1) + \cdots + \lambda^f(B_n) < \infty$
3. If $\int_B |f(z)|\nu(dz) < \infty$
$$\mathbb{E}[N_t^f(B)] = t \int_B f(z)\nu(dz)$$
4. If $\int_B |f(z)|^2\nu(dz) < \infty$
$$\mathbb{V}[N_t^f(B)] = t \int_B |f(z)|^2\nu(dz)$$
5. For $f : H \to \mathbb{R}$ such that the Laplace transform satisfies for any $u > 0$
$$\mathbb{E}\Big[\exp(-uN_t^f)\Big] = \exp\Big(- t \int_H (1 - e^{uf(z)})\nu(dz)\Big).$$
6. For $f : H \to \mathbb{R}$ such that $\int_H (e^{u_0 f(z)} - 1)\nu(dz) < \infty$ for some $u_0 > 0$ we have the **exponential moment** of for all $u \leqslant u_0$ that
$$\mathbb{E}\Big[\exp(uN_t^f)\Big] = \exp\Big(t \int_H (e^{uf(z)} - 1)\nu(dz)\Big).$$

Exponential moments for general integrals: We will apply the following result.

Lemma 2.3.11. *For a measurable function $f : [0,\infty) \times H \to [0,\infty)$ satisfying*
$$\int_0^t \int_H (e^{u_0 f(s,z)} - 1)\nu(dz)ds < \infty \qquad \text{for some } u_0 > 0.$$
we have
$$\mathbb{E}\Big[\exp(u \int_0^t \int_H f(s,z)N(ds,dz))\Big] = \exp\Big(\int_0^t \int_H (e^{uf(s,z)} - 1)\nu(dz)ds\Big).$$

F) Construction of the Poisson random integral from a given compound Poisson process: Consider a compound Poisson process $(C_t)_{t\geqslant 0}$ with $C_1 \sim \mathrm{Cpp}(\lambda,\nu)$. By definition there are a Poisson process $\pi = (\pi_t)_{t\geqslant 0}$ and an i.i.d. family $(Z_k)_{k\in\mathbb{N}}$ with $Z_k \sim \nu$ satisfying $\pi \perp Z$ such that
$$C_t = \sum_{k=1}^{\pi_t} Z_k, \qquad \mathbb{P}\text{-a.s.}, t \geqslant 0.$$

Denote the arrival times of π by $(T_k)_{k\in\mathbb{N}}$. Then the waiting times of π between the arrival times $t_k = T_k - T_{k-1}$ are an i.i.d. family $(t_k)_{k\in\mathbb{N}}$ of random variables with distribution $\mathrm{Exp}(\lambda)$. That is the event $\{\pi_t = k\} = \{T_k \leqslant t < T_{k+1}\}$ and by definition $Z_k = C_{T_k} - C_{T_k-} = \Delta_{T_k} C$. Therefore given a compound Poisson process $(C_t)_{t\geqslant 0}$ we write for $t > 0$

$$C_t = \sum_{k=1}^{\pi_t} \Delta_{T_k} C = \sum_{k=1}^{\pi_t} (\Delta_{T_k} C) \mathbf{1}_H(\Delta_{T_k} C) = \sum_{k=1}^{\pi_t} (\Delta_{T_k} C)\ \delta_{(\Delta_{T_k} C)}(H).$$

By construction (2.8) we have for the corresponding Poisson random measure N given by

$$N((a,b] \times A) := \sum_{k=\pi_a+1}^{\pi_b} \delta_{(\Delta_{T_k} C)}(A)$$

on $(H, \mathcal{B}(H))$ that

$$C_t = \int_H z N([0,t], dz) = N_t^{id}(H).$$

2.3.4 Paths of a Lévy process: the Lévy-Itô decomposition

We consider given an adapted Lévy process $(X_t)_{t\geqslant 0}$ with values in H on $(\Omega, \mathcal{A}, \mathbb{P})$ with càdlàg paths and characteristic triplet (b, Q, ν).

We shall remove more and more jumps of X and construct from these jumps a new pure jump process $\widetilde{X}$ with known characteristic function. Comparing it with the Lévy-Khinchin decomposition of X we can identify the process X path-by-path.

A) Subtracting large jumps: By Lemma 2.1.48 we have for any $\kappa > 0$

$$\#\{s \in [0,T] \mid |\Delta_s X| > \kappa\} < \infty, \qquad \mathbb{P} - \text{a.s.}$$

We consider the compound Poisson process defined by the jumps beyond the threshold κ by

$$\int_{|z|>\kappa} z N([0,t], dz) = \sum_{0\leqslant r\leqslant t} (\Delta_r X) \mathbf{1}_{\{|\Delta_r X| > \kappa\}}$$

and remove the large jumps of the process X by subtracting them

$$X_t^\kappa := X_t - \int_{|z|>\kappa} z N([0,t], dz).$$

B) The remainder process $(X_t^\kappa)_{t\geqslant 0}$ is still a Lévy process: We check stationarity and independence of the increments: Since

$$
\begin{aligned}
& X_t^\kappa - X_s^\kappa \\
&= X_t - \sum_{0\leqslant r\leqslant t} (\Delta_r X)\mathbf{1}_{\{|\Delta_r X|>\kappa\}} - X_s - \sum_{0\leqslant r\leqslant s} (\Delta_r L)\mathbf{1}_{\{|\Delta_r X|>\kappa\}} \\
&= X_t - X_s + \sum_{s<r\leqslant t} (\Delta_r X)\mathbf{1}_{\{|\Delta_r X|>\kappa\}} \\
&\overset{d}{=} X_{t-s} + \sum_{0<r\leqslant t-s} (\Delta_r X)\mathbf{1}_{\{|\Delta_r X|>\kappa\}} = X_{t-s}^\kappa .
\end{aligned}
$$

Hence it is measurable with respect to the σ-algebra

$$\sigma(\{X_r - X_q \mid s \leqslant q \leqslant r \leqslant t\}),$$

and independent from $\mathcal{F}_s := \sigma(\{X_r \mid 0 \leqslant r \leqslant s\})$. Therefore we have obtained a Lévy process and by construction it has bounded jumps $|\Delta_t X^\kappa| \leqslant \kappa$. The processes

$$(X_t^\kappa)_{t\geqslant 0} \quad \text{and} \quad \left(\int_{|z|>\kappa} zN([0,t],dz)\right)_{t\geqslant 0} \quad \text{are independent.}$$

The consequence of a Lévy processes with bounded jumps having any moment: By Proposition (2.2.46) we know that X^κ has any moment of order $p > 0$ for any $t \geqslant 0$ finite: $\mathbb{E}[|X_t^\kappa|^p] < \infty$.
In can be shown more generally, that for any submultiplicative function $f : H \to \mathbb{R}$ the existence of moments $\mathbb{E}\Big[|f(L_t)|\Big] < \infty$ is equivalent to the integrability of the tail

$$\int_{|z|>\kappa} |f(z)|\nu(dz) < \infty.$$

See Sato [61].
Now, since any Lévy measure ν on $B_\kappa^c(0)$ is a finite measure by the Lévy-Khinchin representation and for measures with additional bounded support we have that they exhibit finite polynomial and even exponential moments the result is somehow less surprising.

Nevertheless we have encountered here a breach between the "small" jump integrable part of a Lévy process and the "large" jump part, which we will exploit in the third part of these lecture notes. More precisely we can always divide the Lévy measure ν in a part $\nu\big|_{B_\kappa(0)}$ and $\nu\big|_{B_\kappa^c(0)}$ with very different behavior. While the first one is an infinite measure, however, with exponential moments, the second one is a finite measure with only some or no moments at all. For a given pure jump Lévy processes L this corresponds to the fact that we can rewrite it as the sum $L_t = L_t^\kappa + C_t^\kappa$ of a process L^κ with infinitely many κ-bounded jumps from above with exponential moments and a compound Poisson process C^κ with κ-bounded jumps from below with only some moments.

C) Compensated Poisson random measure We define the compensated Poisson measure for any $A \in \mathcal{B}(H)$ such that $\nu(A) < \infty$ by

$$\tilde{N}([0,t], A) := N([0,t], A) - t\nu(A)$$

and for any $f : H \to H$ measurable such that $\int_A |f(z)|\nu(dz) < \infty$ we define

$$\widetilde{N}_t^f(A) := \int_A f(z)\ \tilde{N}([0,t], dz) = \int_A f(z) N([0,t], dz) - t\int_A f(z)\nu(dz).$$

Properties of compensated Poisson random measures:

1. For disjoint $B_1, \ldots, B_n$ with $\lambda^f(B_1) + \cdots + \lambda^f(B_n) < \infty$, $\lambda^f(B) := \lambda(f^{-1} \circ f(B_i))$ we have that the family of processes $\Big((\widetilde{N}_t^f(B_1))_{t\geqslant 0}, \ldots, (\widetilde{N}_t^f(B_n))_{t\geqslant 0}\Big)$ is independent.

2. $(\widetilde{N}([0,t], A)_{t\geqslant 0}$ defines a compensated Poisson process at intensity $\lambda_A = \nu(A) < \infty$.

3. Since $\mathbb{E}[N_t^f(A)] = t\int_A f(z)\nu(dz)$, whenever the right-hand side is finite, we have that $\int_A |f(z)|\nu(dz) < \infty$ implies that $\mathbb{E}[\widetilde{N}_t^f] = 0$.

4. In addition, if $\nu(A) < \infty$ the $(\widetilde{N}([0,t], A))_{t\geqslant 0}$ is a square integrable martingale with respect to its augmented, right-continuous natural filtration with finite variation paths.

5. "Itô's Isometry": If $\int_B |z|^2\nu(dz) < \infty$

$$\begin{aligned}
\mathbb{V}\Big(\int_B z\ \widetilde{N}([0,t],dz)\Big) &= \mathbb{E}\left[\Big|\int_B z\ \widetilde{N}([0,t],dz)\Big|^2\right] \\
&= \mathbb{E}\left[\Big|\int_B z\ N([0,t],dz) - t\int_B z\nu(dz)\Big|^2\right] \\
&= \mathbb{E}\left[\Big|\int_B z\ N([0,t],dz) - \mathbb{E}\Big[\int_B z\ N([0,t],dz)\Big]\Big|^2\right] \\
&= \mathbb{V}\left[\int_B z\ N([0,t],dz)\right] \\
&= t\int_B |z|^2\nu(dz).
\end{aligned}$$

6. Similarly, we calculate its characteristic function

$$\begin{aligned}
\mathbb{E}\left[e^{i\langle u,\int_B f(z)\widetilde{N}([0,t],dz)\rangle}\right] &= \mathbb{E}\left[e^{i\langle u,\int_B f(z)N([0,t],dz)-t\int_B f(z)\nu(dz)\rangle}\right] \\
&= \mathbb{E}\left[e^{i\langle u,\int_B f(z)N([0,t],dz)}\right] e^{-i\langle u,t\int_B f(z)\nu(dz)\rangle} \\
&= e^{t\int_H (e^{i\langle u,z\rangle}-1)\nu(dz)} e^{-i\langle u,t\int_B f(z)\nu(dz)\rangle} \\
&= \exp(t\int_B (e^{i\langle u,f(z)\rangle} - 1 - i\langle u, f(z)\rangle)\nu(dz)).
\end{aligned}$$

D) Remove more and more compensated jumps Recenter L^κ

$$\widetilde{X}_t^\kappa := X_t^\kappa - \mathbb{E}[X_t^\kappa].$$

Add up compensated small jumps: For $\varepsilon_0 = \kappa$ and $\varepsilon_n \searrow 0$ the convergence in L^2 is obvious since by the Itô isometry

$$\begin{aligned}
&\mathbb{E}\left[\Big(\sum_{n=1}^m \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz)\Big)^2\right] \\
&= \sum_{n=1}^m \mathbb{E}\left[\Big(\int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz)\Big)^2\right] \\
&= t\sum_{n=1}^m \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} |z|^2\nu(dz) \leqslant t\int_{B_\kappa(0)} |z|^2\nu(dz) < \infty.
\end{aligned}$$

The right-hand side is independent of m and we can pass to the limit $m \to \infty$.

The $\mathbb{P}$-a.s. convergence: Denote by

$$\eta_m := \Big(\sum_{n=m+1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} |z|^2\nu(dz)|\Big)^{\frac{1}{4}} \searrow 0, m \to \infty.$$

Hence the Chevbyshev inequality yields

$$\begin{aligned}
&\mathbb{P}(|\sum_{n=m+1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz)| > \eta_m) \\
&\leqslant \frac{1}{\eta_m^2}\mathbb{E}\left[(\sum_{m=1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz))^2 \right] \\
&\leqslant \frac{1}{\eta_m^2} t \sum_{n=m+1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} |z|^2\nu(dz) \\
&= t\eta_m^2 \longrightarrow 0 \qquad \text{as } m \to \infty.
\end{aligned}$$

Now, since $|z|^2\nu(dz)$ is a finite measure we may choose $\varepsilon_n \searrow 0$ converging fast enough such that

$$\sum_{m=1}^{\infty} \eta_m < \infty.$$

The Borel-Cantelli lemma then yields

$$\sum_{m=1}^{\infty} \mathbb{P}(|\sum_{n=m+1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz)| > \eta_m) \leqslant \sum_{m=1}^{\infty} \eta_m < \infty,$$

such that

$$\begin{aligned}
0 &= \mathbb{P}\Big(\limsup_{m\to\infty} \{|\sum_{n=m+1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz)| > \eta_m\}\Big) \\
&= \mathbb{P}\Big(\#\{|\sum_{n=m+1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz)| > \eta_m\} = \infty \Big).
\end{aligned}$$

In other words we have a random variable $m^* : \Omega \to \mathbb{N}$ such that $\mathbb{P}$-a.s.

$$\begin{aligned}
&\lim_{m\to\infty} |\sum_{n=1}^{m} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz)| \\
&\leqslant |\sum_{n=1}^{m^*} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz)| + \eta_{m^*} < \infty.
\end{aligned}$$

Therefore we obtain $\mathbb{P}$-a.s.

$$\bar{X}_t^\kappa = \sum_{n=1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz) = \sum_{n=1}^{\infty} \int_{\varepsilon_{n-1}<|z|\leqslant\varepsilon_n} z\widetilde{N}([0,t],dz).$$

Subtracting all remaining small jumps $\bar{X}_t^\kappa$ of $\tilde{X}^\kappa$

$$X_t^c := \widetilde{X}_t^\kappa - \bar{X}_t^\kappa,$$

we know by construction that X^c is a continuous process. With the same stationarity and independence argument for the increments as for the independence of $X - \int_{|z|>\kappa} yN([0,t],dz) \perp \int_{|z|>\kappa} zN([0,t],dz)$ we have that it is a Lévy process. In addition, we have $X^c \perp \widetilde{X}^d$.

E) X_t^c is a Brownian motion with drift We now identify the law of X^c via the Lévy-Khinchin representation. Since $\widetilde{X}^d \perp \widetilde{X} - X^d$ and

$$\begin{aligned}\mathbb{E}\left[e^{i\langle u,\widetilde{X}_t^\kappa\rangle}\right] &= \exp\left(t\sum_{n=1}^{\infty}\int_{\varepsilon_{n+1}<|z|\leqslant\varepsilon_n}(e^{i\langle u,z\rangle}-1-i\langle u,z\rangle)\nu(dz)\right)\\ &= \exp\left(t\int_{0<|z|\leqslant 1}(e^{i\langle u,z\rangle}-1-i\langle u,z\rangle)\nu(dz)\right). \qquad (2.9)\end{aligned}$$

Looking back what have we achieved? The process

$$X_t^c = \underbrace{\underbrace{X_t - \int_{|z|>\kappa} zN([0,t],H)}_{X_t^\kappa} - \mathbb{E}[X_t^\kappa]}_{\widetilde{X}_t^\kappa} - \int_{|z|\leqslant\kappa} z\widetilde{N}([0,t]\times dz)$$

is a continuous Lévy process and

$$\begin{aligned}&\mathbb{E}\left[e^{i\langle u,X_t-\int_{|z|>\kappa} yN([0,t],dz)-\mathbb{E}[X_t^\kappa]-\int_{|z|\leqslant\kappa} z\widetilde{N}([0,t],dz)\rangle}\right]\\ &= \mathbb{E}\left[e^{i\langle u,X_t-t\mathbb{E}[X_1^\kappa]-(\int_{|z|>\kappa} yN([0,t],dz)+\int_{|z|\leqslant\kappa} z\widetilde{N}([0,t],dz))\rangle}\right]\\ &= e^{t(i\langle b-\mathbb{E}[X_1^\kappa],u\rangle-\frac{1}{2}\langle u,Qu\rangle)}.\end{aligned}$$

Hence X_t^c is a Brownian motion with drift.

The Lévy-Itô representation of a Lévy process: As a consequence, we have proved the following result.

Theorem 2.3.12 (Lévy-Itô representation of a Lévy process)**.** *For a given Levy process (X_t) on $(\Omega, \mathcal{A}, \mathbb{P})$ in H with canonical triplet (b, Q, ν) there is*

- *a vector* $\widetilde{b} = \mathbb{E}[X_1 - \int_{|z|\leqslant 1} zN([0,1], dz)]$
- *a Q-Brownian motion* $(B_t)_{t\geqslant 0}$ *and*
- *a pure jump process* $(\widetilde{X}_t)_{t\geqslant 0}$ *corresponding to the Lévy measure* ν

such that

$$X_t = \widetilde{b}t + B_t + \widetilde{X}_t \qquad \text{for all } t \geqslant 0 \qquad \mathbb{P}\text{-a.s.}$$

with

$$\widetilde{X}_t := \int_{|z|\leqslant 1} z\widetilde{N}([0,t], dz) + \int_{|z|>1} zN([0,t], dz)$$

In addition, the processes $\widetilde{X}$ *and* B *are independent.*

2.4 Quadratic variation and Burkholder's inequality

2.4.1 The quadratic variation of a compensated Poisson random integral:

A) The total and the quadratic variation of a Lévy process: Denote for $T > 0$ by $\Pi([0,T])$ the **set of all finite partitions** $0 = t_0 < t_1 < \dots t_n \leqslant T$, $n \in \mathbb{N}$.
The **total variation of a function** $f : [0,T] \to H$ is given as

$$|f|_T := \sup_{\pi \in \Pi([0,T])} \sup_{t_i \in \pi} |f(t_{i+1}) - f(t_i)| \in [0, \infty]$$

and describes the total "length" of the image $f([0,T])$ in H and can well be infinite.
The **quadratic variation of a function** $f : [0,T] \to H$ is given as

$$[f]_T := \sup_{\pi \in \Pi([0,T])} \sum_{t_i \in \pi} |f(t_{i+1}) - f(t_i)|^2 \in [0, \infty]$$

EXERCISE 2.4.1. *Check that $f(s) := bs$ for some $b \in H$ we have*

$$|f|_T = |b|T \qquad \textit{and} \qquad [f]_T = 0.$$

Remark 2.4.2. *The quadratic or more generally the p-variation for $p > 1$ measures different types of "infinite" path lengths.*

Remark 2.4.3. *For any Lévy process X the quadratic variation is finite $\mathbb{P}$-a.s.*

Theorem 2.4.4 (Lévy)**.** *The quadratic variation of a Q-Brownian motion in H satisfies*

$$[B]_t := \sup_{\pi \in \Theta([0,t])} \sum_{t_i \in \pi} |B_{t_{i+1}} - B_{t_i}|^2 = t \,\mathrm{trace}(Q) \qquad \mathbb{P}- \ \textit{a.s. for all } t \geqslant 0.$$

See for instance [57].

Example 2.4.5. *For a pure jump process this is easier to calculate. For a compound Poisson process $C_t = \int_0^t \int_H z N(ds, dz)$ with $C_1 \sim Cpp(\lambda, \mu)$ we have*

$$|C|_t = \sum_{k=1}^{\pi_t} |Z_k| < \infty.$$

and

$$[C]_t = \sum_{k=1}^{\pi_t} |Z_k|^2 < \infty.$$

Note that these representations themselves are Poisson random integrals:

$$|C|_t = \int_0^t \int_H |z| N(ds, dz) \qquad \textit{and} \qquad |C|_t = \int_0^t \int_H |z|^2 N(ds, dz).$$

For a pure jump Lévy process X in H we have

$$X_t := \int_{|z|\leqslant 1} z\widetilde{N}([0,t], dz) + \int_{|z|>1} zN([0,t], dz)$$

we have

$$[X]_t \leqslant [\int_{|z|\leqslant 1} z\widetilde{N}([0,\cdot], dz)]_t + [\int_{|z|>1} zN([0,\cdot], dz)]_t$$

For the second term we know by the preceding example that

$$[\int_{|z|>1} zN([0,\cdot], dz)]_t = \int_0^t \int_H |z|^2 N(ds, dz).$$

For the first term we see that in case of $\int_{|z|\leqslant 1} |z|\nu(dz) < \infty$ we can write

$$\int_0^t \int_{|z|\leqslant 1} z\widetilde{N}(ds, dz) = \int_0^t \int_{|z|\leqslant 1} zN(ds, dz) - \int_0^t \int_{|z|\leqslant 1} z\nu(dz)ds$$

and hence

$$\begin{aligned}
[\int_0^\cdot \int_{|z|\leqslant 1} z\widetilde{N}(ds, dz)]_t &= [\int_0^\cdot \int_{|z|\leqslant 1} zN(ds, dz) - \int_0^t \int_{|z|\leqslant 1} z\nu(dz)ds]_t \\
&= [\int_0^\cdot \int_{|z|\leqslant 1} zN(ds, dz)]_t \\
&= \int_0^t \int_{|z|\leqslant 1} |z|^2 N(ds, dz).
\end{aligned}$$

The right-hand side is finite $\mathbb{P}$-a.s. since

$$\mathbb{E}\Big[\int_0^t \int_{|z|\leqslant 1} |z|^2 N(ds, dz)\Big] = \int_0^t \int_{|z|\leqslant 1} |z|^2 \nu(dz)ds < \infty$$

for any Lévy measure. By (2.9) we have for a general pure jump Lévy process for any $m \in \mathbb{N}$ and $0 < \varepsilon_{n+1} < \varepsilon_n < \ldots < \kappa$ with $\varepsilon_n \searrow 0$ as $n \to \infty$ that

$$\begin{aligned}
&\mathbb{E}\left[\big[\sum_{n=1}^{m} \int_{\varepsilon_{n+1} < |z| \leqslant \varepsilon_n} z\widetilde{N}([0,\cdot], dz) \big]_t \right] \\
&= \mathbb{E}\left[\sum_{n=1}^{m} \big[\int_{\varepsilon_{n+1} < |z| \leqslant \varepsilon_n} z\widetilde{N}([0,\cdot], dz) \big]_t \right] \\
&= \sum_{n=1}^{m} \mathbb{E}\left[\big[\int_{\varepsilon_{n+1} < |z| \leqslant \varepsilon_n} z\widetilde{N}([0,\cdot], dz) \big]_t \right] \\
&= t \sum_{n=1}^{m} \int_{\varepsilon_{n+1} < |z| \leqslant \varepsilon_n} |z|^2 \nu(dz) \leqslant t \int_{B_\kappa(0)} z|^2 \nu(dz) < \infty.
\end{aligned}$$

The right-hand side is independent of m and we can pass to the limit and obtain that

$$\mathbb{E}\left[\big[\int_{0 < |z| \leqslant \kappa} z\widetilde{N}([0,\cdot], dz) \big]_t \right] = t \sum_{n=1}^{\infty} \int_{\varepsilon_{n+1} < |z| \leqslant \varepsilon_n} |z|^2 \nu(dz) = t \int_{B_\kappa(0)} |z|^2 \nu(dz).$$

For later purpose in Part III we need the following natural generalization. For a pure jump Lévy process X

$$X_t := \int_{|z| \leqslant \kappa} z\widetilde{N}([0,t], dz)$$

and a measurable function $f : [0,\infty) \times H \to [0,\infty)$ we obtain

Lemma 2.4.6. *Under the previous assumptions we obtain for all $\mathbb{P}$-a.s. for all $t \geqslant 0$*

$$[\int_0^{\cdot} \int_{|z| \leqslant \kappa} f(s,z) \widetilde{N}(ds,dz)]_t = \int_0^t \int_{|z| \leqslant \kappa} |f(s,z)|^2 N(ds,dz).$$

2.4.2 The Burkholder-Davis-Gundy inequality

The following result links the expected the supremum of a Poisson random integral to its the expectation of the square root of its quadratic variation.

Theorem 2.4.7. *Let X be a pure jump Lévy process with values in a separable Hilbert space H given as*

$$X_t = \int_0^t \int_{|z| \leqslant 1} z\widetilde{N}(ds,dz), \qquad \mathbb{P} - a.s. \text{ for all } t \geqslant 0.$$

and a measurable function $f : [0,\infty) \times H \to H \to H$. *Then there is a constant* $C > 0$ *such that for any* $p \geqslant 1$ *such that for all* $T \geqslant 0$

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |\int_0^t \int_{|z|\leqslant 1} f(s,z)\widetilde{N}(ds,dz)|^p\Big]$$
$$\leqslant C^p \mathbb{E}\Big[[\int_0^{\cdot} \int_{|z|\leqslant 1} f(s,z)\widetilde{N}(ds,dz)]_T^{\frac{p}{2}}\Big]$$
$$= C^p \mathbb{E}\Big[\Big(\int_0^T \int_{|z|\leqslant 1} |f(s,z)|^2 N(ds,dz)\Big)^{\frac{p}{2}}\Big].$$

Remark 2.4.8. *This result is much more general than stated here, for a proof see for instance [55]. In particular, the inverse inequality is also true for another constant. In addition it remains true if* T *is replaced by a* $(\mathcal{F}_t)_{t\geqslant 0}$*-stopping time* τ.

Remark 2.4.9. *In particular for* $p = 2$ *we obtain the useful inequality*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |\int_0^t \int_{|z|\leqslant 1} f(s,z)\widetilde{N}(ds,dz)|^2\Big] \leqslant C^2 \int_0^T \int_{|z|\leqslant 1} |f(s,z)|^2 \nu(dz)ds.$$

Recently the field of a pathwise approach of stochastic integration associated to the integration theory of rough paths has given rise to many interesting insights. This theory is associated to Terry Lyons, Martin Hairer (Fields medal 2014) and Massimo Gubinelli.
In 2015 Siorpaes and Beiglböck [62] proved a path-by-path version of the Burkholder Davis Gundy inequality, which can be formulated in our language of compensated Poisson random measures as follows.

Theorem 2.4.10 (Siorpaes/Beiglböck)**.** *Given a pure jump* $(\mathcal{F}_t)$*-adapted Lévy process* $X = (X_t)_{t\geqslant 0}$ *over* $(\Omega, \mathcal{A}, \mathbb{P})$ *with values in* H *and with Lévy-Itô decomposition*

$$X_t = \int_0^t \int_{|z|\leqslant \rho} z\widetilde{N}(ds,dz), \qquad t \geqslant 0,$$

for some $\rho > 0$ *and* $(\Psi_t)_{t\geqslant 0}$ *an* $(\mathcal{F}_t)$*-adapted càdlàg process. Then the stochastic process*

$$M_t := \int_0^t \int_{|z|\leqslant \rho} e^{as}\langle \Psi_{s-}, z\rangle \widetilde{N}(ds,dz)$$

satisfies $\mathbb{P}$-*a.s. for any* $T > 0$

$$\begin{aligned}\sup_{t\in[0,T]} |M_t| &\leqslant 6\sqrt{[M]_T} + \int_0^T H_{s-} dM_s \\ &= 6\sqrt{\int_0^T \int_{|z|\leqslant\rho} e^{2as}\langle\Psi_{s-}, z\rangle^2 N(dsdz)} \\ &\qquad + \int_{|z|\leqslant\rho} H_{s-} e^{as}\langle\Psi_{s-}, z\rangle \widetilde{N}(ds, dz),\end{aligned}$$

where $H_t := M_t / \sqrt{[M]_t + \sup_{s\leqslant t} |M_s|^2} \in [-1, 1]$.

Part II

The Allen-Cahn equation

Chapter 3

The deterministic Allen-Cahn equation and its dynamics

3.1 The state space

We consider the following function spaces over interval $J = (0,1)$, which will serve as the state space. The "spatial variable" will be throughout out the text denoted by $\zeta \in J$.

$$\mathcal{L}^2(J) := \{x : J \to \mathbb{R} \mid \text{measurable and } \int_J x(\zeta)^2 d\zeta < \infty\},$$

equipped with the seminorm

$$|x|_\sim := \Big(\int_J x(\zeta)^2 d\zeta\Big)^{\frac{1}{2}}.$$

We denote by $\sim$ the equivalence relation

$$x_1 \sim x_2 \quad :\Longleftrightarrow \quad x_1(\zeta) = x_2(\zeta) \qquad \text{for Lebesgue almost all } \zeta \in J.$$

The Lebesgue space of square integrable functions is then defined by

$$L^2(J) := \mathcal{L}^2(J) \Big/ \sim$$

and

$$|x| := |\widetilde{x}|_\sim, \qquad \text{for any representative } \widetilde{x} \text{ of the equivalence class of } x.$$

We define the following bilinear form on $L^2(J)$

$$\langle x, y\rangle := \int x(\zeta)y(\zeta)d\zeta, \qquad x, y \in L^2(J).$$

EXERCISE 3.1.1. *1. Show that $|\cdot|_\sim$ defines seminorms in $\mathcal{L}^2(J)$.*

2. Show that $|\cdot|_2$ define a norm in $L^2(0,1)$.

3. Show that $\langle\cdot,\cdot\rangle$ defines an inner product with associated norm $|\cdot|$ and verify the parallelogram identity

$$|x+y|^2 + |x-y|^2 = 2|x|^2 + 2|y|^2, \qquad x, y \in L^2(J).$$

4. Show that the normed space $(L^2(J), |\cdot|)$ with inner product $\langle\cdot,\cdot\rangle$ is complete, in other words, it is a Hilbert space.

