SMOLUCHOWSKI-KRAMERS APPROXIMATION FOR STOCHASTIC EQUATIONS WITH LÉVY-NOISE

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Songfu Zhang

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2008

Purdue University

West Lafayette, Indiana

SMOLUCHOWSKI-KRAMERS APPROXIMATION FOR STOCHASTIC EQUATIONS WITH LÉVY-NOISE

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Songfu Zhang

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2008

Purdue University

West Lafayette, Indiana

ACKNOWLEDGMENTS

I would like to thank my advisor Prof. Dr. Michael Röckner for the numerous engaging mathematical discussions I had with him and for his guidance on this thesis. His lectures, which I enjoyed the most among all the courses, gave me a deep insight into stochastic analysis.

In addition, I would like to extend my thanks to Dr. Walter Hoh for checking my whole thesis patiently and for his helpful comments.

TABLE OF CONTENTS

ABSTRACT

Zhang Songfu Ph.D., Purdue University, December 2008. Smoluchowski-Kramers Approximation for Stochastic Equations with Lévy-Noise. Major Professor: Michael Röckner.

A generalization of Smoluchowski-Kramers approximation to Lévy processes is given. It is proved that an analogue of the result in the classic Brownian motion case holds. A momentum model is proposed by applying this result to the financial market. Finally, a partial result of the Smolcuchowski-Kramers approximation in the infinite dimensional case is given.

1. Introduction

In this PhD thesis we will present some results on the Smoluchowski-Kramers approximation.

The motion of a particle of mass μ in a force field $b(X_t) + \sigma(X_t)\dot{W}_t$ with the friction proportional to the velocity (without loss of generality, let the friction coefficient be 1) is governed by the Newton law:

$$\mu\frac{d^2X_t^\mu}{dt^2} = b(X_t^\mu) + \sigma(X_t^\mu)\frac{dW_t}{dt} - \frac{dX_t^\mu}{dt}, X_0^\mu = x_0 \in \mathbb{R}^{d_1}, \frac{dX_0^\mu}{dt} = y_0 \in \mathbb{R}^{d_1} \tag{1.1}$$

where $W_t, t \in [0,T]$ is a Wiener process taking values in $\mathbb{R}^{d_1}$.

The Smoluchowski-Kramers approximation (of X_t^μ by X_t) says that for any $0 \leq T < \infty$, $\delta > 0$ and $x_0, y_0 \in \mathbb{R}^{d_1}$,

$$\lim_{\mu\downarrow 0} P\{\sup_{0\leq t\leq T} |X_t^\mu - X_t| > \delta\} = 0 \tag{1.2}$$

Here X is the solution of the following stochastic differential equation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, X_0 = x_0 \tag{1.3}$$

The Smoluchowski-Kramers approximation is the justification for using the first order equation (1.3) to describe the motion of a small particle disturbed by a Wiener process instead of using the Newton equation (1.1).

There are a number of papers on this subject. For example, Narita [35] proved the Smoluchowski-Kramers approximation for the stochastic Liénard equation with mean-field. In a subsequent paper [36], he showed that by a change of time and displacement, the velocity process $\frac{\partial X^\mu}{\partial t}$ converges to a one-dimensional Ornstein-Uhlenbeck process. Boufoussi and Tudor [8] showed the Smoluchowski-Kramers Approximation for fractional Brownian motion. Mark Freidlin included a proof for the classical situation (1.1) in his paper [21], which was fully elaborated by Ramona Westermann [46]. She also set up a financial model called momentum model by applying the approximation to the financial market. Cerrai and Freidlin [9] extended the above result to the infinite dimensional case. They showed that the solution of the semi-linear stochastic damped wave equations

$$\mu u_{tt}(t,x) = \Delta u(t,x) - u_t(t,x) + b(x, u(t,x)) + Q\dot{W}(t), u(0) = u_0, u_t(0) = v_0 \quad (1.4)$$

endowed with Dirichlet boundary conditions, converges as μ goes to zero to the solution of the semi-linear stochastic heat equation

$$u_t(t,x) = \Delta u(t,x) + b(x, u(t,x)) + Q\dot{W}(t), u(0) = u_0, \quad (1.5)$$

endowed with Dirichlet boundary conditions. They further generalized the results to the multiplicative noise case in [10].

This PhD thesis is to further generalize the Smoluchowski-Kramer approximation. Specifically, first we generalize the approximation to Lévy processes in finite dimension. So we replace $W_t, t \in [0,T]$ by $L_t, t \in [0,T]$, which is a Levy process taking values in $\mathbb{R}^{d_1}$. The equations become

$$\begin{cases} dX_t^\mu = Y_t^\mu dt, X_0^\mu = x_0 \\ dY_t^\mu = \frac{1}{\mu} b(t, X_t^\mu)dt + \frac{1}{\mu}\sigma(t, X_t^\mu)dL_t - \frac{1}{\mu}dX_t^\mu, Y_0^\mu = y_0 \end{cases} \quad (1.6)$$

and the first order equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dL_t, X_0 = x_0 \quad (1.7)$$

We show that if $L(t)$ has finite moments of n-th order for $n = 2^m, m \in \mathbb{N}$, then

$$\sup_{t\in[0,T]} E|X_t^\mu - X_t|^n \to 0 \text{ as } \mu \to 0 \quad (1.8)$$

Moreover, if the equations (1.6) and (1.7) are driven by a poisson random measure with finite intensity measure, that is,

$$\begin{cases} dX_t^\mu = Y_t^\mu dt, X_0^\mu = x_0 \\ dY_t^\mu = \frac{1}{\mu} b(t, X_t^\mu)dt + \frac{1}{\mu}\int_{\mathbb{R}^{d_1}} \sigma(t, X_t^\mu)N(dt, dx) - \frac{1}{\mu} dX_t^\mu, Y_0^\mu = y_0 \end{cases} \tag{1.9}$$

and

$$dX_t = b(t, X_t)dt + \int_{\mathbb{R}^{d_1}} \sigma(t, X_t)N(dt, dx), X_0 = x_0 \tag{1.10}$$

We show that

$$\lim_{\mu \to 0} P(\|X^\mu - X\|_{D([0,T];\mathbb{R}^d)} > \epsilon) = 0. \tag{1.11}$$

where $\|\cdot\|_{D([0,T];\mathbb{R}^d)}$ denote the Skorokhod metric. We prove this by first showing the tightness of the solution X^{μ_n} of equations (1.9) with $\mu_n \to 0$, and then followed by an argument used by Gyöngy and Krylov in [23].

We then apply the approximation to the financial market to establish a momentum model following the idea of Westermann [46]. The differences are that we define the parameter μ to measure the price delay with which a firm's stock price responds to information. This kind of measure is intended to capture all sources of price delay, which may include the friction of the market, the size and analyst coverage of the individual stock itself. We also choose our driving process to be the Generalized Hyperbolic Lévy motion, which is arguably better than Brownian motion by empirical observations. Readers are referred to Eberlein [17] for more detailed argument.

In the infinite dimensional case, we want to prove the Smoluchowski-Kramers approximation for $\frac{\partial}{\partial x}$, while Cerrai [9] proved it for the Laplacian operator Δ_x. We consider the solution of the following equation

$$\begin{cases} \mu \frac{\partial^2 u}{\partial t^2}(t,x) = \frac{\partial u}{\partial x}(t,x) + \Delta_x \frac{\partial u}{\partial t}(t,x) - \frac{\partial u}{\partial t}(t,x) + f(x, u(t,x)) \\ +b(x, u(t,x))\frac{\partial W^Q}{\partial t}(t,x), t > 0, x \in [0, \infty) \\ u(0,x) = u_0, \frac{\partial u}{\partial t}(0,x) = v_0 \end{cases} \tag{1.12}$$

where $W^Q(t,x)$ is a Gaussian mean zero random field, δ-correlated in time and the operator Q characterizes the correlation in the space variables.

We add a remark that this kind of wave equation is difficult to solve for the first order derivative operator. So we add a strong damping term, which is $\Delta_x \frac{\partial u}{\partial t}$.

We fit the above problem in the framework of mild solutions by rewriting it into an abstract stochastic differential equations on a separable (infinite dimensional) Hilbert space H

$$\begin{cases} dX(t) = [AX(t) + F(X(t))]dt + B(X(t))dW(t), t \in [0,T] \\ X(0) = \xi \end{cases} \tag{1.13}$$

where $W(t), t \geq 0$, is a cylindrical Q-Wiener process in a separable Hilbert space U. A mild solution of problem (1.13) is a predictable process $X(t), t \in [0,T]$, such that

$$X(t) = S(t)\xi + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)B(X(s))dW(s) \text{ P-a.s.} \tag{1.14}$$

We prove the existence of a unique mild solution of problem (1.12). Further, under some condition on Q, we will show the solution of the linear equation is δ-Hölder continuous with respect to t for any $\delta < \frac{1}{2}$. Moreover, the momenta of the δ-Hölder norms of the solutions are bounded uniformly in μ. But so far we could not show the convergence, that is, the solution of the equation (1.12) converges to the solution of the equation

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) = \frac{\partial u}{\partial x}(t,x) + f(x,u(t,x)) + b(x,u(t,x))\frac{\partial W^Q}{\partial t}(t,x), t > 0, x \in [0,\infty) \\ u(0) = u_0 \end{cases} \tag{1.15}$$

in probability, i.e.

$$P(\|u^\mu - u\|_{C([0,T],H)} > \epsilon) \to 0 \text{ as } \mu \to 0. \tag{1.16}$$

The reason is that we could not show the tightness of the solutions, which is one of the main step of our approach.

Filipović introduced a weighted Sobolev space for the solutions of the first order equation with the same operator $\frac{\partial}{\partial x}$ in [20]. This operator generates a semigroup which has a lot of nice properties on the Sobolev space. We also considered this space. But it turns out that we could not gain much by switching to this space, because in our case the equations are of second order. The detail is included in Appendix A.

The chapters are summarized as the following. Chapter 2 we recall the concept of Lévy processes, quadratic variation process and stochastic integration with respect to a càdlàg martingale. In chapter 3 we prove the approximation in finite dimension with the Lévy noise. We then establish the momentum model in Chapter 4. The chapter 5 begins with introduction of semigroups and mild solutions, followed by the existence and uniqueness of the mild solutions of equation (1.12). We include the existence and uniqueness result on the weighted Sobolev space in the Appendix A.

2. Lévy Processes and Stochastic Integration

In this chapter we review the theory of Lévy Processes.

2.1 Lévy Processes

Let $X = (X(t), t \geq 0)$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, P)$. We say that X is Lévy process if:

- X(0)=0 a.s.;
- X has independent increments, i.e. for each $n \in N$ and each $0 \leq t_1 < t_2 \leq \cdots < t_{n+1} < \infty$ the random variables $(X(t_{j+1}) - X(t_j), 1 \leq j \leq n)$ are independent;
- X has stationary increments, i.e. $X(t) - X(s)$ has the same distribution as $X(t-s)$ for all $0 \leq s < t < \infty$
- X is stochastically continuous, i.e. for all $a > 0$ and for all $s \geq 0$

$$\lim_{t \to s} P(|X(t) - X(s)| > a) = 0$$

Brownian motion, the Poisson process and the compound Poisson process are examples of Lévy processes.

For a Lévy process X the jump at time t is given by $\Delta X(t) = X(t) - X(t-)$ for each $t \geq 0$.

Lemma 1 *If X is a Lévy process, then for fixed $t > 0, \Delta X(t) = 0$ a.s..*

Proof (cf. Applebaum [2] Lemma 2.3.2) ∎

Let $0 \leq t < \infty$ and $A \in \mathscr{B}(\mathbb{R}^{d_1} - \{0\})$. Define

$$N(t, A) = \#\{0 \leq s \leq t; \Delta X(s) \in A\} = \sum_{0 \leq s \leq t} \chi_A(\Delta X(s)) \tag{2.1}$$

We define

$$\nu(A) = \mathbb{E}(N(1, A)) \tag{2.2}$$

and call it the intensity measure associated with X. We say that $A \in \mathscr{B}(\mathbb{R}^{d_1} - \{0\})$ is bounded below if $0 \notin \bar{A}$.

Proposition 2 *1. If A is bounded below, then $(N(t, A), t \geq 0)$ is a Poisson process with intensity $\nu(A)$.*

2. If $A_1, ..., A_m \in \mathscr{B}(\mathbb{R}^d - \{0\})$ are disjoint, then the random variables $N(t, A_1), ..., N(t, A_m)$ are independent.

Proof (cf. Applebaum [2] Theorem 2.3.5).

■

Remark 3 *It follows immediately that $\nu(A) < \infty$ whenever A is bounded below, hence the measure ν is σ-finite.*

Proof (cf. Applebaum [2] Remark 1 P.89).

■

Definition 4 *A Poisson random measure η on a measurable space $(S, \mathcal{S})$ is a collection of random variables $(\eta(B), B \in \mathcal{S})$ such that:*

1. *$\eta(\varnothing) = 0$;*

2. *(σ-additivity) given any sequence $(B_n, n \in \mathbb{N})$ of mutually disjoint sets in $\mathcal{S}$,*

$$\eta(\bigcup_{n \in \mathbb{N}} B_n) = \sum_{n \in \mathbb{N}} \eta(B_n) \ \ a.s.;$$

3. *(independently scattered property) for each disjoint family $(B_1, ..., B_n)$ in $\mathcal{S}$, the random variables $\eta(B_1), ..., \eta(B_n)$ are independent;*

4. *For each $B \in \mathcal{S}$ such that $\mathbb{E}\eta(B)$ is finite, $\eta(B)$ is a Poisson random variable with parameter $\mathbb{E}\eta(B)$.*

Proposition 5 *Given a σ-finite measure λ on a measurable space $(S, \mathcal{S})$, there exists a Poisson random measure η on a probability space $(\Omega, \mathcal{F}, P)$ such that $\lambda(B) = \mathbb{E}(\eta(B))$ for all $B \in \mathcal{S}$.*

Proof (cf. Ikeda and Watanabe [26] p.42).

■

Suppose that $S = \mathbb{R}^+ \times U$, where $(U, \mathscr{B}(U))$ is a measurable space. Let $\mathcal{S} = \mathscr{B}(\mathbb{R}^+) \bigotimes \mathscr{B}(U)$. Let $p = (p(t), t \geq 0)$ be an adapted process taking values in U such that η is a Poisson random measure on S, where $\eta([0,t) \times A) = \#\{0 \leq s < t; p(s) \in A\}$ for each $t \geq 0, A \in \mathscr{B}(U)$. In this case we say that p is a Poisson point process and η is its associated Poisson random measure.

Let $U = \mathbb{R}^{d_1} - \{0\}$. Let X be a Lévy process; then ΔX is a Poisson point process and N is its associated Poisson random measure. For each $t \geq 0$ and A bounded below, we define the compensated Poisson random measure by $\tilde{N}(t, A) = N(t, A) - t\nu(A)$.

Let f be a Borel measurable function from $\mathbb{R}^{d_1}$ to $\mathbb{R}^{d_1}$ and let A be bounded below; then for each $t > 0$, $\omega \in \Omega$, we define the Poisson integral of f as a random finite sum by

$$\int_A f(x)N(t,dx)(\omega) = \sum_{x \in A} f(x)N(t,\{x\})(\omega) = \sum_{0 \leq u \leq t} f(\Delta X(u))\chi_A(\Delta X(u)) \qquad (2.3)$$

Then:

Proposition 6 1. *for each $t \geq 0, \int_A f(x)N(t,dx)$ has a compound Poisson distribution such that, for each $u \in \mathbb{R}^{d_1}$,*

$$\mathbb{E}(\exp[i(u, \int_A f(x)N(t,dx))]) = \exp[t \int_A (e^{i(u,x)} - 1)\nu \circ f^{-1}(dx)]$$

2. *if* $f \in L^1(A, \nu(A))$, *we have* $\mathbb{E}(\int_A f(x)N(t,dx)) = t\int_A f(x)\nu(dx)$;

3. *if* $f \in L^2(A, \nu(A))$, *we have* $Var(|\int_A f(x)N(t,dx)|) = t\int_A |f(x)|^2\nu(dx)$

Proof (cf. Applebaum [2] Theorem 2.3.8).

■

For each $f \in L^1(A, \nu_A), t \geq 0$, define the compensated Poisson integral by

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

Then $(\int_A f(x)\tilde{N}(t,dx), t \geq 0)$ is a martingale, which can be verified straightforwardly by the definition of martingale.

