Itô's formula for finite variation Lévy processes

Jimi Kotkajuuri

Master's Thesis in Mathematics

University of Jyväskylä
Department of Mathematics and Statistics
Spring 2023

Tiivistelmä: Jimi Kotkajuuri, *Itô's formula for finite variation Lévy processes*, matematiikan pro gradu -tutkielma, 50 s., Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, kevät 2023.

Tämän tutkielman tarkoituksena on tarkastella erästä versiota stokastisen integroinnin avaintuloksesta nimeltään Itôn kaava, jolla on tärkeä rooli niin stokastiikan teorian kuin sen erinäisten sovellusten kannalta. Itôn kaavoja voidaan johtaa perustuen useille eri oletuksille sekä tilanteille. Tässä tutkimuksessa oletamme päätuloksena esittettävän Itôn kaavassa käytettävän stokastisen prosessin olevan Lévy-prosessi, joka toteuttaa rajallisen vaihtelun ehdon ja vastaavasti kaavassa käytettävän funtion oletamme jatkuvaksi ja heikosti derivoituvaksi.

Tulemme käsittelemään oleellisimmat stokastiikan sekä analyysin esitiedot päätuloksena olevaa Itôn kaavan todistamista varten. Stokastisten prosessien osalta käsittelemme yleisimpiä esimerkkejä Lévy-prosesseista sekä esittelemme niiden tärkeimpiä perusominaisuuksia. Määrittelemme myös Poisson satunnaismitan, jonka tärkeänä erikoistapauksena on muun muuassa hyppymitta. Lisäksi esittelemme joitain kuuluisia stokastiikan tuloksia kuten Lévy-Itô-hajotelma sekä Lévy-Khintchine-kaava.

Lisäksi tärkeänä osana Itôn kaavaa määrittelemme ja konstruoimme tarkasti stokastisen integraalin alkaen yksinkertaisista prosesseista ja lopulta yleistäen sen koskemaan laajempaa osaa prosesseista. Jatkona stokastiseen integrointiin tarkastelemme vielä lähemmin erästä stokastisen integraalin laajennusta Poisson satunnaismitan suhteen. Lopuksi esittelemme ja todistamme erään version Itôn kaavasta, joka käyttää oletuksinaan rajallisen vaihtelun ehdon toteuttavaa prosessia, mutta päätulosesta poiketen olettaa funktioiden olevan heikosti derivoituvuuden sijaan ainoastaa jatkuvasti differentioituvia.

Johtuen erityisesti heikosti derivoituvuuden oletuksesta käymme lisäksi läpi joitain reaali- ja funktionaalianalyysin perustuloksia. Erityisenä huomion kohteena ovat tulokset koskien distribuutioteoriaa ja heikkoa derivoituvuutta. Lopuksi näitä esitietoja käyttäen ja oletukset tarkasti määritellen todistamme yksityiskohtaisesti tutkielman päätuloksena olevan version Itôn kaavasta tapauksessa, jossa dimensio on 1.

Abstract: Jimi Kotkajuuri, *Itô's formula for finite variation Lévy processes*, Master's Thesis in Mathematics, 50 s., University of Jyväskylä, Department of Mathematics and Statistics, Spring 2023.

In this thesis we examine a version of the integral result of stochastic integration called Itô's formula which plays an important role both in terms of theory of stochastic and also its various applications. Itô's formulas can be derived based on several different circumstances and situations. In this thesis, we assume that the stochastic process used in Itô's formula presented as the main result is a Lévy process, which fulfills the condition of finite variation, and in addition to this we assume the function used in the formula to be continuous and weakly differentiable.

We will introduce the most essential stochastic and analysis prerequisites for proving the Itô formula as the main result. Regarding stochastic processes, we discuss the most common examples of Lévy processes and present their most important basic properties. We also define the Poisson random measure, whose important special case is jump measure. In addition, we present some famous stochastic results such as the Lévy-Itô decomposition and the Lévy-Khintchine formula.

Furthermore, as an important part of Itô's formula, we precisely define and construct the stochastic integral starting from simple processes and finally generalizing it into a wider range of processes. As a continuation of stochastic integration, we will take a closer look at an extension of the stochastic integral in terms of the Poisson random measure. Finally, we present and prove a version of Itô's formula, which uses as its assumptions a process fulfilling the finite variation condition, but, in contrast to the main result, assumes that the functions are weakly differentiable instead of only continuously differentiable.

Due to the assumption of weak differentiability, we also review some of the basic results of real and functional analysis. Particularly important for the main result are the results regarding distribution theory and weak differentiability. Finally, using this preliminary information and precisely specifying the assumptions, we prove in detail the version of Itô's formula, which is the main result of the thesis, in the case where the dimension is 1.

Contents

1. Introduction

Itô's formula which is named after Japanese mathematician Kiyosi Itô (1915-2008) is a fundamental part of stochastic integration. There exists several extensions of the Itô's formula to different kind of stochastic processes.