Analogously we define the space

$$\mathcal{L}^\infty(J) := \{x : J \to \mathbb{R} \mid \text{ measurable and } \sup_{\zeta\in J}|x(\zeta)| < \infty\}$$

equipped with the seminorm

$$|x|_{\infty,\sim} := \sup_{\zeta\in J}|x(\zeta)|.$$

We set

$$L^\infty(J) := \mathcal{L}^\infty(J)\Big/\sim$$

and for the same equivalence relation $\sim$ as before we set

$$|x|_\infty := |\widetilde{x}|_{\infty,\sim}, \qquad \text{for any representative } \widetilde{x} \text{ of the equivalence class of } x.$$

EXERCISE 3.1.2. *1. Show that $|\cdot|_{\infty,\sim}$ defines a seminorm in $\mathcal{L}^\infty(J)$.*

2. Show that $|\cdot|_\infty$ defines a norm in $L^\infty(J)$.

3. Show that the normed space $(L^\infty(J), |\cdot|_\infty)$ is complete, in other words, it is a Banach space.

However, the true state space will be a subspace of $L^2(J)$ with "smoother" elements. Define

$$\mathcal{C}_0^1(J) := \{x \in \mathcal{C}^1(J,\mathbb{R}) \cap \mathcal{C}(\bar{J},\mathbb{R}) | x(0) = x(1) = 0\}.$$

This space is a normed space with the norm

$$|x|_1 := \sup_{\zeta\in J}|x(\zeta)| + \sup_{\zeta\in J}|x'(\zeta)|$$

EXERCISE 3.1.3.

Show that $(\mathcal{C}_0^1(J), |\cdot|_1)$ *is a normed space.*

Show that $(\mathcal{C}_0^1(J), |\cdot|_1)$ *is complete, in other words, it is a Banach space.*

Since all continuous functions $\mathcal{C}(\bar{J}, \mathbb{R})$ are square integrable, we have for all $x, y \in \mathcal{C}_0^1(J)$ the integration by parts formula

$$\langle x, y \rangle = x(1)Y(1) - x(0)Y(0) - \langle x', Y \rangle = \langle x', Y \rangle, \tag{3.1}$$

where $Y(\zeta) = \int_0^\zeta y(r)dr + C$ for some constant $C \in \mathbb{R}$. Therefore we can define the space of all functions $x \in L^2(J)$ which have some "weak" derivative $x' \in L^2(J)$ such that equation (3.1) is satisfied for all "test functions" $y \in \mathcal{C}_0^1(J)$:

$$H := H_0^1(J) := \{x \in L^2(J) \mid \exists x' \in L^2(J) : \forall \psi \in \mathcal{C}_c^\infty(J) \mid \langle \psi', x \rangle = -\langle x', \psi \rangle\}.$$

We consider on this space the bilinear form

$$\langle\langle x, y \rangle\rangle_0 := \langle x, y \rangle + \langle x', y' \rangle.$$

EXERCISE 3.1.4.

1. *Show that* $(H, \|\cdot\|_0)$, $\|x\|_0 := \sqrt{\langle\langle x, x \rangle\rangle_0}$ *defines a normed space with the inner product* $\langle\langle \cdot, \cdot \rangle\rangle_0$.
2. *Show that* $(H, \|\cdot\|_0)$ *is complete, that is it forms together with the inner product* $\langle\langle \cdot, \cdot \rangle\rangle_0$ *a Hilbert space.*

EXERCISE 3.1.5.

1. *Use the fundamental theorem of calculus and Hölder's inequality to show that for all* $\mathcal{C}_0^1(J)$ *we have the so-called Poincaré inequality*
$$|x|^2 \leqslant |x'|^2 \tag{3.2}$$
and infer that $H_0^1(J) \hookrightarrow L^2(J)$.
2. *Show that*
$$\|x\| := \sqrt{\langle x', x' \rangle}, \qquad x \in H_0^1(J)$$
is an equivalent norm on H *to* $\|\cdot\|_0$ *and hence*
$$(H, \|\cdot\|, \langle\langle \cdot, \cdot \rangle\rangle)$$
is a Hilbert space.

A special case of the Sobolev embedding yields $H \subset \mathcal{C}_0(\bar{J})$.

EXERCISE 3.1.6. *Show that for all* $x \in H$ *we have*

$$|x|_\infty \leqslant \|x\|.$$

and infer $H_0^1(J) \hookrightarrow \mathcal{C}_0(\bar{J})$. *It works analogously to inequality (3.2).*

We summarize our results:

$$(H_0^1(J), \|\cdot\|, \langle\langle \cdot, \cdot \rangle\rangle) \hookrightarrow (\mathcal{C}_0(\bar{J}), |\cdot|_\infty) \hookrightarrow (L^\infty(J), |\cdot|_\infty) \hookrightarrow (L^2(J), |\cdot|, \langle \cdot, \cdot \rangle).$$

3.2 The linear heat equation

A) The linear heat equation: In this section we provide the results for the scalar linear heat equation with Dirichlet boundary conditions and initial values in $x \in H$ or $x \in L^2(0,1)$

$$\begin{cases} \frac{d}{dt}u(t,\zeta) = \frac{\partial^2}{\partial\zeta^2}u(t,\zeta) & t \geqslant 0, \zeta \in (0,1) \\ u(0,\zeta) = x(\zeta) & \zeta \in (0,1) \\ u(t,0) = u(t,1) = 0 & t \geqslant 0. \end{cases} \tag{3.3}$$

The operator $A := -\frac{\partial^2}{\partial\zeta^2}$ is an unbounded operator on a densely defined domain $A : D(A) \to H$, with $D(A) := \mathcal{C}_0^2(J)$, where $\mathcal{C}_0^2(J) := \mathcal{C}^2(J,\mathbb{R}) \cap \{x \in \mathcal{C}(\bar{J},\mathbb{R}) | x(0) = x(1) = 0\}$.
The sequence of eigenvalues $(\lambda_k)_{k\in\mathbb{N}}$ and eigenvectors $(w_k)_{n\in\mathbb{N}}$ of A in $L^2(0,1)$ are given as

$$\lambda_k = (\pi k)^2, \qquad \text{and} \qquad w_k(\zeta) = \sqrt{2}\sin(k\pi\zeta), \qquad \text{for all } k \in \mathbb{N}.$$

Note that the eigenvectors are extremely regular, that is $w_k \in \mathcal{C}_0^\infty(J) := \mathcal{C}^\infty(J) \cap \mathcal{C}_0(\bar{J})$ for all $k \in \mathbb{N}$.

EXERCISE 3.2.1. *1. Verify that $(\lambda_k)_{k\in\mathbb{N}}$ and $(w_k)_{k\in\mathbb{N}}$ are the sequence of eigenvalues and eigenvectors in $\mathcal{C}_0^2$.*

2. The sequence $(w_n)_{n\in\mathbb{N}}$ is an orthonormal basis in $L^2(J)$.

B) The semigroup: Since $\langle Ax, x\rangle = \langle x', x'\rangle$ for any $x \in \mathcal{C}_0^2$ we have by (3.2) that there is a constant Λ_0 such that

$$\langle Ax, x\rangle \leqslant \Lambda_0 |x|^2, \qquad x \in H, x \neq 0$$

for any $t \geqslant 0$.
Solving equation (3.3) for initial condition w_k we obtain the ordinary linear differential equation

$$\frac{d}{dt}u(t; w_k) = -\lambda_k u(t; w_k), \qquad u(0; w_k) = w_k$$

That is $u(t; w_k) = e^{-\lambda_k t} w_k$. Since $(w_k)_{k\in\mathbb{N}}$ is an orthonormal basis in $L^2(J)$ and the linearity of equation (3.3) we have vor any $x \in H$ that

$$u(t;x) = u(t; \sum_{k\in\mathbb{N}} \langle x, w_k\rangle w_k) = \sum_{k\in\mathbb{N}} \langle x, w_k\rangle u(t; w_k) = \sum_{k\in\mathbb{N}} e^{-\lambda_k t}\langle x, w_k\rangle w_k.$$

Since the operator A acts as $-\lambda_k$ on the eigenspace $\mathrm{Lin}(w_k)$ we may rewrite $u(t;x)$ as

$$u(t;x) := e^{At}x := \sum_{k\in\mathbb{N}} e^{-\lambda_k t}\langle x, w_k\rangle w_k.$$

In addition we note that $\frac{d}{d\zeta}x = x' = A^{\frac{1}{2}}x$. Therefore the orthogonality of $(w_k)_{k\in\mathbb{N}}$ implies for any $x \in H$ that

$$\|x\|^2 = |A^{\frac{1}{2}}x|^2 = |\sum_{k\in\mathbb{N}} \lambda_k^{\frac{1}{2}}\langle x, w_k\rangle w_k|^2 = \sum_{k\in\mathbb{N}} \lambda_k |\langle x, w_k\rangle|^2 < \infty.$$

Remark 3.2.2. *The solution operator $H \ni x \mapsto u(t;x) \in H$ defines a strongly continuous semigroup on $L^2(0,1)$ and H, called the* ***heat semigroup****. For notational convenience we shall denote $S(t) := e^{At}$. See for instance [22, 65].*

For an construction of the semigroup S starting with the problem on $\mathbb{R}$ instead of $J = (0,1)$ by the so-called method of images we refer to section 1.3 and 1.4 of [11]. In the sequel we shall stick to this setting.

Properties of the semigroup S: The semigroup S is contracting on $L^2(0,1)$ and H

$$|S(t)x| \leqslant e^{-\Lambda_0 t}|x|, \qquad x \in L^2(0,1) \tag{3.4}$$

$$\|S(t)x\| \leqslant e^{-\Lambda_0 t}\|x\|, \qquad x \in H, \tag{3.5}$$

and as a consequence the operator norm satisfies

$$\|S(t)\| \leqslant e^{-\Lambda_0 t}, \qquad t \geqslant 0. \tag{3.6}$$

EXERCISE 3.2.3. *Prove (3.4), (3.5) and (3.6) with the help of the basis $(w_k)_k$.*

The semigroup is very regularizing in the sense that between between the spaces $H \subset \mathcal{C}_0([0,1]) \cap L^\infty(0,1) \subset L^2(0,1)$ as follows. There is a constant $C > 0$ such that for all $t > 0$

$$\|S(t)x\| \leqslant C\frac{e^{-\Lambda_0 t}}{t^{1/2}}|x|, \qquad x \in L^2(0,1)$$

$$|S(t)x|_\infty \leqslant C\frac{e^{-\Lambda_0 t}}{t^{1/4}}|x|, \qquad x \in L^2(0,1).$$

We refer to Proposition A.12 in [22].

3.3 The deterministic Allen-Cahn equation

A) The Allen-Cahn equation: We consider following nonlinear partial differential equation. For $t \geqslant 0$, $x \in H$ and $\zeta \in J$ consider

$$\begin{cases} \frac{\partial}{\partial t} u(t,\zeta) = \dfrac{\partial^2 u}{\partial \zeta^2}(t,\zeta) + f(u(t,\zeta)) & \\ u(t,0) = u(t,1) = 0 & \text{Dirichlet boundary conditions} \\ u(0,\zeta;x) = x(\zeta) & \text{initial value } x, \end{cases} \tag{3.7}$$

where the non-linearity f is given for some $\lambda > 0$ by

$$f(r) := -\lambda(r^3 - r), \qquad r \geqslant 0. \tag{3.8}$$

B) Weak solutions: There is some freedom of choice for the space of the initial conditions x. Our aim is the establishment of a dynamical system in H, since formally if $u \in H$ we obtain most conveniently

$$\langle \frac{\partial^2 u}{\partial \zeta^2}, u \rangle = -\langle \frac{\partial u}{\partial \zeta}, \frac{\partial u}{\partial \zeta} \rangle = -\|u\|^2.$$

In [65], III.1.1 it is shown via Faedo-Galerkin approximations that equation (3.7) has a unique weak solution in the sense that

$$\frac{d}{dt}\langle u, v\rangle + \langle\langle u, v\rangle\rangle - \langle f(u), v\rangle = 0, \qquad \text{for all } v \in H \cap L^4(J).$$

For this prove it is used mainly the polynomial structure of the nonlinearity. In addition it is shown there that for all $x \in H$ any $T > t_0 > 0$

$$u \in \mathcal{C}([0,T], H) \cap L^2((0,T], H^2(J)).$$

C) Mild solutions: In the sequel we shall establish mild solutions. For this aim we note the following:

Lemma 3.3.1. *For any $\chi > 0$ we have constants $K_{1,\chi}, K_{2,\chi} > 0$ such that*

$$\begin{aligned} |f(r_1) - f(r_2)| &\leqslant K_{1,\chi}|r_1 - r_2| \qquad && r_1, r_2 \in \mathbb{R} \textit{ with } |r_1|, |r_2| \leqslant \chi \\ \|f(x) - f(y)\| &\leqslant K_{2,\chi}\|x - y\| && x, y \in H \textit{ with } \|x\|, \|y\| \leqslant \chi. \end{aligned}$$

Proof. The proof is found in [63], Chapter 5.1.1., and in [20], Lemma 2.1. The first inequality is obvious for polynomials. We show the second inequality.

Step 1: f is locally Lipschitz from $L^2(0,1)$ to H. We start with $x, y \in B_\chi(0) \subset H \hookrightarrow L^\infty(0,1)$. Due to $|x|_\infty \leqslant |\nabla x|$ for all $x \in H$, we have $|x|_\infty, |v|_\infty \leqslant \chi$. In particular for each $\zeta \in (0,1)$, $\theta \in [0,1]$

$$
\begin{aligned}
|f'(x(\zeta) + \theta(y(\zeta) - x(\zeta)))| &\leqslant \sup_{y \in B_\chi(0)} |f'(y)|_\infty =: K_{1,\chi} < \infty, \qquad (3.9)\\
|f''(x(\zeta) + \theta(y(\zeta) - x(\zeta)))| &\leqslant \sup_{y \in B_\chi(0)} |f''(y)|_\infty =: K_{2,\chi} < \infty.
\end{aligned}
$$

Hence due to the mean value theorem

$$
\begin{aligned}
|f(x) - f(y)|_{L^2}^2 &= \int_0^1 (f(x(\zeta)) - f(y(\zeta)))^2 \, \mathrm{d}\zeta \\
&= \int_0^1 \big| \Big(\int_0^1 f'(x(\zeta) + \theta(y(\zeta) - x(\zeta))) \, \mathrm{d}\theta \Big) (x(\zeta) - y(\zeta)) \big|^2 \, \mathrm{d}\zeta \\
&\leqslant K_{1,\chi}^2 |x - y|_{L^2}^2 \leqslant K_{1,\chi}^2 \|x - y\|^2.
\end{aligned}
$$

Step 2: f is locally Lipschitz from H to H For $x, y \in B_\chi(0) \subset H$ we may calculate

$$
\begin{aligned}
&\|f(x) - f(y)\|^2 \\
&= |f'(x)x' - f'(y)y'|_{L^2}^2 \\
&\leqslant 2|f'(x)x' - f'(y)x'|_{L^2}^2 + 2|f'(y)x' - f'(y)y'|_{L^2}^2 \\
&\leqslant 2|f'(x) - f'(y)|^2 \|x\|^2 + 2|f'(y)|^2 \|x - y\|^2 \\
&\leqslant 2{K'}_{2,\chi}^2 R^2 |x - y|^2 + 2K_{1,\chi}^2 \|x - y\|^2 \leqslant \underbrace{2\left((K_{1,\chi}\chi)^2 + {K'}_{2,\chi}^2\right)}_{=: K_{\chi,R}} \|x - y\|^2.
\end{aligned}
$$

□

A **mild solution of the Allen-Cahn equation** (3.7) satisfies the following variation of constants formula in the space H.

$$
u(t;x) = S(t)x + \int_0^t S(t-s) f(u(s;x)) ds.
$$

This equation has unique and well-posed weak and mild solutions in $L^2(0,1)$ and H (cf. [20, 65]). The solutions are most regular for any $t > 0$ and $x \in L^2(0,1)$, that is $u(t;x) \in \mathcal{C}^\infty(0,1) \cap \mathcal{C}_0[0,1]$.

3.4 The dynamics of the Allen-Cahn equation

A) Gradient structure: It is well-known that equation (3.7) enjoys the nonnegative potential function $\mathcal{V}(x) = \int_J \big((\nabla x(\zeta))^2 + F(x(\zeta))\big)d\zeta$ on H for $F(r) = \int_{r_0}^r f(s)ds$ for some r_0 and is rewritten as the following gradient system

$$\frac{\partial}{\partial t}u(t,\zeta) = -(D\mathcal{V})(u(t,\zeta)) \qquad \text{with} \qquad u(0,\zeta;x) = x(\zeta) \qquad \text{for } x \in H.$$

See for instance Brassesco [11]. As a consequence, $\mathcal{V}$ serves as a Lyapunov function and yields the following result (cf. [32, 34]). The level sets of $\mathcal{V}$ remain bounded in H and positive invariant under the dynamical system (3.7).

$$\mathcal{U}^r := \{x \in H \mid \mathcal{V}(x) \leqslant d(r)\}, \qquad d(r) := \inf\{r' > 0 \mid B_r(0) \subseteq \mathcal{U}^{r'}\}, \qquad r > 0.$$

Proposition 3.4.1. *Denote by $\mathcal{P} \subseteq H$ the set of fixed points of (3.7). Then $0 < |\mathcal{P}| < \infty$ and for any $x \in H$ there exists a stationary state $\phi \in \mathcal{P}$ of the system (3.7) such that* $\lim_{t\to\infty} u(t;x) = \phi$.

B) Generic parameters and domains of attraction: In [14, 33] it is shown that the dynamical system u has the **Morse-Smale property**, whenever $\lambda \neq (2\pi n)^2$, that is there are only finitely many critical points $\phi \in H$

$$\frac{\partial^2\phi}{\partial\zeta^2}(\zeta) + f(\phi(\zeta)) = 0, \qquad \zeta \in (0,1)$$

and the (unbounded) linearization of the right-hand side

$$\psi \mapsto \frac{\partial^2\psi}{\partial\zeta^2} + f'(\phi)\psi, \qquad \zeta \in (0,1)$$

has all eigenvalues different from zero. In addition, the infinite-dimensional stable and the finite-dimensional unstable manifolds are known to intersect transversally.

For $\phi \in \mathcal{P}$ we denote the **domain of attraction**

$$D(\phi) := \{x \in H \mid \lim_{t\to\infty} u(t;x) = \phi\}.$$

The subset $\mathcal{P}^-$ of all $\phi \in \mathcal{P}$ such that $D(\phi)$ contains an open ball in H is the **set of stable states**. For $\phi^\pm \in \mathcal{P}^-$ we denote its domain of attraction $D^\pm = D(\phi^\pm)$ and the separating manifold between them by $\mathcal{S} := H \setminus \bigcup_\pm D^\pm$. For $\lambda \neq (2\pi n)^2$, $n \in \mathbb{N}$, the Morse-Smale property implies that $\mathcal{S}$ is a closed $\mathcal{C}^1$-manifold without boundary in H of codimension 1 separating all elements of $(D^\pm)_{\phi^\pm\in\mathcal{P}^-}$ called **separatrix** and containing all unstable fixed points $\mathcal{P}\setminus\mathcal{P}^-$ (cf. [59]).

C) The dissipativity: We shall see that the solution of equation (3.7) is **uniformly absorbed** in a large ball. That is, there are constants $\kappa^* > 0$ and $r^* > 0$ such that such that $t \geqslant \kappa^*$ implies that the solution satisfies $\sup_{x \in H} \|u(t;x)\| \leqslant r^*$.
In order to see this we have to recall the following non-linear version of Gronwall's lemma.

Lemma 3.4.2. *Let $a : (0,\infty) \to (0,\infty)$ be a measurable and absolutely continuous function satisfying*

$$\frac{da}{dt} + \gamma a^p \leqslant \delta$$

for some $p > 1$ and $\delta > 0$. Then for $t \geqslant 0$ we have

$$a(t) \leqslant \left(\frac{\delta}{\gamma}\right)^{1/p} + \frac{1}{(\gamma(p-1)t)^{1/p-1}}.$$

Remark 3.4.3. *Note that the right-hand side does not depend on $a(0)$.*

EXERCISE 3.4.4. *Prove the Lemma 3.4.2.*

Consider the ordinary differential equation

$$\frac{d}{dt}u = -\lambda(u^3 - u), \qquad u(0) = x \in \mathbb{R}.$$

Then multiplying the equation with $\frac{1}{2}u$ we obtain

$$\frac{1}{2}(\frac{d}{dt}u)u = \frac{d}{dt}u^2 = -\lambda(u^4 - u^2) = -\lambda\big((u^2)^2 - u^2\big).$$

Hence for all initial values $x \in \mathbb{R}$ such that $x^2 \geqslant 2$ we have

$$\frac{d}{dt}u^2 = -\lambda(u^4 - u^2) \leqslant -\frac{\lambda}{2}u^4 + C. \tag{3.10}$$

EXERCISE 3.4.5. *Check inequality (3.10) and determine the constant $C > 0$.*

Lemma 3.4.2 applied to inequality (3.10) for the values $p = 2$, $\delta = C$, $\gamma = \frac{\lambda}{2}$ and $a(t) = u(t)^2$ yields

$$u^2(t) \leqslant \left(\frac{2C}{\lambda}\right)^{1/2} + \frac{2}{\lambda t} \qquad \forall\, t > 0.$$

This argument is the core of the same argument given in the space H. This argument can be made rigorous for $u(t, x(\zeta))$ for Lebesgue all $\zeta \in J$ and hence leads to the same result in $L^\infty(J)$. Now we can use the mild solution and the regularization from $L^\infty(J)$ to H and obtain the following dissipativity result.

Proposition 3.4.6. *There exist constants $\kappa^*, r^* > 0$ such that for all $t \geqslant \kappa^*$ and $x \in H$ we have*

$$\|u(t;x)\| \leqslant r^*.$$

A complete proof is given in Section 4.1 of [20]. Another point of view on the dissipativity is the following.

EXERCISE 3.4.7 (Dissipativity in finite dimensions = blow up in finite backward time)**.**

1) Solve the ordinary differential equation

$$\frac{du}{dt} = u^2, \qquad u(0) = x \in \mathbb{R}$$

explicitly (separation of variables). What can you say about the life time of the solution $t \mapsto u(t;x)$?

The time reversal $t \mapsto -t$ leads to the equation

$$\frac{dv}{dt} = -v^2, \qquad v(0) = x \in \mathbb{R}.$$

2) What does the first result imply for the convergence of the solution $t \mapsto v(t;x)$ for different initial values to the unit ball?

D) The fine dynamics justifies reduced domains of attractions: We summarize our knowledge so far.

Dissipativity: There are a finite time $\kappa > 0$ and a radius $r^* > 0$ such that uniformly for all initial values $x \in H$ the solution $t \mapsto u(t;x)$ has entered a ball of $B_{r^*}(0)$ before time $t = \kappa$.

Gradient structure: Inside this ball there all finitely many but at least one critical state. For $\lambda > \pi^2$ there are at least three critical states, precisely or two of them (denoted by ϕ^+ and ϕ^-) are stable and all the others (the zero solution and always pairs) are unstable. The deterministic Allen-Cahn equation is a gradient system, that is no state visited once can be visited again, since the "energy" along the trajectory $t \mapsto \mathcal{V}(u(t;x))$ decreases excluding loops. The stable states $\phi^\pm$ are the only local minima. All other critical states are saddle points.

The global picture of the state space: The state space H is divided completely into two open domains of attraction of the stable states ϕ^+ and ϕ^- separated by a smooth unbounded manifold without borders of codimension 1, which is the complement of the domains of attraction. While the stable states are trivially inside the respective domains of attraction, the unstable states all sit on the separating manifold.

The Morse-Smale property: The Allen-Cahn equation has generically the Morse-Smale property, that is, excluding the values $\lambda \neq (\pi n)^2$, $n \in \mathbb{N}$ all critical points have linearizations with eigenvalues whose real parts are zero, that is there are no purely imaginary eigenvalues. In addition, the eigenvectors of the eigenspaces with positive real part and negative part of the linearization linearly independent in H.

The local dynamics close to a stable state: The dynamics close (that is in a small ball of radius $\delta > 0$, say) to the stable state is topologically equivalent to the dynamics of the linearization with only negative eigenvalues. Hence we encounter an exponential attraction to the stable state with an exponential rate given by the largest negative eigenvalue of the linearization.

Therefore we obtain the following upper bound starting in a ball of radius δ centered in the stable state but out side a ball of radius ε^γ for some $\varepsilon, \gamma > 0$ ($\delta > \varepsilon^\gamma$) to reach the ball $B_{\varepsilon^\gamma}(\phi^\pm)$. There is a constant $\kappa' > 0$ such that for all $x \in B_\delta(\phi^\pm) \setminus B_{\varepsilon^\gamma}(\phi^\pm)$ the condition $t \geqslant \kappa'\gamma|\ln(\varepsilon)|$ yields that $u(t;x) \in B_{\varepsilon^\gamma}(\phi^\pm)$.

The local dynamics close to an unstable state inside a d.o.a.: This dynamics is a bit more subtle. As a consequence of the Morse-Smale property, the linearization of any unstable state has exactly one eigenvector, which transversally points into the domain of attraction of the stable states (in one of the domain of attraction, while by symmetry its negative points to the opposed domain of attraction). More precisely, the projection of the unstable eigenvector in normal direction to the separating manifold is non zero.

Starting close to an unstable state (that is in a ball of fixed radius $\delta > 0$, say) but inside of a domain of attraction (that is not sitting the separating manifold) we obtain the repulsion of the solution $t \mapsto u(t;x)$ from the unstable state in the following sense.

Let $\bar{n}$ be the normal vector of the (smooth) separating manifold at an unstable state ϕ and $\Lambda^* > 0$ be the eigenvalue of the unstable eigenvector which is inward pointing to the domain of a stable attraction.

Starting inside the small ball $B_\delta(\phi)$ but outside a small ball $B_{\varepsilon^\delta}(\phi) \subset B_\delta(\phi)$ of radius ε^γ, $\varepsilon, \gamma > 0$ with $\delta > \varepsilon^\gamma$ we have that

$$\|u(t;x) - \phi\| \geqslant e^{\frac{\Lambda^*}{2}t} |\langle\langle x - \pi_{\mathcal{S}}(x), \bar{n}\rangle\rangle| \geqslant e^{\frac{\Lambda^*}{2}t}\varepsilon^\gamma.$$

Hence $\delta \geqslant e^{\frac{\Lambda^*}{2}t}\varepsilon^\gamma$ implies the existence of a constant $\kappa'' > 0$ such that for all $t \geqslant \kappa''\gamma|\ln(\varepsilon)|$ we have that $u(t;x) \notin B_\delta(\phi)$.

Note that we are interested in the exit from a domain of attraction of a stable state $D^\pm$. Taking into account the mentioned ingredients we will reduce $D^\pm$ slightly to all initial values $x \in D^\pm$ such that for times $t \geqslant 0$ we have $\operatorname{dist}(u(t;x), \mathcal{S}) > \varepsilon^\gamma$ for arbitrarily small $\varepsilon > 0$. This set will be denoted $D_1^\pm(\varepsilon^\gamma)$. By construction it is a subset of $D^\pm$ and it is positive in variant under the solution flow, that is $x \in D_1^\pm(\varepsilon^\gamma)$ implies $u(t;x) \in D_1^\pm(\varepsilon^\gamma)$ for all times $t \geqslant 0$. Its formal definition is given in Definition 3.4.9.
Now, we are in the position to sketch the following "worst case" trajectory $t \mapsto u(t;x)$.

Global relaxation to a bounded set: Starting in some point $x \in D_1^\pm(\varepsilon^\gamma)$ the solution enters the large ball $B_{r^*}(0) \cap D_1^\pm(\varepsilon^\gamma)$ in at most a time $\kappa > 0$ inside of which all critical states are found.

Traveling from ball to ball between the unstable states: We fix once and for all a sufficiently small radius $\delta > 0$ and consider all δ-balls centered in the (finitely many) critical states. Now we count the times the trajectory $u(t;x)$ needs at most to visit the maximal amount of these balls inside the positive invariant set $D_1^\pm(\varepsilon^\gamma)$.

Due to the fact that the solution operator of $x \mapsto u(t;x)$ is regularizing in the sense that bounded sets are sent to compact sets, there is a finite time $t_1(\delta)$ until u enters the δ-ball around the "first" unstable state. By our discussion of the local dynamics of close to unstable states u leaves this ball in at least time $\kappa'\gamma|\ln(\varepsilon)|$ but staying inside the positive invariant (reduced) domain of attraction $D_1^\pm(\varepsilon^\gamma)$.

There exists a uniform time $t_2(\delta)$ until u enters the "second" ball of radius δ around the another unstable state. This ball is left before another time $\kappa'\gamma|\ln(\varepsilon)|$ again staying inside the positive invariant (reduced) domain of attraction $D_1^\pm(\varepsilon^\gamma)$.

However, this procedure can be repeated only finitely many times since there are only finitely many unstable states. Hence we take the maximum of the "local" constants $t_2(\delta)$ and κ'. In addition, having a gradient system prevents that the solution "returns" to a previously visited unstable state.

Travel between the "last" unstable state to the stable state: Finally there is an absolute time $t_3(\delta)$ that u travels from any of the balls of radius δ around the unstable states to the δ-ball around the stable state of the respective domain of attraction.

Local convergence to the stable state: Having reached the δ-ball around the stable state $\phi^\pm$ we need a time at most $\kappa''\gamma|\ln(\varepsilon)|$ until reach a ball of ε^γ around the stable state.