Let $(\epsilon_n, n \in \mathbb{N})$ be a sequence that decreases monotonically to zero such that $\epsilon_1 < 1$. Let

$$A_n = \{x \in \mathbb{R}^{d_1}, \epsilon_{n+1} \leq |x| \leq \epsilon_1\}$$

Proposition 7 $(\int_{A_n} x\tilde{N}(t,dx), t \geq 0)$ *is a Cauchy sequence in martingale space which converges to* $(\int_{|x|<1} x\tilde{N}(t,dx), t \geq 0)$.

Proof (cf. Applebaum [2] Theorem 2.4.11).

■

A very important representation for Lévy processes is:

Theorem 8 *(Lévy-Itô decomposition) If X is a Lévy process, then there exists $b \in \mathbb{R}^{d_1}$, a Brownian motion B_Q with covariance matrix Q and an independent Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R}^{d_1} - \{0\})$ such that, for each $t \geq 0$,*

$$X(t) = bt + B_Q(t) + \int_{|x|<1} x\tilde{N}(t,dx) + \int_{|x|\geq 1} xN(t,dx) \tag{2.4}$$

Here the Poisson random measure N and the intensity measure ν are defined as in (2.1) and (2.2). Moreover ν satisfies

$$\int_{\mathbb{R}^{d_1}} (|y|^2 \wedge 1)\nu(dy) < \infty \tag{2.5}$$

Proof (cf. Applebaum [2] Theorem 2.4.16).

∎

Corollary 9 *If X is a Lévy process then for each $u \in \mathbb{R}^{d_1}, t \geq 0$,*

$$\mathbb{E}(e^{i(u,X(t))}) = \exp(t\{i(b,u) - \frac{1}{2}(u,Au) + \int_{\mathbb{R}^{d_1}} [e^{i(u,y)} - 1 - i(u,y)\chi_B(y)]\nu(dy)\}) \quad (2.6)$$

where $B = \{x \in \mathbb{R}^{d_1} | |x| < 1\}$.

Proof (cf. Applebaum [2] Corollary 2.4.20).

∎

2.2 Stochastic Integration

In this section, we follow the procedure and notation of the unpublished Prof. Jin Ma's lecture notes "Stochastic Analysis and Stochastic Differential Equations".

Denote

- $\mathcal{M}^2 = \{$all martingales such that $\sup_{t\in[0,\infty)} \mathbb{E}[|M_t|^2] < \infty\}$;
- $\mathcal{M}_0^2 = \{M \in \mathcal{M}^2 | M_0 = 0$ a.s.$\}$
- $\mathcal{M}_0^{2,c} = \{M \in \mathcal{M}^2 | t \mapsto M_t$ is continuous and $M_0 = 0$ a.s.$\}$
- $\mathcal{M}_{loc}^2 = \{$all local martingales M such that $\sup_{t\in[0,T]} \mathbb{E}[|M_t|^2] < \infty, \forall T > 0\}$
- $\mathcal{M}_{0,loc}^2 = \{M \in \mathcal{M}_{loc}^2 | M_0 = 0$ a.s.$\}$

We also denote the space of all adapted processes that have finite variation paths to be FV, and FV_0 to be the subspace of FV consists of all FV-process null at 0. For any càdlàg function f, define $\Delta f(t) \triangleq f(t) - f(t-), \forall t$.

Definition 10 • *A process Z is said to be of Class (D) if the family of random variables $\{X_T : T$ a finite stopping time $\}$ is uniformly integrable;*

- *A process Z is said to be of Class (DL) if for each $a > 0$, the family of random variables $\{X_T : T$ a stopping time such that $T \leq a,$ a.s. $\}$ is uniformly integrable.*

It is direct to verify that if a submartingale X can be represented as

$$X_t = M_t + A_t \tag{2.7}$$

where M is a martingale and A is an integrable increasing process, then X is of class (DL). The converse is given by the Doob-Meyer decomposition.

Proposition 11 *(Doob-Meyer's Decomposition Theorem) Suppose that X is a submartingale of class (DL), then there is a unique predictable integrable increasing process A, such that X can be written as (2.7).*

Proof (cf. Robert J. Elliott [18], Theorem 8.9).

■

Proposition 12 *Let $M \in \mathcal{M}_0^{2,c}$. Then there exists a unique continuous increasing process $\langle M \rangle$ null at 0 such that $M^2 - \langle M \rangle$ is a uniformly integrable martingale.*

Proof The existence follows from the Doob-Meyer decomposition. The continuity of $\langle M \rangle$ follows from Robert J. Elliott [18], Remarks 8.21.

■

Definition 13 *A subspace $\mathcal{N}$ of $\mathcal{M}_0^2$ is called stable if*

- *$\mathcal{N}$ is a closed subspace;*

- *$\mathcal{N}$ is stable under stopping, that is, if $N \in \mathcal{N}$, then so is the stopped process N^T, for any stopping time T.*

Proposition 14 *Let $\mathcal{N}$ be a stable subspace of $\mathcal{M}_0^2$. Denote its stable orthogonal complement*

$$\mathcal{N}^\perp = \{Z \in \mathcal{M}_0^2 | \mathbb{E}(Z_t N_t) = 0, \forall t \geq 0, \forall N \in \mathcal{N}\}$$

Then every $M \in \mathcal{M}_0^2$ has a unique decomposition

$$M = N + Z, N \in \mathcal{N}, Z \in \mathcal{N}^\perp$$

Proof (cf. Robert J. Elliott [18], Corollary 9.17).

■

Since $\mathcal{M}_0^{2,c}$ is a stable subspace of $\mathcal{M}_0^2$, we denote the stable orthogonal complement by $\mathcal{M}_0^{2,d}$. We call it the "subspace of purely discontinuous martingale". By the above proposition, every element of $M \in \mathcal{M}_0^2$ can be uniquely decomposed as

$$M = M^c + M^d, M^c \in \mathcal{M}_0^{2,c}, M^d \in \mathcal{M}_0^{2,d}$$

Lemma 15 *Let $M \in \mathcal{M}_0^{2,d}$. Then $M_t^2 - \sum_{0 \leq s \leq t} (\Delta M_s)^2, t \geq 0$ is a uniformly integrable martingale.*

Proof (cf. Rogers, L.C.G. and Williams, D. [40]).

■

Thus we define for $M \in \mathcal{M}_0^2$,

$$[M]_t \triangleq \langle M^c \rangle_t + \sum_{0 \leq s \leq t} (\Delta M_s)^2 \tag{2.8}$$

It follows from Elliott [18] Remarks 10.3 that $M^2 - [M]$ is a uniformly integrable martingale. Moreover, for $M, N \in \mathcal{M}_0^2$, define the process $[M, N]$ by polarization formula

$$[M, N] \triangleq \frac{1}{4}([M + N] - [M - N]) \tag{2.9}$$

For $M, N \in \mathcal{M}_{0,loc}^2$, Since for any stopping time T, it holds that $[M^T, N^T] = [M, N]^T$ (cf. Robert J. Elliott [18], Corollary 10.10), there exists a FV process $[M, N]$ null at 0 such that $MN - [M, N]$ is a local martingale.

Definition 16 *A process X is called a semimartingale if X is a $\{\mathcal{F}_t\}$-adapted càdlàg process, such that is can be written as*

$$X = X_0 + M + A$$

where $M \in \mathcal{M}_{0,loc}$, and $A \in FV_0$.

Let X be a semimartingale with decomposition $X = X_0 + M + A$, since the decomposition is not unique, we define the quadratic variation process of X as the follows. We write $M = M^c + M^d$, then since the process $[M^c]$ is unique, we define

$$[X]_t \triangleq [M^c]_t + \sum_{0 \leq s \leq t} (\Delta X_s)^2 \tag{2.10}$$

By extending the above formula by polarization we obtain $[X, Y]$.

Now we are ready to define the stochastic integral against a square-integrable martingale. We define the space $\mathcal{H}_b$ to be all predictable processes of the form

$$H_t = \sum_i Z_{i-1} \chi_{(T_{i-1}, T_i]}$$

where T_i's are stopping times, and $Z_i \in \mathcal{F}_{T_i}$, $i = 1, 2, \cdots$, and are uniformly bounded. Let $M \in \mathcal{M}_0^2$. Then for any $H \in \mathcal{H}_b$, we define the stochastic integral of H against M to be

$$I(H)_t \triangleq (H \cdot M)_t = \sum_i Z_{i-1} \{M_{T_i \wedge t} - M_{T_{i-1} \wedge t}\}, t \geq 0 \tag{2.11}$$

Then it is easy to check that $I(H)$ is an $\{\mathcal{F}_t\}$-martingale and

$$E(H \cdot M)_\infty^2 = E \sum_i Z_{i-1}^2 [M_{T_i} - M_{T_{i-1}}]^2 = E \sum_i Z_{i-1}^2 \{[M]_{T_i} - [M]_{T_{i-1}}\} = E \int_0^\infty H_s^2 d[M]_s \tag{2.12}$$

where the right side above is understood as Lebesgue-Stieltjes integral.

Define:

$$L^2(M) \triangleq \{ \text{ all predictable process } H \text{ such that } \|H\|_M < \infty\} \tag{2.13}$$

where

$$\|H\|_M^2 \triangleq E\int_0^\infty H_s^2 d[M]_s \tag{2.14}$$

Now consider the mapping $\Phi : \mathcal{H}_b \to L^2(\Omega, \mathcal{F}_\infty)$, where

$$\Phi(H) = I(H)_\infty = (H \cdot M)_\infty$$

Since the norm in $\mathcal{M}_0^2$ is $\mathbb{E}(M_\infty^2)$, we see that I is a linear isometry between $\mathcal{H}_b$ and $\mathcal{M}_0^2$. On the other hand, by simple truncation and a monotone class argument, $\mathcal{H}_b$ is a dense subset of $L^2(M)$. Consequently, the mapping I can be extended to all processes in $L^2(M)$.

Definition 17 *Let $M \in \mathcal{M}_0^2$ and $H \in L^2(M)$. Then the stochastic integral of H with respect to M, denoted by $H \cdot M$, is the image of H under the isometry mapping $I : \mathcal{H}_b \to \mathcal{M}_0^2$, extended to $L^2(M)$.*

Proposition 18

$$[H \cdot M, N] = H \cdot [M, N] = \int H_s d[M, N]_s, dt \times dP\text{-}a.e. \tag{2.15}$$

Proof (cf. Robert J. Elliott [18], Corollary 11.21).

■

For $M \in \mathcal{M}_{0,loc}$, the stochastic integral of H against M is defined similarly after localization. We still denote it by $H \cdot M$. If X is a semimartingale with decomposition $X = M + A$, where $M \in \mathcal{M}_{0,loc}$ and A is an FV process null at zero. We define

$$H \cdot X = H \cdot M + H \cdot A$$

where the second integral is the usual (pathwise) Stieltjes integral. Let $H \in \mathcal{L}_b$, X, Y be two semimartingales, we have

$$[H \cdot X, Y] = \int H d[X, Y]. \tag{2.16}$$

(cf. Robert J. Elliott [18], Theorem 11.44, Theorem 12.9).

Proposition 19 *(Integration-by-parts formula) Let X and Y be semimartingales. Then*

$$X_tY_t - X_0Y_0 = \int_{0+}^{t} X_{s^-}dY_s + \int_{0+}^{t} Y_{s^-}dX_s + [X,Y]_t \tag{2.17}$$

Proof (cf. Rogers, L.C.G. and Williams, D. [40]).

■

Proposition 20 *(Itô's formula) Let $f : \mathbb{R}^n \to \mathbb{R}$ be C^2, and suppose $X = (X^1, \cdots, X^n)$ is an n-dimensional semimartingale. Then*

$$f(X_t) = f(X_0) + \int_{(0,t]} \sum_{i=1}^{n} D_i f(X_{s^-})dX_s^i + \frac{1}{2}\int_0^t \nabla_{ij}^2 f(X_s)d[(X^i)^c, (X^j)^c]_t \tag{2.18}$$

$$+ \sum_{0<s\le t} \{f(X_s) - f(X_{s^-}) - D_i f(X_{s^-})\Delta X_s^i\}. \tag{2.19}$$

Proof (cf. Rogers, L.C.G. and Williams, D. [40]).

■

Proposition 21 *(Burkholder-Davis-Gundy's inequality.) If $(M_t, t \ge 0)$ is a càdlàg martingale, then*

$$c_p\mathbb{E}([M,M]_t)^{\frac{p}{2}} \le \mathbb{E}|\sup_{s\le t} M_s|^p \le C_p\mathbb{E}([M,M]_t)^{\frac{p}{2}} \tag{2.20}$$

for any $p \ge 1$.

Proof (cf. Dellacherie and Meyer [16]).

■

3. Smoluchowski-Kramers Approximation with Lévy Noise

In this chapter we prove some convergence results. Section 3.1 we deal with a general Lévy noise, then in the second section we prove a stronger result when driving by a Poisson random measure.

3.1 S-K Approximation Driven by a General Lévy Noise

Let a Lévy process with characteristics (b, Q, ν) be given such that the corresponding Lévy measure ν satisfies

$$\int_{|x|\geq 1} |x|^2 \nu(dx) < \infty \tag{3.1}$$

Remark 22 *In this case, as seen in Knäble [33], since $\int_{|x|\geq 1} xN(t,dx) = \int_{|x|\geq 1} x\tilde{N}(t,dx) + \int_{|x|\geq 1} x\nu(dx)$, the Lévy-Itô decomposition can be further rewritten as*

$$L(t) = mt + B_Q(t) + \int_{\mathbb{R}^{d_1}} x\tilde{N}(t,dx) \tag{3.2}$$

where $m = b + \int_{|x|\geq 1} x\nu(dx)$.

Define $M(t,A) = \int_{A-\{0\}} x\tilde{N}(t,dx), A \in \mathscr{B}(\mathbb{R}^{d_1})$, then $(M(\cdot,A))_t$ is a pure jump type martingale with $[M(t,A)] = \int_{A-\{0\}} x^2 N(t,dx)$ as seen from equation (2.8).

We consider solutions of the following equation:

$$\begin{cases} dX_t^\mu = Y_t^\mu dt, X_0^\mu = x_0 \\ dY_t^\mu = \frac{1}{\mu} b(t, X_t^\mu)dt + \frac{1}{\mu}\sigma(t, X_t^\mu)dL_t - \frac{1}{\mu}dX_t^\mu, Y_0^\mu = y_0 \end{cases} \tag{3.3}$$

and the first order equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dL_t, X_0 = x_0 \tag{3.4}$$

where $L(t)$ is a Lévy process taking values on $\mathbb{R}^{d_1}$ with ν satisfies equation (3.1). Thus the integration with respect to L is understood as integration with respect to dt, a Brownian motion and a martingale $M(\cdot, \mathbb{R}^{d_1})$.

We assume $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^{d \times d_1}$ are measurable functions. Suppose that $b(t, x)$ and $\sigma(t, x)$ are bounded uniformly in t and x and satisfy the global Lipschitz condition:

$$|b(t,x) - b(t,y)| + \|\sigma(t,x) - \sigma(t,y)\| \leq K|x-y|, \forall x, y \in \mathbb{R}^{d_1}, t \in [0,T] \qquad (3.5)$$

Then there exist unique solutions X^μ and X for these two equations (cf. Ikeda and Watanabe [26] Theorem 4.9.1).

Proposition 23 *Assume that* $\int_{|x| \geq 1} |x|^n \nu(dx) < \infty$ *for some* $n = 2^m$, $m \in \mathbb{N}$. *Then for every* $T \in [0, \infty)$ *we have*

$$\sup_{t \in [0,T]} E|X_t^\mu - X_t|^n \to 0 \text{ as } \mu \to 0 \qquad (3.6)$$

Proof We have $dX_t^\mu = Y_t^\mu dt$. We solve for Y_t^μ from the second equation of (3.3), and then integrate to get X_t^μ.

$$\begin{aligned} Y_t^\mu &= e^{-\frac{t}{\mu}} y_0 + e^{-\frac{t}{\mu}} \frac{1}{\mu} \int_0^t e^{\frac{s}{\mu}} (b(s, X_s^\mu) + \sigma(s, X_s^\mu) m) ds \\ &+ e^{-\frac{t}{\mu}} \frac{1}{\mu} \int_0^t e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) dB_Q(s) \\ &+ e^{-\frac{t}{\mu}} \frac{1}{\mu} \int_0^t \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) M(ds, dx) \end{aligned} \qquad (3.7)$$

By the integration-by-parts formula in Proposition 19.

$$
\begin{aligned}
& \int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) M(dr, dx) ds \\
= & \int_0^t \int_0^s \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) M(dr, dx) d(-e^{-\frac{s}{\mu}}) \\
= & \int_0^t \int_0^{s-} \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) M(dr, dx) d(-e^{-\frac{s}{\mu}}) \\
= & \int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_s^\mu) M(ds, dx) - e^{-\frac{t}{\mu}} \int_0^t \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) M(ds, dx)
\end{aligned}
$$

where we used $[\int_0^{\cdot} \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) M(dr, dx), -e^{-\frac{\cdot}{\mu}}]_s = 0$ above.

Similarly, for the Brownian stochastic integral and Lebesgue integral we have

$$
\int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s e^{\frac{r}{\mu}} f(r) dr ds = \int_0^t f(s) ds - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} f(s) ds \tag{3.8}
$$

$$
\int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s e^{\frac{r}{\mu}} f(r) dB(r) ds = \int_0^t f(s) dB(s) - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} f(s) dB(s) \tag{3.9}
$$

Applying these three formulas in the second equality below, we obtain

$$
\begin{aligned}
X_t^\mu = & \; x_0 + \mu(1 - e^{-\frac{t}{\mu}}) y_0 + \int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s e^{\frac{r}{\mu}} (b(r, X_r^\mu) + \sigma(r, X_r^\mu) m) dr ds \\
& + \int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) dB(r) ds \\
& + \int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) M(dr, dx) ds \\
= & \; x_0 + \mu(1 - e^{-\frac{t}{\mu}}) y_0 + \int_0^t (b(s, X_s^\mu) + \sigma(s, X_s^\mu) m) ds \\
& - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} (b(s, X_s^\mu) + \sigma(s, X_s^\mu) m) ds \\
& + \int_0^t \sigma(s, X_s^\mu) dB(s) - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) dB(s) \\
& + \int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_s^\mu) M(ds, dx) - e^{-\frac{t}{\mu}} \int_0^t \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) M(ds, dx)
\end{aligned} \tag{3.10}
$$

$$\begin{aligned}X_t^\mu - X_t &= \mu(1-e^{-\frac{t}{\mu}})y_0\\&+\int_0^t [b(s,X_s^\mu)-b(s,X_s)+(\sigma(s,X_s^\mu)-\sigma(s,X_s^\mu))m]ds\\&-e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}(b(s,X_s^\mu)+\sigma(s,X_s^\mu)m)ds\\&+\int_0^t[\sigma(s,X_s^\mu)-\sigma(s,X_s)]dB(s)-e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)dB(s)\\&+\int_0^t\int_{\mathbb{R}^{d_1}}[\sigma(s,X_s^\mu)-\sigma(s,X_s)]M(ds,dx)\\&-e^{-\frac{t}{\mu}}\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)M(ds,dx) \qquad (3.11)\end{aligned}$$

Then

$$\begin{aligned}E|X_t^\mu - X_t|^n &\leq C_1\mu|y_0|^n\\&+C_1E|\int_0^t[b(s,X_s^\mu)-b(s,X_s)+(\sigma(s,X_s^\mu)-\sigma(s,X_s^\mu))m]ds|^n\\&+C_1E|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}(b(s,X_s^\mu)+\sigma(s,X_s^\mu)m)ds|^n\\&+C_1E|\int_0^t[\sigma(s,X_s^\mu)-\sigma(s,X_s)]dB(s)|^n\\&+C_1E|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)dB(s)|^n\\&+C_1E|\int_0^t\int_{\mathbb{R}^{d_1}}[\sigma(s,X_s^\mu)-\sigma(s,X_s)]M(ds,dx)|^n\\&+C_1E|e^{-\frac{t}{\mu}}\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)M(ds,dx)|^n\\&= I_1+I_2+I_3+I_4+I_5+I_6+I_7 \qquad (3.12)\end{aligned}$$

Let us consider the terms separately, $I_1 = C_1\mu|y_0|^n$.

$$\begin{aligned}I_2 &= C_1E|\int_0^t[b(s,X_s^\mu)-b(s,X_s)+(\sigma(s,X_s^\mu)-\sigma(s,X_s^\mu))m]ds|^n\\&\leq C_1T^{n-1}E\int_0^t|b(s,X_s^\mu)-b(s,X_s)+(\sigma(s,X_s^\mu)-\sigma(s,X_s^\mu))m|^n ds\\&\leq C_1T^{n-1}2^nK^n(1+m^n)\int_0^t E|X_s^\mu - X_s|^n ds \qquad (3.13)\end{aligned}$$

$$\begin{aligned} I_3 &= C_1 E|e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} (b(s, X_s^\mu) + \sigma(s, X_s^\mu) m) ds|^n \\ &= C_1 (\|b\| + \|\sigma\| m)^n |e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} ds|^n \\ &\leq C_1 (\|b\| + \|\sigma\| m)^n |\mu|^n \end{aligned} \tag{3.14}$$

$$\begin{aligned} I_4 &= C_1 E| \int_0^t [\sigma(s, X_s^\mu) - \sigma(s, X_s)] dB(s)|^n \\ &\leq C_1 C_n E(\int_0^t |\sigma(s, X_s^\mu) - \sigma(s, X_s)|^2 ds)^{\frac{n}{2}} \\ &\leq C_1 C_n C_T \int_0^t E|\sigma(s, X_s^\mu) - \sigma(s, X_s)|^n ds \\ &\leq C_1 C_n C_T K^n \int_0^t E|X_s^\mu - X_s|^n ds \end{aligned} \tag{3.15}$$

$$\begin{aligned} I_5 &= C_1 E|e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) dB(s)|^n \\ &= C_1 e^{-\frac{nt}{\mu}} E| \int_0^t e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) dB(s)|^n \\ &\leq C_1 C_n e^{-\frac{nt}{\mu}} E(\int_0^t e^{\frac{2s}{\mu}} |\sigma(s, X_s^\mu)|^2 ds)^{\frac{n}{2}} \\ &\leq C_1 C_n \|\sigma\|^n e^{-\frac{nt}{\mu}} (\int_0^t e^{\frac{2s}{\mu}} ds)^{\frac{n}{2}} \\ &\leq C_1 C_n \|\sigma\|^n \mu^{\frac{n}{2}} \end{aligned} \tag{3.16}$$

To estimate the rest two terms, we will apply the famous BDG inequality Proposition 21. First we need to find the quadratic variation of $\int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_s^\mu) M(ds, dx)$. By Proposition 18,

$$[\int_0^\cdot \int_{\mathbb{R}^{d_1}} \sigma(s, X_s^\mu) M(ds, dx)]_t = \int_0^t |\sigma(s, X_s^\mu)|^2 d[M(\cdot, \mathbb{R}^{d_1})]_s \tag{3.17}$$

Then by formula (2.8), since $M(s, \mathbb{R}^{d_1})$ is a pure jump process,

$$[M(\cdot, \mathbb{R}^{d_1})]_s = \int_0^s \int_{\mathbb{R}^{d_1}} |x|^2 N(dr, dx) \tag{3.18}$$

Thus,

$$[\int_0^{\cdot}\int_{\mathbb{R}^{d_1}}\sigma(s,X_s^\mu)M(ds,dx)]_t=\int_0^t\int_{\mathbb{R}^{d_1}}|\sigma(s,X_s^\mu)|^2|x|^2N(ds,dx) \tag{3.19}$$

Below we apply BDG inequality repeatedly to reduce the power of $N(ds,dx)$ integration by a factor of 2. We denote $M_n=\int_{\mathbb{R}^{d_1}}|x|^n\nu(dx)$ for $n\in\mathbb{N}$. We also use C_n and $C_n^{'}$ to keep track the constants depending on n, but they may vary line by line.

$$\begin{aligned}
I_6 &= C_1E|\int_0^t\int_{\mathbb{R}^{d_1}}[\sigma(s,X_s^\mu)-\sigma(s,X_s)]M(ds,dx)|^n \\
&\le C_1C_nE\{|\int_0^t\int_{\mathbb{R}^{d_1}}|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^2|x|^2N(ds,dx)|^{\frac{n}{2}}\} \\
&\le C_1C_nE\{|\int_0^t\int_{\mathbb{R}^{d_1}}|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^2|x|^2\tilde{N}(ds,dx)|^{\frac{n}{2}}\} \\
&\quad +C_1C_nE\{|\int_0^t\int_{\mathbb{R}^{d_1}}|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^2|x|^2\nu(dx)ds|^{\frac{n}{2}}\} \\
&\le C_1C_nC_n^{'}E\{|\int_0^t\int_{\mathbb{R}^{d_1}}|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^4|x|^4N(ds,dx)|^{\frac{n}{4}}\} \\
&\quad +C_1C_nE\{|\int_0^t|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^2dsM_2|^{\frac{n}{2}}\} \\
&\le C_1C_nC_n^{'}E\{|\int_0^t\int_{\mathbb{R}^{d_1}}|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^4|x|^4N(ds,dx)|^{\frac{n}{4}}\} \\
&\quad +C_1C_nC_TM_2^{\frac{n}{2}}E\{\int_0^t|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^nds\} \\
&\le C_1C_nC_n^{'}E\{|\int_0^t\int_{\mathbb{R}^{d_1}}|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^4|x|^4N(ds,dx)|^{\frac{n}{4}}\} \\
&\quad +C_1C_nC_TM_2^{\frac{n}{2}}K^n\int_0^tE|X_s^\mu-X_s|^nds
\end{aligned}$$

The last step is:

$$\begin{aligned}
&E|\int_0^t\int_{\mathbb{R}^{d_1}}|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^n|x|^nN(ds,dx)| \\
=&\int_0^t\int_{\mathbb{R}^{d_1}}E|\sigma(s,X_s^\mu)-\sigma(s,X_s)|^n|x|^n\nu(dx)ds \\
\le& M_nK^n\int_0^tE|X_s^\mu-X_s|^nds
\end{aligned}$$

Thus

$$I_6 \leq C(n,T,K,\nu)\int_0^t E|X_s^\mu - X_s|^n ds \tag{3.20}$$

$$\begin{aligned}
I_7 &= C_1 E|e^{-\frac{t}{\mu}}\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)M(ds,dx)|^n \\
&\leq C_1 e^{-\frac{nt}{\mu}}E\{|\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{2s}{\mu}}|\sigma(s,X_s^\mu)|^2|x|^2N(ds,dx)|^{\frac{n}{2}}\} \\
&= C_1 e^{-\frac{nt}{\mu}}E\{|\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{2s}{\mu}}|\sigma(s,X_s^\mu)|^2|x|^2\tilde{N}(ds,dx) \\
&\quad +\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{2s}{\mu}}|\sigma(s,X_s^\mu)|^2|x|^2\nu(dx)ds|^{\frac{n}{2}}\} \\
&\leq C_1C_n e^{-\frac{nt}{\mu}}E|\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{2s}{\mu}}|\sigma(s,X_s^\mu)|^2|x|^2\tilde{N}(ds,dx)|^{\frac{n}{2}} \\
&\quad +C_1C_n e^{-\frac{nt}{\mu}}E|\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{2s}{\mu}}|\sigma(s,X_s^\mu)|^2|x|^2\nu(dx)ds|^{\frac{n}{2}} \\
&\leq C_1C_nC_n' e^{-\frac{nt}{\mu}}E\{|\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{4s}{\mu}}|\sigma(s,X_s^\mu)|^4|x|^4N(ds,dx)|^{\frac{n}{4}}\} \\
&\quad +C_1C_nM_2^{\frac{n}{2}}e^{-\frac{nt}{\mu}}\|\sigma\|^n|\int_0^t e^{\frac{2s}{\mu}}ds|^{\frac{n}{2}} \\
&\leq C_1C_nC_n' e^{-\frac{nt}{\mu}}E\{|\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{4s}{\mu}}|\sigma(s,X_s^\mu)|^4|x|^4N(ds,dx)|^{\frac{n}{4}}\} + C_1C_nM_2^{\frac{n}{2}}\|\sigma\|^n\mu^{\frac{n}{2}}
\end{aligned}$$

for the same reason as I_6, we have

$$I_7 \leq C(n,\nu)\mu + C(n,\nu)\mu^{\frac{n}{2}} \tag{3.21}$$

So in all we have

$$E|X_t^\mu - X_t|^n \leq \mu C(|y_0|,n,T,K,\|\sigma\|,\|b\|,\nu,m) + C(T,K,n,m,\nu)\int_0^t E|X_s^\mu - X_s|^n ds \tag{3.22}$$

By Gronwall's lemma, the claim follows. ∎

3.2 S-K Approximation Driven by a Poisson Random Measure

Now consider the case that the driving process is the Poisson random measure N over $\mathbb{R}^+ \times \mathbb{R}^{d_1}$ with intensity measure $dt \bigotimes \nu(dx)$, where ν is a finite measure on $\mathbb{R}^{d_1}$.

Thus we consider solutions of the following equations

$$\begin{cases} dX_t^\mu = Y_t^\mu dt, X_0^\mu = x_0 \\ dY_t^\mu = \frac{1}{\mu} b(t, X_t^\mu)dt + \frac{1}{\mu} \int_{\mathbb{R}^{d_1}} \sigma(t, X_t^\mu) N(dt, dx) - \frac{1}{\mu} dX_t^\mu, Y_0^\mu = y_0 \end{cases} \tag{3.23}$$

and

$$dX_t = b(t, X_t)dt + \int_{\mathbb{R}^{d_1}} \sigma(t, X_t) N(dt, dx), X_0 = x_0 \tag{3.24}$$

We again assume $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^{d \times d_1}$ are measurable functions. Suppose that $b(t, x)$ and $\sigma(t, x)$ are bounded uniformly in t and x and satisfy the global Lipschitz condition (3.5). Again it is well known that equation (3.23) and (3.24) admit a unique (strong) solution (cf. Ikeda and Watanabe [26] Theorem 4.9.1).

We fix $T \in [0, \infty)$, and take the space $D([0, T], \mathbb{R}^d)$ to be the space of càdlàg functions on $[0, T]$ with the Skorokhod topology, which makes it a polish space.

Lemma 24 *Suppose the Lévy measure ν is a finite measure on $\mathbb{R}^{d_1}$, then*

$$\sup_{0 \le t \le T} E|X^\mu(t) - X(t)| \to 0 \ \textit{as}\ \mu \to 0 \tag{3.25}$$

Proof

$$Y_t^\mu = e^{-\frac{t}{\mu}} y_0 + e^{-\frac{t}{\mu}} \frac{1}{\mu} \int_0^t e^{\frac{s}{\mu}} b(s, X_s^\mu) ds + e^{-\frac{t}{\mu}} \frac{1}{\mu} \int_0^t \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) N(ds, dx) \tag{3.26}$$

By the integration by parts formula in Proposition 19, and using the fact that $[\int_0^\cdot \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) N(dr, dx), -e^{-\frac{\cdot}{\mu}}]_s = 0$

$$
\begin{aligned}
& \int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) N(dr, dx) ds \\
= & \int_0^t \int_0^s \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) N(dr, dx) d(-e^{-\frac{s}{\mu}}) \\
= & \int_0^t \int_0^{s-} \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) N(dr, dx) d(-e^{-\frac{s}{\mu}}) \\
= & \int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_s^\mu) N(ds, dx) - e^{-\frac{t}{\mu}} \int_0^t \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) N(ds, dx) \quad (3.27)
\end{aligned}
$$

Applying the above equation to the second equality below:

$$
\begin{aligned}
X_t^\mu &= x_0 + \mu(1 - e^{-\frac{t}{\mu}}) y_0 + \int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s e^{\frac{r}{\mu}} b(r, X_r^\mu) dr ds \\
&\quad + \int_0^t e^{-\frac{s}{\mu}} \frac{1}{\mu} \int_0^s \int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}} \sigma(r, X_r^\mu) N(dr, dx) ds \\
&= x_0 + \mu(1 - e^{-\frac{t}{\mu}}) y_0 + \int_0^t b(s, X_s^\mu) ds - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} b(s, X_s^\mu) ds \\
&\quad + \int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_s^\mu) N(ds, dx) - e^{-\frac{t}{\mu}} \int_0^t \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) N(ds, dx) \quad (3.28)
\end{aligned}
$$

Thus

$$
\begin{aligned}
E|X_t^\mu - X_t| &\le \mu|y_0| + E\left|\int_0^t [b(s, X_s^\mu) - b(s, X_s)] ds\right| + E\left|e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} b(s, X_s^\mu) ds\right| \\
&\quad + E\left|\int_0^t \int_{\mathbb{R}^{d_1}} [\sigma(s, X_s^\mu) - \sigma(s, X_s)] N(ds, dx)\right| \\
&\quad + E\left|e^{-\frac{t}{\mu}} \int_0^t \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) N(ds, dx)\right| \\
&\le \mu|y_0| + K \int_0^t E|X_s^\mu - X_s| ds + \mu\|b\| \\
&\quad + K\nu(\mathbb{R}^{d_1}) \int_0^t E|X_s^\mu - X_s| ds + \mu\|\sigma\|\nu(\mathbb{R}^{d_1}) \quad (3.29)
\end{aligned}
$$

The assertion follows from Gronwall's lemma.