Remark 3.4.8. *Summing up the upper bounds of the time we obtain the following. Let $m = |\mathcal{P}|$ then for $x \in D^\pm(\varepsilon^\gamma)$ we obtain the rather crude but qualitatively correct (with respect to its logarithmic growth for small ε) bound that*

$$t \geqslant \kappa + t_1(\delta) + m(t_2 + \kappa'\gamma|\ln(\varepsilon)|) + t_3 + \kappa''\gamma|\ln(\varepsilon)|$$

implies that $u(t;x) \in B_{\varepsilon^\gamma}(\phi^\pm)$. All the previous heuristics can be made rigorous via the rather technical construction of the stable and the unstable manifold. For the proofs we refer to Chapter 2 in [20].

Definition 3.4.9. *Define for $\delta_1 > 0$*

$$D_1^\pm(\delta_1) := \{x \in D^\pm \mid \bigcup_{t\geqslant 0} B_{\delta_1}(u(t;x)) \subseteq D^\pm\}. \tag{3.11}$$

In finite dimensions, where the dynamical system $t \mapsto u(t;x)$ also exists for negative times, these reductions have a much more natural shape, see Appendix in [36].
In the light of Remark 3.4.8 we obtain the following result.

Proposition 3.4.10. *For any generic choice of f such that (3.7) is Morse Smale there exist constants $\kappa_0, \kappa_1 > 0$ (depending only on f) which satisfy the following. For any monotonically growing function $\gamma_\cdot : (0,1] \to (0,1)$ with $\lim_{\varepsilon\to 0}\gamma_\varepsilon = 0$ there is a constant $\varepsilon_0 \in (0,1]$ such that for each $\varepsilon \in (0,\varepsilon_0]$ the conditions $t \geqslant \kappa_0 + \kappa_1|\ln\gamma_\varepsilon|$ and $x \in D_1^\pm(\gamma_\varepsilon)$ imply $\|u(t;x) - \phi^\pm\| < \frac{1}{4}\gamma_\varepsilon$.*

E) Further reduction of the domains of attraction: We have seen that in order to capture the main dynamics of u and steer clear for all times from sticky unstable fixed point it is appropriate to introduce the reduced domain of attraction $D_1^\pm(\delta_1)$.
In order to account also for small perturbations of u we introduce the following nested further reduced domains of attraction with positive invariance properties. For perturbations $\psi \in \mathbb{D} \cap L^\infty([0,\infty), H)$ and $x \in H$ we consider the

unique càdlàg mild solution v_ψ of the equation

$$\begin{cases} \dfrac{\partial v_\psi}{\partial t}(t,\zeta) = \dfrac{\partial^2 v_\psi}{\partial \zeta^2}(t,\zeta) + f(v_\psi(t,\zeta)+\psi(t,\zeta)) & \text{for } \ t \geqslant 0, \quad \zeta \in (0,1) \\ v_\psi(t,0) = v_\psi(t,1) = 0 & \text{Dirichlet b. c.} \\ v_\psi(0,\zeta) = x(\zeta) & \text{initial condition } x. \end{cases} \tag{3.12}$$

We define for $\delta_i > 0$, $i = 1, \ldots, 4$ and $\chi > 0$ sufficiently large the following reductions of $D^\pm$ (and $D_1^\pm(\delta_1)$):

$$\begin{aligned} D_1^\pm(\delta_1, \chi) &:= \{x \in D^\pm \mid \bigcup_{t\geqslant 0} B_{\delta_1}(u(t;x)) \subseteq D^\pm \cap \mathcal{U}^\chi\}, \\ D_2^\pm(\delta_1,\delta_2,\chi) &:= \{x \in D_1^\pm(\delta_1,\chi) \mid \ \forall\, \psi \in \mathbb{D}([0,\infty),H) \text{ with } \sup_{t\geqslant 0}\|\psi(t)\| \leqslant \delta_2 : \\ &\qquad \bigcup_{t\geqslant 0} B_{\delta_2}(v_\psi(t;x)) \subseteq D_1^\pm(\delta_1,\chi)\ \}, \\ D_3^\pm(\delta_1,\delta_2,\delta_3,\chi) &:= \{x \in D_2^\pm(\delta_1,\delta_2,\chi) \mid \text{for all } \psi \in \mathbb{D}([0,\infty),H) \\ &\qquad \text{with } \sup_{t\geqslant 0}\|\psi(t)\| \leqslant \delta_3 : \ \bigcup_{t\geqslant 0} B_{\delta_3}(v_\psi(t;x)) \subseteq D_2^\pm(\delta_1,\delta_2,\chi)\ \}, \\ D_4^\pm(\delta_1,\delta_2,\delta_3,\delta_4,\chi) &:= \{x \in D_3^\pm(\delta_1,\delta_2,\delta_3,\chi) \mid v \in B_{\delta_4}(x) \in D_3^\pm(\delta_1,\delta_2,\delta_3,\chi)\ \}. \end{aligned} \tag{3.13}$$

The reduced domains of attractions are nested by construction and for all $\delta \in (0,1)$ we have $D^\pm = \bigcup_{\delta\in(0,1]} D_3^\pm(\delta,\delta,\delta,\delta)$ (cf. [19]). For any $\delta_0 \in (0,1]$ such that $\phi^\pm \in D^\pm(\delta_0)$ and $\delta \in (0,\delta_0]$ we have the positive invariance of the reduced domains of attraction under the dynamical system (3.7). For convenience we set $D_2^\pm(\delta) := D_2^\pm(\delta,\delta)$, $D_3^\pm(\delta) := D_3^\pm(\delta,\delta,\delta)$ and $D_4^\pm(\delta)$ analogously.
These properties are shown in [19] Chapter 2.

EXERCISE 3.4.11 (Properties of the associated scalar system)**.** *We study for a parameter $\lambda > 0$ the ordinary differential equation*

$$\frac{du}{dt} = -\lambda(u^3 - u), \qquad u(0) = x \in \mathbb{R}.$$

1) *What can you say about existence and uniqueness of the solution $t \mapsto u(t;x)$?*

2) *Calculate the stationary states and determine its stability.*

3) *Calculate a lower bound κ_0 such that $t \geqslant \kappa_0$ implies for any $x \in \mathbb{R}$ that $u(t;x) \in B_2(0)$.*

4) What can you say about the role of the parameter $\lambda > 0$*?*

5) Consider a stable state ϕ *of the previous system. Consider* $0 < \varepsilon < \delta$ *for* $\delta < \frac{1}{10}$*, say. Compare the original system to the linearized system and calculate a lower bound* $\kappa > 0$ *such that* $t \geqslant \kappa$ *and* $x \in B_\delta(\phi)$ *implies* $u(t;x) \in B_\varepsilon(\phi)$*. How does that constant grow in terms of* ε*?*

Chapter 4

The Allen-Cahn equation perturbed by additive pure jump Lévy noise

4.1 The stochastic convolution

Given a filtered probability space $\mathbf{\Omega} = (\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geqslant 0})$ satisfying the usual conditions in the sense of Protter [57] let $L = (L(t))_{t\geqslant 0}$ be a càdlàg version of a pure jump Lévy process in $(H, \mathcal{B}(H))$, that is the characteristic triplet satisfies $b = 0$ and $Q = 0$ and $0 \neq \nu \in \mathcal{M}_0(H)$. By the Lévy-Itô decomposition given in Theorem 2.3.12 we have seen that there exists an associated Poisson random measure N on $\mathbf{\Omega}$ such that

$$L(t) = \int_0^t \int_{\|z\|\leqslant 1} z\widetilde{N}(ds,dz) + \int_0^t \int_{\|z\|>1} zN(ds,dz), \qquad \mathbb{P}-\text{a.s. for all } t \geqslant 0. \tag{4.1}$$

The stochastic convolution w.r.t. the square-integrable martingale part of (4.1): We denote the first term of $L(t)$ by $\widetilde{L}(t)$ and consider the linear equation

$$dX(t) = \frac{\partial^2}{\partial \zeta^2} X(t)dt + \varepsilon d\widetilde{L}(t), \qquad X(0) = x \in H. \tag{4.2}$$

Theorem 4.1.1. *The stochastic differential equation (4.2) has a unique mild* *solution satisfying*

$$\widetilde{X}(t) = S(t)x + \int_0^t \int_{\|z\|\leqslant 1} S(t-s)\varepsilon z \widetilde{N}(ds,dz), \qquad \mathbb{P}-a.s.\ t \geqslant 0, \tag{4.3}$$

where S is the heat semigroup given in Section 3.2. In addition, the property (3.6) implies that the process $\widetilde{X}$ has a càdlàg version in H. The transition probabilities of the process $\widetilde{X}$ define a Feller family and imply the strong Markov property in H w.r.t. the enhanced natural filtration $(\mathcal{F}_t)_{t\geqslant 0}$. That is for any bounded continuous function $f: H \to \mathbb{R}$ and any $(\mathcal{F}_t)_{t\geqslant 0}$-stopping times $\sigma \leqslant \tau$ $\mathbb{P}$-a.s. we have

$$\mathbb{E}\Big[f(\widetilde{X}(\tau)) \mid \mathcal{F}_\sigma\Big] = \mathbb{E}\Big[f(\widetilde{X}(\tau)) \mid X(\sigma)\Big], \qquad s \leqslant t, \mathbb{P}-\ a.s.$$

The interlacing of the compound Poisson part of (4.1): Note that the second term, which we will denote by η in (4.1) $\eta_t := \int_0^t \int_{\|z\|>1} zN(ds,dz)$ is a compound Poisson process, that is we know

$$\eta_t = \sum_{k=1}^{\infty} W_k \mathbf{1}\{t_1 + \cdots + t_k < t\}$$

for an i.i.d. family $(W_k)_{k\in\mathbb{N}}$ of random vectors with distribution $\frac{\nu(\cdot \cap B_1^c(0))}{\nu(B_1^c(0))}$ and $(t_k)_{k\in\mathbb{N}}$ the i.i.d. family of waiting times between the jumps with distribution $\mathrm{Exp}_{\nu(B_1^c(0))}$. Then the solution of

$$dX(t) = \frac{\partial^2}{\partial\zeta^2} X(t)dt + \varepsilon dL(t), \qquad X(0) = x \in H. \tag{4.4}$$

is constructed as follows. Set the arrival times $T_i := t_1 + \cdots + t_i$.

- On $[0, T_1)$ we have $X(t) = \widetilde{X}(t) = \widetilde{X}^1(t)$ that is it follows the dynamic by (4.2).
- At the moment $t = T_1$ we set $X(T_1) = \widetilde{X}^1(T_1) + W_1$.
- On (T_1, T_2), the system starts in $X(T_1)$ evolves according to the dynamics in (4.2) however, with respect to the noise $L_{t+T_1} - L_{T_1}$. We denote this process by $\widetilde{X}^2$.
- At the moment $t = T_2$ we set $X(T_2) = \widetilde{X}^2_{T_2} + W_2$
- On (T_1, T_2), the system starts in $X(T_2)$ evolves according to the dynamics in (4.2) however, with respect to the noise $L_{t+T_2} - L_{T_2}$. We denote this process by $\widetilde{X}^3$.

- At the moment $t = T_3$ we set $X(T_3) = \widetilde{X}^3_{T_3} + W_3$
- etc.

Remark 4.1.2. *Since the processes $\widetilde{L}$ and η are independent, the strong Markov property and the càdlàg property remain untouched.*

An estimate of the stochastic Lévy convolution: In the spirit of the inequality by Siorpaez [62] there are path-by-path estimates of the stochastic convolution (4.3) given recently by Salavati and Zangeneh in Theorem 6 of [64].

Theorem 4.1.3 (Salavati and Zangeneh). *Fix $\rho > 0$ and let $L = (L_t)_{t\geqslant 0}$ be the H-valued càdlàg version of a Lévy process given by*

$$L_t = \int_0^t \int_{|z|\leqslant\rho} z\widetilde{N}(ds, dz).$$

Under the previous assumptions on S and H we denote the process

$$X(t) = S(t)x + \int_0^t \int_{\|z\|\leqslant\rho} S(t-s)z\widetilde{N}(ds, dz), \qquad \mathbb{P} - a.s.\ t \geqslant 0. \tag{4.5}$$

Then we have $\mathbb{P}$-a.s. for all $t \geqslant 0$

$$\|X(t)\|^2 \leqslant e^{-2\Lambda_0 t}\|x\|^2 + 2\int_0^t \int_{\|z\|\leqslant\rho} e^{-2\Lambda_0(t-s)}\langle\langle X_{s-}, z\rangle\rangle\widetilde{N}(ds, dz)$$
$$+ \int_0^t \int_{\|z\|\leqslant\rho} e^{-2\Lambda_0}\|z\|^2 N(ds, dz).$$

4.2 The solution of the stochastic reaction diffusion equation

Given a filtered probability space $\mathbf{\Omega} = (\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geqslant 0})$ satisfying the usual conditions in the sense of Protter [57] let $L = (L(t))_{t\geqslant 0}$ be a càdlàg version of a pure jump Lévy process in $(H, \mathcal{B}(H))$.
Consider the formal stochastic reaction diffusion equation for $t > 0, x \in H, \zeta \in J$ and $\varepsilon \in (0, 1]$

$$\begin{aligned} dX^\varepsilon(t,\zeta) &= \big(\frac{\partial^2}{\partial\zeta^2}X^\varepsilon(t,\zeta) + f(X^\varepsilon(t,\zeta))\big)dt + \varepsilon dL(t,\zeta)) \\ &\text{with} \quad X^\varepsilon(t,0) = X^\varepsilon(t,1) = 0 \quad \text{and} \quad X^\varepsilon(0,\zeta) = x(\zeta), \end{aligned} \tag{4.6}$$

in the sense that

$$dL(t,\zeta) = \int_{|z|\leqslant 1} z(\zeta))\widetilde{N}(dtdz) + \int_{|z|>1} z(\zeta)N(dtdz).$$

Proposition 4.2.1. *Then for any mean zero càdlàg $L^2(\mathbb{P};H)$-martingale $\xi = (\xi(t))_{t\geqslant 0}$, $T > 0$, and initial value $x \in H$ equation (4.6) driven by $\varepsilon d\xi$ instead of εdL has a unique càdlàg mild solution $(Y^\varepsilon(t;x))_{t\in[0,T]}$. The transition kernels of the solution process Y^ε induce a homogeneous Markov family satisfying the Feller property and hence the strong Markov property.*

The proof relies on the local Lipschitz continuity of $f : H \to H$. A proof for dissipative polynomials f is given in [58], Chapter 10, and for Allen-Cahn in [19]. By interlacing of large jumps this notion of solution is extended to the heavy-tailed process L.

Corollary 4.2.2. *For $x \in H$ equation (4.6) has a global càdlàg mild solution $(X^\varepsilon(t;x))_{t\geqslant 0}$, which satisfies the strong Markov property.*

Part III

The First Exit Problem

Chapter 5

The specific hypotheses and the main results

5.1 The asymptotic first exit problem

The main object of study: For $\gamma, \varepsilon \in (0,1]$ and $\chi > 0$ sufficiently large, $x \in D_2^{\pm}(\varepsilon^\gamma, \chi)$ and the càdlàg mild solution $(X^\varepsilon(t;x))_{t\geqslant 0}$ of (4.6) we define the **first exit time from the reduced domain of attraction** $D_2^{\pm}(\varepsilon^\gamma, \chi)$

$$\tau_x^{\pm}(\varepsilon) := \inf\{t > 0 \mid X^\varepsilon(t;x) \notin D_2^{\pm}(\varepsilon^\gamma, \chi)\}.$$

Specific Hypotheses:

(D) *The function f satisfies $\lambda \neq (\pi n)^2$ and $\lambda > \pi^2$.*

(S.1) *The Lévy measure $\nu \in \mathcal{M}_0(H)$ is regularly varying with index $-\alpha$, $\alpha > 0$ and limit measure $\mu \in \mathcal{M}_0(H)$.* We denote h the regularly varying function $h(r) = r^{-\alpha}\ell(r)$ such that

$$\lim_{r\to\infty} \frac{\nu(rB_1^c(0))}{h(r)} = \mu(B_1^c(0)).$$

We refer to Subsection 1.2.5 for further explanations.

The limit measure of relevant increments: We define the set of **essential increment vectors** $z \in H$ sending $x \in H$ to the set $U \in \mathcal{B}(H)$ as

$$\mathcal{J}^U(x) := \{z \in H \mid x + z \in U\}, \qquad x \in H,$$

and the **asymptotic essential increment measure**

$$m^{\pm}(A) := \mu(\mathcal{J}^{A}(\phi^{\pm})) \qquad \text{for } A \in \mathcal{B}(H) \text{ with } 0 \notin \bar{A}$$

and

$$h_\varepsilon := h(\frac{1}{\varepsilon}) \text{ for } \varepsilon \in (0,1].$$

We further assume the joint non-degeneracy of μ and the dynamical system.

(S.2) *For all $\phi^{\pm} \in \mathcal{P}^-$ we have $m^{\pm}((D^{\pm})^{\mathsf{c}}) > 0$.*

(S.3) *For all $\phi^{\pm} \in \mathcal{P}^-$ and $\eta > 0$ there are $\delta, \chi > 0$ such that*

$$m^{\pm}(H \setminus (D_4^{+}(\delta,\chi) \cup D_4^{-}(\delta,\chi)) < \eta.$$

We define the **characteristic exit rate** $\lambda_\varepsilon^{\pm}$ of system (4.6) from $D^{\pm}$ by

$$\lambda_\varepsilon^{\pm} := \nu\Big(\frac{1}{\varepsilon}\mathcal{J}^{(D^{\pm})^{\mathsf{c}}}(\phi^{\pm})\Big), \qquad \varepsilon \in (0,1]. \tag{5.1}$$

Then (S.1) implies $\frac{\lambda_\varepsilon^{\pm}}{h_\varepsilon} = \frac{\lambda_\varepsilon^{\pm}}{\varepsilon^{\alpha}\ell(\frac{1}{\varepsilon})} \stackrel{\varepsilon\to 0+}{\longrightarrow} m^{\pm}((D^{\pm})^{\mathsf{c}})$.

The exit time result: The main exit time result theorem reads as follows.

Theorem 5.1.1. *Let Hypotheses (D) and (S.1-3) be satisfied. Then there is a* EXP(1)*-distributed family of random variables $(s^{\pm}(\varepsilon))_{\varepsilon\in(0,1]}$ on Ω satisfying the following. For any $C > 0$ and $\theta \in (0,1)$ there are $\chi > 0$ and $\varepsilon_0, \gamma \in (0,1]$ such that $\varepsilon \in (0,\varepsilon_0]$ implies*

$$\sup_{x\in D_3^{\pm}(\varepsilon^\gamma,\chi)} \mathbb{E}\Big[e^{\theta|\lambda_\varepsilon^{\pm}\tau_x^{\pm}(\varepsilon,\chi) - s^{\pm}(\varepsilon)|}\Big] \leqslant 1 + C.$$

As a consequence, we have the convergence of all moments

$$\lim_{\varepsilon\to 0} \sup_{x\in D_3^{\pm}(\varepsilon^\gamma,\chi)} \mathbb{E}[|\lambda_\varepsilon^{\pm}\tau_x^{\pm}(\varepsilon,\chi)|^n] \in [n! - C, n! + C]$$

and the following polynomial behavior

$$\begin{aligned}\sup_{x\in D_3^{\pm}(\varepsilon^\gamma,\chi)} \mathbb{E}\Big[\tau_x^{\pm}(\varepsilon,\chi)\Big] &\in [\frac{1-C}{\lambda_\varepsilon^{\pm}}, \frac{1+C}{\lambda_\varepsilon^{\pm}}] \\ &\subseteq [\frac{1-2C}{\varepsilon^{\alpha}\ell(\frac{1}{\varepsilon})m^{\pm}((D^{\pm})^{\mathsf{c}})}, \frac{1+2C}{\varepsilon^{\alpha}\ell(\frac{1}{\varepsilon})m^{\pm}((D^{\pm})^{\mathsf{c}})}].\end{aligned}$$

The supremum in the previous expressions can be changed to the infimum.

In terms of [10] the memorylessness of $s(\varepsilon)$ describes the “unpredictability” of the exit times with a “polynomial” loss of memory as opposed to a Gaussian “exponential” loss of memory.

The flow decomposition: For the statement of the main result about the exit loci we write $\Delta_t L := L(t) - L(t-)$, $t \geqslant 0$ for L given as the pure jump Lévy process in Section 4.2 and (S.1). For $\rho \in (0,1)$ and $\varepsilon \in (0,1]$ we construct large jump times of L by

$$T_0(\varepsilon) := 0, \qquad T_k(\varepsilon) := \inf\left\{t > T_{k-1}(\varepsilon) \;\middle|\; \|\Delta_t L\| > \varepsilon^{-\rho}\right\}, \quad k \geqslant 1, \tag{5.2}$$

and large jump increments by $W_k(\varepsilon) := \Delta_{T_k(\varepsilon)} L$, $k \in \mathbb{N}$. The family $(W_k(\varepsilon))_{k\in\mathbb{N}}$ is i.i.d. with

$$\mathbb{P}(\varepsilon W_k \in A) = \frac{\nu(A \cap B^{\mathsf{c}}_{\varepsilon^{-\rho}}(0))}{\nu(B^{\mathsf{c}}_{\varepsilon^{-\rho}}(0))}, \quad \text{where } \beta_\varepsilon := \nu(B^{\mathsf{c}}_{\varepsilon^{-\rho}}(0))$$

$$\text{behaves as } \frac{\beta_\varepsilon}{(\rho^\varepsilon)^{-\alpha}\ell(\rho^\varepsilon)} \stackrel{\varepsilon\to 0+}{\longrightarrow} m^{\pm}(B^{\mathsf{c}}_1(0)).$$

The exit locus result:

Theorem 5.1.2. *Let Hypotheses (D) and (S.1-3) be satisfied. Then there is a family of random variables $(k^*(\varepsilon))_{\varepsilon\in(0,1]}$ on Ω with $k^*(\varepsilon)$ being $\mathrm{GEO}(\frac{\lambda^{\pm}_\varepsilon}{\beta_\varepsilon})$ distributed and satisfying the following. For any $C > 0$ and $0 < p < \alpha$ there are $\chi > 0$ and $\gamma, \rho, \varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0, \varepsilon_0]$ implies*

$$\sup_{x\in D_3^{\pm}(\varepsilon^\gamma,\chi)} \mathbb{E}\Big[\|X^\varepsilon(\tau_x^{\pm}(\varepsilon,\chi);x) - (\phi^{\pm} + \varepsilon W_{k^*(\varepsilon)})\|^p\Big] \leqslant C$$

and in particular for any $U \in \mathcal{B}(H)$ with $m^{\pm}(U) > 0$ and $m^{\pm}(\partial U) = 0$ we have

$$\sup_{x\in D_3^{\pm}(\varepsilon^\gamma,\chi)} \left|\mathbb{P}(X^\varepsilon(\tau^{\pm}(\varepsilon,\chi);x) \in A) - \frac{m^{\pm}(A \cap (D^{\pm})^c)}{m^{\pm}((D^{\pm})^c)}\right| \leqslant C.$$

The second result follows from the first one since $\lim_{\varepsilon\to 0}\mathbb{P}(\varepsilon W_{k^*(\varepsilon)}(\varepsilon) \in A) = \frac{\mu(A\cap(D^{\pm})^c)}{\mu((D^{\pm})^c)}$ for any $A \in \mathcal{B}(H)$. In fact we will only prove in the first result the convergence in probability. In order to infer convergence in L^p, $p \in (0,\alpha)$ it is only necessary to prove uniform integrability. This is carried out in the more general setting of multiplicative noise in [35].

5.2 Understanding the models of the exit times and exit loci

We now construct on $(\Omega, \mathcal{A}, \mathbb{P})$ the random variables $(s^{\pm}(\varepsilon))_{\varepsilon\in(0,1]}$ of Theorem 5.1.1 and $(k^*(\varepsilon))_{\varepsilon\in(0,1]}$ of Theorem 5.1.2.

Definition 5.2.1. *For given scales $\rho^{\cdot}$ and $\gamma_{\cdot}$ in (6.1), $B_j^{\diamond}(\varepsilon) := \{\varepsilon W_j \in \mathcal{J}^{(D^{\pm})^c}(\phi^{\pm})\}$ and the arrival times T_k of W_k given in (6.2) we define*

$$s^{\pm}(\varepsilon) := \sum_{k=1}^{\infty} T_k \prod_{j=1}^{k-1} (1 - \mathbf{1}(B_j^{\diamond}))\mathbf{1}(B_k^{\diamond}), \qquad \varepsilon \in (0,1],$$

and

$$k^*(\varepsilon) := \sum_{k=1}^{\infty} (k-1) \prod_{j=1}^{k-1} (1 - \mathbf{1}(B_j^{\diamond}))\mathbf{1}(B_k^{\diamond}).$$

Lemma 5.2.2. *For given scale $\rho^{\cdot}$ in (6.1) the random variables $s^{\pm}(\varepsilon)$ is exponentially distributed with rate $\lambda_{\varepsilon}^{\pm}$ and any of the random variables $k^*(\varepsilon)$ is geometrically distributed with rate $\mathbb{P}(B^{\diamond}) = \frac{\lambda_{\varepsilon}^{\pm}}{\beta_{\varepsilon}}$. In particular $\bar{s}^{\pm}(\varepsilon) := \lambda_{\varepsilon}^{\pm} s^{\pm}(\varepsilon)$ is exponentially distributed with rate 1.*

Proof. Since the family $(W_k)_{k \in \mathbb{N}}$ is i.i.d. and $B_k^{\diamond} = \{\varepsilon W_k \in (D^{\pm})\}$ we have that $k^*(\varepsilon)$ is geometrically distributed by construction with rate $\mathbb{P}(B^{\diamond}) = \frac{\lambda_{\varepsilon}^{\pm}}{\beta_{\varepsilon}}$. Let $\theta > 0$. We calculate the Laplace transform of $s^{\pm}(\varepsilon)$

$$\begin{aligned}
\mathbb{E}\left[e^{-\theta s^{\pm}(\varepsilon)}\right] &= \mathbb{E}\left[e^{-\theta \sum_{k=1}^{\infty} T_k \prod_{j=1}^{k-1}(1-\mathbf{1}(B_j^{\diamond}))\mathbf{1}(B_k^{\diamond})}\right] \\
&= \mathbb{E}\left[\prod_{k=1}^{\infty} e^{-\theta T_k \prod_{j=1}^{k-1}(1-\mathbf{1}(B_j^{\diamond}))\mathbf{1}(B_k^{\diamond})}\right] \\
&= \sum_{k=1}^{\infty} \mathbb{E}\left[e^{-\theta T_k} \prod_{j=1}^{k-1} (1 - \mathbf{1}(B_j^{\diamond}))\mathbf{1}(B_k^{\diamond})\right] \\
&= \sum_{k=1}^{\infty} \mathbb{E}\left[\prod_{j=1}^{k-1} e^{-\theta t_j}(1 - \mathbf{1}(B_j^{\diamond}))e^{-\theta t_k}\mathbf{1}(B_k^{\diamond})\right].
\end{aligned}$$

Exploiting the independence of $(W_k)_{k \in \mathbb{N}}$ and $(T_k)_{k \in \mathbb{N}}$ as well as the stationarity of $(W_k)_{k\in\mathbb{N}}$ each summand takes the form

$$\begin{aligned}
&\mathbb{E}\left[\prod_{j=1}^{k-1} e^{-\theta t_j}(1-\mathbf{1}(B_j^\diamond))e^{-\theta t_k}\mathbf{1}(B_k^\diamond)\right] \\
&= \prod_{j=1}^{k-1} \mathbb{E}\left[e^{-\theta t_j}(1-\mathbf{1}(B_j^\diamond))\right]\mathbb{E}\left[e^{-\theta t_k}\mathbf{1}(B_k^\diamond)\right] \\
&= \left(\mathbb{E}\left[e^{-\theta t_1}\right](1-\mathbb{P}(B_1^\diamond))\right)^{k-1}\mathbb{E}\left[e^{-\theta t_1}\right]\mathbb{P}(B_1^\diamond) \\
&= \left(\frac{\beta_\varepsilon}{\theta+\beta_\varepsilon}(1-\frac{\lambda_\varepsilon^\pm}{\beta_\varepsilon})\right)^{k-1}\frac{\beta_\varepsilon}{\theta+\beta_\varepsilon}\frac{\lambda_\varepsilon^\pm}{\beta_\varepsilon}.
\end{aligned}$$

Finally we conclude

$$\begin{aligned}
\mathbb{E}\left[e^{-\theta s^\pm(\varepsilon)}\right] &= \sum_{k=1}^{\infty}\left(\frac{\beta_\varepsilon}{\theta+\beta_\varepsilon}(1-\frac{\lambda_\varepsilon^\pm}{\beta_\varepsilon})\right)^{k-1}\frac{\beta_\varepsilon}{\theta+\beta_\varepsilon}\frac{\lambda_\varepsilon^\pm}{\beta_\varepsilon} \\
&= \frac{\beta_\varepsilon}{\theta+\beta_\varepsilon}\frac{\lambda_\varepsilon^\pm}{\beta_\varepsilon}\frac{1}{1-\frac{\beta_\varepsilon}{\theta+\beta_\varepsilon}(1-\frac{\lambda_\varepsilon^\pm}{\beta_\varepsilon})} = \frac{\lambda_\varepsilon^\pm}{\beta_\varepsilon}\frac{1}{\frac{\theta+\beta_\varepsilon}{\beta_\varepsilon}-(1-\frac{\lambda^\pm(\varepsilon)}{\beta_\varepsilon})} \\
&= \frac{\lambda_\varepsilon^\pm}{\theta+\lambda_\varepsilon^\pm} = \widehat{\mathrm{EXP}(\lambda_\varepsilon)}(\theta).
\end{aligned}$$

□

Chapter 6

Deviations of the small noise solution

This section is devoted to a large deviations type estimate for the stochastic convolution on the time scale of the waiting time between consecutive large jumps $[T_i, T_{i+1})$. Due to the strong Markov property it turns out that it is sufficient to treat the time interval $[0, T_1)$. In this section we quantify the fact, that in the time interval strictly between two adjacent large jumps the solution of (4.6) is perturbed by only the small noise component and deviates from the solution of the deterministic equation by only a small ε-dependent quantity, with probability converging to 1 exponentially in the small noise limit $\varepsilon \to 0$.