■

Proposition 25 *Let $\{\mu_n | n \in \mathbb{N}\}$ be a sequence of positive real numbers converging to 0, and X^{μ_n} the solution of the second order equation (3.23) corresponding to μ_n, sup-*

pose that the Lévy measure ν is a finite measure on $\mathbb{R}^{d_1}$, then the family of probability measures $\{\mathcal{L}(X^{\mu_n})\}_{n\in\mathbb{N}}$ is tight on $D([0,T],\mathbb{R}^d)$.

Proof To show the tightness of X^{μ_n}, we apply Aldous' tightness criterion, which consists of two conditions. Condition 1:

$$\lim_{a\to\infty}\overline{\lim_{n\to\infty}}P[\sup_{t\le T}|X^{\mu_n}(t)|\ge a]=0 \text{ for each } T>0 \tag{3.30}$$

Condition 2: For each $\epsilon,\eta,T>0$, there exist a $\delta_0>0$ and an $n_0\in\mathbb{N}$ such that, if $\delta\le\delta_0$ and $n\ge n_0$, and if τ is a X^{μ_n}-stopping time satisfying $\tau\le T$, then

$$P[|X^{\mu_n}_{\tau+\delta}-X^{\mu_n}_{\tau}|\ge\epsilon]\le\eta \tag{3.31}$$

To verify the first condition, we use the first representation of equation (3.28):

$$\begin{aligned}X^{\mu}_t &= x_0+\mu(1-e^{-\frac{t}{\mu}})y_0+\int_0^t e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s e^{\frac{r}{\mu}}b(r,X^{\mu}_r)drds\\ &\quad+\int_0^t e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s\int_{\mathbb{R}^{d_1}}e^{\frac{r}{\mu}}\sigma(r,X^{\mu}_r)N(dr,dx)ds\end{aligned} \tag{3.32}$$

Then

$$
\begin{aligned}
E\sup_{t\le T}|X_t^\mu| &\le |x_0|+\mu|y_0|+E\sup_{t\le T}|\int_0^t e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s e^{\frac{r}{\mu}}b(r,X_r^\mu)drds| \\
&\quad +E\sup_{t\le T}|\int_0^t e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s\int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}}\sigma(r,X_r^\mu)N(dr,dx)ds| \\
&\le |x_0|+\mu|y_0|+E\int_0^T e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s e^{\frac{r}{\mu}}|b(r,X_r^\mu)|drds \\
&\quad +E\int_0^T e^{-\frac{s}{\mu}}\frac{1}{\mu}|\int_0^s\int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}}\sigma(r,X_r^\mu)N(dr,dx)|ds \\
&\le |x_0|+\mu|y_0|+E\int_0^T e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s e^{\frac{r}{\mu}}|b(r,X_r^\mu)|drds \\
&\quad +E\int_0^T e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s\int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}}|\sigma(r,X_r^\mu)|N(dr,dx)ds \\
&\le |x_0|+\mu|y_0|+\|b\|\int_0^T e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s e^{\frac{r}{\mu}}drds \\
&\quad +\int_0^T e^{-\frac{s}{\mu}}\frac{1}{\mu}\int_0^s\int_{\mathbb{R}^{d_1}} e^{\frac{r}{\mu}}E|\sigma(r,X_r^\mu)|\nu(dx)drds \\
&\le |x_0|+\mu|y_0|+\|b\|T+\|\sigma\|\nu(\mathbb{R}^{d_1})T \\
&\le C(|x_0|,|y_0|,\|b\|,\|\sigma\|,\nu(\mathbb{R}^{d_1}),T)
\end{aligned}
$$

Plugging in μ_n, it is easy to see that the first condition is fulfilled. For the second condition, we use the second representation of equation (3.28), namely

$$
\begin{aligned}
X_t^\mu &= x_0+\mu(1-e^{-\frac{t}{\mu}})y_0 \\
&\quad +\int_0^t b(s,X_s^\mu)ds-e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}b(s,X_s^\mu)ds \\
&\quad +\int_0^t\int_{\mathbb{R}^{d_1}}\sigma(s,X_s^\mu)N(ds,dx)-e^{-\frac{t}{\mu}}\int_0^t\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx)
\end{aligned}
\tag{3.33}
$$

$$\begin{aligned}
E|X^\mu_{\tau+\delta} - X^\mu_\tau| &\leq E|\mu(1-e^{-\frac{\tau+\delta}{\mu}})y_0 - \mu(1-e^{-\frac{\tau}{\mu}})y_0| \\
&+E|\int_0^{\tau+\delta} b(s,X^\mu_s)ds - \int_0^\tau b(s,X^\mu_s)ds| \\
&+E|e^{-\frac{\tau}{\mu}}\int_0^\tau e^{\frac{s}{\mu}} b(s,X^\mu_s)ds - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau+\delta} e^{\frac{s}{\mu}} b(s,X^\mu_s)ds| \\
&+E|\int_0^{\tau+\delta}\int_{\mathbb{R}^{d_1}} \sigma(s,X^\mu_s)N(ds,dx) - \int_0^\tau\int_{\mathbb{R}^{d_1}} \sigma(s,X^\mu_s)N(ds,dx)| \\
&+E|e^{-\frac{\tau}{\mu}}\int_0^\tau\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X^\mu_s)N(ds,dx) \\
&-e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau+\delta}\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X^\mu_s)N(ds,dx)| \\
&\leq 2\mu|y_0| + I_2 + I_3 + I_4 + I_5
\end{aligned}$$

Considering the terms separately,

$$I_2 = E|\int_0^{\tau+\delta} b(s,X^\mu_s)ds - \int_0^\tau b(s,X^\mu_s)ds| \leq E\{\int_0^T \chi_{[\tau,\tau+\delta]}(s)|b(s,X^\mu_s)|ds\} = \|b\|\delta$$

$$\begin{aligned}
I_3 &= E|e^{-\frac{\tau}{\mu}}\int_0^\tau e^{\frac{s}{\mu}} b(s,X^\mu_s)ds - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau+\delta} e^{\frac{s}{\mu}} b(s,X^\mu_s)ds| \\
&\leq E|e^{-\frac{\tau}{\mu}}\int_0^\tau e^{\frac{s}{\mu}} b(s,X^\mu_s)ds - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau} e^{\frac{s}{\mu}} b(s,X^\mu_s)ds| \\
&\quad +E|e^{-\frac{\tau+\delta}{\mu}}\int_0^\tau e^{\frac{s}{\mu}} b(s,X^\mu_s)ds - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau+\delta} e^{\frac{s}{\mu}} b(s,X^\mu_s)ds| \\
&= J_1 + J_2
\end{aligned}$$

$$
\begin{aligned}
J_1 &= E|e^{-\frac{\tau}{\mu}}\int_0^\tau e^{\frac{s}{\mu}}b(s,X_s^\mu)ds - e^{-\frac{\tau+\delta}{\mu}}\int_0^\tau e^{\frac{s}{\mu}}b(s,X_s^\mu)ds| \\
&= E|(e^{-\frac{\tau}{\mu}} - e^{-\frac{\tau+\delta}{\mu}})\int_0^\tau e^{\frac{s}{\mu}}b(s,X_s^\mu)ds| \\
&= E|e^{-\frac{\tau}{\mu}}(1-e^{-\frac{\delta}{\mu}})\int_0^\tau e^{\frac{s}{\mu}}b(s,X_s^\mu)ds| \\
&\leq E\{e^{-\frac{\tau}{\mu}}\frac{\delta}{\mu}\int_0^\tau e^{\frac{s}{\mu}}|b(s,X_s^\mu)|ds\} \\
&\leq \|b\|\frac{\delta}{\mu}E\{e^{-\frac{\tau}{\mu}}\int_0^\tau e^{\frac{s}{\mu}}ds\} \\
&\leq \delta\|b\|
\end{aligned}
$$

$$
\begin{aligned}
J_2 &= E|e^{-\frac{\tau+\delta}{\mu}}\int_0^\tau e^{\frac{s}{\mu}}b(s,X_s^\mu)ds - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau+\delta} e^{\frac{s}{\mu}}b(s,X_s^\mu)ds| \\
&= E|e^{-\frac{\tau+\delta}{\mu}}\int_\tau^{\tau+\delta} e^{\frac{s}{\mu}}b(s,X_s^\mu)ds| \\
&\leq E\{e^{-\frac{\tau+\delta}{\mu}}\int_\tau^{\tau+\delta} e^{\frac{s}{\mu}}|b(s,X_s^\mu)|ds\} \\
&\leq \|b\|E(\int_\tau^{\tau+\delta} ds) \\
&= \|b\|\delta
\end{aligned}
$$

$$
\begin{aligned}
I_4 &= E|\int_0^{\tau+\delta}\int_{\mathbb{R}^{d_1}}\sigma(s,X_s^\mu)N(ds,dx) - \int_0^\tau\int_{\mathbb{R}^{d_1}}\sigma(s,X_s^\mu)N(ds,dx)| \\
&\leq E\{\int_0^T\int_{\mathbb{R}^{d_1}}\chi_{[\tau,\tau+\delta]}(s)|\sigma(s,X_s^\mu)|N(ds,dx)\} \\
&= \int_0^T\int_{\mathbb{R}^{d_1}}E(\chi_{[\tau,\tau+\delta]}(s)|\sigma(s,X_s^\mu)|)\nu(dx)ds \\
&\leq \|\sigma\|\nu(\mathbb{R}^{d_1})\int_0^T E(\chi_{[\tau,\tau+\delta]}(s))ds \\
&= \|\sigma\|\nu(\mathbb{R}^{d_1})E\int_0^T \chi_{[\tau,\tau+\delta]}(s)ds \\
&= \|\sigma\|\nu(\mathbb{R}^{d_1})\delta
\end{aligned}
$$

$$
\begin{aligned}
I_5 &= E|e^{-\frac{\tau}{\mu}}\int_0^\tau\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx) - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau+\delta}\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx)| \\
&\leq E|e^{-\frac{\tau}{\mu}}\int_0^\tau\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx) - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau}\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx)| \\
&\quad +E|e^{-\frac{\tau+\delta}{\mu}}\int_0^\tau\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx) - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau+\delta}\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx)| \\
&= J_3 + J_4
\end{aligned}
$$

$$
\begin{aligned}
J_3 &= E|e^{-\frac{\tau}{\mu}}\int_0^\tau\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx) - e^{-\frac{\tau+\delta}{\mu}}\int_0^{\tau}\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx)| \\
&= E|(e^{-\frac{\tau}{\mu}} - e^{-\frac{\tau+\delta}{\mu}})\int_0^\tau\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx)| \\
&= E|(e^{-\frac{\tau}{\mu}})(1 - e^{-\frac{\delta}{\mu}})\int_0^\tau\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}\sigma(s,X_s^\mu)N(ds,dx)| \\
&\leq E(e^{-\frac{\tau}{\mu}}\frac{\delta}{\mu}\int_0^\tau\int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}}|\sigma(s,X_s^\mu)|N(ds,dx)) \\
&= \frac{\delta}{\mu}E(\int_0^T\int_{\mathbb{R}^{d_1}} \chi_{[0,\tau]}e^{-\frac{\tau-s}{\mu}}|\sigma(s,X_s^\mu)|N(ds,dx)) \\
&= \frac{\delta}{\mu}\int_0^T\int_{\mathbb{R}^{d_1}} E|\chi_{[0,\tau]}e^{-\frac{\tau-s}{\mu}}\sigma(s,X_s^\mu)|\nu(dx)ds \\
&= \frac{\delta}{\mu}\|\sigma\|\nu(\mathbb{R}^{d_1})\int_0^T E(\chi_{[0,\tau]}e^{-\frac{\tau-s}{\mu}})ds \\
&= \frac{\delta}{\mu}\|\sigma\|\nu(\mathbb{R}^{d_1})E(e^{-\frac{\tau}{\mu}}\int_0^\tau e^{\frac{s}{\mu}}ds) \\
&\leq \delta\|\sigma\|\nu(\mathbb{R}^{d_1})
\end{aligned}
$$

$$\begin{aligned}
J_4 &= E|e^{-\frac{\tau+\delta}{\mu}} \int_0^\tau \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) N(ds, dx) - e^{-\frac{\tau+\delta}{\mu}} \int_0^{\tau+\delta} \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) N(ds, dx)| \\
&= E|e^{-\frac{\tau+\delta}{\mu}} \int_\tau^{\tau+\delta} \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X_s^\mu) N(ds, dx)| \\
&\leq E(e^{-\frac{\tau+\delta}{\mu}} \int_\tau^{\tau+\delta} \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} |\sigma(s, X_s^\mu)| N(ds, dx)) \\
&\leq \|\sigma\| E(\int_\tau^{\tau+\delta} \int_{\mathbb{R}^{d_1}} N(ds, dx)) \\
&= \|\sigma\| \int_0^T \int_{\mathbb{R}^{d_1}} E(\chi_{[\tau,\tau+\delta]}) \nu(dx) ds \\
&= \|\sigma\| \nu(\mathbb{R}^{d_1}) \int_0^T E(\chi_{[\tau,\tau+\delta]}) ds \\
&= \|\sigma\| \nu(\mathbb{R}^{d_1}) E(\int_0^T \chi_{[\tau,\tau+\delta]} ds) \\
&= \|\sigma\| \nu(\mathbb{R}^{d_1}) \delta
\end{aligned}$$

Adding all up, we have

$$\begin{aligned}
E|X_{\tau+\delta}^\mu - X_\tau^\mu| &\leq I_1 + I_2 + I_3 + I_4 + I_5 \\
&\leq I_1 + I_2 + J_1 + J_2 + I_4 + J_3 + J_4 \\
&\leq 2\mu|y_0| + 3\|b\|\delta + 3\|\sigma\|\nu(\mathbb{R}^{d_1})\delta
\end{aligned}$$

which can be made arbitrary small by letting μ and δ small. Thus condition 2 follows, hence the tightness.

■

Lemma 26 *If X is a process with càdlàg sample paths in $D([0,T];\mathbb{R}^d)$, then the complement in $[0,T]$ of*

$$D(X) \triangleq \{0 \leq t \leq T : P(X(t) = X(t-)) = 1\} \tag{3.34}$$

is at most countable.

Proof (cf. S.N.Ethier and T.G.Kurtz [19] Lemma 3.7.7).

■

Theorem 27 *Assume the Lévy measure ν is a finite measure on $\mathbb{R}^{d_1}$. Let μ_n be a sequence of positive real numbers converging to 0 as n goes to ∞. Let X^{μ_n} be the solution of the second order equation (3.23) corresponding to μ_n and z the solution of the first order equation (3.24). Then, for any $T > 0$ and for any $\epsilon > 0$ we have*

$$\lim_{n\to\infty} P(\|X^{\mu_n} - z\|_{D([0,T];\mathbb{R}^d)} > \epsilon) = 0. \tag{3.35}$$

Proof We base our proof on an observation by Gyöngy and Krylov in [23]. Let Z_n be a sequence of random elements in a Polish space (E, ρ) equipped with the Borel σ-algebra. Then Z_n converges in probability to an E-valued random element if and only if for every pair of subsequences Z_{n_l} and $Z_{n_l'}$ there exists a subsequence $v_k := (Z_{n_{l(k)}}, Z_{n_{l(k)}'})$ converging weakly to a random element v supported on the diagonal $\{(x, y) \in E \times E : x = y\}$.

Now take any pair of subsequences μ_{n_l} and $\mu_{n_l'}$, then $(X^{\mu_{n_l}}, X^{\mu_{n_l'}})$ is a tight family of processes in $D([0,T];\mathbb{R}^{2d})$, due to the tightness of the sequence $\{\mathcal{L}(X^{\mu_n})\}$ in $D([0,T];\mathbb{R}^d)$. By the Skorokhod representation theorem, there exists a sequence of random elements $\{v_k\} := \{(X_1^k, X_2^k, \int_0^\cdot \int_{\mathbb{R}^{d_1}} \hat{N}^k(ds, dx))\}$ in $D([0,T];\mathbb{R}^d)^2 \times D([0,T],\mathbb{R})$, defined on some probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$, such that the law of v_k coincides with the law of $(X^{\mu_{n_{l(k)}}}, X^{\mu_{n_{l(k)}'}}, \int_0^\cdot \int_{\mathbb{R}^{d_1}} N(ds, dx))$, for each k. And v_k converges $\hat{P}$-a.s. to some random element $v := (X_1, X_2, \int_0^\cdot \int_{\mathbb{R}^{d_1}} \hat{N}(ds, dx)) \in D([0,T];\mathbb{R}^d)^2 \times D([0,T],\mathbb{R})$, that is,

$$\|X_i^k - X_i\|_{D([0,T];\mathbb{R}^d)} \to 0, i = 1, 2 \text{ as } k \to \infty$$

$$\| \int_0^\cdot \int_{\mathbb{R}^{d_1}} \hat{N}^k(ds, dx) - \int_0^\cdot \int_{\mathbb{R}^{d_1}} \hat{N}(ds, dx)\|_{D([0,T];\mathbb{R})} \to 0 \text{ as } k \to \infty$$

Now we want to show $X_1 = X_2$. If that is true, then by Gyöngy and Krylov, there exists some $z \in D([0,T], \mathbb{R}^d)$ such that X^{μ_n} converges to z in probability.

Define:

$$\hat{\mathcal{F}}_t^k \triangleq \sigma(X_1^k(s), X_2^k(s), \int_0^s \int_{\mathbb{R}^{d_1}} \hat{N}^k(dr, dx) : s \leq t)$$

$$\hat{\mathcal{F}}_t \triangleq \sigma(X_1(s), X_2(s), \int_0^s \int_{\mathbb{R}^{d_1}} \hat{N}(dr, dx) : s \leq t)$$

Then it is easy to see that for every k the process $(\hat{N}^k, \hat{\mathcal{F}}_t^k)$ and $(\hat{N}, \hat{\mathcal{F}}_t)$ are Poisson random measures with the same distribution as N. Define the set

$$\begin{aligned}
U \quad \triangleq \quad & \{(f, g, h) \in D([0, T], \mathbb{R}^d)^2 \times D([0, T], \mathbb{R}) | \\
& f(t) = x_0 + \mu(1 - e^{-\frac{t}{\mu}})y_0 + \int_0^t b(s, f(s))ds - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} b(s, f(s))ds \\
& + \int_0^t \sigma(s, f(s))dh(s) - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} \sigma(s, f(s))dh(s); \\
& g(t) = x_0 + \mu(1 - e^{-\frac{t}{\mu}})y_0 + \int_0^t b(s, g(s))ds - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} b(s, g(s))ds \\
& + \int_0^t \sigma(s, g(s))dh(s) - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} \sigma(s, g(s))dh(s); \forall t \in [0, T]\}
\end{aligned}$$

Then since

$$\hat{P}((X_1^k, X_2^k, \int_0^\cdot \int_{\mathbb{R}^{d_1}} \hat{N}^k(ds, dx) \in U, \forall k) = P((X^{\mu_{n_{l(k)}}}, X^{\mu_{n'_{l(k)}}}, \int_0^\cdot \int_{\mathbb{R}^{d_1}} N(ds, dx)) \in U, \forall k)$$

and the right hand side of the above equation equals to 1, so the left hand side also equals to 1, that is, both X_1^k and X_2^k verify formula (3.32) with N replaced by $\hat{N}^k$, X by X_1 and X_2 and μ by $\mu_{n_{l(k)}}$ and $\mu_{n'_{l(k)}}$ respectively. We define R_1^k and R_2^k obtained from the expression R below replacing X by X_1 and X_2, μ by $\mu_{n_{l(k)}}$ and $\mu_{n'_{l(k)}}$ and N with $\hat{N}^k$ respectively.

$$\begin{aligned}
R(t) \quad = \quad & \mu(1 - e^{-\frac{t}{\mu}})y_0 - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} b(s, X(s))ds \\
& -e^{-\frac{t}{\mu}} \int_0^t \int_{\mathbb{R}^{d_1}} e^{\frac{s}{\mu}} \sigma(s, X(s))N(ds, dx)
\end{aligned} \tag{3.36}$$

Then $R_1^k(t)$ and $R_2^k(t)$ converge to 0 in $L^1(\hat{\Omega})$, as $\mu_{n_{l(k)}}$ and $\mu_{n'_{l(k)}}$ go to 0. Thus, for a subsequence, they converge $\hat{P}$-a.s. to 0.

Since X_1 and X_2 are processes with sample paths in $D([0,T],\mathbb{R}^d)$, the complement of $D(X_i)$ (cf. Lemma 26) is at most countable. So for $t \in D(X_i)$, $X_i^k(t) \to X_i(t)$ a.s. follows from the convergence in Skorokhod topology. We also have:

$$\int_0^t b(s, X_i^k(s))ds \to \int_0^t b(s, X_i(s))ds \text{ a.s.} \tag{3.37}$$

by Lebesgue dominated convergence theorem. And

$$\begin{aligned}
&\hat{\mathbb{E}}\{|\int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_i^k(s))\hat{N}^k(ds,dx) - \int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_i(s))\hat{N}(ds,dx)|\} \\
\leq\ & \int_0^t \int_{\mathbb{R}^{d_1}} \hat{\mathbb{E}}|\sigma(s, X_i^k(s)) - \sigma(s, X_i(s))|\nu(dx)ds \\
=\ & \int_{\mathbb{R}^{d_1}} \hat{\mathbb{E}} \int_0^t |\sigma(s, X_i^k(s)) - \sigma(s, X_i(s))|ds\nu(dx) \to 0
\end{aligned} \tag{3.38}$$

by Lebesgue dominated convergence theorem with respect to the product measure $\hat{P} \times dx$. Thus possibly for a subsequence,

$$\int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_i^k(s))\hat{N}^k(ds,dx) \to \int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_i(s))\hat{N}(ds,dx) \text{ a.s.} \tag{3.39}$$

So for $t \in D(X_i), i = 1, 2$, we have

$$X_i(t) = x_0 + \int_0^t b(s, X_i(s))ds + \int_0^t \int_{\mathbb{R}^{d_1}} \sigma(s, X_i(s))\hat{N}(ds,dx) \text{ a.s.}, i = 1,2 \tag{3.40}$$

For t outside of $D(X_i)$, we can use right-continuity of the sample path to prove the same equality. Thus both X_1 and X_2 coincide with the solution of equation (3.24) perturbed by the noise $\hat{N}$, which is unique.

It follows that X^{μ_n} converges in probability to some random variable $z \in D([0,T],\mathbb{R}^d)$. But following the above argument again shows that z solves equation (3.24).

■

4. Application in Finance: The Momentum Model

In this chapter we will extend the famous Black-Scholes model and study this extended model using the results in Chapter 3.

The Black-Scholes model assumes:

- A frictionless market, that is, no transactions costs, no taxes, trading being continuous. All securities are perfectly divisible. There are no penalties to short selling.
- The risk-free rate of interest, r, is constant over time, i.e.

$$dB_t = rB_t dt \tag{4.1}$$

- The underlying stock price is a Geometric Brownian Motion, i.e.

$$\frac{dX_t}{X_t} = bdt + \sigma dW_t \tag{4.2}$$

We concentrate on two unreasonable assumptions in the B-S model showed by empirical evidence. One is that the market is assumed to be frictionless, while the other being that the driving process is Brownian motion.

This chapter is organized as follows: We use the first two sections to discuss how to get around the two assumptions, respectively. We introduce a parameter μ to capture the impact of the market frictions in the first section. Then we introduce Generalized hyperbolic Lévy processes to be the driving process. After that - in the third section - we define the momentum model combining the two extensions discussed above.

4.1 Market Friction, Price Delay, Momentum

Ample empirical evidence demonstrates the existence of sizeable market frictions. Researchers in behavior finance have discussed the importance of many market frictions, such as incomplete information (Merton [34], Shapiro [43]), liquidity (Amihud and Mendelson [1]), short sale constraints (Chen, Hong, Stein [12]), taxes (Constantinides [13]). Those frictions can delay the process of information incorporation, and hence there is a delay in asset price response to news. For example, Arbel, Carvell, and Strebel [3] argue institutional forces and transactions costs can delay the process of information incorporation for less visible, segmented firms. In addition, Peng [38] shows that information capacity constraints can cause a delay in asset price responses to news.

Price delay may also result from lack of liquidity of an asset's shares, which can potentially arise from many sources. First, according to Hicks' [24] "liquidity preference" notion, which says investors hold financial assets not only for their returns but also to facilitate adjustments to changes in economic conditions. So when risk-averse investors anticipate a recession, they prefer to invest in less risky, more liquid assets. Secondly, firms themselves can also cause illiquidity. Small and high book-to-market stocks are less liquid.

There are a number of papers document this kind of delayed reaction to news. For example, McQueen, Pinegar, and Thorley [29] find that small stocks react to good common news more slowly than large stocks, but not for bad common news. Hong, Lim, and Stein [30] provide evidence to show that firm-specific information diffuses only gradually across the investing public. This shows that in reality the market is not efficient.

On the other hand, some researchers find that, at medium-term horizons ranging from three to twelve months, stock returns exhibit momentum. For example, Jegadeesh and Titman [32] , using a U.S. sample of NYSE/AMEX stocks over the period from 1965 to 1989, find that past winners on average continue to outperform past losers. The result appears to be robust: Rouwenhorst [41] obtains very similar numbers in a sample of 12 European countries over the period from 1980 to 1995.

The frictions of information diffusion in the market and momentum in stock prices may highly correlate with the underreaction. For example, Chan [11] relates the evidence on momentum to the evidence on the market's underreaction to earnings-related information. Hong [31] assume that if information diffuses gradually across the population, prices underreact in the short run, which means that the momentum traders can make profit by trend-chasing.

So one natural question is: how do we incorporate the impact of those potential frictions on the price process of a stock and capture the momentum effects?

We introduce a parameter μ to measure price delay with which a firm's stock price responds to information. We standardized it such that $0 \leq \mu < 1$, where $\mu = 0$ represents the case that the stock is "infinitely efficient" to incorporate news, and $\mu = 1$ the case that the stock is "infinitely delayed". Delayed firms tend to be small, volatile, less visible, and neglected by many market participants.

The parameter μ should be estimated using empirical data. For example, In Hou [25], "Market Frictions, Price Delay, and the Cross-Section of Expected Returns", several measures are introduced to capture the average delay with which a firm's stock price responds to information. The liquidity measure defined by Liu [28] is also a good candidate.

In order to define the notion of momentum we interpret the price process of a financial product as a motion. Considering the analogue between the mass in physics and the price delay in finance, it is natural to define the momentum as the momentum of the motion of a particle in physics:

Definition 28 *Let $X_t = X_t(\omega)$ be the price process of a financial product, μ is the measure of price delay, as introduced above. The momentum of X_t is defined as μY_t, where $Y_t \triangleq \frac{d}{dt}X_t$, if it exists.*

So the momentum is just the time derivative of the price process multiplied with the price delay μ. We note that in the real market, where price is given tick by tick, the momentum always exist. If we model the price process by a continuous price process, then this will also be the case.

4.2 Generalized Hyperbolic Lévy Motions

It is well known that the normal distribution fit poorly to log returns of most financial assets such as stocks or indices. Empirical densities of log returns show that tiny price movements and big changes occur with higher frequency, small and middle sized movements are less more frequent than predicted by the normal law. In order to achieve a better fit, it is preferable to replace Brownian motion by a Lévy process. In the context of Lévy process, we mention the Meixner process which was introduced in Schoutens, W. and Teugels, J.L. [42]. Barndorf-Nielsen, O.E. proposed the Normal Inverse Gaussian Lévy process in [6]. Eberlein and Keller proposed the Hyperbolic Lévy motion.

For the sake of completeness, we will briefly recall the Generalized Hyperbolic Lévy motions. Readers are referred to Eberlein [17].

Generalized hyperbolic distributions were introduced by Barndorff-Nielsen [4]. Their Lebesgue densities are given by

$$d_{GH}(x;\lambda,\alpha,\beta,\delta,\mu) = a(\lambda,\alpha,\beta,\delta)(\delta^2+(x-\mu)^2)^{(\lambda-\frac{1}{2})/2} K_{\lambda-\frac{1}{2}}(\alpha\sqrt{\delta^2+(x-\mu)^2})\exp(\beta(x-\mu)) \quad (4.3)$$

where

$$a(\lambda,\alpha,\beta,\delta) = \frac{(\alpha^2-\beta^2)^{\lambda/2}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^\lambda K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}$$

is the normalizing constant and K_ν denotes the modified Bessel function of the third kind with index ν. One key integral representation for Bessel functions of the third kind is

$$K_\nu(x) = \frac{1}{2}\int_0^\infty u^{\nu-1}\exp\left[-\frac{1}{2}x\left(u+\frac{1}{u}\right)\right]du$$

The densities above depend on five parameters: $\alpha > 0$ determines the shape, β with $0 \leq |\beta| < \alpha$ the skewness and $\mu \in \mathbb{R}$ the location. $\delta > 0$ is a scaling parameter comparable to σ in the normal distribution, while $\lambda \in \mathbb{R}$ characterizes certain subclasses. By changing λ, we can essentially modify the heaviness of the tails.

There are certain subclasses of interest. The case when $\lambda = 1$ corresponds to the subclass of hyperbolic distributions. The normal inverse Gaussian distribution results when $\lambda = -1/2$.

Generalized hyperbolic distributions have a number of nice analytic properties. Their moment-generating function is given by

$$M_{GH}(u) = e^{\mu u}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^{\lambda/2}\frac{K_\lambda(\delta\sqrt{\alpha^2-(\beta+u)^2})}{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})} \quad (4.4)$$

for $|\beta+u| < \alpha$. From this formula we can see that moments of all integer order exist. This also gives the Lévy-Khintchin representation of the characteristic function of generalized hyperbolic distributions.

$$\ln(\phi_{GH}(u)) = iuE[GH] + \int_{-\infty}^{+\infty}(e^{iux}-1-iux)g(x)dx \quad (4.5)$$

where the density of the Lévy measure is

$$g(x) = \frac{e^{\beta x}}{|x|}\Big(\int_0^\infty \frac{exp(-\sqrt{2y+\alpha^2}|x|)}{\pi^2 y(J_\lambda^2(\delta\sqrt{2y}) + Y_\lambda^2(\delta\sqrt{2y}))dy + \lambda e^{-\alpha|x|}}\Big) \text{ if } \lambda \geq 0 \tag{4.6}$$

and

$$g(x) = \frac{e^{\beta x}}{|x|}\int_0^\infty \frac{exp(-\sqrt{2y+\alpha^2}|x|)}{\pi^2 y(J_{-\lambda}^2(\delta\sqrt{2y}) + Y_{-\lambda}^2(\delta\sqrt{2y}))dy} \text{ if } \lambda < 0 \tag{4.7}$$

Here J_λ and Y_λ are the Bessel functions of the first and second kind, respectively.

Barndorff-Nielsen and Halgreen [5] showed that generalized hyperbolic distributions are infinitely divisible, and as such each member of this family generates a Lévy process. By choosing a càdlàg version, we have the Generalized hyperbolic Lévy motion. We note that since the Lévy-Khintchin representation has only a drift and a jump term, thus Generalized hyperbolic Lévy motions do not have a continuous Gaussian component.

Analyzing the behavior of the densities g of the Lévy measure for $x \to 0$ shows that the Lévy measures have infinte mass in every neighborhood of the origin. This means that the process has an infinite number of small jumps in every finite time interval.

Since there is no continuous component in the Lévy-Ito decomposition, Generalized hyperbolic Lévy motions X_t can be written in the form

$$X_t = bt + \int_{\mathbb{R}} x\tilde{N}(t, dx) \tag{4.8}$$

where $\tilde{N}$ is the compensated Poisson random measure associated with the process $(X_t)_{t\geq 0}$. The compensator of N is deterministic and of the form $dt\nu(dx)$, since $(X_t)_{t\geq 0}$ is a Lévy process.

4.3 The Momentum Model

Now we are ready to introduce the momentum model:

- There exist certain market frictions, captured by the parameter μ.

- The risk-free rate of interest r is constant over time, i.e.

$$dB_t = rB_t dt \tag{4.9}$$

- The price process of the risky asset X_t^μ follows:

$$\begin{aligned} dX_t^\mu &= Y_t^\mu dt \\ \frac{dX_t^\mu}{X_t^\mu} &= bdt + \sigma \int_{\mathbb{R}} x\tilde{N}(dt, dx) - \frac{\mu dY_t^\mu}{X_t^\mu} \end{aligned} \tag{4.10}$$

  where $\tilde{N}$ is the compensated Poisson random measure associated to a generalized hyperbolic Lévy process.

We note that μ is introduced in the model and the driving process becomes a generalized hyperbolic Lévy process. Besides that, in the system (4.10), the first equation states the relationship between the price process and the momentum given the price delay μ according to Definition 28. The second equation models the "law of motion" for the financial product. Comparing to the Black-Scholes model, we subtract $\frac{\mu dY_t^\mu}{X_t^\mu}$ representing the incremental momentum divided by the price of the financial product.