Time scales: For the correct understanding of the role of the different scales we should avoid cancelations and therefore formulate and prove our results for the following abstract scale functions

$$\begin{aligned}
&\rho^{\cdot} : (0,1] \to [1,\infty), \qquad \lim_{\varepsilon\to 0+} \rho^{\varepsilon} = \infty, \qquad \lim_{\varepsilon\to 0+} \varepsilon\rho^{\varepsilon} = 0 \\
&\gamma_{\cdot} : (0,1] \to (0,1], \qquad \lim_{\varepsilon\to 0+} \gamma_{\varepsilon} = 0+ \\
&T^{\cdot} : (0,1] \to [1,\infty), \qquad \lim_{\varepsilon\to 0+} T^{\varepsilon} = \infty.
\end{aligned} \tag{6.1}$$

before choosing them numerically in Proposition 6.2.1. We assume that all expressions appearing in (6.1) are monotonic.

In this section we shall use the language of Poisson random measures introduced in Section 2.3.

The flow decomposition of L: We recall the following notation for the abstract scales of (6.1)

$$T_0 := 0, \quad T_k := \inf\left\{t > T_{k-1} \;\middle|\; \|\Delta_t L\| > \rho^\varepsilon\right\}, \quad W_k := \Delta_{T_k} L, \quad k \geqslant 1. \tag{6.2}$$

We define the compound Poisson η^ε process of L which consists only of large jumps with intensity $\beta_\varepsilon := \nu\left(B^{\mathsf{c}}_{\rho^\varepsilon}(0)\right)$ and the jump probability measure by $\mathbb{P}(W_k \in A) = \nu(A \cap B^{\mathsf{c}}_{\rho^\varepsilon}(0))/\beta_\varepsilon$, in other words

$$\eta^\varepsilon_t := \int_0^t \int_{\|z\|>\rho^\varepsilon} zN(ds,dz) = \sum_{k=1}^\infty W_k \mathbf{1}_{\{T_k<t\}}, \quad t \geqslant 0,$$

The complementary small jumps process $\xi^\varepsilon := L - \eta^\varepsilon$ has the following shape

$$\begin{aligned}\xi^\varepsilon_t &= \int_0^t \int_{\|z\|\leqslant 1} z\widetilde{N}(ds,dz) + \int_0^t \int_{1<\|z\|\leqslant\rho^\varepsilon} zN(ds,dz)\\ &= \int_0^t \int_{\|z\|\leqslant\rho^\varepsilon} z\widetilde{N}(ds,dz) + \int_0^t \int_{1<\|z\|\leqslant\rho^\varepsilon} z\nu(dz)ds.\end{aligned}$$

Due to its uniformly bounded jump size the process ξ^ε has exponential moments for any $\varepsilon \in (0,1]$ and $\xi^\varepsilon - t\int_{1<\|z\|\leqslant\rho^\varepsilon} z\nu(dz)$ is a mean zero $\mathcal{F}$-martingale in H.
The i.i.d. family of $\mathrm{EXP}(\beta_\varepsilon)$-distributed waiting times between successive large jumps of η^ε_t is given by $t_0 = 0$ and $t_k := T_k - T_{k-1}$, for $k \geqslant 1$. Denote by the process L between the waiting times $\xi^{\varepsilon,k}(t) := L_{t+T_{k-1}} - L_{T_{k-1}}$ for $t \in [0,t_k)$. In particular, the i.i.d. families $(t_k)_{k\in\mathbb{N}}$, $(W_k)_{k\in\mathbb{N}}$, $(\xi^{\varepsilon,k}(t))_{t\in[0,t_k),k\in\mathbb{N}}$ are independent.

The small jumps solution before the first large jump T_1: Denote for $\varepsilon \in (0,1]$, $y \in H$, $t \geqslant 0$ and $\zeta \in J$ by Y^ε the mild solution of

$$\begin{cases} dY^\varepsilon(t,\zeta) = \left(\frac{\partial^2 Y^\varepsilon}{\partial\zeta^2}(t,\zeta) + f(Y^\varepsilon(t,\zeta))\right)dt + \varepsilon d\xi^\varepsilon(t,\zeta) & t \geqslant 0, \zeta \in (0,1)\\ Y^\varepsilon(t,0;y) = Y^\varepsilon(t,1;y) = 0 & \text{Dirichlet b.c.}\\ Y^\varepsilon(0,\zeta;y) = y(\zeta) & \text{initial value } y \in H.\end{cases} \tag{6.3}$$

In the following two subsections we derive all results on the stochastic convolution w.r.t. to ξ^ε up to the hitting time of Y^ε leaving a large ball. We shall get rid of that artificial time horizon in the proof of Proposition 6.2.3 showing that the process Y^ε at time σ is inside the large ball on the event of small noise convolution. For $\chi > 0$, $\varepsilon \in (0,1]$ and $y \in D_2^{\pm}(\gamma_\varepsilon,\chi)$ set

$$\sigma^1 := \sigma^1_{\chi,y}(\varepsilon) := \inf\{t > 0 \mid Y^\varepsilon(t;y) \notin \mathcal{U}^\chi\}. \tag{6.4}$$

6.1 Exponential estimate of a stochastic convolution with bounded jumps

In this subsection we show that on a time interval $[0, T^\varepsilon]$, $T^\varepsilon \nearrow \infty, \varepsilon \searrow 0$, the stochastic convolution with respect to $\varepsilon d\xi^\varepsilon(s)$ is very small (i.e. of order $\leqslant \gamma_\varepsilon^q$ for some $q \geqslant 1$) with exponential probability in terms of γ_ε tending to 1 as long as Y^ε and the stochastic convolution remain inside the sublevel set $\mathcal{U}^\chi$ containting a (large) ball of radius $\chi > 0$.

The stochastic small noise convolution: For the solution Y^ε of (6.3) with $y \in D_2^\pm(\gamma_\varepsilon, \chi)$ we consider the stochastic convolution process

$$\begin{aligned}\Psi_t^\varepsilon &:= \int_0^t S(t-s)d\big(\varepsilon\xi^\varepsilon(s)\big)\\ &= \int_0^t \int_{0<\|z\|\leqslant\rho^\varepsilon} S(t-s)\varepsilon z\widetilde{N}(ds,dz) + \int_0^t \int_{1<\|z\|\leqslant\rho^\varepsilon} S(t-s)\varepsilon z\nu(dz)ds\\ &=: \widetilde{\Psi}_t^\varepsilon + b_t^\varepsilon.\end{aligned}$$

The process $(\Psi_t^\varepsilon)_{t\geqslant 0}$ is a $\mathcal{F}$-adapted càdlàg process. For $\chi > 0$, $\varepsilon \in (0,1]$ and $y \in D_2^\pm(\gamma_\varepsilon, \chi)$ set

$$\sigma^2 := \sigma_\chi^2(\varepsilon) := \inf\{t > 0 \mid \Psi_t^\varepsilon \notin \mathcal{U}^\chi\} \qquad \text{and} \qquad \sigma := \sigma^1 \wedge \sigma^2. \tag{6.5}$$

Proposition 6.1.1. *Let the Hypotheses (D) and (S.1-3) be satisfied. The functions $\rho^\cdot, \gamma_\cdot$ and $T^\cdot$ given by (6.1) satisfy for some $q \geqslant 1$ the limit relation*

$$\lim_{\varepsilon\to 0}\Gamma(\varepsilon) = 0, \qquad \textit{where} \qquad \Gamma(\varepsilon) := \frac{\varepsilon\rho^\varepsilon}{\gamma_\varepsilon^{4q+3}}T^\varepsilon \textit{ is monotonic }. \tag{6.6}$$

Then there are $\chi > 0$ and $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0, \varepsilon_0]$ implies

$$\mathbb{P}(\sup_{s\in[0,\sigma\wedge T^\varepsilon]} \|\Psi_s^\varepsilon\| > \gamma_\varepsilon^q) \leqslant \exp(-(5\gamma_\varepsilon)^{-1}). \tag{6.7}$$

Proof. The plan of the proof is as follows. First we get rid of the drift b^ε (Step 0). In order to control $\widetilde{\Psi}^\varepsilon$ we start with the exponential Kolmogorov inequality where we introduce the free parameters λ and c. We estimate the stochastic convolution by a result of Salavati and Zangeneh and derive an exponential version of the Burkholder-Davis-Gundy inequality (Step 1). Then we optimize over the free parameters and use the Campbell representation of the Laplace transform of the quadratic variation of Poisson random integrals and a Campbell type estimate given in Lemma 6.1.2. This allows for a comparison principle for the characteristic exponent of Ψ^ε (Step 2) and allows to conclude (Step 3).

Step 0: Drift estimate. We show there are $\chi > 0$ and $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0, \varepsilon_0]$ implies

$$\|b^\varepsilon_{\sigma\wedge T^\varepsilon}\| < \frac{1}{2}\gamma_\varepsilon^q.$$

The triangular inequality and the norm estimate of the heat semigroup S yield

$$\begin{aligned}
\|b^\varepsilon_{\sigma\wedge T^\varepsilon}\| &\leqslant \| \int_0^{\sigma\wedge T^\varepsilon} \int_{1\leqslant\|z\|\leqslant\rho^\varepsilon} S(\sigma\wedge T^\varepsilon - s)\varepsilon z \nu(dz)ds\| \\
&\leqslant \int_0^{\sigma\wedge T^\varepsilon} \int_{1\leqslant\|z\|\leqslant\rho^\varepsilon} \|S(\sigma\wedge T^\varepsilon - s)\varepsilon z\|\nu(dz)ds \\
&\leqslant \int_0^{\sigma\wedge T^\varepsilon} \int_{1\leqslant\|z\|\leqslant\rho^\varepsilon} e^{-\Lambda_0(\sigma\wedge T^\varepsilon - s)}\|\varepsilon z\|\nu(dz)ds \\
&= \varepsilon \int_0^{\sigma\wedge T^\varepsilon} e^{-\Lambda_0(\sigma\wedge T^\varepsilon - s)}ds \int_{1\leqslant\|z\|\leqslant\rho^\varepsilon} \|z\|\nu(dz) \\
&\leqslant \frac{\nu(B_1^{\mathsf{c}}(0))}{\Lambda_0}\frac{\varepsilon\rho^\varepsilon}{\gamma_\varepsilon} \leqslant C_0\frac{\varepsilon\rho^\varepsilon}{\gamma_\varepsilon}.
\end{aligned}$$

Note that the right side is independent of the size of σ. We assume $C_0 \geqslant 1$ without loss of generality. The limit (6.6) implies the existence of a constant $\varepsilon_0 \in (0,1]$ such that for $\varepsilon \in (0,\varepsilon_0]$ we have $\varepsilon\rho^\varepsilon \leqslant \frac{C_1}{2C_0}\gamma_\varepsilon^3 \leqslant \frac{C_1}{2}\gamma_\varepsilon^2$ and hence satisfies the claim.

We start the proof of main estimate.

Step 1: Exponential estimate of the stochastic convolution: For ε_0 of Step 0 and $\varepsilon \in (0,\varepsilon_0]$ we have

$$\begin{aligned}
\mathbb{P}(\sup_{t\leqslant\sigma\wedge T^\varepsilon} \|\Psi^\varepsilon_t\| > \gamma_\varepsilon^q) &\leqslant \mathbb{P}(\sup_{t\leqslant\sigma\wedge T^\varepsilon} \|\widetilde{\Psi}^\varepsilon_t\| > \frac{1}{2}\gamma_\varepsilon^q) + \mathbb{P}(\|b^\varepsilon_\sigma\| > \frac{1}{2}\gamma_\varepsilon^q) \\
&= \mathbb{P}(\sup_{t\leqslant\sigma\wedge T^\varepsilon} \|\widetilde{\Psi}^\varepsilon_t\| > \frac{1}{2}\gamma_\varepsilon^q).
\end{aligned}$$

Kolmogorov's exponential inequality yields for the free parameter $\lambda > 0$

$$\begin{aligned}
\mathbb{P}(\sup_{t\leqslant\sigma\wedge T^\varepsilon} \|\widetilde{\Psi}^{\varepsilon,y}_t\| > \frac{1}{2}\gamma_\varepsilon^q) &= \mathbb{P}(\lambda\sup_{t\leqslant\sigma\wedge T^\varepsilon} \|\widetilde{\Psi}^{\varepsilon,y}_t\|^2 > \lambda\frac{1}{4}\gamma_\varepsilon^{2q}) \\
&\leqslant \exp(-\lambda\frac{1}{4}\gamma_\varepsilon^{2q})\mathbb{E}\Big[\exp(\lambda\sup_{t\leqslant\sigma\wedge T^\varepsilon} \|\widetilde{\Psi}^{\varepsilon,y}_t\|^2)\Big]. \qquad (6.8)
\end{aligned}$$

Note that $(M^{(1)}_t)_{t\wedge\sigma}$ defined as $M^{(1)}_t := \varepsilon\xi^\varepsilon_t = \int_0^t\int_{0<\|z\|\leqslant\rho^\varepsilon}\varepsilon z\widetilde{N}(ds,dz)$ is a $\mathcal{F}$-martingale. The pathwise estimate of the stochastic convolution shown by

Salavati and Zangeneh in Theorem 6 of [64] yields the $\mathbb{P}$-a.s. inequality ($\widetilde{\Psi}_0^\varepsilon = 0$) for any $n \in \mathbb{N}$ and $t \geqslant 0$ with multiplicative integrand

$$
\begin{aligned}
\|\widetilde{\Psi}_t^\varepsilon\|^{2n} &= \Big(\|\widetilde{\Psi}_t^\varepsilon\|^2\Big)^n \\
&\leqslant \bigg(2\int_0^t e^{-2\Lambda_0(t-s)}\langle\langle\widetilde{\Psi}_{s-}^\varepsilon, dM_s^{(1)}\rangle\rangle \\
&\qquad + \sum_{0\leqslant s\leqslant t} e^{-2\Lambda_0(t-s)}\big(\|\widetilde{\Psi}_s^\varepsilon\|^2 - \|\widetilde{\Psi}_{s-}^\varepsilon\|^2 - 2\langle\langle\widetilde{\Psi}_{s-}^\varepsilon, \Delta_s M^{(1)}\rangle\rangle\big)\bigg)^n \\
&= e^{-2n\Lambda_0 t}\bigg(2\int_0^t\int_{\|z\|\leqslant\rho^\varepsilon} e^{2\Lambda_0 s}\langle\langle\widetilde{\Psi}_{s-}^\varepsilon, \varepsilon z\rangle\rangle\widetilde{N}(ds,dz) \\
&\qquad\qquad + \int_0^t\int_{\|z\|\leqslant\rho^\varepsilon} e^{2\Lambda_0 s}\|\varepsilon z\|^2 N(ds,dz)\bigg)^n \\
&=: e^{-2n\Lambda_0 t}\big(M_t^{(2)} + M_t^{(3)}\big)^n.
\end{aligned}
$$

Therefore we obtain

$$
\begin{aligned}
\sup_{t\in[0,\sigma\wedge T^\varepsilon]}\|\widetilde{\Psi}_t^\varepsilon\|^{2n} &\leqslant e^{-2n\Lambda_0(\sigma\wedge T^\varepsilon)}\Big(\sup_{t\in[0,\sigma\wedge T^\varepsilon]}|M_t^{(2)}| + \sup_{t\in[0,\sigma\wedge T^\varepsilon]}|M_t^{(3)}|\Big)^n \\
&= e^{-2n\Lambda_0(\sigma\wedge T^\varepsilon)}\Big(\sup_{t\in[0,\sigma\wedge T^\varepsilon]}|M_t^{(2)}| + M_{\sigma\wedge T^\varepsilon}^{(3)}\Big)^n.
\end{aligned}
$$

Using the pathwise Burkholder-Davis-Gundy inequality by Siorpaes [62] we continue for $H_s := M_s^{(2)}/\big(\sqrt{[M^{(2)}]_s + \sup_{r\in[0,s]}(M_r^{(2)})^2}\,\big)^{-1}$, $s \geqslant 0$ (which satisfies $\mathbb{P}$-a.s. $|H_s| \leqslant 1$) the inequality

$$
\begin{aligned}
&\Big(\sup_{t\in[0,\sigma\wedge T^\varepsilon]}|M_t^{(2)}| + M_{\sigma\wedge T^\varepsilon}^{(3)}\Big)^n \\
&\qquad\leqslant \bigg(6\sqrt{[M^{(2)}]_{\sigma\wedge T^\varepsilon}} - \int_0^{\sigma\wedge T^\varepsilon} H_{s-}dM_s^{(2)} + M_{\sigma\wedge T^\varepsilon}^{(3)}\bigg)^n
\end{aligned}
$$

and furthermore

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,\sigma\wedge T^\varepsilon]}\|\widetilde{\Psi}_t^\varepsilon\|^{2n}\Big] \\
&\leqslant \mathbb{E}\Big[e^{-2n\Lambda_0(\sigma\wedge T^\varepsilon)}\Big(6\sqrt{[M^{(2)}]_{\sigma\wedge T^\varepsilon}} - \int_0^{\sigma\wedge T^\varepsilon} H_{s-}dM_s^{(2)} + M_{\sigma\wedge T^\varepsilon}^{(3)}\Big)^n\Big].
\end{aligned}
$$

We continue the estimate of the exponential factor on the right side of (6.8). With the help of the monotone convergence theorem and the elementary estimate $abcd \leqslant \frac{1}{4}(a^4+b^4+c^4+d^4)$ we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\exp(\lambda \sup_{t\leqslant \sigma\wedge T^\varepsilon} \|\widetilde{\Psi}^\varepsilon_t\|^2)\Big]\\
&= \mathbb{E}\Big[\sum_{n=0}^\infty \frac{\lambda^n}{n!} \sup_{t\leqslant \sigma\wedge T^\varepsilon} \|\widetilde{\Psi}^\varepsilon_t\|^{2n}\Big]\\
&= \sum_{n=0}^\infty \frac{\lambda^n}{n!}\mathbb{E}\Big[\sup_{t\leqslant \sigma\wedge T^\varepsilon} \|\widetilde{\Psi}^\varepsilon_t\|^{2n}\Big]\\
&\leqslant \sum_{n=0}^\infty \mathbb{E}\Big[\frac{\lambda^n}{n!} e^{-2n\Lambda_0(\sigma\wedge T^\varepsilon)}\Big(6\sqrt{[M^{(2)}]_{\sigma\wedge T^\varepsilon}} - \int_0^{\sigma\wedge T^\varepsilon} H_{s-}dM^{(2)}_s + M^{(3)}_{\sigma\wedge T^\varepsilon}\Big)^n\Big]\\
&= \mathbb{E}\Big[\exp\Big(\lambda e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}\Big(6\sqrt{[M^{(2)}]_{\sigma\wedge T^\varepsilon}} - \int_0^{\sigma\wedge T^\varepsilon} H_{s-}dM^{(2)}_s + M^{(3)}_{\sigma\wedge T^\varepsilon}\Big)\Big)\Big]\\
&\leqslant \mathbb{E}\Big[\exp\Big(\Big(\frac{3\lambda}{c^2} e^{-4\Lambda_0(\sigma\wedge T^\varepsilon)}[M^{(2)}]_{\sigma\wedge T^\varepsilon}\Big) + 3\lambda c^2\\
&\qquad\qquad - \lambda e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}\Big(\int_0^{\sigma\wedge T^\varepsilon} H_{s-}dM^{(2)}_s - e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}M^{(3)}_{\sigma\wedge T^\varepsilon}\Big)\Big)\Big]\\
&\leqslant \frac{1}{4}\mathbb{E}\Big[\exp\Big(\frac{12\lambda}{c^2} e^{-4\Lambda_0(\sigma\wedge T^\varepsilon)}[M^{(2)}]_{\sigma\wedge T^\varepsilon}\Big)\Big] + \frac{1}{4}\exp\Big(12\lambda c^2\Big)\\
&\quad + \frac{1}{4}\mathbb{E}\Big[\exp\Big(-4\lambda e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}\int_0^{\sigma\wedge T^\varepsilon} H_{s-}dM^{(2)}_s\Big)\Big]\\
&\quad + \frac{1}{4}\mathbb{E}\Big[\exp\Big(4\lambda e^{-4\Lambda_0(\sigma\wedge T^\varepsilon)} M^{(3)}_{\sigma\wedge T^\varepsilon}\Big)\Big]\\
&=: J_1(\varepsilon) + J_2(\varepsilon) + J_3(\varepsilon) + J_4(\varepsilon).
\end{aligned}
$$

Step 2: Campbell's formula and Optimization over the free parameters. The idea is now to choose the free parameters λ and c as ε-dependent functions $\lambda_\varepsilon := \frac{1}{\gamma_\varepsilon^{2q+1}}$ and $c_\varepsilon = \gamma_\varepsilon^{q+1}$ such that on the right side of (6.8) the term structure we obtain reads

$$\exp(-\lambda_\varepsilon \frac{1}{4}\gamma_\varepsilon^{2q})\Big(J_1(\varepsilon) + J_2(\varepsilon) + J_3(\varepsilon) + J_4(\varepsilon)\Big),$$

from which we extract the desired estimate. Note that

$$\exp(-\lambda_\varepsilon \frac{1}{4}\gamma_\varepsilon^{2q}) = \exp(-\frac{1}{4}\frac{1}{\gamma_\varepsilon})$$

which gives the desired convergence in ε as long as the terms $J_i(\varepsilon)$ remain uniformly bounded. We estimate the terms $J_1 - J_4$ one by one.

J_1**:** Note that $\mathbb{P}$-a.s. we have

$$\begin{aligned}
[M^{(2)}]_{\sigma\wedge T^\varepsilon} &= \Big[2\int_0^\cdot \int_{\|z\|\leqslant \rho^\varepsilon} e^{2\Lambda_0 s}\Big(\langle\langle \widetilde{\Psi}^\varepsilon_{s-}, \varepsilon z\rangle\rangle\Big)\widetilde{N}(ds,dz)\Big]_{\sigma\wedge T^\varepsilon}\\
&= 4\int_0^{\sigma\wedge T^\varepsilon}\int_{\|z\|\leqslant \rho^\varepsilon} e^{4\Lambda_0 s}\Big(\langle\langle \widetilde{\Psi}^\varepsilon_{s-}, \varepsilon z\rangle\rangle\Big)^2 N(ds,dz)\\
&\leqslant C_1\int_0^{\sigma\wedge T^\varepsilon}\int_{\|z\|\leqslant \rho^\varepsilon} e^{4\Lambda_0 s}\|\widetilde{\Psi}^\varepsilon_{s-}\|^2\|\varepsilon z\|^2 N(ds,dz)\\
&\leqslant C_2\int_0^{T^\varepsilon}\int_{\|z\|\leqslant \rho^\varepsilon} e^{4\Lambda_0 s}\|\varepsilon z\|^2 N(ds,dz).
\end{aligned}$$

Campbell's formula for Poisson random measures yields

$$\begin{aligned}
&\mathbb{E}\Big[\exp\big(\frac{\lambda_\varepsilon}{c_\varepsilon^2}e^{-4\Lambda_0(\sigma\wedge T^\varepsilon)}12[M^{(2)}]_{\sigma\wedge T^\varepsilon}\big)\Big]\\
&\leqslant \mathbb{E}\Big[\exp\big(\frac{C_3}{\gamma_\varepsilon^{4q+3}}\int_0^{\sigma\wedge T^\varepsilon}\int_{\|z\|\leqslant \rho^\varepsilon} e^{4\Lambda_0(s-\sigma\wedge T^\varepsilon)}\|\varepsilon z\|^2 N(ds,dz)\big)\Big]\\
&\leqslant \mathbb{E}\Big[\exp\big(\frac{C_3}{\gamma_\varepsilon^{4q+3}}\int_0^{T^\varepsilon}\int_{\|z\|\leqslant \rho^\varepsilon} \|\varepsilon z\|^2 N(ds,dz)\big)\Big]\\
&= \mathbb{E}\Big[\exp\big(T^\varepsilon\int_{\|z\|\leqslant \rho^\varepsilon}\big(\exp\big(\frac{C_3\|\varepsilon z\|^2}{\gamma_\varepsilon^{4q+3}}\big)-1\big)\nu(dz)ds\big)\Big].
\end{aligned}$$

The limit (6.6) implies then

$$\sup_{s\in[0,\sigma\wedge T^\varepsilon]}\sup_{\|z\|\leqslant\rho^\varepsilon}\frac{\|\varepsilon z\|^2}{\gamma_\varepsilon^{4q+3}} \leqslant \frac{(\varepsilon\rho^\varepsilon)^2}{\gamma_\varepsilon^{4q+3}} \longrightarrow 0, \qquad \text{as } \varepsilon\to 0.$$

Hence for $\varepsilon_0 \in (0,1]$ such that $\varepsilon\in(0,\varepsilon_0]$ implies $\frac{(\varepsilon\rho^\varepsilon)^2}{\gamma_\varepsilon^{4q+3}} \leqslant 1$ and $\rho^\varepsilon \geqslant \int_{\|z\|\leqslant 1}\|z\|^2\nu(dz)/\nu(B_1^\complement(0))$ using the estimate $(e^r-1)\leqslant(e-1)r$ for all $r\in[0,1]$

we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\exp\Big(T^\varepsilon\int_{\|z\|\leqslant\rho^\varepsilon}\Big(\exp\big(\frac{C_3\|\varepsilon z\|^2}{\gamma_\varepsilon^{4q+3}}\big)-1\Big)\nu(dz)ds\Big)\Big]\\
&\leqslant\mathbb{E}\Big[\exp\Big((e-1)T^\varepsilon\int_{\|z\|\leqslant\rho^\varepsilon}\frac{C_3\|\varepsilon z\|^2}{\gamma_\varepsilon^{4q+3}}\nu(dz)\Big)\Big]\\
&\leqslant\mathbb{E}\Big[\exp\Big(C_3(e-1)T^\varepsilon\frac{\varepsilon^2}{\gamma_\varepsilon^{4q+3}}\int_{\|z\|\leqslant\rho^\varepsilon}\|z\|^2\nu(dz)\Big)\Big]\\
&\leqslant\mathbb{E}\Big[\exp\Big(C_3(e-1)\frac{\varepsilon^2T^\varepsilon}{\gamma_\varepsilon^{4q+3}}\Big(\int_{\|z\|\leqslant1}\|z\|^2\nu(dz)+\rho^\varepsilon\nu(B_1^{\mathsf{c}}(0))\Big)\Big)\Big]\\
&\leqslant\exp\Big(2C_3C_4\frac{(\varepsilon\rho^\varepsilon)^2T^\varepsilon}{\gamma_\varepsilon^{4q+3}}\Big)=\exp\Big(2C_3C_4\Gamma(\varepsilon)\Big)
\end{aligned}
$$

for $C_4=\int_{\|z\|\leqslant1}\|z\|^2\nu(dz)+\nu(B_1^{\mathsf{c}}(0))$. The limit (6.6) yields the result. Note that the right-hand side is estimated only by monotonicity. We have thus obtained for $\varepsilon\in(0,\varepsilon_0]$

$$
\begin{aligned}
\exp(-\lambda_\varepsilon\frac{1}{4}\gamma_\varepsilon^{2q})J_1(\varepsilon)&\leqslant\exp(-\frac{1}{\gamma_\varepsilon}\frac{1}{4})J_1(\varepsilon)\\
&\leqslant\frac{1}{4}\exp(-\frac{1}{\gamma_\varepsilon}\frac{1}{4}+\Gamma(\varepsilon))\leqslant\frac{1}{4}\exp(-\frac{1}{\gamma_\varepsilon}\frac{1}{5}).
\end{aligned}
$$

J_2**:** Trivially, there is $\varepsilon_0\in(0,1]$ such that for $\varepsilon\in(0,\varepsilon_0]$ we have

$$
\begin{aligned}
\exp(-\lambda_\varepsilon\frac{1}{4}\gamma_\varepsilon^{2q})J_2(\varepsilon)&\leqslant\frac{1}{4}\exp(-\frac{1}{\gamma_\varepsilon}\frac{1}{4}+6\lambda_\varepsilon c_\varepsilon^2)\\
&=\frac{1}{4}\exp(-\frac{1}{\gamma_\varepsilon}\frac{1}{4}+6\gamma_\varepsilon)\leqslant\frac{1}{4}\exp(-\frac{1}{\gamma_\varepsilon}\frac{1}{5}).
\end{aligned}
$$

J_3**:** Recall that

$$
\begin{aligned}
J_3(\varepsilon)&=\frac{1}{4}\mathbb{E}\Big[\exp\Big(-4\lambda_\varepsilon e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}\int_0^{\sigma\wedge T^\varepsilon}H_{s-}dM_s^{(2)}\Big)\Big]\\
&\leqslant\frac{1}{4}\mathbb{E}\Big[\exp\Big(2e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}\lambda_\varepsilon^2\big(\int_0^{\sigma\wedge T^\varepsilon}H_{s-}dM_s^{(2)}\big)^2+2\Big)\Big]\\
&\leqslant2\mathbb{E}\Big[\exp\Big(2e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}\lambda_\varepsilon^2\big(\int_0^{\sigma\wedge T^\varepsilon}H_{s-}dM_s^{(2)}\big)^2\Big)\Big]
\end{aligned}
$$

and

$$
M_t^{(2)}=2\int_0^t\int_{\|z\|\leqslant\rho^\varepsilon}e^{2\Lambda_0s}\langle\langle\widetilde{\Psi}_{s-}^\varepsilon,\varepsilon z\rangle\rangle\widetilde{N}(ds,dz).
$$

For the function $h(s-,\varepsilon z) := e^{2\Lambda_0 s-} H_{s-}\langle\langle \widetilde{\Psi}^\varepsilon_{s-}, \varepsilon z\rangle\rangle$ we have

$$\int_0^{\sigma\wedge T^\varepsilon} H_{s-} dM_s^{(2)} = 2\int_0^{\sigma\wedge T^\varepsilon}\int_{\|z\|\leqslant\rho^\varepsilon} h(s-,\varepsilon z)\widetilde{N}(ds,dz).$$

Note that for $s\in[0,\sigma\wedge T^\varepsilon]$, $z\in H$ and $\varepsilon\in(0,1]$ we have

$$|h(s-,\varepsilon z)| \leqslant d(\chi) e^{2\Lambda_0 s}\varepsilon\|z\|$$

and hence the limit (6.6) yields $\mathbb{P}$-a.s. as $\varepsilon\to 0+$

$$\sup_{s\in[0,\sigma\wedge T^\varepsilon]}\sup_{\|z\|\leqslant\rho^\varepsilon} 2\lambda_\varepsilon^2 e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}|h(s-,\varepsilon z)| \leqslant 2d(\chi)\frac{\varepsilon\rho^\varepsilon}{\gamma_\varepsilon^{2(2p+1)}} \to 0.$$

Lemma 6.1.2 yields that there is $\varepsilon_0\in(0,1]$ such that $\varepsilon\in(0,\varepsilon_0]$ implies

$$\mathbb{E}\Big[\exp\Big(2e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)}\lambda_\varepsilon^2\big(\int_0^{\sigma\wedge T^\varepsilon} H_{s-} dM_s^{(2)}\big)^2\Big)\Big] \leqslant 2.$$

and hence

$$J_3(\varepsilon)\leqslant 4.$$

This yields for slightly reduced $\varepsilon_0\in(0,1)$ and $\varepsilon\in(0,\varepsilon_0]$ the estimate $\exp(-\lambda_\varepsilon\frac{1}{4}\gamma_\varepsilon^{2q})J_3(\varepsilon)\leqslant\frac{1}{4}\exp(-\frac{1}{\gamma_\varepsilon}\frac{1}{5})$.