Motivated by considerations from behavioral finance as introduced in Section 4.1, when the measure of price delay μ is not extremely small, investors would like to model the asset price in terms of the momentum model which is essentially the second order Newton equation. This momentum model helps to capture the momentum effect.

Corollary 29 *The system of stochastic differential equations for the the risky asset in the momentum model (4.10) exhibits a unique strong solution* (X_t^μ, Y_t^μ).

Proof (cf. Ikeda and Watanabe [26] Theorem 4.9.1).

■

Corollary 30 *Assume that* $\int_{|x|\geq 1} |x|^n \nu(dx) < \infty$ *for some* $n = 2^m$, $m \in \mathbb{N}$. *Let* X_t *denote the solution of*

$$\frac{dX_t}{X_t} = bdt + \sigma \int_{\mathbb{R}} x\tilde{N}(dt, dx) \tag{4.11}$$

Then for every $T \in [0, \infty)$ *we have*

$$\sup_{t\in[0,T]} E|X_t^\mu - X_t|^n \to 0 \text{ as } \mu \to 0 \tag{4.12}$$

Proof The proof is a direct implication of Proposition 23 by setting $b(t,x) = bx, \sigma(t,x) = \sigma x$.

■

The economic interpretation of the above corollary is what by considering financial products with very little frictions, the price process can be described by an equation of "Black-Scholes type" but with generalized hyperbolic Lévy process as the disturbing noise.

We note that there are a lot of possible future research points for this momentum model, such as:

- What are pricing formulae for our momentum model?
- Can this model be evaluated with the help of empirical research?
- How should the model be calibrated?
- What further economical and mathematical interpretations and heuristics are possible or reasonable?
- Any more parallels between physics and finance? And what consequences do they exhibit?

5. Smoluchowski-Kramers Approximation in Infinite Dimension

In this chapter we will present some remarks on the Smoluchowski-Kramers approximation in infinite dimension.

Section 5.1 states some basic concepts and definitions from semigroup theory. In section 5.2 we introduce a kind of weak solution, the mild solution. Using these tools we establish the existence and uniqueness of the solution for the generator $\frac{\partial}{\partial x}$ in section 5.3.

5.1 Semigroups and Generators

Definition 31 *Let B be a Banach space. A strongly continuous semigroup of bounded linear operators or a C_0-semigroup is a family $(S(t), t \geq 0)$ of bounded linear operators on B with*

- $S(0) = I$*, where I is the identity operator on B;*
- $S(t+s) = S(t)S(s)$ *for every $t, s \geq 0$ (the semigroup property);*
- $\lim_{t \downarrow 0} S(t)u = u$ *for all $u \in B$ (strong continuity).*

Lemma 32 *Let $(S(t), t \geq 0)$ be a C_0-semigroup. Then exist constant $\beta \geq 0$ and $M \geq 1$ such that*

$$\|S(t)\| \leq Me^{\beta t} \text{ for all } t \geq 0 \tag{5.1}$$

Proof (cf. Pazy [37], Theorem 1.2.2). ■

Definition 33 *Given a strongly continuous semigroup* $(S(t), t \geq 0)$ *on* B*, the linear operator* $(A, D(A))$ *on* B *defined by*

$$D(A) \triangleq \{u \in B | \lim_{t \downarrow 0} \frac{1}{t}(S(t)u - u) \text{ exists}\} \tag{5.2}$$

$$Au \triangleq \lim_{t \downarrow 0} \frac{1}{t}(S(t)u - u), u \in D(A) \tag{5.3}$$

is called the (infinitesimal) generator of the semigroup $S(t)$.

Proposition 34 *Let* $S(t)$ *be a* C_0*-semigroup with infinitesimal generator* A*. Then*

- *For* $u \in B$,
$$\lim_{h \to 0} \frac{1}{h} \int_t^{t+h} S(s)uds = S(t)u. \tag{5.4}$$

- *For* $u \in B$, $\int_0^t S(s)uds \in D(A)$ *and*
$$A\left(\int_0^t S(s)uds\right) = S(t)u - u. \tag{5.5}$$

- *For* $u \in D(A)$, $S(t)u \in D(A)$ *and*
$$\frac{d}{dt}S(t)u = AS(t)u = S(t)Au. \tag{5.6}$$

- *For* $u \in D(A)$,
$$S(t)u - S(s)u = \int_s^t S(\tau)Aud\tau = \int_s^t AS(\tau)ud\tau. \tag{5.7}$$

Proof (cf. Pazy [37], Theorem 1.2.4). ■

Proposition 35 *Let* $S(t)$ *and* $T(t)$ *be* C_0 *semigroups of bounded linear operators with infinitesimal generators* A *and* B *respectively. If* $A = B$*, then* $T(t) = S(t)$ *for* $t \geq 0$.

Proof (cf. Pazy [37], Theorem 1.2.6). ■

Proposition 36 *(Hille-Yosida).A linear operator* A *is the infinitesimal generator of a* C_0 *semigroup* $S(t)$ *satisfying* $\|S(t)\| \leq Me^{\beta t}$*, if and only if*

1. *A is closed and $D(A)$ is dense in B;*

2. *The resolvent set $\rho(A)$ of A contains the ray (β, ∞) and the resolvent of A, $R(\lambda : A) = (\lambda I - A)^{-1}, \lambda \in \rho(A)$ satisfy*

$$\|R(\lambda : A)^n\| \leq \frac{M}{(\lambda - \beta)^n} \text{ for } \lambda > \beta, n = 1, 2, ... \tag{5.8}$$

Proof (cf. Pazy [37], Theorem 1.5.3). ∎

Let B^* be the dual of B. For every $u \in B$ we define the duality set $F(u) \subset B^*$ by

$$F(u) = \{f \in B^* | f(u) = \|u\|^2 = \|f\|^2\} \tag{5.9}$$

Definition 37 *A linear operator A is dissipative if for every $u \in D(A)$ there is a $f \in F(u)$ such that $f(Au) \leq 0$.*

Proposition 38 *A linear operator A is dissipative if and only if*

$$\|(\lambda I - A)u\| \geq \lambda \|u\| \text{ for all } u \in D(A) \text{ and } \lambda > 0 \tag{5.10}$$

Proof (cf. Pazy [37], Theorem 1.4.2).

∎

If $B = H$ is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, $A : D(A) \subset B \to B$ is dissipative if and only if

$$\langle A(u), u \rangle \leq 0, \forall u \in D(A) \tag{5.11}$$

Proposition 39 *(Lumer-Phillips). Let A be a linear operator with dense domain $D(A)$ in B.*

- *If A is dissipative and $(\lambda_0 - A)(D(A)) = B$, for some $\lambda_0 > 0$, then A is the infinitesimal generator of a C_0 semigroup of contractions on B;*

- *If A is the infinitesimal generator of a C_0 semigroup of contractions on B, then $(\lambda - A)(D(A)) = B$ for all $\lambda > 0$ and A is dissipative. Moreover, for every $u \in D(A)$, and every $f \in F(u)$, $f(A(u)) \leq 0$.*

Proof (cf. Pazy [37], Theorem 1.4.3).

■

Remark 40 *If A is dissipative and $(\lambda_0 - A)(D(A)) = B$, for some $\lambda_0 > 0$, then A is closed.*

Proof (cf. Pazy [37], Theorem 1.4.3).

■

The following definition of analytic semigroups is essentially from Da Prato and Zabczyk [14], Appendix A.4.

For any $\omega \in \mathbb{R}$ and $\theta \in (0, \pi)$ we denote by $S_{\omega,\theta}$ the sector in $\mathbb{C}$:

$$S_{\omega,\theta} = \{\lambda \in \mathbb{C} - \{\omega\} : |\arg(\lambda - \omega)| \leq \theta\}.$$

Assume A is a linear closed operator satisfies the following condition (**M**):

- $\exists \omega \in \mathbb{R}, \theta_0 \in (\frac{\pi}{2}, \pi) : \rho(A) \supset S_{\omega,\theta_0}$,
- $\exists M > 0$ such that
$$R(\lambda, A) \leq \frac{M}{|\lambda - \omega|}, \forall \lambda \in S_{\omega,\theta_0}.$$

then we can define a semigroup $S(\cdot)$ of bounded linear operators in E by setting $S(0) = I$ and

$$S(t) = \frac{1}{2\pi i} \int_{\gamma_{\epsilon,\theta}} e^{\lambda t} R(\lambda, A) d\lambda, t > 0 \tag{5.12}$$

where $\theta \in (\frac{\pi}{2}, \theta_0)$ and $\gamma_{\epsilon,\theta}$ is the following, oriented counterclockwise, path in $\mathbb{C}$

$$\gamma_{\epsilon,\theta} = \gamma^+_{\epsilon,\theta} \cup \gamma^-_{\epsilon,\theta} \cup \gamma^0_{\epsilon,\theta},$$

$$\gamma^{\pm}_{\epsilon,\theta} = \{z \in \mathbb{C} : z = \omega + re^{\pm i\theta}, r \geq \epsilon\}$$

$$\gamma^0_{\epsilon,\theta} = \{z \in \mathbb{C} : z = \omega + re^{\pm i\eta}, |\eta| \leq \theta\}$$

Proposition 41 *Assume that A fulfils condition ($\mathbf{M}$) and let $S(\cdot)$ be defined by (5.12). Then the following statements hold.*

1. *The mapping $S : (0, +\infty) \to \mathcal{L}(B), t \mapsto S(t)$ is analytic. Moreover for any $u \in B, t > 0$ and $n = 1, 2, ...,$ $S(t)u \in D(A^n)$, and*

$$S^{(n)}(t)u = A^n S(t)u.$$

2. *$S(t+s) = S(t)S(s)$ for all $t, s \geq 0$.*

3. *$S(\cdot)u$ is continuous at 0 if and only if $u \in \overline{D(A)}$.*

4. *$\exists M, N > 0$ such that*

$$\|S(t)\| \leq Me^{\omega t}, t \geq 0 \tag{5.13}$$

$$\|AS(t)\| \leq e^{\omega t}(\frac{N}{t} + \omega M), \forall t > 0. \tag{5.14}$$

5. *$S(\cdot)$ can be extended to an analytic $\mathcal{L}(B)$-valued function in $S_{0,\theta_0-\frac{\pi}{2}}$.*

Proof (cf. Da Prato and Zabczyk [14], Appendix A.4). ■

Because of property 5, we say that $S(\cdot)$ is an analytic semigroup.

Now we consider an important subclass of analytic semigroups.

Definition 42 *Let H be a Hilbert space. A linear operator $A : D(A) \subset H \to H$ is variational if*

- *there exists a Hilbert space V densely embedded in H and a continuous bilinear form $a : V \times V \to \mathbb{R}$ such that*

$$-a(v, v) \geq c_1\|v\|_V^2 - c_2\|v\|_H^2 \tag{5.15}$$

  for suitable constants $c_1 > 0, c_2 \geq 0$;

- $D(A) = \{u \in V | a(u, \cdot)$ *is continuous in the topology of* $H\}$*;*

- $a(u, v) = \langle Au, v\rangle_H, \forall u \in D(A), \forall v \in V$*.*

Proposition 43 *Let A be a variational operator in H. Then A generates an analytic semigroup $S(\cdot)$. Moreover, if A is symmetric then A is self-adjoint.*

Proof (cf. Tanabe [45]). ■

5.2 Mild Solutions

This section is essentially taken out from Knoche and Frieler [22], Chapter 3.

Let $(U, \|\|_U)$ and $(H, \|\|_H)$ be separable Hilbert spaces. Let $Q = I_U$ and fix a cylindrical Q-Wiener process $W(t), t \geq 0$, in U on a probability space $(\Omega, \mathcal{F}, P)$ with a normal filtration $\mathcal{F}_t$, $t \geq 0$. For a fixed $T > 0$, we consider the following type of stochastic differential equations in H

$$\begin{cases} dX(t) = [AX(t) + F(X(t))]dt + B(X(t))dW(t), t \in [0, T] \\ X(0) = \xi \end{cases} \tag{5.16}$$

where

- $A : D(A) \to H$ is the infinitesimal generator of a C_0-semigroup $S(t), t \geq 0$, of linear operators on H;

- $F : H \to H$ is $\mathscr{B}(H)/\mathscr{B}(H)$-measurable;

- $B : H \to L(U, H)$;

- ξ is a H-valued, $\mathcal{F}_0$-measurable random variable.

Definition 44 *(Mild solution). An H-valued predictable process $X(t), t \in [0, T]$, is called a mild solution of the above problem if*

$$X(t) = S(t)\xi + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)B(X(s))dW(s) \ \textit{P-a.s.} \tag{5.17}$$

for each $t \in [0, T]$.

For each $T > 0$ and $p \geq 2$, denote by $\mathcal{H}^p(T, H)$ the Banach space of all (equivalence classes) of predictable H-valued processes Y with the norm

$$\|Y\|_{H^p} \triangleq \sup_{t\in[0,T]} (E(\|Y(t)\|^p))^{\frac{1}{p}} < \infty \tag{5.18}$$

Proposition 45 *Assume*

- $F : H \to H$ *is Lipschitz continuous, i.e. that there exists a constant* $C > 0$ *such that*

  $$\|F(x) - F(y)\| \leq C\|x - y\| \text{ for all } x, y \in H;$$

- $B : H \to L(U, H)$ *is strongly continuous, i.e. that the mapping*

  $$x \mapsto B(x)u$$

  is continuous from H *to* H *for each* $u \in U$;

- *For all* $t \in (0, T]$ *and* $x \in H$ *we have* $S(t)B(x) \in L_2(U, H)$;

- *There is a square integrable mapping* $K : [0, T] \mapsto [0, \infty)$ *such that*

  $$\|S(t)(B(x) - B(y))\|_{L_2} \leq K(t)\|x - y\|$$

  and

  $$\|S(t)B(x)\|_{L_2} \leq K(t)(1 + \|x\|)$$

  for all $t \in (0, T]$ *and* $x, y \in H$.

Then for any $T > 0$ *and* $p \geq 2$ *there exists a unique mild solution* $X(\xi) \in \mathcal{H}^p(T, H)$ *with initial condition*

$$\xi \in L^p(\Omega, \mathcal{F}_0, P; H) \triangleq L_0^p.$$

In addition we even obtain that the mapping

$$X : L_0^p \to \mathcal{H}^p(T, H)$$

$$\xi \mapsto X(\xi)$$

is Lipschitz continuous with Lipschitz constant $L_{T,p}$.

Proof (cf. Frieler and Knoche [22], Theorem 3.2).

■

Remark 46 *It follows from the Lipschitz continuity of* X *that there exists a constant* $C_{T,p}$ *independent of* $\xi \in L_0^p$ *such that*

$$\|X(\xi)\|_{H^p} \leq C_{T,p}(1 + \|\xi\|_{L^p})$$

(cf. Frieler and Knoche [22], Remark 3.1).

5.3 S-K Approximation for the Operator $\frac{\partial}{\partial x}$

We consider the following stochastic partial differential equation

$$\begin{cases} \mu \frac{\partial^2 u}{\partial t^2}(t,x) = \frac{\partial u}{\partial x}(t,x) + \Delta_x \frac{\partial u}{\partial t}(t,x) - \frac{\partial u}{\partial t}(t,x) + f(x,u(t,x)) \\ +b(x,u(t,x)) \frac{\partial W^Q}{\partial t}(t,x), t > 0, x \in [0,\infty) \\ u(0,x) = u_0, \frac{\partial u}{\partial t}(0,x) = v_0 \end{cases} \tag{5.19}$$

where we assume

- $f : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ is measurable and

$$\sup_{x \in [0,\infty)} |f(x,\sigma) - f(x,\rho)| \leq K_f |\sigma - \rho|, \sigma, \rho \in \mathbb{R}$$

  for some positive constant K_f. Moreover

$$\sup_{x \in [0,\infty)} |f(x,0)| \triangleq f_0 < \infty.$$

- $b : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ is measurable and

$$\sup_{x \in [0,\infty)} |b(x,\sigma) - b(x,\rho)| \leq K_b |\sigma - \rho|, \sigma, \rho \in \mathbb{R}$$

  for some positive constant K_b. Moreover

$$\sup_{x \in [0,\infty)} |b(x,0)| \triangleq b_0 < \infty.$$

- The bounded linear operator $Q : L^2([0,\infty)) \to L^2([0,\infty))$ is symmetric, non-negative and with finite trace.