J_4**:** This case resembles the one of $J_1(\varepsilon)$. Since we have only positive jumps we have $\mathbb{P}$-a.s.

$$M^{(3)}_{\sigma\wedge T^\varepsilon} = \int_0^{\sigma\wedge T^\varepsilon}\int_{\|z\|\leqslant\rho^\varepsilon} e^{2\Lambda_0 s}\|\varepsilon z\|^2 N(ds,dz)$$

leading to

$$\begin{aligned}&\mathbb{E}\Big[\exp\Big(4\lambda_\varepsilon e^{-2\Lambda_0(\sigma\wedge T^\varepsilon)} M^{(3)}_{\sigma\wedge T^\varepsilon}\Big)\Big]\\ &= \mathbb{E}\Big[\exp\Big(\int_0^{\sigma\wedge T^\varepsilon}\int_{\|z\|\leqslant\rho^\varepsilon}\big(\exp(4\lambda_\varepsilon e^{-2\Lambda_0((\sigma\wedge T^\varepsilon)-s)}\|\varepsilon z\|^2)-1\big)\nu(dz)ds\Big)\Big].\end{aligned}$$

Analogously to $J_1(\varepsilon)$ we obtain with the help of the limit (6.6) that for $\varepsilon\to 0+$ we have $\mathbb{P}$-a.s.

$$\sup_{m\in\mathbb{N}}\sup_{s\in[0,\sigma\wedge T^\varepsilon]}\sup_{\|z\|\leqslant\rho^\varepsilon}\lambda_\varepsilon e^{-2\Lambda_0((\sigma\wedge T^\varepsilon)-s)}\|\varepsilon z\|^2 \leqslant \lambda_\varepsilon(\varepsilon\rho^\varepsilon)^2 \leqslant \frac{(\varepsilon\rho^\varepsilon)^2}{\gamma_\varepsilon^{2q+1}} \to 0.$$

The additional choice of $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0,\varepsilon_0]$ implies $2\frac{(\varepsilon\rho^\varepsilon)^2}{\gamma_\varepsilon^{2q+1}} \leqslant 1$ and $\rho^\varepsilon \geqslant \int_{\|z\|\leqslant 1} \|z\|^2 \nu(dz)/\nu(B_1^{\mathsf{c}}(0))$ ensures

$$
\begin{aligned}
J_4(\varepsilon) &\leqslant \frac{1}{4}\mathbb{E}\Big[\exp\Big(\int_0^{\sigma\wedge T^\varepsilon}\int_{\|z\|\leqslant\rho^\varepsilon}\Big(\exp\big(2\lambda_\varepsilon e^{-2\Lambda_0((\sigma\wedge T^\varepsilon)-s)}\|\varepsilon z\|^2\big)-1\Big)\nu(dz)ds\Big)\Big]\\
&\leqslant \frac{1}{4}\mathbb{E}\Big[\exp\Big((e-1)2C_6\int_0^{\sigma\wedge T^\varepsilon} e^{-2\Lambda_0((\sigma\wedge T^\varepsilon)-s)}ds\lambda_\varepsilon(\varepsilon\rho^\varepsilon)^2\Big)\Big]\\
&\leqslant \frac{1}{4}\exp\Big(\frac{(e-1)C_6}{\Lambda_0}\frac{(\varepsilon\rho^\varepsilon)^2}{\gamma_\varepsilon^{2q+1}}\Big)\longrightarrow\frac{1}{4}, \qquad \text{as } \varepsilon\to 0+.
\end{aligned}
$$

Step 3: We conclude for $\varepsilon_0 \in (0,1]$ chosen sufficiently small that $\varepsilon \in (0,\varepsilon_0]$ yields

$$
\begin{aligned}
\mathbb{P}(\sup_{t\in[0,\sigma\wedge T^\varepsilon]}\|\Psi_t^{\varepsilon,y}\| > \gamma_\varepsilon^q) &\leqslant \exp(-\lambda_\varepsilon\frac{1}{4}\gamma_\varepsilon^{2q})\mathbb{E}\Big[\exp(\lambda_\varepsilon\sup_{t\leqslant\sigma\wedge T^\varepsilon}\|\widetilde{\Psi}_t^{\varepsilon,y}\|^2)\Big]\\
&\leqslant \exp(\frac{1}{5\gamma_\varepsilon}).
\end{aligned}
$$

□

Lemma 6.1.2 (A Campbell type estimate)**.** *Under the notation of Step 2 of the proof of the preceding theorem there is $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0,\varepsilon_0]$ implies*

$$
\mathbb{E}\Big[\exp(\lambda_\varepsilon e^{-2\Lambda_0(t\wedge\sigma)}Z_{t\wedge\sigma}^2)\Big] \leqslant 2.
$$

Proof. Recall the notation $h(s-,\varepsilon z) := e^{2\Lambda_0 s-}H_{s-}\langle\langle\widetilde{\Psi}_{s-}^{\varepsilon,x}, G(Y(s-),\varepsilon z)\rangle\rangle$ and

$$
\int_0^{\sigma\wedge T^\varepsilon} H_{s-}dM_s^{(2)} = \int_0^{\sigma\wedge T^\varepsilon}\int_{\|z\|\leqslant\rho^\varepsilon} h(s-,\varepsilon z)\widetilde{N}(dsdz).
$$

Consider for any $z \in H$ the process an $(\mathcal{F}_t)_{t\geqslant 0}$-predictable process $(H(t,z))_{t\geqslant 0}$

$$
Z_t := \int_0^t\int_{\|z\|\leqslant\rho^\varepsilon} h(s-,\varepsilon z)\widetilde{N}(ds,dz).
$$

Then Itô's formula for Poisson random measures yields $\mathbb{P}$-a.s.

$$
\begin{aligned}
&\exp(\lambda_\varepsilon e^{-2\Lambda_0 t} Z_t^2) \\
&= 1 + \int_0^t \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \lambda_\varepsilon 2 Z_{s-} e^{-\Lambda_0 s-} dZ_s \\
&\quad - 2\Lambda_0 \int_0^t \exp(\lambda_\varepsilon e^{-2\Lambda_0 s} Z_s^2) e^{-\Lambda_0 s} \lambda_\varepsilon Z_s^2 ds \\
&\quad + \sum_{0<s\leqslant t} \Big(\exp(\lambda_\varepsilon e^{-2\Lambda_0 s} Z_s^2) - \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \\
&\qquad\qquad - \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \lambda_\varepsilon e^{-2\Lambda_0 s-} 2 Z_{s-} \Delta_s Z \Big) \\
&= 1 + \int_0^t \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \lambda_\varepsilon e^{-2\Lambda_0 s-} 2 Z_{s-} h(s-, \varepsilon z) \widetilde{N}(dsdz) \\
&\quad + \int_0^t \exp(\lambda_\varepsilon e^{-2\Lambda_0 s} Z_s^2) \lambda_\varepsilon e^{-2\Lambda_0 s} (-2\Lambda_0) Z_s^2 ds \\
&\quad + \int_0^t \int_{\|z\|\leqslant \rho^\varepsilon} \Big(\exp(\lambda_\varepsilon e^{-2\Lambda_0 s} (Z_{s-} + h(s-, \varepsilon z))^2) \\
&\qquad\qquad - \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \Big) \widetilde{N}(dsdz) \\
&\quad + \int_0^t \int_{\|z\|\leqslant \rho^\varepsilon} \Big(\exp(\lambda_\varepsilon e^{-2\Lambda_0 s} (Z_{s-} + h(s-, \varepsilon z))^2) - \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \\
&\qquad\qquad - \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \lambda_\varepsilon e^{-2\Lambda_0 s-} 2 Z_{s-} h(s-, \varepsilon z) \Big) \nu(dz) ds.
\end{aligned}
$$

Let σ be the $(\mathcal{F}_t)_{t\geqslant 0}$-stopping time σ defined in (6.4) and (6.5) then the optimal stopping theorem yields

$$
\begin{aligned}
&\mathbb{E}\Big[\exp(\lambda_\varepsilon e^{-2\Lambda_0 (t\wedge\sigma)} Z_{t\wedge\sigma}^2) \Big] \\
&= 1 - 2\Lambda_0 \lambda_\varepsilon \mathbb{E}\Big[\int_0^{t\wedge\sigma} \exp(\lambda_\varepsilon e^{-2\Lambda_0 s} (Z_{s-} + h(s-, \varepsilon z))^2) e^{-2\Lambda_0 s} Z_s^2 ds \Big] \\
&+ \mathbb{E}\Big[\int_0^{t\wedge\sigma} \int_{\|z\|\leqslant \rho^\varepsilon} \Big(\exp(\lambda_\varepsilon e^{-2\Lambda_0 s} (Z_{s-} + h(s-, \varepsilon z))^2) - \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \\
&\qquad + \exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z_{s-}^2) \lambda_\varepsilon e^{-2\Lambda_0 s-} 2 Z_{s-} h(s-, \varepsilon z) \Big) \nu(dz) ds \Big].
\end{aligned}
$$

Due to the nonnegativity of the second term we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\exp(\lambda_\varepsilon e^{-2\Lambda_0(t\wedge\sigma)} Z^2_{t\wedge\sigma})\Big] \\
&\leqslant 1+\mathbb{E}\Big[\int_0^{t\wedge\sigma}\int_{\|z\|\leqslant\rho^\varepsilon}\Big(\exp(\lambda_\varepsilon e^{-2\Lambda_0 s}(Z_{s-}+h(s-,\varepsilon z))^2)-\exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z^2_{s-}) \\
&\qquad +\exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z^2_{s-})\lambda_\varepsilon e^{-2\Lambda_0 s-}2Z_{s-}h(s-,\varepsilon z)\Big)\nu(dz)ds\Big].
\end{aligned}$$

Taylor's formula yields

$$\begin{aligned}
&\Big(\exp(\lambda_\varepsilon e^{-2\Lambda_0 s}(Z_{s-}+h(s-,\varepsilon z))^2)-\exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z^2_{s-}) \\
&\qquad +\exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z^2_{s-})\lambda_\varepsilon e^{-\Lambda_0 s-}2Z_{s-}h(s-,\varepsilon z)\Big) \\
&\leqslant 2(1+2\lambda_\varepsilon e^{-2\Lambda_0 s}\|Z_{s-}\|^2)\lambda_\varepsilon e^{-2\Lambda_0 s}\exp(\lambda_\varepsilon e^{-2\Lambda_0 s}(Z_{s-})^2)\frac{1}{2}\|h(s-,\varepsilon z)\|^2.
\end{aligned}$$

On the event $s\leqslant\sigma$ and having in mind that $\lambda_\varepsilon\varepsilon\rho^\varepsilon\longrightarrow 0$ by the limit (6.6) as $\varepsilon\to 0$ we have

$$\begin{aligned}
&\Big(\exp(\lambda_\varepsilon e^{-2\Lambda_0 s}(Z_{s-}+h(s-,\varepsilon z))^2)-\exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z^2_{s-}) \\
&\qquad +\exp(\lambda_\varepsilon e^{-2\Lambda_0 s-} Z^2_{s-})\lambda_\varepsilon e^{-2\Lambda_0 s-}2Z_{s-}h(s-,\varepsilon z)\Big) \\
&\leqslant 4d(\chi)^2 g_1^2(\chi)\exp(\lambda_\varepsilon e^{-2\Lambda_0 s}(Z_{s-})^2)\ \lambda_\varepsilon^2\varepsilon^2\|z\|^2.
\end{aligned}$$

Hence

$$\begin{aligned}
&\mathbb{E}\Big[\exp(\lambda_\varepsilon e^{-2\Lambda_0(t\wedge\sigma)} Z^2_{t\wedge\sigma})\Big] \\
&\leqslant 1+4d(\chi)^2 g_1^2(\chi)\mathbb{E}\Big[\int_0^{t\wedge\sigma}\int_{\|z\|\leqslant\rho^\varepsilon}\exp(\lambda_\varepsilon e^{-2\Lambda_0 s}(Z_{s-})^2)\ \lambda_\varepsilon^2\varepsilon^2\|z\|^2\nu(dz)ds\Big] \\
&\leqslant 1+4d(\chi)^2 g_1^2(\chi)\lambda_\varepsilon^2\varepsilon^2\int_{\|z\|\leqslant\rho^\varepsilon}\|z\|^2\nu(dz)\int_0^t\mathbb{E}\Big[\exp(\lambda_\varepsilon e^{-2\Lambda_0 s}(Z_{s-\wedge\sigma})^2)\Big]ds \\
&\leqslant 1+4d(\chi)^2 g_1^2(\chi)\ (\int_{\|z\|\leqslant 1}\|z\|^2\nu(dz)+\nu(B_1^c(0))(\lambda_\varepsilon\varepsilon\rho^\varepsilon)^2 \\
&\qquad \cdot\int_0^t\mathbb{E}\Big[\exp(\lambda_\varepsilon e^{-2\Lambda_0 s}(Z_{s-\wedge\sigma})^2)\Big]ds.
\end{aligned}$$

For $\phi(t)=\mathbb{E}\Big[\exp(\lambda_\varepsilon e^{-2\Lambda_0(t\wedge\sigma)} Z^2_{t\wedge\sigma})\Big]$ we obtain Gronwall-Bellman inequality

$$\phi(t)\leqslant 1+K_\varepsilon\int_0^t\phi(s)ds,$$

and hence

$$\phi(t) \leqslant \exp(tK_\varepsilon).$$

For $t \leqslant T^\varepsilon$ with $T^\varepsilon(\varepsilon\rho^\varepsilon\lambda_\varepsilon)^2 \to 0$ there is $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0,\varepsilon_0]$ implies

$$\phi(T^\varepsilon) \leqslant \exp(T^\varepsilon K_\varepsilon) \leqslant 2.$$

□

6.2 Exponential estimates of the deviations of the small jump equation

For $\varepsilon, \gamma \in (0,1]$, $y \in H$, $t \geqslant 0$ we define the event of a small stochastic convolution

$$\mathcal{E}_{t,y}(\gamma,\varepsilon) := \{ \sup_{s\in[0,t]} \|\Psi_s^{\varepsilon,y}\| \leqslant \gamma\} \tag{6.9}$$

$$E_y(\gamma,\varepsilon) := \{ \sup_{s\in[0,T_1]} \|Y^\varepsilon(s;y) - u(s;y)\| \leqslant \gamma\}. \tag{6.10}$$

We suppress the dependence on $\varepsilon \in (0,1]$. This subsection is dedicated to the proof the following estimate used in the proof of the main result.

Proposition 6.2.1. *Let the Hypotheses (D) and (S.1) be satisfied and the functions $\gamma_\cdot, \rho^\cdot$ given by (6.1). Then there exists a constant $q \geqslant 1$ such that if $\gamma_\cdot, \rho^\cdot$ satisfy in condition (6.6) for q we obtain there is $\chi >_0$ such that for any $\theta \in (0,1)$ there is a constant $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0,1]$ satisfies*

$$\sup_{x\in D_2^\pm(\gamma_\varepsilon,\chi)} \mathbb{E}\left[e^{\theta\lambda^\pm(\varepsilon)T_1}\mathbf{1}(E_x^{\mathsf{c}})\right] \leqslant 2\exp(\frac{1}{5\gamma_\varepsilon}). \tag{6.11}$$

In addition, the scales can be chosen as follows

$$\gamma_\varepsilon := \varepsilon^{\gamma'}, \quad \rho^\varepsilon := \varepsilon^{-\rho'}, \quad \beta_\varepsilon = \nu(\rho^\varepsilon B_1^{\mathsf{c}}(0)) = O((\rho^\varepsilon)^{-\alpha})_{\varepsilon\to 0} T^\varepsilon := e^{-\widetilde{\theta}} \tag{6.12}$$

satisfying $(1+2\alpha)\rho' + (4q+3)\gamma' <$ and $\widetilde{\theta} = 2\alpha\rho'$.

Remark 6.2.2. *In this case the scales satisfy (6.6) and $T^\varepsilon\beta_\varepsilon \nearrow \infty$ as $\varepsilon \searrow 0$. Without loss of generality we set $\gamma' < \alpha\rho'$.*

The proof strategy consists in the estimate the event of E_y by $\mathcal{E}_{y,T_1}$. For this effect we introduce the nonlinear residuum R^ε of the randomness in Y^ε

$$R_t^{\varepsilon,x} := Y^\varepsilon(t;x) - u(t;x) - \Psi_t^{\varepsilon,x}, \qquad t \geqslant 0,\ x \in D_2^{\pm}(\gamma_\varepsilon, \chi), \quad \varepsilon \in (0,1]. \tag{6.13}$$

The quantity we have to control in E_x has the shape $Y^\varepsilon - u = \Psi^\varepsilon + R^\varepsilon$. By Proposition 6.1.1 we have a good estimate of Ψ^ε. It is therefore natural to control R^ε in terms of Ψ^ε, which is done first for large initial values (of Y^ε) on small time scales and then for initial values (of Y^ε) close to the stable state and large time scales.

Proposition 6.2.3. *Let the Hypotheses (D) and (S.1) be satisfied and the functions $\gamma_\cdot, \rho^\cdot, T^\cdot$ given by (6.1). Then there exists a constant $q \geqslant 1$ such that if $\gamma_\cdot, \rho^\cdot$ satisfy in condition (6.6) for q such that there are constants $\chi > 0$ and $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0, \varepsilon_0]$ and $x \in D_2^{\pm}(\gamma_\varepsilon, \chi)$ imply*

$$\mathcal{E}_{T_1,x}(\gamma_\varepsilon^q) \subseteq \{\sup_{t\in[0,T_1]} \|Y^\varepsilon(t;x) - u(t;x)\| \leqslant (1/2)\gamma_\varepsilon\}, \qquad \mathbb{P}\text{-}a.s. \tag{6.14}$$

Proof. of Proposition 6.2.1: Due to the independence of Y^ε and T_1 and Proposition 6.2.3 we estimate

$$\begin{aligned}
&\sup_{x\in D_2^{\pm}(\varepsilon^\gamma,\chi)} \mathbb{E}\left[e^{\theta\lambda_\varepsilon^{\pm} T_1}\mathbf{1}(E_x^{\mathsf{c}})\right] \\
&\leqslant \int_0^{T^\varepsilon} \sup_{x\in D_2^{\pm}(\gamma_\varepsilon,\chi)} \mathbb{P}(\sup_{t\leqslant s}\|\Psi_t^{\varepsilon,x}\| > \gamma_\varepsilon)\beta_\varepsilon \varepsilon^{-\beta_\varepsilon s} ds + \int_{T^\varepsilon}^\infty \beta_\varepsilon \varepsilon^{-\beta_\varepsilon s} ds \\
&\leqslant \sup_{x\in D_2^{\pm}(\gamma_\varepsilon,\chi)} \mathbb{P}(\sup_{t\in[0,T^\varepsilon]} \|\Psi_t^{\varepsilon,x}\| > \gamma_\varepsilon^q) + e^{-T^\infty\beta_\varepsilon} \leqslant \exp(-\frac{1}{5\gamma_\varepsilon}) + \exp(-\frac{1}{\varepsilon^{\alpha\rho}}) \\
&\leqslant 2\exp(-\frac{1}{5\gamma_\varepsilon})
\end{aligned}$$

□

Lemma 6.2.4. *Let the Hypotheses (D) and (S.1) be satisfied and the functions $\gamma_\cdot, \rho^\cdot$ given by (6.1). Then are $\chi > 0$ and $\varepsilon_0 \in (0,1]$ and $q > 0$ such that for any $K > 0$, $\varepsilon \in (0, \varepsilon_0]$ and $x \in D^{\pm}(\gamma_\varepsilon, \chi)$ and for all functions $T : (0,1] \to \mathbb{R}$ monotonically increasing and satisfying $T^\varepsilon \to \infty$ as $\varepsilon \searrow 0$ we have the following. The bound $T^\varepsilon \leqslant \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|$ (given in Proposition 3.4.10) implies on the event $\mathcal{E}_{T^\varepsilon\wedge\sigma}(\gamma_\varepsilon^q)$ that*

$$\sup_{t\in[0,T^\varepsilon\wedge\sigma]} \|R_t^{\varepsilon,x}\| \leqslant K\gamma_\varepsilon. \tag{6.15}$$

Proof. Fix $\chi > 0$ sufficiently large and and $\varepsilon \in (0,1]$ and $x \in D^{\pm}(\gamma_\varepsilon, \chi)$. Then the positive invariance of $\mathcal{U}^\chi$ and $D^{\pm}(\gamma_\varepsilon, \chi) \subset \mathcal{U}^\chi$ yield

$$\sup_{x\in D_2^{\pm}(\gamma_\varepsilon,\chi)} \sup_{t\geqslant 0} \|u(t;x)\| \leqslant \sup_{y\in\mathcal{U}^\chi} \|y\| = d(\chi).$$

The process $R^{\varepsilon,x}_t$ satisfies formally for all $x \in H$ and $\varepsilon \in (0,1]$

$$\frac{dR^{\varepsilon,x}_t}{dt} = \Delta R^{\varepsilon,x}_t + f(Y^\varepsilon(t;x)) - f(u(t;x)), \qquad R^{\varepsilon,x}_0 = 0. \tag{6.16}$$

Recall that $f : H \to H$ is locally Lipschitz continuous. Hence it is globally Lipschitz continuous on $\mathcal{U}^\chi$ and we obtain for $y, u \in H$

$$\|f(y) - f(u)\| \leqslant \ell(y,u)\|y-u\|,$$

for $\ell : H \times H \to (0,\infty), (y,u) \mapsto \ell(y,u)$ being some continuous function. The mild formulation of (6.16) and the identity $Y^\varepsilon(T;x) - u(t;x) = R^{\varepsilon,x}_t + \Psi^{\varepsilon,x}_t$ imply the estimate

$$\begin{aligned}\|R^{\varepsilon,x}_t\| &\leqslant \int_0^t e^{-\Lambda_0(t-s)}\|f(Y^\varepsilon(s;x)) - f(u(s;x))\|\, ds\\ &\leqslant \int_0^t e^{-\Lambda_0(t-s)}\ell(Y^\varepsilon(s;x), u(s;x))\,\|R^{\varepsilon,x}_s - \Psi^{\varepsilon,x}_s\|ds.\end{aligned}$$

We define the stopping time $\sigma'(\varepsilon) := \inf\{t > 0 \mid \|R^{\varepsilon,x}_t\| > 1\} \wedge \sigma$. Then we obtain on the events $\{t \leqslant T^\varepsilon \wedge \sigma'\} \cap \mathcal{E}_{T^\varepsilon \wedge \sigma'}(\gamma^q_\varepsilon)$ for some fixed but unknown $q \geqslant 1$ and which we determine at the end of the proof the boundedness

$$\|Y^\varepsilon(t;x)\| \leqslant \|u(t;x)\| + \|\Psi^{\varepsilon,x}_t\| + \|R^{\varepsilon,x}_t\| \leqslant d(\chi) + 2.$$

Due to $x \in \mathcal{U}^\chi$ the positive invariance of $u(t;x) \in \mathcal{U}^\chi$ for all $t \geqslant 0$ the events $\mathcal{E}_{T^\varepsilon\wedge\sigma'}(\gamma^q_\varepsilon)$ and $t \leqslant T^\varepsilon \wedge \sigma'$ imply for $\ell_\chi := \sup_{(y,u)\in(B_{d(\chi)+2}(0))^2} \ell(y,u) < \infty$ the estimate

$$e^{\Lambda_0 t}\|R^{\varepsilon,x}_t\| \leqslant \ell_\chi\Big(\int_0^t e^{\Lambda_0 s}\|R^{\varepsilon,x}_s\|ds - \int_0^t e^{\Lambda_0 s}\|\Psi^{\varepsilon,x}_s\|ds\Big).$$

Hence Gronwall's lemma with $e^{\Lambda_0 0}R^{\varepsilon,x}_0 = 0$ yields on $\mathcal{E}_{\sigma'}(\gamma^q_\varepsilon)$ and $t \leqslant T^\varepsilon \wedge \sigma$

$$\begin{aligned}e^{\Lambda_0 t}\|R^{\varepsilon,x}_t\| &\leqslant \int_0^t e^{\ell_\chi(t-s)}\int_0^s e^{\Lambda_0 r}\|\Psi^{\varepsilon,x}_r\|drds\\ &\leqslant \sup_{r\in[0,t]}\|\Psi^{\varepsilon,x}_r\|\int_0^t\int_0^s e^{\ell_\chi(t-s)}e^{\Lambda_0 r}drds.\end{aligned}$$

The elementary calculus

$$\int_0^t\int_0^s e^{\ell_\chi(t-s)}e^{\Lambda_0 r}drds = \frac{e^{\ell_\chi t}}{\ell_\chi(\ell_\chi - \Lambda_0)} + \frac{1}{\Lambda_0\ell_\chi} - \frac{e^{\Lambda_0 t}}{\Lambda_0(\ell_\chi - \Lambda_0)}$$

yields for $\kappa := \ell_\chi - \Lambda_0 > 0$ the estimate

$$\|R_t^{\varepsilon,x}\| \leqslant \frac{e^{\kappa t}}{\kappa^2} \sup_{r\in[0,t]} \|\Psi_r^{\varepsilon,x}\|$$

for $t \leqslant T^\varepsilon \wedge \sigma'$ on $\mathcal{E}_{T^\varepsilon\wedge\sigma'}(\gamma_\varepsilon^q)$. Setting $q := \kappa_1\kappa + 2$ we obtain for any $K > 0$ a value $\varepsilon_0 \in (0,1]$ sufficiently small such that $\varepsilon \in (0,\varepsilon_0]$ implies for $T^\varepsilon := \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|$ on the event $\mathcal{E}_{T^\varepsilon\wedge\sigma'}(\gamma_\varepsilon^{a\kappa+1})$ the desired estimate

$$\sup_{t\in[0,T^\varepsilon\wedge\sigma]} \|R_t^{\varepsilon,x}\| \leqslant \frac{e^{\kappa T^\varepsilon}}{\kappa^2}\gamma_\varepsilon^{a\kappa+2} \leqslant K\gamma_\varepsilon. \tag{6.17}$$

If $\varepsilon_0 \in (0,1]$ is additionally small enough such that $K\gamma_\varepsilon < 1$ for $\varepsilon \in (0,\varepsilon_0]$ we have on the event $\mathcal{E}_{T^\varepsilon\wedge\sigma'}(\gamma_\varepsilon^{q\kappa+1})$

$$\inf\{t > 0 \mid \|R_t^{\varepsilon,x}\| > 1\} > T^\varepsilon \wedge \sigma,$$

which implies (6.15). □

Lemma 6.2.5. *Let the Hypotheses (D) and (S.1) be satisfied and the functions $\gamma_\cdot, \rho^\cdot$ given by (6.1). Then for all $\phi^\pm \in \mathcal{P}^-$ there exist constants $\chi, \delta_0, \delta_1, K_0 > 0$ such that for all $x \in B_{\delta_0}(\phi^\pm)$ and $\varepsilon \in (0,1]$ we have on the event $\mathcal{E}_\sigma(\delta_2)$ the inequality*

$$\sup_{t\in[0,\sigma]} \|R_t^{\varepsilon,x}\| \leqslant K_0 \sup_{r\in[0,\sigma]} \|\Psi_r^{\varepsilon,x}\|. \tag{6.18}$$

Proof. As before we fix an arbitrary time scale $T^\cdot : (0,1] \to (0,\infty)$ satisfying w.l.o.g. $T^\varepsilon \to \infty$ monotonically as $\varepsilon \to 0$.
The stability of $\phi^\pm$ yields that the linearization $\Delta v + f'(\phi^\pm)v$ of $\Delta u + f(u)$ in $\phi^\pm$ has strictly negative maximal eigenvalues $-\Lambda_1 < 0$, in that $\langle \Delta v + f'(\phi^\pm)v, v\rangle \leqslant -\Lambda_1|v|^2$ for $v \in H$. We fix $\delta_0 \in (0,1)$ such that additionally

$$\sup_{v,w\in B_{\delta_0}(\phi^\pm)} \|f'(v) - f'(w)\| \leqslant \frac{\Lambda_1}{4} \quad \text{and} \quad \sup_{v\in B_{\delta_0}(\phi^\pm)} \|f'(v)\| \leqslant 2\|f'(\phi^\pm)\| =: C'.$$

The stability also implies the existence of $\delta_1 \in (0,1)$ such that for $x \in B_{\delta_1}(\phi^\pm)$

$$u(t;x) \in B_{\frac{\delta_0}{4}}(\phi^\pm) \qquad t \geqslant 0.$$

Denote for $x \in B_{\delta_1}(\phi^\pm)$

$$\sigma' := \inf\{t > 0 \mid \|R_t^{\varepsilon,x}\| > \frac{\delta_0}{4}\} \wedge \sigma.$$

With the help of the decomposition (6.13) the mean value theorem applied to (6.16) yields

$$
\begin{aligned}
&\frac{dR_t^{\varepsilon,x}}{dt} = \Delta R_t^{\varepsilon,x} + f(Y^\varepsilon(t;x)) - f(u(t;x))\\
&= \Delta R_t^{\varepsilon,x} + \int_0^1 f'(u(t;x) + \theta(R_t^{\varepsilon,x} + \Psi_t^{\varepsilon,x}))d\theta(R_t^{\varepsilon,x} + \Psi_t^{\varepsilon,x})\\
&= \Delta R_t^{\varepsilon,x} + f'(\phi^\pm)R_t^{\varepsilon,x} + \int_0^1 \big(f'(u(t;x) + \theta(R_t^{\varepsilon,x} + \Psi_t^{\varepsilon,x})) - f'(\phi^\pm)\big)d\theta R_t^{\varepsilon,x}\\
&\qquad + \int_0^1 f'(u(t;x) + \theta(R_t^{\varepsilon,x} + \Psi_t^{\varepsilon,x}))d\theta \Psi_t^{\varepsilon,x}.
\end{aligned}
$$

Multiplying with $R_t^{\varepsilon,x}$ in $L^2(0,1)$, integration by parts and Young's inequality yield for any $\delta_1 < \frac{\delta_0}{4}$ on $\{t \leqslant T^\varepsilon \wedge \sigma'\} \cap \mathcal{E}_{T^\varepsilon \wedge \sigma'}(\delta_1)$

$$
\frac{1}{2}\frac{d}{dt}|R_t^{\varepsilon,x}|^2 + \Lambda_1|R_t^{\varepsilon,x}|^2 \leqslant \frac{\Lambda_1}{4}|R_t^{\varepsilon,x}|^2 + C_0|R_t^{\varepsilon,x}||\Psi_t^{\varepsilon,x}| \leqslant \frac{\Lambda_1}{2}|R_t^{\varepsilon,x}|^2 + \frac{(C_0)^2}{\Lambda_0}|\Psi_t^{\varepsilon,x}|^2
$$

such that

$$
\frac{d}{dt}|R_t^{\varepsilon,x}|^2 + \Lambda_1|R_t^{\varepsilon,x}|^2 \leqslant \frac{2(C_0)^2}{\Lambda_1}|\Psi_t^{\varepsilon,x}|^2.
$$

Hence Gronwall's lemma with the initial condition $R_0^{\varepsilon,x} = 0$ yields

$$
|R^\varepsilon(s;x)|^2 \leqslant \frac{2(C_0)^2}{\Lambda_1}|\Psi_t^{\varepsilon,x}|^2 \tag{6.19}
$$

on $\{t \leqslant \sigma' \wedge T^\varepsilon\} \cap \mathcal{E}_{\sigma' \wedge T^\varepsilon}(\delta_1)$. In order to obtain an estimate in H we use the smoothing property of the heat semigroup S and the mean value theorem as well as (6.19) on $\{t \leqslant T^\varepsilon \wedge \sigma'\} \cap \mathcal{E}_{T^\varepsilon \wedge \sigma'}(\delta_1)$

$$
\begin{aligned}
\|R_t^{\varepsilon,x}\| &\leqslant C_1 \int_0^t \frac{e^{-\Lambda_0(t-s)}}{\sqrt{t-s}}|f(Y_s^{\varepsilon,x}) - f(u(s;x))|ds\\
&\leqslant C_1(C_0 + \frac{\Lambda_0}{4})\int_0^t \frac{e^{-\Lambda_0(t-s)}}{\sqrt{t-s}}(|R_s^{\varepsilon,x}| + |\Psi_s^{\varepsilon,x}|)ds\\
&\leqslant C_1(C_0 + \frac{\Lambda_0}{4})\Big(2\frac{(C_0)^2}{\Lambda_0} + 1\Big)\int_0^t \frac{e^{-\Lambda_0(t-s)}}{\sqrt{t-s}}ds \sup_{r\in[0,t]}|\Psi_r^{\varepsilon,x}|\\
&\leqslant C_2 \sup_{r\in[0,t]}\|\Psi_r^{\varepsilon,x}\|,
\end{aligned}
$$

where $K_1 = C_2 = C_1(C_0 + \frac{\Lambda_0}{4})\Big(\frac{2(C')^2}{\Lambda_0} + 1\Big)\int_0^\infty \frac{\exp(-\Lambda_0 r)}{\sqrt{r}}dr < \infty$. If in addition $\delta_1 < \frac{1}{K_1}$ we get $\inf\{t > 0 \mid \|R_t^{\varepsilon,x}\| > \frac{\delta_0}{4}\} > T^\varepsilon \wedge \sigma$ on $\mathcal{E}_{T^\varepsilon \wedge \sigma}(\delta_1)$. In addition, we have not used any specific property of T^ε. This finishes the proof. □

We combine the previous two lemmas.

Lemma 6.2.6. *Let the Hypotheses (D) and (S.1) be satisfied and the functions $\gamma_\cdot, \rho^\cdot$ given by (6.1). For the constant $q \geqslant 1$ obtained in Lemma 6.2.4 let $\gamma_\cdot, \rho^\cdot$ additionally satisfy condition (6.6).*
Then there are constants $\chi > 0$ and $\varepsilon_0 \in (0,1]$ such that for any $\varepsilon \in (0, \varepsilon_0]$ and $x \in D_2^{\pm}(\gamma_\varepsilon, \chi)$ we have on the event $\mathcal{E}_\sigma(\gamma_\varepsilon^q)$ the inequality

$$\sup_{t\in[0,\sigma]} \|R_t^{\varepsilon,x}\| \leqslant \frac{1}{4}\gamma_\varepsilon. \tag{6.20}$$

Proof. As in the previous lemmas we prove the result for an arbitrary scale $T^\cdot : (0,1] \to [1,\infty)$ with $\lim_{\varepsilon\to 0} T^\varepsilon = \infty$ monotonically. Recall the constants $\kappa_0, \kappa_1 > 0$ from Proposition 3.4.10 and denote by

$$s^\varepsilon := \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|.$$

Assume $\varepsilon_0 \in (0,1]$ sufficiently small such that $\gamma_\varepsilon \leqslant \delta_0$ and $\gamma_\varepsilon^q < \delta_1$ given in Lemma 6.2.5 . For $T^\varepsilon \leqslant s^\varepsilon$ for all $\varepsilon \in (0,1]$ the result follows immediately by Lemma 6.2.4 for $K = \frac{1}{4}$. For $T^\varepsilon > s^\varepsilon$ for all $\varepsilon \in (0,1]$. Fix $\varepsilon_0 \in (0,1]$ sufficiently small such that $\varepsilon \in (0, \varepsilon_0]$ implies for $x \in D_2^{\pm}(\gamma_\varepsilon, \chi)$ on $\{t \leqslant T^\varepsilon \wedge \sigma\} \cap \mathcal{E}_{T^\varepsilon\wedge\sigma}(\gamma_\varepsilon^q)$ both

$$\|u(t;x) - \phi^{\pm}\| \leqslant \frac{1}{4}\gamma_\varepsilon \qquad \text{for } t \geqslant s^\varepsilon \qquad \text{and} \tag{6.21}$$

$$\sup_{s\in[0,t]} \|R_s^{\varepsilon,x}\| \leqslant \frac{1}{9}\gamma_\varepsilon \qquad \text{for } t \leqslant s^\varepsilon. \tag{6.22}$$

Then we obtain by Lemma 6.2.4 for $K = \frac{1}{9}$ on the event $\mathcal{E}_{T^\varepsilon\wedge\sigma}(\gamma_\varepsilon^q)$

$$\|Y^\varepsilon(s^\varepsilon;x) - \phi^{\pm}\| \leqslant \|u(s^\varepsilon;x) - \phi^{\pm}\| + \|R_{s^\varepsilon}^{\varepsilon,x}\| + \|\Psi_{s^\varepsilon}^{\varepsilon,x}\| \leqslant \Big(\frac{1}{4} + \frac{1}{9}\Big)\gamma_\varepsilon + \gamma_\varepsilon^q \leqslant \frac{1}{2}\gamma_\varepsilon.$$

By Lemma 6.2.5 for all $x \in B_{\delta_0}(\phi)$ the solution $u(t;x) \in B_{\delta_1}(\phi)$ for all $t \geqslant 0$. In addition, there is $\ell_0 \in (0,1]$ such that

$$\|u(t;x) - u(t;y)\| \leqslant \ell_0 \|x - y\|, \qquad \text{for all } x, y \in B_{\delta_0}(\phi), t \geqslant 0.$$

Hence we have for $\varepsilon \in (0, \varepsilon_0]$ and $x \in D_2^{\pm}(\gamma_\varepsilon, \chi)$ on the event $\{t \leqslant T^\varepsilon \wedge \sigma\} \cap \mathcal{E}_{T^\varepsilon\wedge\sigma}(\gamma_\varepsilon^q)$

$$\begin{aligned}\|u(t;x) - u(t - s^\varepsilon; Y^\varepsilon(s^\varepsilon;x))\| &\leqslant \|u(s^\varepsilon;x) - Y^\varepsilon(s^\varepsilon;x)\| \\ &\leqslant \|R_{s^\varepsilon}^{\varepsilon,x}\| + \|\Psi_{s^\varepsilon}^{\varepsilon,x}\| \leqslant \frac{1}{9}\gamma_\varepsilon + \gamma_\varepsilon^q.\end{aligned}$$

and thus

$$\begin{aligned}\|R_t^{\varepsilon,x}\| &= \|Y^\varepsilon(t-s^\varepsilon, s^\varepsilon, Y^\varepsilon(s^\varepsilon;x)) - u(t-s^\varepsilon; u(s^\varepsilon;x)) - \Psi_t^{\varepsilon,x}\| \\ &\leqslant \|Y^\varepsilon(t-s^\varepsilon, s^\varepsilon, Y^\varepsilon(s^\varepsilon;x)) - u(t-s^\varepsilon; Y^\varepsilon(s^\varepsilon;x)) - \Psi_{t-s^\varepsilon,s^\varepsilon}^{\varepsilon,Y^\varepsilon(s^\varepsilon;x)}\| \\ &\quad + \|u(t-s^\varepsilon; u(s^\varepsilon;x)) - u(t-s^\varepsilon; Y^\varepsilon(s^\varepsilon;x))\| + \|\Psi_t^{\varepsilon,x}\| + \|\Psi_{t-s^\varepsilon,s^\varepsilon}^{\varepsilon,Y^\varepsilon(s^\varepsilon;x)}\| \\ &\leqslant \sup_{s^\varepsilon\leqslant t\leqslant T^\varepsilon\wedge\sigma} \|R_t^{\varepsilon,x}\| + \sup_{t\in[0,s^\varepsilon\wedge\sigma]} \|\Psi_t^{\varepsilon,x}\| + 2\gamma_\varepsilon^q \\ &\leqslant \frac{2}{9}\gamma_\varepsilon + 3\gamma_\varepsilon^q \leqslant \frac{1}{4}\gamma_\varepsilon.\end{aligned}$$

Since we have not used any specific property of T^ε this finishes the proof. □

Proof. (of Proposition 6.2.3). Let the assumptions of Lemma 6.2.6 be satisfied for some $\chi > 0$ and $q \geqslant 1$ given by Lemma 6.2.4. We consider a time scale $T^\cdot : (0,1] \to \mathbb{R}$ satisfying $\lim_{\varepsilon\to 0} T^\varepsilon = \infty$ monotonically. Without loss of generality we assume $T^\varepsilon \geqslant s^\varepsilon$. By Lemma 6.2.6 there exists $\varepsilon_0 \in (0,1]$ such that for all $\varepsilon \in (0,\varepsilon_0]$ and $x \in D_2^\pm(\gamma_\varepsilon,\chi)$ we have

$$\begin{aligned}&\{\sup_{t\in[0,T^\varepsilon\wedge\sigma]} \|Y^\varepsilon(t;x) - u(t;x)\| \geqslant \frac{\gamma_\varepsilon}{2}\} \\ &= \{\sup_{t\in[0,T^\varepsilon\wedge\sigma]} \|R_t^{\varepsilon,x} + \Psi_t^{\varepsilon,x}\| \geqslant \frac{\gamma_\varepsilon}{2}\} \\ &\subseteq \{\sup_{t\in[0,T^\varepsilon\wedge\sigma]} \|R_t^{\varepsilon,x}\| \geqslant \frac{\gamma_\varepsilon}{4}\} \cup \{\sup_{t\in[0,T^\varepsilon\wedge\sigma]} \|\Psi_t^{\varepsilon,x}\| \geqslant \frac{\gamma_\varepsilon}{4}\} \\ &\subseteq \{\sup_{t\in[0,T^\varepsilon\wedge\sigma]} \|R_t^{\varepsilon,x}\| \geqslant \frac{\gamma_\varepsilon}{4}\} \cup \mathcal{E}^{\mathsf{c}}_{T^\varepsilon\wedge\sigma,x}(\gamma_\varepsilon^q) = \mathcal{E}^{\mathsf{c}}_{T^\varepsilon\wedge\sigma,x}(\gamma_\varepsilon^q).\end{aligned}$$

Note that by limit (6.6) we have for $\varepsilon_0 \in (0,1]$ sufficiently small, $\varepsilon \in (0,\varepsilon_0]$ and $x \in D_2^\pm(\gamma_\varepsilon,\chi) \subset \mathcal{U}^{\chi-2\gamma_\varepsilon}$ that $\|\varepsilon\Delta_\sigma\xi^\varepsilon\| \leqslant \|\varepsilon\Delta_\sigma\xi^\varepsilon\| \leqslant \varepsilon\rho^\varepsilon \leqslant \gamma_\varepsilon$. Choose in addition $\varepsilon_0 \in (0,1]$ small enough such that $2\gamma_\varepsilon < \chi - \gamma_\varepsilon$ for all $\varepsilon \in (0,\varepsilon_0]$. Then by definition of σ, since $u(t;x) \in B_{\frac{1}{2}\gamma_\varepsilon}(\phi^\pm)$ for $t \geqslant s^\varepsilon$ and $D_2^\pm(\gamma_\varepsilon,\chi)$ this implies that

$$\begin{aligned}\mathcal{E}_{T^\varepsilon\wedge\sigma}(\gamma_\varepsilon^q) \cap \{\sigma < T^\varepsilon\} = \mathcal{E}_\sigma(\gamma_\varepsilon^q) &\cap \{\sup_{t\in[0,\sigma]} \|Y^\varepsilon(t;x) - u(t;x)\| \leqslant (1/2)\gamma_\varepsilon\} \\ &\cap \{Y^\varepsilon(t;x) \in D_1^\pm(\gamma_\varepsilon,\chi) \text{ for all } t \in [0,\sigma)\} \\ &\cap \{Y^\varepsilon(t;x) \in B_{\gamma_\varepsilon}(\phi) \text{ for all } t \in [0,\sigma)\} \\ &\cap \{Y^\varepsilon(\sigma-;x) \in B_{2\gamma_\varepsilon}(\phi^\pm)\} \\ &\cap \{Y^\varepsilon(\sigma-;x) + \varepsilon\Delta_\sigma\xi^\varepsilon \in (D^\pm \cap \mathcal{U}^\chi) \setminus D_1^\pm(\gamma_\varepsilon,\chi)\} \\ &\subseteq \{Y^\varepsilon(\sigma;x) \in B_{2\gamma_\varepsilon}(\phi^\pm)\} \cap \{Y^\varepsilon(\sigma;x) \notin B_{\chi-\gamma_\varepsilon}(0)\} = \emptyset.\end{aligned} \tag{6.23}$$

In addition, $\mathcal{E}_{T^\varepsilon \wedge \sigma, x}(\gamma_\varepsilon^q) \subseteq \mathcal{E}_{T^\varepsilon, x}(\gamma_\varepsilon^q + \frac{1}{2}\gamma_\varepsilon^q) \subseteq \{\Psi_s^{\varepsilon, x} \in \mathcal{U}^\chi$ for all $s \in [0, T^\varepsilon]\}$ by definition. Together with (6.23) this implies on the event $\mathcal{E}_{T^\varepsilon \wedge \sigma}(\gamma_\varepsilon^q)$ that $\mathbb{P}$-a.s. we have $\sigma \geqslant T^\varepsilon$. Since T^ε was chosen arbitrary this finishes the proof. $\square$

Chapter 7

The large jump dynamics

7.1 Exit events and their estimates

Recall the arrival times $T_k = t_1 + \cdots + t_k$ of W_k from (6.2). The following events are the building blocks of the first exit times. For $j \in \mathbb{N}$, $x \in H$, $\chi > 0$ sufficiently large and a given rate $\gamma : (0,1) \to (0,1)$ with $\gamma_\varepsilon \to 0$ as $\varepsilon \to 0$ we define

$$A_x := \{Y^\varepsilon(t;x) \in D_2^\pm(\gamma_\varepsilon, \chi) \text{ for } t \in [0, t_1], \quad Y^\varepsilon(t_1;x) + \varepsilon\Delta_{t_j} L \in D_2^\pm(\gamma_\varepsilon, \chi)\}$$

$$A_x^- := \{Y^\varepsilon(t;x) \in D_2^\pm(\gamma_\varepsilon, \chi) \text{ for } t \in [0, t_1], \quad Y^\varepsilon(t_1;x) + \varepsilon\Delta_{t_j} L \in D_3^\pm(\gamma_\varepsilon, \chi)\}$$

$$B_x := \{Y^\varepsilon(t;x) \in D_2^\pm(\gamma_\varepsilon, \chi) \text{ for } t \in [0, t_1], \quad Y^\varepsilon(t_1;x) + \varepsilon\Delta_{t_j} L \in D_2^\pm(\gamma_\varepsilon, \chi)\}$$

$$C_x := \{Y^\varepsilon(t;x) \notin D_2^\pm(\gamma_\varepsilon, \chi) \text{ for some } t \in [0, t_1)\}.$$

The **shift operator** $\Theta_s : \mathbb{D}([0,\infty), H) \to \mathbb{D}([0,\infty), H)$ by $s > 0$ is defined as $\psi(t) \circ \Theta_s := \psi(t+s)$ for $s, t > 0$. It is applied to the event A_x^j by

$$A_x^j \circ \Theta_s = \{Y^\varepsilon(t+s; Y^\varepsilon(s;x)) \in D_2^\pm(\gamma_\varepsilon, \chi) \text{ for all } t \in (s, t_j + s) \quad \text{and}$$
$$Y^\varepsilon(t_j + s; Y^\varepsilon(s;x)) + \varepsilon\Delta_{t_j+s} L) \in D_2^\pm(\gamma_\varepsilon, \chi)\}.$$

In particular, since $t_j \circ \Theta_{T_{j-1}} = T_j$ we obtain

$$\begin{aligned} A_x^j &:= A_x \circ \Theta_{T_{j-1}} \\ &= \{Y^\varepsilon(t, Y^\varepsilon(T_{j-1};x)) \circ \Theta_{T_{j-1}} \in D_2^\pm(\gamma_\varepsilon, \chi) \text{ for all } t \in (T_{j-1}, T_j) \quad \text{and} \\ &\quad Y^\varepsilon(T_j;x) + \varepsilon\Delta_{T_j} L \in D_2^\pm(\gamma_\varepsilon, \chi)\} \end{aligned} \tag{7.1}$$

and the analogous expressions for B^j_x, C^j_x and E^j_x defined in (6.10) by the shift (7.1) respectively. By construction we have the representations

$$\{\tau_x = T_k\} = \bigcap_{j=1}^{k-1} A^j_x \cap B^k_x \qquad \text{and} \qquad \{\tau_x \in (T_{k-1}, T_k)\} = \bigcap_{j=1}^{k-1} A^j_x \cap C^k_x. \tag{7.2}$$

7.2 Proof of Theorem 5.1.1

In this section we prove two results which are not congruent to Theorem 5.1.1 and Theorem 5.1.2. In Proposition 7.2.1 we show the statement of Theorem 5.1.1 and additionally the convergence in probability of the first exit loci of Theorem 5.1.2.

Proposition 7.2.1. *For any $\theta \in (0,1)$ there are $\varepsilon_0, \gamma \in (0,1]$ and $\chi > 0$ such that $\varepsilon \in (0, \varepsilon_0]$ implies for any $U \in \mathcal{B}(H)$ with $m^\pm(\partial U) = 0$ that*

$$\sup_{x \in D_3^\pm(\varepsilon^\gamma, \chi)} \mathbb{E}\Big[e^{\theta|\lambda^\pm_\varepsilon \tau^\pm_x(\varepsilon,\chi) - \bar{s}^\pm(\varepsilon)|} \big(1 + |\mathbf{1}\{X^\varepsilon(\tau; x) \in U\} - \mathbf{1}\{W_{k^*(\varepsilon)} \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap(D^\pm)^{\mathsf{c}}}(\phi^\pm)\}|\big)\Big] \leqslant 1 + C.$$

The statement of Proposition 7.2.1 directly implies the statement of Theorem 5.1.1.

Proof. The proof is organized in four consecutive steps. First, the strong Markov property reduces the main expression to four geometric sums, whose limit consists of expressions involving certain events, which are estimated in step 2. As a next step we estimate the resulting expressions using all the previous results available.

Step 0: Conventions and assumptions. Fix $C' \in (0, \frac{1}{2}(1-\theta))$, $\chi > 0$ large enough and $\delta \in (0,1]$ sufficiently small such that $\mathcal{P} \subseteq \mathcal{U}^\chi$ and satisfying

$$m^\pm\big(D \setminus D_4^\pm(\delta, \chi)\big) < \frac{C'}{5}. \tag{7.3}$$

This is possible due to Hypothesis (S.3). We keep the scales (6.12), which satisfy (6.6). In addition assume $\varepsilon_0 \in (0,1]$ sufficiently small such that $\gamma_\varepsilon \leqslant \delta$. Due to the ubiquitous dependence of all quantities of ε, $\pm$ and χ we drop these dependencies. For convenience we write $D_i = D_i^\pm(\gamma_\varepsilon, \chi)$.

Step 1: Reduction to expressions based on events on $(0, T_1]$. We identify

$$\sup_{x\in D_3} \mathbb{E}\Big[e^{\theta|\lambda_\varepsilon \tau_y - \bar s|}\big(1 + |\mathbf{1}\{X^\varepsilon(\tau; x) \in U\} - \mathbf{1}\{W_{k^*} \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^c}(\phi)\}|\big)\Big] \\ \leqslant S_{11} + S_{12} + S_2 + S_3,$$

where

$$\begin{aligned}
S_{11} &:= \sum_{k=1}^\infty \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon|\tau_y - T_k|}\mathbf{1}\{\tau_y = T_k\} \cap \{s = T_k\} \\
&\qquad \big(1 + |\mathbf{1}\{X^\varepsilon(\tau; y) \in U\} - \mathbf{1}\{W_{k^*} \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^c}(\phi)\}|\big)\Big], \\
S_{12} &:= 2\sum_{k=1}^\infty \sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon|\tau_y - T_k|}\mathbf{1}\{\tau_y \in (T_{k-1}, T_k)\} \cap \{s = T_k\}\Big], \\
S_2 &:= 2\sum_{k=1}^\infty \sum_{\ell=1}^{k-1} \sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon|\tau_y - T_k|}\mathbf{1}\{\tau_y \in (T_{\ell-1}, T_\ell]\} \cap \{s = T_k\}\Big], \\
S_3 &:= 2\sum_{k=1}^\infty \sum_{\ell=k+1}^{\infty} \sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon|\tau_y - T_k|}\mathbf{1}\{\tau_y \in (T_{\ell-1}, T_\ell]\} \cap \{s = T_k\}\Big].
\end{aligned}$$

In the sequel we estimate the preceding expressions using the representations in (7.2) and the strong Markov property with respect to the $\mathcal{F}$-stopping times T_k.

The main term $\mathbf{S_{11}}$: This expression treats the case that the large jumps of X^ε and of the model are totally synchronized, that either both do not trigger an exit or both do. It is served first since it is the only one of order $O(1)_{\varepsilon\to 0}$ all other expressions are $o(1)_{\varepsilon\to 0}$. We denote the symmetric difference $E_1 \triangle E_2 := (E_1 \setminus E_2) \cup (E_2 \setminus E_1)$ for events E_1, E_2. In the sequel we repeatedly use strong

Markov estimates of the following type

$$
\begin{aligned}
&\mathbb{E}\Big[\mathbf{1}\big(\{\tau_y = T_k\} \cap \bigcap_{j=1}^{k-1} A_j^\diamond \cap B_k^\diamond\big) \\
&\qquad\qquad \big(1 + |\mathbf{1}\{X^\varepsilon(T_k; y) \in U\} - \mathbf{1}\{W_k \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^\mathsf{c}}(\phi)\}|\big)\Big] \\
&\leqslant \mathbb{E}\Big[\mathbf{1}\Big(\bigcap_{j=1}^{k-1} A_y^j \cap A_j^\diamond \cap B_y^k \cap B_k^\diamond\Big) \\
&\qquad\qquad \Big(1 + \mathbf{1}\{X^\varepsilon(T_k; y) \in U\} \triangle \{W_k \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^\mathsf{c}}(\phi)\}\Big)\Big] \\
&= \mathbb{E}\Big[\mathbf{1}\Big(\bigcap_{j=1}^{k-1} A_y^j \cap A_j^\diamond\Big) \\
&\quad\ \mathbb{E}\Big[\mathbf{1}\Big(B_y^k \cap B_k^\diamond\Big)\Big(1 + \mathbf{1}\{X^\varepsilon(T_k; y) \in U\} \triangle \{W_k \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^\mathsf{c}}(\phi)\}\Big) \mid \mathcal{F}_{T_{k-1}}\Big]\Big] \\
&= \mathbb{E}\Big[\mathbf{1}\Big(\bigcap_{j=1}^{k-1} A_y^j \cap A_j^\diamond\Big) \\
&\quad\ \mathbb{E}_{X^\varepsilon(T_{k-1}; y)}\Big[\mathbf{1}\big(B_\cdot^k \cap B_k^\diamond\big)\big(1 + \mathbf{1}\{X^\varepsilon(T_k; y) \in U\} \triangle \{W_k \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^\mathsf{c}}(\phi)\}\big)\Big]\Big] \\
&\leqslant \mathbb{E}\Big[\mathbf{1}\Big(\bigcap_{j=1}^{k-1} A_y^j \cap A_j^\diamond\Big)\Big] \\
&\qquad \sup_{y\in D_2} \mathbb{E}\Big[\mathbf{1}\big(B_y \cap B^\diamond\big)\big(1 + \mathbf{1}\{X^\varepsilon(T_k; y) \in U\} \triangle \{W_k \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^\mathsf{c}}(\phi)\}\big)\Big].
\end{aligned}
$$

The $(k-1)$-fold iteration of this argument yields

$$
\begin{aligned}
S_{11} &\leqslant \sum_{k=1}^{\infty} \sup_{y\in D_2} \mathbb{P}(A_y \cap A^\diamond)^{k-1} \\
&\qquad \sup_{y\in D_2} \mathbb{E}\Big[\mathbf{1}\big(B_y \cap B^\diamond\big)\big(1 + \mathbf{1}\{X^\varepsilon(T_k; y) \in U\} \triangle \{W_k \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^\mathsf{c}}(\phi)\}\big)\Big] \\
&= \frac{\sup_{y\in D_2} \mathbb{E}\Big[\mathbf{1}\big(B_y \cap B^\diamond\big)\big(1 + \mathbf{1}\{X^\varepsilon(T_k; y) \in U\} \triangle \{W_k \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^\mathsf{c}}(\phi)\}\big)\Big]}{1 - \sup_{y\in D_2} \mathbb{P}(A_y \cap A^\diamond)}.
\end{aligned}
\tag{7.4}
$$

The diagonal error $\mathbf{S_{12}}$: This term describes that the large jumps of the solution X^ε and the model are synchronized up to the last jump, where the error occurs. The remaining diagonal term is estimated as follows

$$S_{12} \leqslant 2\sum_{k=1}^{\infty} \sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon t_k}\mathbf{1}\Big(\bigcap_{j=1}^{k-1} (A_y^j \cap A_j^\diamond) \cap \big(C_y^k \cap B_k^\diamond\big)\Big)\Big].$$

For $k = 1$ we obtain the term

$$\sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\big(C_y \cap B^\diamond\big)\Big].$$