It follows that there exists an orthonormal basis $e_k, k \in \mathbb{N}$ of $L^2([0,\infty))$ such that

$$Qe_k = \lambda_k e_k$$

And $W^Q(t)$ is formally defined as

$$W^Q(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k$$

Then W^Q is the noise with covariance given by

$$\mathbb{E}\langle W^Q(t), h\rangle_{L^2([0,\infty))}\langle W^Q(t), k\rangle_{L^2([0,\infty))} = (t \wedge s)\langle Qh, k\rangle_{L^2([0,\infty))}.$$

We add a remark here that usually it is difficult to solve this kind of wave equation for the first order derivative operator. So we add a strong damping term, which is $\Delta_x \frac{\partial u}{\partial x}$.

Let

$$V \triangleq H_0^{1,2}([0,\infty)) \times H_0^{1,2}([0,\infty)), H \triangleq H_0^{1,2}([0,\infty)) \times L^2([0,\infty)) \tag{5.20}$$

For any $\mu > 0$, $(h,k) \in H$, we define

$$A_\mu \begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{1}{\mu}\frac{\partial}{\partial x} & \frac{1}{\mu}\frac{\partial^2}{\partial x^2} - \frac{1}{\mu} \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \tag{5.21}$$

$$F_\mu(h,k)(x) = \frac{1}{\mu}(0, f(x,h(x)), x \in [0,\infty), \tag{5.22}$$

$$B_\mu(h,k)(x) = \frac{1}{\mu}(0, b(x,h(x)), x \in [0,\infty), \tag{5.23}$$

$$\mathcal{Q}_\mu(h,k) = \frac{1}{\mu}(0, Qk), \tag{5.24}$$

Setting $z \triangleq (u, \frac{\partial u}{\partial t})$, we rewrite equation (5.19) in the following equivalent abstract form

$$dz(t) = [A_\mu z(t) + F_\mu(z(t))]dt + B_\mu(z(t))dW^{Q_\mu}(t), z(0) = z_0 \tag{5.25}$$

Proposition 47 *For each $\mu < 1$, A_μ generates an analytic semigroup $S_\mu(\cdot)$ on H with $\|S_\mu(t)\|_H \leq e^{(\frac{1}{2}+\frac{1}{2\mu})t}$.*

Proof note that

$$V^* := H_0^{1,2}([0,\infty)) \times H^{-1}([0,\infty)) \tag{5.26}$$

We consider A_μ with initial domain $C_0^\infty((0,\infty)) \times C_0^\infty((0,\infty))$ as an operator taking values in V^*. So for $u_1, v_1, u_2, v_2 \in C_0^\infty((0,\infty))$:

$$\begin{aligned}
& |_{V^*}\langle A_\mu \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \rangle_V| \\
= \;& |_{V^*}\langle \begin{pmatrix} v_1 \\ \frac{1}{\mu}u_1^{'} + \frac{1}{\mu}v_1^{''} - \frac{1}{\mu}v_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \rangle_V| \\
= \;& |\langle \begin{pmatrix} v_1 \\ \frac{1}{\mu}u_1^{'} + \frac{1}{\mu}v_1^{''} - \frac{1}{\mu}v_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \rangle_H| \\
\leq \;& \|v_1^{'}\|_{L^2}\|u_2^{'}\|_{L^2} + \|v_1\|_{L^2}\|u_2\|_{L^2} + \frac{1}{\mu}\|u_1^{'}\|_{L^2}\|v_2\|_{L^2} + \frac{1}{\mu}\|v_1^{'}\|_{L^2}\|v_2^{'}\|_{L^2} + \frac{1}{\mu}\|v_1\|_{L^2}\|v_2\|_{L^2} \\
\leq \;& (\|u_2\|_{1,2} + \|v_2\|_{1,2})(\|v_1^{'}\|_{L^2} + \|v_1\|_{L^2} + \frac{1}{\mu}\|u_1^{'}\|_{L^2} + \frac{1}{\mu}\|v_1^{'}\|_{L^2} + \frac{1}{\mu}\|v_1\|_{L^2}) \\
\leq \;& (\frac{1}{\mu} + 1)\| \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} \|_V \| \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \|_V
\end{aligned}$$

This shows that A_μ can be extended (uniquely) to a bounded linear operator $A_\mu : V \to V^*$. And we have:

$$\|A_\mu \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} \|_{V^*} \leq (\frac{1}{\mu} + 1)\| \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} \|_V \tag{5.27}$$

Let $\mu < 1$,

$$
\begin{aligned}
&{}_{V^*}\langle A_\mu \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \rangle_V \\
= \ & \int v' u' dx + \int v u dx + \frac{1}{\mu} \int u' v dx - \frac{1}{\mu} \int |v'|^2 dx - \frac{1}{\mu} \int |v|^2 dx \\
\leq \ & \frac{1}{2}\|u'\|_{L^2}^2 + \frac{1}{2}\|v'\|_{L^2}^2 + \frac{1}{2}\|u\|_{L^2}^2 + \frac{1}{2}\|v\|_{L^2}^2 + \frac{1}{2\mu}\|u'\|_{L^2}^2 + \frac{1}{2\mu}\|v\|_{L^2}^2 - \frac{1}{\mu}\|v'\|_{L^2}^2 - \frac{1}{\mu}\|v\|_{L^2}^2 \\
\leq \ & (\frac{1}{2} + \frac{1}{2\mu})(\|u'\|_{L^2}^2 + \|u\|_{L^2}^2) - (\frac{1}{\mu} - \frac{1}{2})\|v'\|_{L^2}^2 - (\frac{1}{2\mu} - \frac{1}{2})\|v\|_{L^2}^2 \\
\leq \ & \frac{1}{\mu}\| \begin{pmatrix} u \\ v \end{pmatrix} \|_H^2 - (\frac{1}{2\mu} - \frac{1}{2})\| \begin{pmatrix} u \\ v \end{pmatrix} \|_V^2
\end{aligned}
$$

Define a bilinear form $a_\mu : V \times V \to \mathbb{R}$ by

$$a_\mu(w_1, w_2) = \langle A_\mu w_1, w_2 \rangle. \tag{5.28}$$

Then a_μ is continuous and satisfy the inequality (5.15).

So if we restrict the domain of A_μ to the set $\{w_1 \in V | a(w_1, \cdot)$ is continuous in the topology of $H\}$, that is

$$D(A_\mu) = \{w_1 \in V | A_\mu w_1 \in H\} \tag{5.29}$$

We still denote the operator by A_μ. Then A_μ is a variational generator. Hence it generates an analytic semigroup $S_\mu(\cdot)$.

For the norm, let $C_\mu = \frac{1}{\mu}$. Let $T_\mu(\cdot)$ be the semigroup generated by $A_\mu - C_\mu$. For $w \in H$,

$$\frac{d}{dt}\|T_\mu(t)w\|_H^2 = 2\langle (A_\mu - C_\mu)T_\mu(t)w, T_\mu(t)w \rangle_H \leq -(\frac{1}{\mu} - 1)\|T_\mu(t)w\|_V^2 \tag{5.30}$$

Then:

$$\|T_\mu(t)w\|_H^2 \leq \|w\|_H^2 - (\frac{1}{\mu} - 1)\int_0^t \|T_\mu(s)w\|_V^2 ds \leq \|w\|_H^2 - (\frac{1}{\mu} - 1)\int_0^t \|T_\mu(s)w\|_H^2 ds \tag{5.31}$$

By Gronwell's lemma, we have $\|T_\mu(t)w\|_H^2 \leq \|w\|_H^2 e^{-(\frac{1}{\mu}-1)t}$. Hence $\|T_\mu(t)\|_H \leq e^{(\frac{1}{2\mu}-\frac{1}{2})t}$. And then $\|S_\mu(t)\|_H = \|e^{C_\mu t}T_\mu(t)\|_H \leq e^{(\frac{1}{2}+\frac{1}{2\mu})t}, \forall\mu$.

■

Proposition 48 *For any initial data $(u_0, v_0) \in H$, problem (5.19) has a unique mild solution.*

Proof It suffices to show the lipschitz continuity of B_μ and F_μ. Then the assertion follows from Proposition 45.

For any $z_1 = (u_1, v_1)$, $z_2 = (u_2, v_2) \in H$, we have

$$\|F_\mu(z_1) - F_\mu(z_1)\|_H = \frac{1}{\mu}\|f(\cdot, u_1) - f(\cdot, u_2)\|_{L^2} \leq \frac{K_f}{\mu}\|u_1 - u_2\|_{L^2} \leq \frac{K_f}{\mu}\|z_1 - z_2\|_H \tag{5.32}$$

$$\|B_\mu(z_1) - B_\mu(z_1)\|_H = \frac{1}{\mu}\|b(\cdot, u_1) - b(\cdot, u_2)\|_{L^2} \leq \frac{K_b}{\mu}\|u_1 - u_2\|_{L^2} \leq \frac{K_b}{\mu}\|z_1 - z_2\|_H \tag{5.33}$$

■

Below we consider the linear equation of problem (5.19) where the trace of Q satisfying certain conditions.

$$\begin{cases} \mu\frac{\partial^2\eta}{\partial t^2}(t,x) = \frac{\partial\eta}{\partial x}(t,x) + \Delta_x\frac{\partial\eta}{\partial t}(t,x) - \frac{\partial u}{\partial t} + \frac{\partial W_\mu^Q}{\partial t}(t,x) \\ \eta(0) = 0, \frac{\partial\eta}{\partial t}(0) = 0 \end{cases} \tag{5.34}$$

We define $\mathcal{Q}_\mu$ as in (5.24). We have

$$Tr(\mathcal{Q}_\mu) = \frac{1}{\mu}Tr(Q_\mu). \tag{5.35}$$

According to Proposition 45, we know there exists a mild solution $W^\mu(t)$ in H, defined as in (5.20).

$$W^\mu(t) = \int_0^t S_\mu(t-s)dW^{\mathcal{Q}_\mu}(s) \tag{5.36}$$

then $\eta^\mu(t) = \Pi_1 W^\mu(t)$.

Lemma 49 *Fix $T > 0$, assume $TrQ_\mu < \mu e^{-\frac{T}{\mu}}$, then the solution $\eta^\mu(t), t \in [0,T]$ is γ-Hölder continuous with respect to t for any $\gamma < \frac{1}{2}$. Moreover, the momenta of the γ-Hölder norms of η^μ are bounded uniformly in μ, that is*

$$\sup_{0<\mu<1} \mathbb{E}\|\eta^\mu\|_{C^\gamma([0,T];H)} := C_{T,p} < \infty \tag{5.37}$$

Proof Let $T \geq t > s > 0$. By Itô's isometry, we have

$$\begin{aligned} \mathbb{E}\|W^\mu(t) - W^\mu(s)\|^2 &= \int_s^t \|S_\mu(t-\sigma) \circ \mathcal{Q}_\mu^{\frac{1}{2}}\|^2 d\sigma \\ &\quad + \int_0^s \|\left(S_\mu(t-\sigma) - S_\mu(s-\sigma)\right) \circ \mathcal{Q}_\mu^{\frac{1}{2}}\|^2 d\sigma \\ &= I_1 + I_2 \end{aligned} \tag{5.38}$$

Plugging (5.35) in the third inequality below:

$$\begin{aligned} I_1 &\leq \int_s^t \|S_\mu(t-\sigma)\|^2 Tr(\mathcal{Q}_\mu) d\sigma \leq \int_s^t e^{2(\frac{1}{2}+\frac{1}{2\mu})(t-\sigma)} Tr(\mathcal{Q}_\mu) d\sigma \\ &\leq e^{(1+\frac{1}{\mu})T} \frac{1}{\mu} \mu e^{-\frac{T}{\mu}} (t-s) \leq e^T (t-s) \end{aligned} \tag{5.39}$$

For I_2, let $(f_k)_{k\in\mathbb{N}}$ be an orthonormal basis of H. Below we use (5.6) in the second line, and (5.14) in the fourth line.

$$
\begin{aligned}
I_2 &= \Sigma_k \int_0^s \|(S_\mu(t-\sigma) - S_\mu(s-\sigma))\mathcal{Q}_\mu^{\frac{1}{2}} f_k\|^2 d\sigma \\
&= \Sigma_k \int_0^s \| \int_{s-\sigma}^{t-\sigma} A_\mu S_\mu(\rho)\mathcal{Q}_\mu^{\frac{1}{2}} f_k d\rho\|^2 d\sigma \\
&\leq \Sigma_k \int_0^s (\int_{s-\sigma}^{t-\sigma} \|A_\mu S_\mu(\rho)\mathcal{Q}_\mu^{\frac{1}{2}} f_k\| d\rho)^2 d\sigma \\
&\leq \Sigma_k \int_0^s (\int_{s-\sigma}^{t-\sigma} C_1 \|\mathcal{Q}_\mu^{\frac{1}{2}} f_k\| \frac{d\rho}{\rho})^2 d\sigma \\
&= C_1 \Sigma_k \|\mathcal{Q}_\mu^{\frac{1}{2}} f_k\|^2 \int_0^s (\int_{s-\sigma}^{t-\sigma} \frac{d\rho}{\rho})^2 d\sigma \\
&= C_1 Tr(\mathcal{Q}_\mu) \int_0^s (\int_{s-\sigma}^{t-\sigma} \frac{d\rho}{\rho})^2 d\sigma \\
&= C_1 \frac{1}{\mu} Tr(Q_\mu) \int_0^s (\int_{s-\sigma}^{t-\sigma} \frac{d\rho}{\rho})^2 d\sigma \\
&\leq C_1 \exp(-\frac{T}{\mu}) \int_0^s (\int_{s-\sigma}^{t-\sigma} \frac{d\rho}{\rho})^2 d\sigma
\end{aligned}
$$

Let $\gamma \in (0, \frac{1}{2})$, note that

$$
\int_{s-\sigma}^{t-\sigma} \rho^{\gamma-1} d\rho = \int_s^t (\rho-\sigma)^{\gamma-1} d\rho \leq \int_s^t (\rho - s)^{\gamma-1} d\rho = \int_0^{t-s} \rho^{\gamma-1} d\rho = \frac{(t-s)^\gamma}{\gamma}
$$

Then,

$$
\begin{aligned}
I_2 &\leq C_1 \exp(-\frac{T}{\mu}) \int_0^s (s-\sigma)^{-2\gamma} |\int_{s-\sigma}^{t-\sigma} \rho^{\gamma-1} d\rho|^2 d\sigma \\
&\leq \frac{C_1 \exp(-\frac{T}{\mu}) T^{1-2\gamma}}{\gamma^2(1-2\gamma)} (t-s)^{2\gamma}
\end{aligned}
$$

Thus for any $\gamma \in (0, \frac{1}{2})$ we have

$$
\mathbb{E}\|W^\mu(t) - W^\mu(s)\|^2 \leq C[(t-s) + \frac{T^{1-2\gamma}}{\gamma^2(1-2\gamma)}(t-s)^{2\gamma}] \tag{5.40}
$$

Since $\eta^\mu(t) = \Pi_1 W^\mu(t)$, we have

$$
\sup_{0<\mu<1} \mathbb{E}\|\eta^\mu(t) - \eta^\mu(s)\|^2 \leq C(t-s)^{2\gamma} \tag{5.41}
$$

By the Garcia-Rodemich-Rumsey theorem,and using the fact that $\eta^\mu(t)$ is gaussian for each t, we have $\forall p \geq 1$

$$
\sup_{0<\mu<1} \mathbb{E}\|\eta^\mu\|^p_{C^\gamma([0,T];H)} := C_{T,p} < \infty \tag{5.42}
$$

■

But so far we could not get the tightness of η^μ.

LIST OF REFERENCES

LIST OF REFERENCES

[1] Amihud, Yakov, and Haim Mendelson,1986, Asset pricing and bid-ask spread, Journal of Financial Economics 17, 223-249.

[2] D. Applebaum: Lèvy Processes and Stochastic Calculus, Cambridge University Press, Cambridge 2004

[3] Arbel, Avner, Steven Carvell, and Paul Strebel, 1983, giraffes, institutions, and neglected Firms, Financial Analysts Journal 2-8.

[4] Barndorf-Nielsen, O.E.: Exponentially decreasing distributions for the logarithm of particle size. Proceedings of the Royal Society London, ser. A, 353:401-19.

[5] Barndorf-Nielsen, O.E., and Halgreen, O.1977. Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 38:309-12.

[6] Barndorf-Nielsen, O.E.: Normal inverse Gaussian distributions and the modeling of stock returns. Research Report No. 300, Department of Theoretical Statistics, Aarhus University, (1995)

[7] P.Billingsley: Convergence of Probability Measures(second edition), Wiley(1999)

[8] B. Boufoussi, C.A. Tudor, Kramers-Smoluchowski approximation for stochastic equations with fractional Brownian motion, Rev. Roumaine Math. Pures Appl. 50 (2) (2005) 125-136.

[9] Cerrai,S., Freidlin, M.: On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom Measures, Probab. Theory Relat. Fields 135, 363-394(2006)

[10] Cerrai,S., Freidlin, M.: Smoluchowski-Kramers approximation for a general class of SPDEs, Probab. Theory Relat. Fields 135, 363-394(2006)

[11] Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, Journal of Finance 51, 1681-1713.

[12] Chen, Joseph, Harrison HOng, and Jeremy C. Stein, 2002, Breadth of ownership and stock returns, forthcoming Journal of Financial Economics.

[13] Constantinides, George, 1984, Optimal stock trading with personal taxes, Journal of Financial Economics 13, 65-89.

[14] Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions, Cambridge University Press, Cambridge 1992

[15] Da Prato, G., Zabczyk, J.: Ergodicity for Infinite Dimensional Systems. London Mathematical Society, Lecture Notes Series 229, Cambridge University Press, Cambridge 1996

[16] C. Dellacherie, P.A. Meyer: Probabilités et Potential: Theorie des Martingales, Hermann(1980)

[17] Eberlein,E., and Keller, U. 1995. Hyperbolic distributions in finance. Bernoulli 1:281-99.

[18] Robert J. Elliott, 1982. Stochastic Calculus and Applications, Springer-Verlag New York Heidelberg Berline

[19] S.N.Ethier, T.G.Kurtz: Markov Processes, Characterisation and Convergence, Wiley(1986)

[20] D. Filipović: Consistency Problems for Heath-Jarrow-Morton Interest Rate Models, Springer(2001)

[21] Freidlin M.: Some remarks on the Smoluchowski-Kramers approximation. Journal of Statistical Physics 117, 617-634(2004)

[22] K. Frieler and C. Knoche, Solutions of stochastic diferential equations in infinite dimensional Hilbert spaces and their dependence on initial data, Diploma Thesis, Bielefeld University, BiBoS-Preprint E02-04-083, 2001.

[23] Gyöngy, I., Krylov, N.V.: Existence of strong solutions for Itô's stochastic equations via approximations, Probab. Theory Relat. Fields 103, 143-158(1996)

[24] Hicks, J., 1967. Critical Essays in Monetary Theory, Oxford University Press, Oxford.

[25] Hou, Kewei and Tobias J. Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns, Review of Financial Studies, 18, 981-1020.

[26] N.Ikeda, S.Watanabe, Stochastic Differential Equations and Diffusion Processes (second edition), North Holland-Kodansha(1989)

[27] Kramers, H.: Brownian motion in a field of force and the diffusion model of chemical reactions. Physica 7, 284-304(1940)

[28] Liu, Weimin, 2006, A liquidity-augmented capital asset pricing model, Journal of financial economics, 82, 631-671.

[29] McQueen, Grant, Michael Pinegar, and steven Thorley, 1996, Delayed reaction to good news and the cross-autocorreclation of portfolio returns, Journal of Finance, 51, 889-919.

[30] Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad news travel slowly: size, analyst coverage, and the profitability of momentum strategies, Journal of Finance, 55, 265-296.

[31] Hong, Harrison and Jeremy C. Stein, 1999, A unified theory of underraction, momentum trading, and overreaction in asset markets, Journal of Finance, 54, 2143-2184.

[32] Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, Journal of Finance 48, 65-91.

[33] Knäble Kristian, Stochastic Evolution Equations with Lévy Noise and Applications to the Heath-Jarrow-Morton Model. Diplomathesis, Universität Bielefeld 2006

[34] Merton, Robert C., A simple model of capital market equilibrium with incomplete information, Journal of Finance 42, 483-510, 1987

[35] NARITA, K. (1991) The Smoluchowski-Kramers approximation for stochastic Lienard equation with mean-field. Adv. Appl. Prob. 23, 303-316.

[36] NARITA, K. (1991) Asymptotic Behavior of Velocity Process in the Smoluchowski-Kramers Approximation for Stochastic Differential Equations. Adv. Appl. Prob. 23, 303-316.

[37] Pazy,A.: Semigroups of linear operators and applications to partial differential equations. Springer-Verlag, New York, 1983

[38] Peng, Lin, 2002, Learning with information capacity constraints, Working paper, Duke University.

[39] C. Prévôt, M. Röckner: A Concise Course on Stochastic Partial Differential Equations Lecture Notes in Mathematics, vol. 1905, Springer, 2007.

[40] Rogers, L.C.G. and Williams, D.: Diffusions, Markov Processes and Martingales, Vol.2: Itô Calculus, Wiley and Sons. (1987)

[41] Rouwenhorst, K. Geert, 1998, International momentum strategies, Journal of Finance 53, 267-284.

[42] Schoutens, W. and Teugels, J.L. , Lévy processes, polynomials and martingales. Commun. Statist.- Stochastic Models 14 (1 and 2),(1998).

[43] Shapiro, Alexander, The investor recoginition hypothesis in a dynamic general equilibrium: theory and evidence, Review of Financial Studies 15, 97-141.

[44] Smoluchowski,M.: Drei Vortage über Diffusion Brownsche Bewegung und Koagulation von Kolloidteilchen.Physik Zeit. 17, 557-585(1916)

[45] Tanabe H.: Equations of evolution, Pitman.(1979)

[46] Westermann, R.: Smoluchowski-Kramers Approximation for Stochastic Differential Equations and Application in Behavioral Finance 2006

APPENDIX

APPENDIX

A. S-K Approximation for $\frac{\partial}{\partial x}$ on Weighted Spaces

As seen from Chapter 5, we could not prove the convergence of the Smoluchowski-Kramers approximation on our chosen space H. In [20], Filipović considered a weighted Sobolev space on the first order equation with the same operator $\frac{\partial}{\partial x}$. We also considered this space. But it turns out to be not helpful, because our equation is of second order. Below is the detail explanation.

We consider the same problem described in section 5.3, that is, we have the following equation

$$\begin{cases} \mu\frac{\partial^2 u}{\partial t^2}(t,x) = \frac{\partial u}{\partial x}(t,x) + \Delta_x\frac{\partial u}{\partial t}(t,x) - \frac{\partial u}{\partial t}(t,x) + f(x,u(t,x)) \\ +b(x,u(t,x))\frac{\partial W^Q}{\partial t}(t,x), t > 0, x \in [0,\infty) \\ u(0,x) = u_0, \frac{\partial u}{\partial t}(0,x) = v_0 \end{cases} \tag{A.1}$$

written into the following abstract form

$$dz(t) = [A_\mu Z(t) + F_\mu(Z(t))]dt + B_\mu(z(t))dW(t), z(0) = z_0 \tag{A.2}$$

with $z \triangleq (u, \frac{\partial u}{\partial t})$. For any $\mu > 0$, $(u,v) \in H$, we define

$$A_\mu z = A_\mu \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{1}{\mu}\frac{\partial}{\partial x} & \frac{1}{\mu}\frac{\partial^2}{\partial x^2} - \frac{1}{\mu} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{A.3}$$

$$F_\mu(u,v)(x) = \frac{1}{\mu}(0, f(x,u(x)), x \in [0,\infty), \tag{A.4}$$

$$B_\mu(u,v)(x) = \frac{1}{\mu}(0, b(x,u(x)), x \in [0,\infty), \tag{A.5}$$

$$\mathcal{Q}_\mu(u,v) = \frac{1}{\mu}(0, Qv), \tag{A.6}$$

We consider A_μ as an operator on the sobolev weighted space $V := H_w^0 \times H_w^0$ with $H := H_w^0 \times L_w^2$. Here

$$H_w := \{h \in L_{loc}^1(\mathbb{R}_+) | \exists h^{'} \in L_{loc}^1(\mathbb{R}_+) \text{ and } \|h\|_w < \infty\} \tag{A.7}$$

with the norm

$$\|h\|_w^2 := \int_0^\infty |h(x)|^2 w(x)dx + \int_0^\infty |h^{'}(x)|^2 w(x)dx \tag{A.8}$$

where $w : \mathbb{R}_+ \rightarrow [1, \infty)$ is a non-decreasing C^1-function such that $w^{-\frac{1}{3}} \in L^1(\mathbb{R}_+)$. In this section, we choose a particular $w(x) = e^{\alpha x}$, for $\alpha \neq 0$.

Let

$$H_w^0 := \{h \in H_w | h(0) = 0\} \tag{A.9}$$

$$L_w^2 := \{f \in L_{loc}^1 | \|f\|_{L_w}^2 = \int_0^\infty |f(x)|^2 w(x)dx < \infty\} \tag{A.10}$$

Remark 50 *If $f \in L^2(\mathbb{R}_+)$ and also $f^{'} \in L^2(\mathbb{R}_+)$ then since*

$$2\int_0^y f(x)f^{'}(x)dx = |f(y)|^2 - |f(0)|^2 \tag{A.11}$$

$|f(x)|^2$ tends to a limit as $x \rightarrow \infty$. And then we can conclude $f(\infty) = 0$, since $f \in L^2(\mathbb{R})$.

Proposition 51 *For each $\mu < 2$, A_μ generates an analytic semigroup $S_\mu(\cdot)$ on H with $\|S_\mu(t)\|_H \leq e^{(\frac{1}{2}+\frac{\alpha^2-1}{2\mu} \vee \frac{1}{2\mu})t}$.*

Proof First we consider A_μ with initial domain $\{f \in H_w^0 | f^{'} \in H_w^0\} \times \{f \in H_w^0 | f^{'} \in H_w^0\}$ as an operator taking values in V^*. Note that this domain is dense in V.

$$
\begin{aligned}
& {}_{V^*}\langle A_\mu \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \rangle_V \\
= \; & {}_{V^*}\langle \begin{pmatrix} 0 & 1 \\ \frac{1}{\mu}\frac{\partial}{\partial x} & \frac{1}{\mu}\frac{\partial^2}{\partial x^2} - \frac{1}{\mu} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \rangle_V \\
= \; & \int vuwdx + \int v^{'}u^{'}wdx + \frac{1}{\mu}\int u^{'}vwdx + \frac{1}{\mu}\int v^{''}vwdx - \frac{1}{\mu}\int vvwdx
\end{aligned}
$$

Integrating by parts, the fourth term becomes:

$$
\frac{1}{\mu}\int v^{''}vwdx = \frac{1}{\mu}(v(\infty)v^{'}(\infty)w(\infty) - v(0)v^{'}(0)w(0) - \int v^{'}v^{'}wdx - \int v^{'}vw^{'}dx) \tag{A.12}
$$

We apply Remark 50 on $f = vw^{\frac{1}{2}}$. First, $\|f\|_{L^2}^2 \leq \|v\|_w^2 < \infty$. On the other hand,

$$
\begin{aligned}
\int_0^\infty |f^{'}|^2 dx &= \int_0^\infty |v^{'}w^{\frac{1}{2}} + \frac{1}{2}vw^{-\frac{1}{2}}|^2 dx \\
&= \int |v^{'}|^2 wdx + \int v^{'}vdx + \frac{1}{4}\int |v|^2 w^{-1} dx < \infty
\end{aligned}
$$

Hence $v(\infty)w^{\frac{1}{2}}(\infty) = 0$. And for the same logic as above, since $v^{'} \in H_w$, we have $v^{'}(\infty)w^{\frac{1}{2}}(\infty) = 0$. Thus $v(\infty)v^{'}(\infty)w(\infty) = 0$.

Now note that $w^{'} = \alpha w$, the fourth term becomes

$$
\begin{aligned}
\frac{1}{\mu}\int v^{''}vwdx &= \frac{1}{\mu}(v(\infty)v^{'}(\infty)w(\infty) - v(0)v^{'}(0)w(0) - \int v^{'}v^{'}wdx - \int v^{'}vw^{'}dx) \\
&= -\frac{1}{\mu}\int v^{'}v^{'}wdx - \frac{1}{\mu}\int v^{'}vw^{'}dx \\
&= -\frac{1}{\mu}\int |v^{'}|^2 wdx - \frac{\alpha}{2\mu}\int (v^2)^{'}wdx \\
&= -\frac{1}{\mu}\int |v^{'}|^2 wdx - \frac{\alpha}{2\mu}(v^2(\infty)w(\infty) - v^2(0)w(0) - \alpha\int |v|^2 wdx) \\
&= -\frac{1}{\mu}\int |v^{'}|^2 wdx + \frac{\alpha^2}{2\mu}\int |v|^2 wdx
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& {}_{V^*}\langle A_\mu \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \rangle_V \\
= & \int vuwdx + \int v^{'}u^{'}wdx + \frac{1}{\mu}\int u^{'}vwdx + \frac{1}{\mu}\int v^{''}vwdx - \frac{1}{\mu}\int vvwdx \\
= & \int vuwdx + \int v^{'}u^{'}wdx + \frac{1}{\mu}\int u^{'}vwdx - \frac{1}{\mu}\int |v^{'}|^2wdx + \frac{\alpha^2}{2\mu}\int |v|^2wdx - \frac{1}{\mu}\int vvwdx \\
= & \int vuwdx + \int v^{'}u^{'}wdx + \frac{1}{\mu}\int u^{'}vwdx - \frac{1}{\mu}\int |v^{'}|^2wdx + \frac{\alpha^2-2}{2\mu}\int |v|^2wdx \\
\leq & \frac{1}{2}\int |u|^2wdx + \frac{1}{2}\int |v|^2wdx + \frac{1}{2}\int |u^{'}|^2wdx + \frac{1}{2}\int |v^{'}|^2wdx \\
& + \frac{1}{2\mu}\int |u^{'}|^2wdx + \frac{1}{2\mu}\int |v|^2wdx - \frac{1}{\mu}\int |v^{'}|^2wdx + \frac{\alpha^2-2}{2\mu}\int |v|^2wdx \\
= & -(\frac{1}{\mu} - \frac{1}{2})\int |v^{'}|^2wdx + \frac{1}{2}\int |u|^2wdx + (\frac{1}{2} + \frac{1}{2\mu})\int |u^{'}|^2wdx + (\frac{1}{2} + \frac{\alpha^2-1}{2\mu})\int |v|^2wdx \\
= & -(\frac{1}{\mu} - \frac{1}{2})\| \begin{pmatrix} u \\ v \end{pmatrix} \|_V + \frac{1}{\mu}\int |u|^2wdx + \frac{3}{2\mu}\int |u^{'}|^2wdx + \frac{\alpha^2+1}{2\mu}\int |v|^2wdx \\
= & -(\frac{1}{\mu} - \frac{1}{2})\| \begin{pmatrix} u \\ v \end{pmatrix} \|_V + (\frac{\alpha^2+1}{2\mu} \vee \frac{3}{2\mu})\| \begin{pmatrix} u \\ v \end{pmatrix} \|_H
\end{aligned}
$$

Now A_μ can be extends(uniquely) to a bounded linear operator $A_\mu : V \to V^*$. By restricting the domain of A_μ as in the proof of Proposition 47, we get a variational generator. We still denote it by A_μ. Then it generates an analytic semigroup $S_\mu(\cdot)$, for each μ. And follow the same calculation in Proposition 47, we have $\|S_\mu(t)\|_H \leq e^{(\frac{1}{2}+\frac{\alpha^2-1}{2\mu}\vee\frac{1}{2\mu})t}$, $\forall \mu < 2$.

■

Proposition 52 *For any initial data $(u_0, v_0) \in H$, problem (A.1) has a unique mild solution.*

Proof It suffices to show the lipschitz continuity of B_μ and F_μ. Then the assertion follows from Proposition 45.

For any $z_1 = (u_1, v_1)$, $z_2 = (u_2, v_2) \in H$, we have

$$\|F_\mu(z_1) - F_\mu(z_1)\|_H = \frac{1}{\mu}\|f(\cdot, u_1) - f(\cdot, u_2)\|_{L^2_w} \le \frac{K_f}{\mu}\|u_1 - u_2\|_{L^2_w} \le \frac{K_f}{\mu}\|z_1 - z_2\|_H \tag{A.13}$$

$$\|B_\mu(z_1) - B_\mu(z_1)\|_H = \frac{1}{\mu}\|b(\cdot, u_1) - b(\cdot, u_2)\|_{L^2_w} \le \frac{K_b}{\mu}\|u_1 - u_2\|_{L^2_w} \le \frac{K_b}{\mu}\|z_1 - z_2\|_H \tag{A.14}$$

■

VITA

VITA

Songfu Zhang was born on November 5, 1978 in Zhejiang, P.R.China and he received his Bachelor's in 1999 from Nanjing University of Aeronautics and Astronautics in China. He came to Purdue University in August 2002 and obtained his Master's in May 2005 and his Ph.D. in December 2008. His thesis advisor is Prof. Dr. Michael Röckner.