For $k \geqslant 2$ we take into account that the event A_y represents the non-exit from D_2. However, in order to dominate C_y we need initial values in D_3 (see estimate 7.11), for which we take the last large jump before. The error made by this jump to stay within D_3 instead of D_2 turns out to be negligible by Hypothesis (S.4). First obtain by the analogous strong Markov arguments arguments as for the term S_{11}

$$\begin{aligned}\sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon t_k}\mathbf{1}\Big(\bigcap_{j=1}^{k-1} (A_y^j \cap A_j^\diamond) \cap \big(C_y^k \cap B_k^\diamond\big)\Big)\Big]&\\ \leqslant \sup_{y\in D_2} \mathbb{P}(A_y \cap A^\diamond)^{k-2} \sup_{y\in D_2} \mathbb{E}\Big[\mathbf{1}(A_y^1 \cap A_1^\diamond)e^{\theta\lambda_\varepsilon t_2}\mathbf{1}(C_y^2 \cap B_2^\diamond)\Big]&,\end{aligned}$$

such that

$$\begin{aligned}S_{12} &\leqslant 2\sup_{D_3}\mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\big(C_y \cap B^\diamond\big)\Big]\\ &\qquad + 2\sum_{k=2}^{\infty}\sup_{D_2}\mathbb{P}(A_y\cap A^\diamond)^{k-2}\sup_{D_2}\mathbb{E}\Big[\mathbf{1}(A_y^1 \cap A_1^\diamond)e^{\theta\lambda_\varepsilon t_2}\mathbf{1}(C_y^2 \cap B_2^\diamond)\Big]\\ &\leqslant 2\sup_{y\in D_3}\mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\big(C_y\big)\Big] + \frac{2\sup_{y\in D_2}\mathbb{E}\Big[\mathbf{1}(A_y^1 \cap A_1^\diamond)e^{\theta\lambda_\varepsilon t_2}\mathbf{1}(C_y^2 \cap B_2^\diamond)\Big]}{1-\sup_{y\in D_2}\mathbb{P}(A_y\cap A^\diamond)}.\end{aligned} \tag{7.5}$$

The off-diagonal error of first kind $\mathbf{S_2}$: This term describes that the error that the jumps the model trigger the first exit before the those of the X^ε. The estimate of $\{\tau_y \in (T_{\ell-1}, T_\ell]\}$ and the representation of $\{s = T_k\}$ yield

$$\begin{aligned}S_2 \leqslant 2\sum_{k=1}^{\infty}\sum_{\ell=1}^{k-1}\sup_{y\in D_3}\mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_\ell+\cdots+t_k)}\mathbf{1}\Big(\bigcap_{j=\ell+1}^{k-1} A_j^\diamond \cap B_k^\diamond\Big)&\\ \mathbf{1}\Big(\bigcap_{j=1}^{\ell-1} (A_y^j \cap A_j^\diamond) \cap \big((B_y^\ell \cup C_y^\ell) \cap A_\ell^\diamond\big)\Big)\Big]&.\end{aligned}$$

For each of the summands $k \in \mathbb{N}$ and $\ell = 1$ we combine the mutual independence of the families $(T_k)_{k\in\mathbb{N}}$ and $(W_k)_{k\in\mathbb{N}}$ with the analogous strong Markov estimate

$$\begin{aligned}\sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_1+\cdots+t_k)}\mathbf{1}\Big(\bigcap_{j=2}^{k-1} A_j^\diamond \cap B_k^\diamond\Big)\mathbf{1}\Big(\big(B_y\cup C_y\big)\cap A^\diamond\Big)\Big] \\ \leqslant (1-\mathbb{P}(A^\diamond)) \sup_{y\in D_2} \mathbb{P}\big((B_y\cup C_y)\cap A^\diamond\big)\mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\Big]^k \mathbb{P}(A^\diamond)^{k-1}.\end{aligned}$$

For $k\in\mathbb{N}$ and $k-1\geqslant \ell\geqslant 2$ we obtain

$$\begin{aligned}&\sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_\ell+\cdots+t_k)}\mathbf{1}\Big(\bigcap_{j=\ell+1}^{k-1} A_j^\diamond \cap B_k^\diamond\Big)\mathbf{1}\Big(\bigcap_{j=1}^{\ell-1}\big(A_y^j\cap A_j^\diamond\big)\cap\big(B_y^\ell\cup C_y^\ell\big)\cap A_\ell^\diamond\Big)\Big] \\ &\leqslant (1-\mathbb{P}(A^\diamond)) \sup_{y\in D_2} \mathbb{P}\big((A_y^1\cap A_1^\diamond)\cap\big((B_y^2\cup C_y^2\big)\big)\ \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]^{k-1}\mathbb{P}(A^\diamond)^{k-2} \\ &\qquad\qquad \sup_{y\in D_2}\mathbb{P}(A_y\cap A^\diamond)^{\ell-2}\Big(\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]\mathbb{P}(A^\diamond)\Big)^{-(\ell-2)}.\end{aligned}$$

Monotonicity yields

$$\sup_{y\in D_2}\mathbb{P}(A_y\cap A^\diamond)\leqslant \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]\mathbb{P}(A^\diamond),$$

such that

$$\sum_{\ell=2}^{k-1}\sup_{y\in D_2}\mathbb{P}(A_y\cap A^\diamond)^{\ell-2}\Big(\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]\mathbb{P}(A^\diamond)\Big)^{-(\ell-2)}\leqslant k-2,$$

and hence

$$\begin{aligned}S_2 &\leqslant 2\mathbb{P}(B^\diamond)\sup_{y\in D_3}\mathbb{P}\big((B_y\cup C_y)\cap A^\diamond\big)\ \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]\sum_{k=1}^\infty \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]^{k-1}\mathbb{P}(A^\diamond)^{k-1} \\ &\quad + 2\mathbb{P}(B^\diamond)\sup_{y\in D_2}\mathbb{P}\big((A_y^1\cap A_1^\diamond)\cap\big((B_y^2\cup C_y^2)\cap A_2^\diamond\big)\big) \\ &\qquad\qquad \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]\sum_{k=1}^\infty (k-1)\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]^{k-2}\mathbb{P}(A^\diamond)^{k-2} \\ &= 2\mathbb{P}(B^\diamond)\frac{\sup_{y\in D_3}\mathbb{P}\big((B_y\cup C_y)\cap A^\diamond\big)}{\big(1-\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]\mathbb{P}(A^\diamond)\big)} \\ &\qquad + 2\mathbb{P}(B^\diamond)\frac{\sup_{y\in D_2}\mathbb{P}\big((A_y^1\cap A_1^\diamond)\cap\big((B_y^2\cup C_y^2)\cap A_2^\diamond\big)\big)}{\big(1-\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]\mathbb{P}(A^\diamond)\big)^2}. \qquad (7.6)\end{aligned}$$

The off-diagonal error of second kind $\mathbf{S_3}$: This term describes that the error that the jumps of X^ε trigger the first exit before the model does. Due to the doubly infinite summation this turns out the most cumbersome case:

$$S_3 \leqslant 2 \sum_{k=1}^{\infty} \sum_{\ell=k+1}^{\infty} \sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_k+\cdots+t_\ell)} \mathbf{1}\Big(\bigcap_{j=1}^{k-1} \big(A_y^j \cap A_j^\diamond\big)\Big)\mathbf{1}\Big(A_y^k \cap B_k^\diamond\Big)\mathbf{1}\Big(\bigcap_{j=k+1}^{\ell-1} A_y^j \cap \big(B_y^\ell \cup C_y^\ell\big)\Big)\Big].$$

The strong Markov estimates as in S_{11} yield for $\ell = k+1$ the estimate

$$\sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_k+t_{k+1})}\mathbf{1}\Big(\bigcap_{j=1}^{k-1} \big(A_y^j \cap A_j^\diamond\big)\Big)\mathbf{1}\big(A_y^k \cap B_k^\diamond\big)\mathbf{1}\big(B_y^k \cup C_y^k\big)\Big]$$
$$\leqslant \sup_{y\in D_2} \mathbb{P}\big(A_y \cap A^\diamond\big)^{k-1} \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_1+t_2)}\Big(\big(A_y^1 \cap B_1^\diamond\big) \cap \big(B_y^2 \cup C_y^2\big)\Big)\Big].$$

For $\ell \geqslant k+2$ we obtain the summands

$$\sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_k+\cdots+t_\ell)}\mathbf{1}\Big(\bigcap_{j=1}^{k-1} \big(A_y^j\cap A_j^\diamond\big)\Big)\mathbf{1}\Big(A_y^k\cap B_k^\diamond\Big)\mathbf{1}\Big(\bigcap_{j=k+1}^{\ell-1} A_y^j\cap\big(B_y^\ell\cup C_y^\ell\big)\Big)\Big]$$
$$\leqslant \sup_{y\in D_2} \mathbb{P}\big(A_y \cap A^\diamond\big)^{k-1} \sup_{y\in D_2} \mathbb{P}\big(A_y \cap B^\diamond\big) \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y)\Big]^{\ell-1-k}$$
$$\sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_2+t_1)}\mathbf{1}(A_y^1)\mathbf{1}(B_y^2 \cup C_y^2)\Big].$$

Assuming $\sup_{y\in D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y)\big] < 1$ for $\varepsilon\in(0,\varepsilon_0]$ for $\varepsilon_0\in(0,1]$ small enough which we verify in Step 3 we obtain

$$
\begin{aligned}
&S_3/2\\
&\leqslant \sum_{k=1}^{\infty}\sup_{D_2}\mathbb{P}\big(A_y\cap A^\diamond\big)^{k-1}\bigg(\sup_{D_2}\mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_1+t_2)}\mathbf{1}\Big(A_y^1\cap B_1^\diamond\cap\big(B_y^2\cup C_y^2\big)\Big)\Big]\\
&\quad+\sup_{D_2}\mathbb{P}\big(A_y\cap B^\diamond\big)\mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_1+t_2)}\mathbf{1}\Big(\big(A_y^1\cap B_1^\diamond\big)\cap\big(B_y^2\cup C_y^2\big)\Big)\Big]\\
&\qquad\qquad\qquad\qquad\qquad\qquad\sum_{\ell=k+2}^{\infty}\sup_{D_2}\mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y)\Big]^{\ell-1-k}\bigg)\\
&=\bigg(\sup_{D_2}\mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_1+t_2)}\mathbf{1}\Big(A_y^1\cap B_1^\diamond\cap\big(B_y^2\cup C_y^2\big)\Big)\Big]\\
&\qquad+\frac{\sup_{D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y)\big]\sup_{D_2}\mathbb{P}(A_y\cap B^\diamond)}{1-\sup_{D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y)\big]}\\
&\qquad\cdot\sup_{y\in D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon(t_1+t_2)}\mathbf{1}\big(A_y^1\cap B_1^\diamond\cap\big(B_y^2\cup C_y^2\big)\big)\big]\bigg)\Big/\Big(1-\sup_{D_2}\mathbb{P}(A_y\cap A^\diamond)\Big)\\
&\leqslant\frac{\sup_{D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon(t_1+t_2)}\mathbf{1}\big(A_y^1\cap B_1^\diamond\cap\big(B_y^2\cup C_y^2\big)\big)\big]}{1-\sup_{D_2}\mathbb{P}(A_y\cap A^\diamond)}\\
&\qquad\qquad\qquad\cdot\bigg(1+\frac{\sup_{D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y)\big]\sup_{D_2}\mathbb{P}(A_y\cap B^\diamond)}{1-\sup_{D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y)\big]}\bigg).
\end{aligned}
\tag{7.7}
$$

Step 2: Precise estimates of the events on $(0, T_1]$**.** First note the following inclusions, which are valid by construction

$$\begin{aligned}
&\mathcal{J}^{D_2}(B_{\gamma_\varepsilon}(\phi)) \subseteq \mathcal{J}^{D}(\phi)\\
&\mathcal{J}^{D_2^{\mathsf{c}}}(B_{\gamma_\varepsilon}(\phi)) \subseteq \mathcal{J}^{D_3^{\mathsf{c}}}(\phi) \quad \text{and} \quad D_3^{\mathsf{c}} \subseteq D^{\mathsf{c}} \cup (D \setminus D_3(\gamma_\varepsilon)) \cup (\mathcal{U}^{\chi - 3\gamma_\varepsilon})^{\mathsf{c}}\\
&\mathcal{J}^{D_2 \setminus D_3}(B_{\gamma_\varepsilon}(\phi)) \subseteq \mathcal{J}^{D \setminus D_4}(\phi).
\end{aligned}$$

We prove the following event estimates.

Claim 1: For $y \in D_2$ we have

$$\mathbf{1}(A_y) \leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^D(\phi)\} + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|\} + \mathbf{1}(E_y^{\mathsf{c}}), \tag{7.8}$$

$$\begin{aligned}
\mathbf{1}(A_y^-) &\leqslant \mathbf{1}(A_y)\\
&\leqslant \mathbf{1}(A_y^-) + \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D \setminus D_4}(\phi)\} + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|\}\\
&\qquad + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}),
\end{aligned} \tag{7.9}$$

$$\begin{aligned}
\mathbf{1}(B_y) &\leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D_3^{\mathsf{c}}}(\phi)\} + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|\}\\
&\qquad + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}).
\end{aligned} \tag{7.10}$$

The stronger condition $y \in D_3$ implies

$$\mathbf{1}(C_y) \leqslant \mathbf{1}(E_y^{\mathsf{c}}). \tag{7.11}$$

The first estimates of $\mathbf{1}(A_y)$ and $\mathbf{1}(B_y)$ since for $y \in D_2$ we have

$$\begin{aligned}
&\mathbf{1}(A_y) \leqslant \mathbf{1}(A_y)\mathbf{1}(E_y) + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\leqslant \mathbf{1}(A_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\} + \mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\leqslant \mathbf{1}(A_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 \geqslant \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} \\
&\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} \\
&\qquad + \mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\leqslant \mathbf{1}(A_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 \geqslant \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} \\
&\qquad + \mathbf{1}(A_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} \\
&\qquad + \mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\leqslant \mathbf{1}\Big(\bigcap_{y \in B_{\gamma_\varepsilon}(\phi)} \{\varepsilon W_1 \in \mathcal{J}^{D_2}(y)\}\Big)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\} \\
&\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} \\
&\qquad + \mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\
&= \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D_2}(B_{\gamma_\varepsilon}(\phi))\} - \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D_2}(B_{\gamma_\varepsilon}(\phi))\}\mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} \\
&\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + \mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D}(\phi)\} + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + \mathbf{1}(E_y^{\mathsf{c}}).
\end{aligned} \tag{7.12}$$

In addition, the reduced event A_y^- satisfies

$$\begin{aligned}
&\mathbf{1}(A_y^-) \leqslant \mathbf{1}(A_y) \\
&\leqslant \mathbf{1}(A_y^-) + \mathbf{1}\{Y^\varepsilon(t;y) \in D_2 \text{ for } t \in [0,T_1] \text{ and } Y^\varepsilon(T_1;y) + \varepsilon W_1 \in D_2 \setminus D_3\} \\
&\leqslant \mathbf{1}(A_y^-) + \mathbf{1}\{Y^\varepsilon(t;y) \in D_2 \text{ for } t \in [0,T_1] \text{ and } Y^\varepsilon(T_1;y) + \varepsilon W_1 \in D_2 \setminus D_3\} \\
&\qquad \cdot \mathbf{1}(E_y) + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\leqslant \mathbf{1}(A_y^-) + \mathbf{1}\{Y^\varepsilon(t;y) \in D_2 \text{ for } t \in [0,T_1] \text{ and } Y^\varepsilon(T_1;y) + \varepsilon W_1 \in D_2 \setminus D_3\} \\
&\qquad \cdot \mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\} + \mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\leqslant \mathbf{1}(A_y^-) + \mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\qquad + \mathbf{1}\{Y^\varepsilon(t;y) \in D_2 \text{ for } t \in [0,T_1] \text{ and } Y^\varepsilon(T_1;y) + \varepsilon W_1 \in D_2 \setminus D_3\} \\
&\qquad \cdot \mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 \geqslant \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} \\
&\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} \\
&\leqslant \mathbf{1}(A_y^-) + \mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\
&\qquad + \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D_2 \setminus D_3}(B_{\gamma_\varepsilon}(\phi))\} + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\}
\end{aligned}$$

Analogously, we obtain

$$
\begin{aligned}
&\mathbf{1}(B_y)\\
&\leqslant \mathbf{1}(B_y)\mathbf{1}(E_y) + \mathbf{1}(E_y^{\mathsf{c}})\\
&= \mathbf{1}(B_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\} + \mathbf{1}(B_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}})\\
&= \mathbf{1}(B_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 \geqslant \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\}\\
&\quad + \mathbf{1}(B_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\}\\
&\quad + \mathbf{1}(B_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 \geqslant \varepsilon^{\Upsilon_1}\}\\
&\quad + \mathbf{1}(B_y)\mathbf{1}(E_y)\mathbf{1}\{|\varepsilon W_1| \leqslant \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}})\\
&\leqslant \mathbf{1}\{Y^\varepsilon(T_1;y) \in B_{\gamma_\varepsilon}(\phi)\}\mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D_2^{\mathsf{c}}}(Y^\varepsilon(T_1;y))\}\\
&\quad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + 0 + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}})\\
&\leqslant \mathbf{1}\Big(\bigcap_{y\in B_{\gamma_\varepsilon}(\phi)} \{\varepsilon W_1 \in \mathcal{J}^{D_2^{\mathsf{c}}}(y)\}\Big)\\
&\quad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}})\\
&\leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D_2^{\mathsf{c}}}(B_{\gamma_\varepsilon}(\phi))\}\\
&\quad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}})\\
&\leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D_3^{\mathsf{c}}}(\phi)\}\\
&\quad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}})
\end{aligned}
\tag{7.13}
$$

For $y \in D_3$ the definition of the reduced domain of attraction yields that the event E_y implies that $Y^\varepsilon(t;y) \in B_{\frac{1}{2}\gamma_\varepsilon}(u(t;y)) \subset D_2(\gamma_\varepsilon)$ for all $t \in [0, T_1]$. That is $E_y \subset C_y^c$.

Claim 2: For $y \in D_2$ and $U \in \mathcal{B}(H)$ we have

$$
\mathbf{1}(A_y \cap B^\diamond) \leqslant \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + \mathbf{1}(E_y^{\mathsf{c}}) \tag{7.14}
$$

$$
\begin{aligned}
\mathbf{1}(B_y \cap A^\diamond) \leqslant{}& \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D\setminus D_3(\delta)}(\phi)\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}})\\
&+ \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\}
\end{aligned}
\tag{7.15}
$$

$$
\begin{aligned}
&\mathbf{1}(B_y \cap B^\diamond)\mathbf{1}\big(\{X(T_1;y) \in U\} \triangle \{\varepsilon W_1 \in \mathcal{J}^U(\phi)\}\big)\\
&\leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{B_{(\ell+1)\gamma_\varepsilon}(\partial U)\cap D^{c}}(\phi)\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\}\\
&\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + \mathbf{1}(E_y^{\mathsf{c}}).
\end{aligned}
\tag{7.16}
$$

The estimate (7.14) is a direct consequence of (7.8) in Claim 1. Using (7.10) we obtain

$$\begin{aligned}\mathbf{1}(B_y \cap A^\diamond) \leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D\setminus D_4(\delta)}(\phi)\} + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|\} \\ + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}).\end{aligned}$$

For (7.16) the inclusion $\mathcal{J}^U(B_{\gamma_\varepsilon}(\phi)) \subseteq \mathcal{J}^U(\phi)$, the Lipschitz continuity of $y \mapsto y + z$ with Lipschitz constant $1 + \ell$ for any z and (7.10) yield

$$\begin{aligned}&\mathbf{1}(B_y \cap B^\diamond)\mathbf{1}\{X(T_1; y) \in U\} \triangle \{\varepsilon W_1 \in \mathcal{J}^U\} \\ &\leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D^{\mathsf{c}}}(\phi) \cap \big(\mathcal{J}^U(B_{\gamma_\varepsilon}(\phi)) \triangle \mathcal{J}^U(\phi)\big\} \\ &\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\ &\leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D^{\mathsf{c}}}(\phi) \cap \big(\mathcal{J}^U(\phi) \setminus \mathcal{J}^U(B_{\gamma_\varepsilon}(\phi))\big\} \\ &\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}) \\ &\leqslant \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{B_{(\ell+1)\gamma_\varepsilon}(\partial U) \cap D^{\mathsf{c}}}(\phi)\} \\ &\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}}).\end{aligned}$$

Step 3: Estimates of the resulting expressions: Step 2 provides the estimates to dominate respectively the term S_{11} by (7.4), S_{12} by (7.5), S_2 by (7.6) and S_3 by (7.7).

Event A_y: Due to (7.8) there is a constant $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0, \varepsilon_0]$ implies

$$\begin{aligned}&\sup_{y \in D_2} \mathbb{P}(A_y \cap A^\diamond) \\ &\leqslant \sup_{y \in D_2} \mathbb{E}\big[e^{\theta \lambda_\varepsilon T_1}\mathbf{1}(A_y)\big] \\ &\leqslant 1 - \frac{(1-\theta)\lambda_\varepsilon}{\beta_\varepsilon - \theta\lambda_\varepsilon} + \frac{\beta_\varepsilon^2}{\beta_\varepsilon - \theta\lambda_\varepsilon}((\kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|)\frac{\nu(\varepsilon^{\Upsilon_1 - 1}B_1^c(0))}{\beta_\varepsilon} + \frac{C_2\beta_\varepsilon}{\beta_\varepsilon - \theta\lambda_\varepsilon} 2e^{-\frac{1}{5\gamma_\varepsilon}} \\ &\leqslant 1 - \Big(\frac{(1-\theta)}{1 - \theta\frac{\lambda_\varepsilon}{\beta_\varepsilon}} - C'\Big)\frac{\lambda_\varepsilon}{\beta_\varepsilon} \\ &\leqslant 1 - \frac{1-\theta}{2}\frac{\lambda_\varepsilon}{\beta_\varepsilon} \\ &\leqslant 1 - (1 - C')\frac{\lambda_\varepsilon}{\beta_\varepsilon} < 1. \end{aligned} \tag{7.17}$$

Event B_y: Using that ν is regularly varying and the initial choice of ε_0 in (7.3) we obtain

$$\lim_{\varepsilon\to 0} \frac{\nu(\frac{1}{\varepsilon}\mathcal{J}^{D\setminus D_4(\gamma_\varepsilon)}(\phi))}{\beta_\varepsilon}\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)^{-1} \leqslant \lim_{\varepsilon\to 0} \frac{\nu(\frac{1}{\varepsilon}\mathcal{J}^{D\setminus D_4(\gamma_{\varepsilon_0})}(\phi))}{\nu(\frac{1}{\varepsilon}\mathcal{J}^{D}(\phi))}$$
$$= \frac{\mu(\mathcal{J}^{D\setminus D_4(\gamma_{\varepsilon_0})}(\phi))}{\mu(\mathcal{J}^{D}(\phi))} \leqslant C'. \qquad (7.18)$$

We use estimate (7.10) yields a constant $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0,\varepsilon_0]$ implies

$$\begin{aligned}
&\sup_{y\in D_2} \mathbb{P}(B_y\cap B^\diamond) \leqslant \sup_{y\in D_2} \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(B_y)\big]\\
&\leqslant \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\big]\mathbb{P}(W_1\in \frac{1}{\varepsilon}\mathcal{J}^{D^{\mathsf{c}}}(\phi)) + \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\{T_1<\kappa_0+\kappa_1|\ln(\gamma_\varepsilon)|\}\big]\frac{\nu(\varepsilon^{\Upsilon_1-1}B_1^c(0))}{\beta_\varepsilon}\\
&\qquad + \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\{T_1<\varepsilon^{\Upsilon_1}\}\big] + \sup_{y\in D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(E_y^{\mathsf{c}})\big]\\
&\leqslant \frac{\beta_\varepsilon}{\beta_\varepsilon-\theta\lambda_\varepsilon}\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon} + \big(1-e^{-(\beta_\varepsilon-\theta\lambda_\varepsilon)(\kappa_0+\kappa_1|\ln(\gamma_\varepsilon)|)}\big)\frac{\nu(\varepsilon^{\Upsilon_1-1}B_1^c(0))}{\beta_\varepsilon} + \beta_\varepsilon\varepsilon^{\Gamma_1}\\
&\qquad + \frac{\beta_\varepsilon-\theta\lambda_\varepsilon}{\beta_\varepsilon}2e^{-(5\gamma_\varepsilon)^{-1}}\Big)\\
&\leqslant (1+2C')\frac{\lambda_\varepsilon}{\beta_\varepsilon}. \qquad (7.19)
\end{aligned}$$

Event C_y: Estimate (7.11) yields a constant $\varepsilon_0 \in (0,1]$ sufficiently small such that $\varepsilon \in (0,\varepsilon_0]$ implies

$$\begin{aligned}
\sup_{y\in D_3}\mathbb{P}(C_y) &\leqslant \sup_{y\in D_3}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\big(C_y\big)\big]\\
&\leqslant \sup_{y\in D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(E_y^{\mathsf{c}})\big]\\
&= 2\exp(-(5\gamma_\varepsilon)^{-1}) \leqslant C'\frac{\lambda_\varepsilon}{\beta_\varepsilon}. \qquad (7.20)
\end{aligned}$$

Events $A_y \cap B^\diamond$ and $B_y\cap A^\diamond$: By (7.14) we obtain directly $\varepsilon_0\in(0,1]$ such that for $\varepsilon\in(0,\varepsilon_0]$ we obtain with the analogous calculations

$$\sup_{y\in D_2}\mathbb{P}(A_y\cap B^\diamond) \leqslant \sup_{y\in D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y\cap B^\diamond)\big] \leqslant C'\frac{\lambda_\varepsilon}{\beta_\varepsilon}. \qquad (7.21)$$

and with the help of (7.15), the regular variation and (7.3) for $\varepsilon\in(0,\varepsilon_0]$

$$\sup_{y\in D_2}\mathbb{P}(B_y\cap A^\diamond) \leqslant \sup_{y\in D_2}\mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(B_y\cap A^\diamond)\big] \leqslant C'\frac{\lambda_\varepsilon}{\beta_\varepsilon}. \qquad (7.22)$$

Exit events after a non-exit first large jump: First note the trivial inclusions

$$\begin{aligned}
&\sup_{y\in D_2} \mathbb{P}\big(A_y \cap \big(B_y^2 \cup C_y^2\big) \cap A_2^\diamond\big)\\
&\leqslant \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon t_2}\mathbf{1}\big(A_y \cap \big(B_y^2 \cup C_y^2\big) \cap A_2^\diamond\big)\Big]\\
&\leqslant \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon (t_1+t_2)}\mathbf{1}\big(A_y \cap \big(B_y^2 \cup C_y^2\big) \cap A_2^\diamond\big)\Big]\\
&\leqslant \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon (t_1+t_2)}\big(A_y \cap B_y^2 \cap A_2^\diamond\big)\Big] + \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon (t_1+t_2)}\mathbf{1}\big(A_y \cap C_y^2\big)\Big].
\end{aligned}$$

The analogous estimate holds obviously true without the presence of $A_2^\diamond$. With the help of the strong Markov property and (7.22) we obtain $\varepsilon_0 \in (0,1]$ such that for $\varepsilon \in (0,\varepsilon_0]$

$$\begin{aligned}
&\sup_{y\in D_2} \mathbb{E}\big[e^{\theta\lambda_\varepsilon (t_1+t_2)}\big(A_y \cap B_y^2 \cap A_2^\diamond\big)\big]\\
&\qquad \leqslant \sup_{y\in D_2} \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\big(A_y\big)\big] \sup_{y\in D_2} \mathbb{E}\big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\big(B_y \cap A^\diamond\big)\big] \leqslant C' \frac{\lambda_\varepsilon}{\beta_\varepsilon}.
\end{aligned} \tag{7.23}$$

In case of $A_2^\diamond$ being replaced by $B_1^\diamond$ we obtain for $\varepsilon \in (0,\varepsilon_0]$

$$\begin{aligned}
&\sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon (t_1+t_2)}\big(A_y \cap B^\diamond \cap B_y^2\big)\Big]\\
&\qquad \leqslant \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\big(A_y \cap B^\diamond\big)\Big] \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\big(B_y\big)\Big]\\
&\qquad \leqslant C'(1+C')\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)^2.
\end{aligned} \tag{7.24}$$

Keeping in mind the convention $A_y = A_y^1$ and $A_y^- = A_y^{1-}$ the estimates (7.9), (7.17), (7.18) and (7.20) yield a constant $\varepsilon_0 \in (0,1]$ sufficiently small such that for $\varepsilon \in (0, \varepsilon_0]$

$$
\begin{aligned}
&\sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_1+t_2)}\mathbf{1}\big(A_y \cap C_y^2\big)\Big]\\
&\leqslant \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon(t_1+t_2)}\Big(\mathbf{1}\big(A_y^-\big)\mathbf{1}\big(C_y^2\big) + \mathbf{1}\{\varepsilon W_1 \in \mathcal{J}^{D\setminus D_4}(\phi)\}\\
&\qquad\qquad + \mathbf{1}\{|\varepsilon W_1| > \varepsilon^{\Upsilon_1}\}\mathbf{1}\{t_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\} + \mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\} + \mathbf{1}(E_y^{\mathsf{c}})\Big)\Big]\\
&\leqslant \sup_{y\in D_2} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(A_y)\Big] \sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(C_y)\Big]\\
&\qquad + \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\Big]\Big(\mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\Big]\mathbb{P}(\varepsilon W_1 \in \mathcal{J}^{D\setminus D_4}(\phi))\\
&\qquad\qquad + \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\{T_1 < \kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|\}\Big]\mathbb{P}(|\varepsilon W_1| > \varepsilon^{\Upsilon_1})\\
&\qquad\qquad + \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}\{T_1 < \varepsilon^{\Upsilon_1}\}\Big] + \sup_{y\in D_3} \mathbb{E}\Big[e^{\theta\lambda_\varepsilon T_1}\mathbf{1}(E_y^{\mathsf{c}})\Big]\Big)\\
&\leqslant \Big(\frac{\beta_\varepsilon}{\beta_\varepsilon - \theta\lambda_\varepsilon}\Big)\Big(C'\frac{\beta_\varepsilon}{\beta_\varepsilon - \theta\lambda_\varepsilon}\frac{\lambda_\varepsilon}{\beta_\varepsilon} + \frac{\beta_\varepsilon}{\beta_\varepsilon - \theta\lambda_\varepsilon}\beta_\varepsilon(\kappa_0 + \kappa_1|\ln(\gamma_\varepsilon)|)\frac{\nu(\varepsilon^{\Upsilon_1-1}B_1^c(0))}{\beta_\varepsilon}\\
&\qquad\qquad + \beta_\varepsilon\varepsilon^{\Upsilon_1} + C'\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)\\
&\leqslant \Big(1 + \theta\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)4C'\frac{\lambda_\varepsilon}{\beta_\varepsilon} \leqslant 6C'\frac{\lambda_\varepsilon}{\beta_\varepsilon}.
\end{aligned}
\tag{7.25}
$$

Step 4: Concluding estimates of the sums of (7.4):

Estimate of $\mathbf{S_{11}}$: Since $m^\pm(\partial U) = \mu(\mathcal{J}^{\partial U}(\phi^\pm)) = 0$ by assumption and the regular variation of ν by Hypothesis (S.1) we have for any $\varepsilon_0 \in (0, 1]$ and $\varepsilon \in (0, \varepsilon_0]$ that

$$
\begin{aligned}
&\lim_{\varepsilon\to 0} \mathbb{P}\big(\varepsilon W_1 \in \mathcal{J}^{B_{(\ell+1)\gamma_\varepsilon}(\partial U)\cap D^{\mathsf{c}}}(\phi)\big)\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)^{-1}\\
&\qquad = \lim_{\varepsilon\to 0} \frac{\nu\big(\frac{1}{\varepsilon}\mathcal{J}^{B_{(\ell+1)\gamma_\varepsilon}(\partial U)\cap D^{\mathsf{c}}}(\phi)\big)}{\nu(\rho^\varepsilon B_1^{\mathsf{c}}(0))}\frac{\nu(\rho^\varepsilon B_1^{\mathsf{c}}(0))}{\nu\big(\frac{1}{\varepsilon}\mathcal{J}^{D^{\mathsf{c}}}(\phi)\big)}\\
&\qquad\qquad \leqslant \frac{\mu\big(\mathcal{J}^{B_{(\ell+1)\delta}(\partial U)\cap D^{\mathsf{c}}}(\phi)\big)}{\mu(\mathcal{J}^{D^{\mathsf{c}}}(\phi))} \leqslant \frac{C'}{5}.
\end{aligned}
\tag{7.26}
$$

Hence (7.16), (7.19) and (7.26) yield

$$\begin{aligned}
&\sup_{y\in D_2} \mathbb{E}\Big[\mathbf{1}\big(B_y \cap B^\diamond\big)\big(1+\mathbf{1}\{X^\varepsilon(T_1;y)\in U\} \,\triangle\, \{\varepsilon W_1 \in \mathcal{J}^{U\cap D^c}\}\Big] \\
&\leqslant \mathbb{P}(\varepsilon W_1 \in \mathcal{J}^{D^c}(\phi)) + \mathbb{P}(\varepsilon W_1 \in \mathcal{J}^{B_{(\ell+1)\gamma_\varepsilon}(\partial U)\cap D^c}(\phi)) \\
&\qquad + 2\mathbb{P}(T_1 < \kappa_0 + \kappa_1 |\ln(\gamma_\varepsilon)|)\mathbb{P}(|\varepsilon W_1| > \varepsilon^{\Upsilon_1}) + \mathbb{P}(T_1 < \varepsilon^{\Upsilon_1}) + 2\sup_{y\in D_2}\mathbb{P}\big(E_y^c\big) \\
&\leqslant (1+C')\frac{\lambda_\varepsilon}{\beta_\varepsilon},
\end{aligned}$$

such that for $\varepsilon \in (0,\varepsilon_0]$ the sum S_{11} given in (7.4) satisfies

$$S_{11} \leqslant \frac{1+C'}{1-C'} \leqslant 1+4C'. \tag{7.27}$$

S_{12} given by (7.5): By (7.17), (7.20) and (7.25) the sum S_{12} given in (7.5) satisfies for $\varepsilon \in (0,\varepsilon_0]$ the estimate

$$S_{12} \leqslant \frac{6C'}{1-C'} \leqslant 12C'. \tag{7.28}$$

S_2 given by (7.6): Using the estimates (7.20), (7.22), (7.23) and (7.25) the sum S_2 given in (7.6) satisfies for $\varepsilon \in (0,\varepsilon_0]$ the estimate

$$\begin{aligned}
S_2 &\leqslant \frac{2(1+3C')C'}{1-\frac{\beta_\varepsilon}{\beta_\varepsilon-\theta\lambda_\varepsilon}(1-\frac{\lambda_\varepsilon}{\beta_\varepsilon})}\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)^2 + 2\frac{6C'}{\big(1-\frac{\beta_\varepsilon}{\beta_\varepsilon-\theta\lambda_\varepsilon}(1-\frac{\lambda_\varepsilon}{\beta_\varepsilon})\big)^2}\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)^2 \\
&= 2(1+3C')C'\frac{\beta_\varepsilon-\theta\lambda_\varepsilon}{(1-\theta)\lambda_\varepsilon}\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)^2 + 12C'\Big(\frac{\beta_\varepsilon-\theta\lambda_\varepsilon}{(1-\theta)\lambda_\varepsilon}\Big)^2\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big)^2 \\
&\leqslant C' + 13C' = 14C'.
\end{aligned} \tag{7.29}$$

S_3 given by (7.7): Using (7.17), (7.24), (7.25) and (7.21) we obtain $\varepsilon_0 \in (0,\varepsilon_0]$ such that $\varepsilon \in (0,\varepsilon_0]$ implies

$$S_3 \leqslant \Big(\frac{2C'(1+C')}{(1-C')}\Big(\frac{\lambda_\varepsilon}{\beta_\varepsilon}\Big) + \frac{6C'}{1-C'}\Big)\Big(1+\frac{C'}{(1-C')}\Big) \leqslant 26C'. \tag{7.30}$$

We finally collect (7.27 - 7.30) and conclude that there is $\varepsilon_0 \in (0,1]$ such that $\varepsilon \in (0,\varepsilon_0]$ implies

$$\sup_{x\in D_2}\mathbb{E}\Big[e^{\theta\lambda_\varepsilon|\tau_x - s(\varepsilon)|}\big(1+|\mathbf{1}\{X^\varepsilon(\tau_x;x)\in U\} - \mathbf{1}\{W_{k^*(\varepsilon)} \in \frac{1}{\varepsilon}\mathcal{J}^{U\cap D^c}(\phi)\}|\big)\Big] \leqslant 1+56C'.$$

Since $C' \in (0, \frac{1-\theta}{2})$ was chosen arbitrary this finishes the proof.

□

Appendix: Conventions and background material

Basics in Probability

In this section we provide some basic notation and result for the completeness of the text. We refer to a proper probability course and classicale text books in probability.

Kolmogorov's probability space:

- $(\Omega, \mathcal{A}, \mathbb{P})$ **probability space** (in general abstract)
- $\Omega \neq \emptyset$ is called the **sample space** and elements $\omega \in \Omega$ are called samples
- Denote by 2^Ω the collection of all subsets of Ω and $\mathcal{A} \subset 2^\Omega$ is the **sigma algebra** of events.

  1. $\Omega \in \mathcal{A}$,
  2. $A \in \mathcal{A} \Rightarrow A^c := \Omega \setminus A \in \mathcal{A}$ and
  3. for any sequence $(A_n)_{n\in\mathbb{N}}$ with $A_n \in \mathcal{A}$ we have $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$)

  The elements of $\mathcal{A}$ are called **events**.

- A **probability measure** or **distribution** $\mathbb{P}$ is a mapping $\mathbb{P} : \mathcal{A} \longrightarrow [0,1]$ satisfying $\mathbb{P}(\emptyset) = 0$ and for any disjoint family $(A_n)_{n\in\mathbb{N}}$ of events $A_n \in \mathcal{A}$ we have
$$\mathbb{P}(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n).$$

Independence of events:

- Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a family $(A_n)_{n\in\mathbb{N}}$ of events is called **independent** w.r.t. $\mathbb{P}$ if for any $N \in \mathbb{N}$ and any choice $N \in \mathbb{N}$, $k_1, \dots, k_N \in \mathbb{N}$, $k_i \neq k_j$, we have

$$\mathbb{P}(A_{k_1} \cap \dots \cap A_{k_N}) = \prod_{\ell=1}^{N} \mathbb{P}(A_{k_\ell}).$$

The size of Probability spaces: Borel-Cantelli

Lemma 7.2.2 (Borel-Cantelli-Lemma)**.**
On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we consider a sequence of events $(A_i)_{i\in\mathbb{N}}$, $A_i \in \mathcal{A}$.

1. *If $\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty$, then*

$$\mathbb{P}(\limsup_{i\to\infty} A_i) = 0,$$

where

$$\limsup_{i\to\infty} A_i = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_i = \{\omega \in \Omega \mid \#\{\omega \in A_i\} = \infty\}.$$

"all elements ω in an infinite overlap of the A_i"

2. *If the family $(A_i)_{i\in\mathbb{N}}$ is independent and $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$, then*

$$\mathbb{P}(\liminf_{i\to\infty} A_i) = 1,$$

where

$$\liminf_{i\to\infty} A_i = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_i = \{\omega \in \Omega \mid \#\{\omega \notin A_i\} < \infty\}.$$

"all elements ω in an eventually complete overlap of the A_i"

Random variables and vectors:

- A **random variable** $X : \Omega \to \mathbb{R}$ is a $(\mathcal{A}, \mathcal{B}(\mathbb{R}))$ measurable mapping, that is the preimage of the sigma algebra (which itself is a sigma algebra!) $X^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{A}$ as a sub-sigma algebra.

Given a separable Hilbert space $(H, |\cdot|, \langle\cdot,\cdot\rangle)$ equipped with its Borel sigma algebra $\mathcal{B}(H)$.

- A **random vector** $X : \Omega \to H$ is a $(\mathcal{A}, \mathcal{B}(H))$-measurable mapping.
- The **distribution of a random vector** $X : \Omega \to H$, denoted by $\mathbb{P}_X$, is a distribution on $\mathcal{B}(H)$ defined by

$$\mathbb{P}_X(A) := \mathbb{P}(X^{-1}(A)) = \mathbb{P}(X \in A), A \in \mathcal{B}(\mathbb{R}).$$

- The **expectation of a random vector** $X : \Omega \to H$ has is defined as

$$\mathbb{E}[X] := \int_\Omega X(\omega) d\mathbb{P}(\omega) = \int_H x d\mathbb{P}_X(x)$$

  as long as $\mathbb{E}[|X|] < \infty$, defined analogously. A random variable is called **centered** if $\mathbb{E}[X] = 0$.
- The **variance of a random vector** is given by $\mathbb{V}(X) := \mathbb{E}[|X - \mathbb{E}[X]|^2] = \mathbb{E}[|X|^2] - |\mathbb{E}[X]|^2$, as long as $\mathbb{E}[|X|^2\Big] < \infty$. Clearly, $\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$ for any $a \in \mathbb{R}$ and $b \in H$.
- Two random vectors $X, Y : \Omega \to H$ are called **independent** (we often write $X \perp Y$) if for all $A, B \in \mathcal{B}(H)$ the distribution of the vector (X, Y) given by $\mathbb{P}_{(X,Y)}(A \times B) := \mathbb{P}((X, Y) \in A \times B)$ satisfies

$$\mathbb{P}_{(X,Y)}(A \times B) = \mathbb{P}_X(A)\mathbb{P}_Y(B),$$

  in other words: $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$.
- A family $(X_n)_{n\in\mathbb{N}}$ of random variables $X_n : \Omega \to \mathbb{R}$ is called **independent** if for any finite $J \subset \mathbb{N}$ we have for any family of events $A_j \in \mathcal{B}(\mathbb{R})$

$$\mathbb{P}(X_j \in A_j, \text{ for all } j \in J) = \prod_{j\in J} \mathbb{P}(X_j \in A_j)$$

Weak convergence or convergence in distribution:

- A family μ_n of probability measures on $(H, \mathcal{B}(H))$ **converges weakly** to a probability measure μ on $(H, \mathcal{B}(H))$ if for any bounded continuous function $f \in \mathcal{C}_b(H, \mathbb{R})$

$$\int_H f(z)\mu_n(dz) \longrightarrow \int_H f(z)\mu(dz)$$

  Equivalently we say that a family of random vectors $X_n : \Omega \to H$ **converges in distribution** to a random variable X ($X_n \xrightarrow{d} X$) iff

$$\mathbb{P}_{X_n} \overset{n\to\infty}{\longrightarrow} \mathbb{P}_X, \qquad \text{weakly.}$$

The sum of independent random vectors: the convolution

- For two independent random variables $X_1, X_2 : \Omega \to \mathbb{R}$ with $\mathbb{P}_{X_1} = \mu_1$ and $\mathbb{P}_{X_2} = \mu_2$ we have that the sum $X_1 + X_2$ has the distribution

$$\mathbb{P}_{X_1+X_2}(A) = \mathbb{P}(X_1 + X_2 \in A) =: \mu_1 * \mu_2(A) = \int_{\mathbb{R}} \mu_1(A - x) d\mu_2(x).$$

This is denoted as the **convolution of the probability measures**. Note that $\mu_1 * \mu_2 = \mathbb{P}_{X_1+X_2} = \mathbb{P}_{X_2+X_1} = \mu_2 * \mu_1$.

Characteristic functions characterize distributions:

- The **characteristic function** $\phi_\mu : H \to \mathbb{C}$ of a probability measure μ on $(H, \mathcal{B}(H))$ defined as

$$\phi_\mu(u) := \int_H e^{i\langle u,z\rangle} \mu(dz), \qquad u \in H.$$

For a random vector $X : \Omega \to H$ we denote

$$\phi_X(u) := \phi_{\mathbb{P}_X}(u), \qquad u \in H.$$

Example 7.2.3. *1. $X \sim \mathcal{N}(m, \sigma^2)$, then*

$$\phi_X(u) = e^{imu - \frac{\sigma^2 u^2}{2}}$$

2. $X \sim \mathrm{Exp}(\lambda)$, then

$$\phi_X(u) = \frac{\lambda}{\lambda - iu}$$

Basic properties of the characteristic function:

Injectivity: If $\phi_{\mu_1}(u) = \phi_{\mu_1}(u)$ for all $u \in H$, then $\mu_1 = \mu_2$, that is the characteristic function "characterizes" the distribution μ.

Boundedness: $|\phi_\mu(u)|_{\mathbb{C}} \leqslant 1$ for all $u \in H$.

Sign change: $\phi(-u) = \bar{\phi}_X(u)$ the conjugate complex $\overline{a + ib} = a - ib$

Symmetry: The symmetry implies $\mathcal{L}(X) = \mathcal{L}(-X) \implies \phi_X(u) \in \mathbb{R}$ for all $u \in H$.

Continuity in 0: $\lim_{u\to 0} \phi_X(u) = 1$

Moments translate into smoothness: If $\mathbb{E}[|X_j|^k] < \infty$ for some $k \in \mathbb{N}$ and $j = 1, \ldots, d$, then ϕ_X is k-times continuously differentiable and we have

$$\mathbb{E}[X_j^k] = \frac{1}{i^k} \frac{\partial^k}{\partial x_j^k} \phi_X(u)\Big|_{u=0}.$$

Weak convergence $\Longrightarrow$ pointwise convergence of the c.f.: For random vectors $X, X_n : \Omega \to H$ with $\phi_n := \phi_{X_n}$ and $\phi := \phi_X$ we have that

$$\lim_{n\to\infty} \phi_n(u) = \phi(u) \qquad \text{for all } u \in H,$$

implies $X_n \xrightarrow{d} X$.

Ptwise conv. + continuity in 0 of the limit $\Longrightarrow$ weak conv.: For a family $(X_n)_{n\in\mathbb{N}}$ of random vectors $X_n : \Omega \to H$ with $\phi_n := \phi_{X_n}$ and a function $\phi : H \to \mathbb{C}$ satisfying $\lim_{u\to 0} \phi(u) = 1$ and

$$\lim_{n\to\infty} \phi_n(u) = \phi(u) \qquad \text{for all } u \in H,$$

then ϕ is the characteristic function of a probability distribution.

Independence implies multiplicativity: The independence of random variables $X \perp Y$ implies

$$\phi_{X+Y}(u) = \phi_{\mathbb{P}_X * \mathbb{P}_Y}(u) = \phi_X(u)\phi_Y(u), \qquad u \in H.$$

Characterization of independence: A family $(X_1, \dots, X_n)$ of random variables $X_k : \Omega \to H$ is independent, iff for all $(u_1, \dots, u_n) \in (H)^n$

$$\phi_{(X_1,\dots,X_n)}((u_1, \dots, u_n)^T) = \phi_{X_1}(u_1) \dots \phi_{X_n}(u_n).$$

Moment generating function: For some nonnegative random variable $X : \Omega \to \mathbb{R}$ the function $r \mapsto \psi(r) := \mathbb{E}[e^{-rX}]$ for $r \geqslant 0$ is called the **moment generating function** of X. It determines the law of X uniquely and satisfies similar properties as the characteristic function. Whenever it exits its relation their relation is given by $\psi_X(r) := \phi(ir)$.

Inversion formula: For a random variable $X : \Omega \to \mathbb{R}$ such that $\int_{\mathbb{R}} |\phi_X(u)| du < \infty$, then X has a bounded, absolutely continuous density f (w.r.t the Lebesgue measure) and

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ixu} \phi_X(u) du.$$

Index

Bibliography

[1] D. Applebaum. Lévy processes and stochastic calculus, volume 116 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, second edition, 2009.

[2] A. Araujo, E. Giné. On tails and domains of attractions of stable measures in Banach spaces. Transactions of the American Mathematical Society., Vol 248, No. 1, 105–119, 1979.

[3] N. Berglund and Gentz B. On the noise-induced passage through an unstable periodic orbit I: Two-level model. J. Stat. Phys., 114(5–6):1577–1618, 2004.

[4] N. Berglund and B. Gentz, The Eyring–Kramers law for potentials with nonquadratic saddles.

[5] ——, Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers' law and beyond, EJP, 18 (2013), pp. no. 24, 58.

[6] P. Billingsley. Convergence of Probability Measures, Wiley Series in Probability and Statistics, second edition, 1999.

[7] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, Metastability in reversible diffusion processes I: Sharp asymptotics for capacities and exit times, J. of the EMS, 6 (2004), pp. 399–424.

[8] A. Bovier, V. Gayrard, and M. Klein, Metastability in reversible diffusion processes II: Precise asymptotics for small eigenvalues, Journal of the European Mathematical Society, 7 (2005), pp. 69–99.

[9] S. Brassesco. Some results on small random perturbations of an infinite dimensional dynamical system. Stochastic Processes and their Applications, 38:33–53, 1991.

[10] S. Brassesco. Unpredictability of an exit time. Stochastic Processes and their Applications, 63:55–65, 1996.

[11] S. Brassesco. Perturbaciones estocásticas de la ecuación de Allen-Cahn. Lecture notes of XXV Escuela Venezolana de Matemáticas EMALCA-Venezuela, 2012.

[12] N. H. Bingham, C. M. Goldie, J. L. Teugels. *Regular variation*, Cambridge University Press, Cambridge, 1987.

[13] A. Budhiraja, P. Nyquist. *Large deviations for multidimensional state-dependent shot-noise processes*, J. Appl. Probab., Vol. 52, Number 4 (2015), 1097-1114.

[14] N. Chafee, E. F. Infante. *A bifurcation problem for a nonlinear partial differential equation of parabolic type*, Applicable Anal. (4), 17-37, 1974/75.

[15] A. Budhiraja, J. Chen, P. Dupuis. *Large Deviations for Stochastic Partial Differential Equations Driven by a Poisson Random Measure*, Stoch. Proc. App. 123 (2013), no 2, 523-560.

[16] A. Budhiraja, P. Dupuis, A. Ganguly. *Moderate Deviation Principles for Stochastic Differential Equations with Jumps*, The Annals of Probability, 2016, Vol. 44, No 3., 1723–1775.

[17] M. V. Day. On the exponential exit law in the small parameter exit problem. Stochastics, 8:297–323, 1983.

[18] M. V. Day. Exit cycling for the Van der Pol oscillator and quasipotential calculations. Journal of Dynamics and Differential Equations, 8(4):573–601, 1996.

[19] A. Debussche, M. Högele, P. Imkeller. *Metastability for the Chafee-Infante equation with small heavy-tailed Lévy noise, Electronic Communications in Probability, 16, 213–225*, 2011.

[20] A. Debussche, M. Högele, P. Imkeller. *The dynamics of nonlinear reaction-diffusion equations with small Lévy noise.* Springer Lecture notes in Mathematics, Vol. 2085, 163p. (2013).

[21] J. D. Deuschel, D. W. Stroock. *Lare Deviations.* American Mathematical Society, 2001.

[22] G. DaPrato, J. Zabzcyk. *Stochastic Equations in Infinite Dimensions. Encyclopedia of Mathematics and Its Applications*, vol. 44, (Cambridge University Press 1992).

[23] A. Dembo and O. Zeitouni. *Large deviation techniques and applications, second edition*, Applications of Mathematics, vol. 38, 1998.

[24] J. Elstrodt. *Maß und Integrationstheorie, 3., erweiterte Auflage*, Springer-Lehrbuch, 2002.

[25] G. W. Faris, G. Jona-Lasinio. *Large fluctuations for a nonlinear heat equation with noise*, J. Phys. A **15** (1982), 3025-3055 .

[26] M. I. Freidlin. Random perturbations of reaction-diffusion equations: the quasideterministic approximation. Transactions of the American Mathematical Society, 305(2):665–697, 1988.

[27] A. D. Ventsel' and M. I. Freidlin. On small random perturbations of dynamical systems. Russian Mathematical Surveys, 25(1):1–55, 1970.

[28] M. I. Freidlin and A. D. Wentzell. Random perturbations of dynamical systems, volume 260 of Grundlehren der Mathematischen Wissenschaften. Springer, second edition, 1998.

[29] A. Galves, E. Olivieri, and M. E. Vares, Metastability for a class of dynamical systems subject to small random perturbations, The Annals of Probability, 15 (1987), pp. 1288–1305.

[30] H.-O. Georgii. Stochastik, Einführung in die Wahrscheinlichkeitstheorie und Statistik. deGruyter, 4. Auflage, 2009.

[31] V. V. Godovanchuk. Asymptotic probabilities of large deviations due to large jumps of a Markov process. Theory of Probability and its Applications, 26:314–327, 1982.

[32] J. Hale. *Some infinite-dimensional Morse-Smale systems defined by parabolic partial differential equations*, CWI Quarterly, **12** (1999), no 3&4, 239-314.

[33] D. Henry. *Some infinite-dimensional Morse-Smale systems defined by parabolic partial differential equations*, J. Differential Equations, **59** (1985), no 2, 165-205.

[34] D. Henry. *Geometric theory of semilinear parabolic equations*, Lecture notes in Mathematics, 840. Springer-Verlag, Berlin-New York, 1981.

[35] M. Högele. The first exit problem and metastability of generic scalar dissipative reaction-diffusion equations subject to multiplicative regularly varying Lévy noise at small intensity. https://arxiv.org/abs/1706.07745 (2017)

[36] M. Högele, I. Pavlyukevich. The exit problem from the neighborhood of a global attractor for heavy-tailed Lévy diffusions. Stochastic Analysis and Applications, 32 (1), 163–190 (2013)

[37] M. Högele, I. Pavlyukevich: Metastability in a class of hyperbolic dynamical systems perturbed by heavy-tailed Lévy type noise. Stoch. and Dyn., 15 (3), 1550019-1 – 1550019-26 (2015)

[38] H. Hulk and F. Lindskog. *Regular variation for measures on metric spaces*, Publ. Inst. Math. (Beograd) (N.S.) **80** (2006), no 94, 121-140.

[39] P. Imkeller and I. Pavlyukevich, *Metastable behaviour of small noise Lévy-driven diffusions*, ESAIM: Prob. Stat., **12** (2008), 412-437.

[40] P. Imkeller and I. Pavlyukevich, *First exit times of SDEs driven by stable Lévy processes*, Stoch. Proc. Appl., **116** (2006), no 4, 611-642.

[41] P. Imkeller, *Mallilavin's Calculus and Applications in stochastic control and finance*, Institute of Mathematics Polish Academy of Sciences, Impan Lecture Notes No. 1, 2008.

[42] P. Imkeller, I. Pavlyukevich, and M. Stauch. First exit times of non-linear dynamical systems in $\mathbb{R}^d$ perturbed by multifractal Lévy noise. J. Stat. Phys., 141(1):94–119, 2010.

[43] O. Kallenberg. Foundations of Modern Probability, 2nd edition (2002), Springer-Verlag New York, Berlin, Heidelberg.

[44] C. Kipnis and C. M. Newman, The metastable behavior of infrequently observed, weakly random, one-dimensional diffusion processes, SIAM Journal on Applied Mathematics, 45 (1985), pp. 972–982.

[45] A. Klenke. Wahrscheinlichkeitstheorie, Springer, 2006.

[46] V. N. Kolokoltsov, Semiclassical analysis for diffusions and stochastic processes, vol. 1724 of Lecture Notes in Mathematics, Springer, 2000.

[47] V. N. Kolokol'tsov and K. A. Makarov, Asymptotic spectral analysis of a small diffusion operator and the life times of the corresponding diffusion process, Russian Journal of Mathematical Physics, 4 (1996), pp. 341–360.

[48] D. Koshnevisan. Probability, Graduate Studies in Mathematics, vol. 80, AMS, 2007.

[49] P. Kotelenez. Stochastic Ordinary and Stochastic Partial Differential Equations: Transition from Microscopic to Macroscopic Equations. Springer, 2007

[50] H. A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, Physica, 7 (1940), pp. 284–304.

[51] Y. V. Martsynyuk. New multivariate central limit theorems in linear structural and functional error-in-variables models. Elec. J. Stat., Vol. 1 (2007), 347–380.

[52] C. Marinelli, C. Prévôt, and M. Röckner, Regular dependence on initial data for stochastic evolution equations with multiplicative Poisson noise, J. Funct. Anal. 258 (2010), no. 2, 616-649.

[53] C. Marinelli and M. Röckner, *Well-posedness and asymptotic behavior for stochastic reaction-diffusion equations with multiplicative Poisson noise* EJP, Vol. 15 (2010), pa. no. 49, p. 1528-1555.

[54] C. Marinelli and M. Röckner, On uniqueness of mild solutions for dissipative stochastic evolution equations, Infin. Dimens. Anal. Quantum Probab. Relat. Top. .

[55] C. Marinelli and M. Röckner, On the maximal inequalities of Burkholder, Davis and Gundy, Expositiones Mathematicae, Vol. 34, 1, 2016, Pages 1-26.

[56] I. Pavlyukevich. First exit times of solutions of stochastic differential equations driven by multiplicative Lévy noise with heavy tails. Stoch. and Dyn., 11(2&3), 2011.

[57] P. E. Protter. Stochastic integration and differential equations, volume 21 of Applications of Mathematics. Springer, second edition, 2004.

[58] S. Peszat and J. Zabczyk, *Stochastic partial differential equations with Lévy noise (an evolution equation approach)*, Cambridge University Press, Cambridge, 2007.

[59] G. Raugel, *Global attractors in partial differential equations*, Fiedler, Bernold (ed.), Handbook of dynamical systems. Vol 2, 885-982, North-Holland, Amsterdam, 2002.

[60] C. Rocha, *Examples of Attractors in Reaction-Diffusion equations*, J. Diff. Equ. 73, 178-195, 1988

[61] K. Sato. Lévy processes and infinitely divisible distributions, volume 68 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

[62] M. Beiglböck, P. Siorpaes Applications of pathwise Burkholder-Davis-Gundy inequalities. Bernoulli, 21 (1), 2015, 360–373 `https://arxiv.org/pdf/1305.6188.pdf`

[63] , S. G. Sell and Y. You. Dynamics of evolutionary equations, Applied Mathematical Sciences 143, Springer, 2002.

[64] E. Salavati and B. Z. Zangeneh. A maximal inequality for the p-th power of stochastic convolution integrals, volume 68 of Journal of Inequalities and Applications. 155: 1–16, 2016.

[65] R. Temam. *Infinite-dimensional dynamical systems in Mechanics and Physics*, Springer, New York, Applied Mathematical Sciences Series, vol. 68, 1988. Second augmented edition, 1997.

[66] J. B. Walsh. An introduction to stochastic partial differential equations Ecole d'Eté de Prob. de St. Flour XIV, 1984, Lect. Notes in Math 1180, Springer Verlag (1986).

[67] D. Werner. *Funktionalanalysis, 3., neu bearbeitete und erweiterte Auflage*, Springer-Lehrbuch, 2000.

La Asociación Matemática Venezolana fue fundada en 1990 como una organización civil sin fines de lucro cuya finalidad es trabajar por el desarrollo de las matemáticas en Venezuela.

Asociación Matemática Venezolana
Apartado 47.898, Caracas 1041-A, Venezuela

IVIC

Gobierno **Bolivariano** de Venezuela

Ministerio del Poder Popular para **Educación Universitaria, Ciencia y Tecnología**

1817 - 2017 ZAMORA UNIÓN CÍVICO MILITAR