Our aim is to work with version of Itô's formula which assume finite variation Lévy process and functions that are both continuous and admits weak derivatives. Lévy processes are a class of stochastic processes named after famous french mathematician Paul Lévy (1886-1971) who was especially productive and respected in the area of probability theory. Most famous examples of Lévy processes are the Brownian motion, Poisson process and Wiener process.

First in section 2 we introduce some basic topics and properties of both probability theory and stochastic processes, especially in the case of Lèvy processes. Moreover we discuss important topics related to the our main result for example Poisson random measure, one of its special cases called jump measure J_X and also widely used Lévy-Itô-decomposition and Levy-Khintchine reprensetation.

Section 3 is mainly dedicated to the construction of stochastic integral with respected martingale-valued measure M, which we are going to do thoroughly. After the introduction of general stochastic integral we will concentrate on integration with respect Poisson random measure and especially jump measure. We are also going to introduce and prove one of the Itô's formula needed later during the proof of the main result.

Section 4 covers various collection of analysis tools. We recall for example some of the well know fundamental theorems like Lebesgue dominated convergence, Lebesgue differentiation theorem and fundamental theorem of Lebesgue integral calculus. We also discussed about important distribution theory topics like convolution, mollifier and weak derivatives and proves a few important results we will use later.

Last section is all about the main result. We are going to discuss profoundly about assumptions of the main theorem and when they holds. Proof of this Itô formula is going to perform detailedly and it use comprehensively previously proven results from this thesis.

2. Stochastic processes

2.1. **Random variables.**

Definition 2.1. Let X be an $\mathbb{R}^d$-valued random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The *law* of X is the probability measure defined by

$$\mu_X(A) = \mathbb{P}(X \in A)$$

and therefore the *characteristic function* $\phi_X : \mathbb{R}^d \to \mathbb{C}$ of the X is defined by

$$\phi_X(u) = \mathbb{E}\left[e^{i(u,X)}\right] = \int_{\mathbb{R}^d} e^{i(u,x)}\, d\mu_X(x)$$

for each $u \in \mathbb{R}^d$.

Next we will introduce another useful concept of probability theory called conditional expectation and some of its key properties.

Theorem 2.2 (Definition and existence of conditional expectation)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space,* X *a random variable such that* $\mathbb{E}[|X|] < \infty$ *and* $\mathcal{G} \subset \mathcal{F}$ *be a sub-*σ*-algebra. Then there exists a random variable* Y *such that*

(i) Y *is* $\mathcal{G}$ *measurable.*
(ii) $\mathbb{E}[|Y|] < \infty$.
(iii) *For every set* $A \in \mathcal{G}$ *we have*

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

A random variable Y *with the properties* $(i)-(iii)$ *is called the* conditional expectation $\mathbb{E}[X|\mathcal{G}]$ *of* X *given* $\mathcal{G}$*, and we write* $Y = \mathbb{E}[X|\mathcal{G}]$ *almost surely.*

Proof. See chapter 9.5 of [13]. □

Proposition 2.3. *Let* $X, X_1, X_2, ...$ *be integrable random variables and* $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ *be a sub-*σ*-algebras of* $\mathcal{F}$*. Then the following assertions are true:*

(i) *If* $\mu, \lambda \in \mathbb{R}$*, then*

$$\mathbb{E}[\lambda X_1 + \mu X_2|\mathcal{G}] = \lambda\mathbb{E}[X_1|\mathcal{G}] + \mu\mathbb{E}[X_2|\mathcal{G}]$$

almost surely.
(ii) *If* $X_1 \leq X_2$ *almost surely, then* $\mathbb{E}[X_1|\mathcal{G}] \leq \mathbb{E}[X_2|\mathcal{G}]$ *almost surely.*
(iii) *If* $X \geq 0$ *almost surely, then* $\mathbb{E}[X|\mathcal{G}] \geq 0$ *almost surely.*
(iv) *One has that* $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}]$ *almost surely.*
(v) *If* X *is* $\mathcal{G}$*-measurable, then* $\mathbb{E}[X|\mathcal{G}] = X$ *almost surely.*
(vi) *Let* $\mathcal{H}$ *be measurable and a sub* σ*-algebra of* $\mathcal{G}$*, then*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{H}]$$

almost surely.
(vii) *If* $Z : \Omega \to \mathbb{R}$ *is* $\mathcal{G}$*-measurable and* $\mathbb{E}[|ZX|] < \infty$*,then*

$$\mathbb{E}[ZX|\mathcal{G}] = Z\mathbb{E}[X|\mathcal{G}]$$

almost surely.
(viii) *If* $\mathcal{G} = \{\emptyset, \Omega\}$*, then* $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.
(ix) *If for all* $B \in \mathcal{B}(\mathbb{R})$ *and all* $A \in \mathcal{G}$*, one has that*

$$\mathbb{P}(\{X \in B\} \cap A) = \mathbb{P}(X \in B)\mathbb{P}(A),$$

i.e. if X *is independent from* $\mathcal{G}$*, then* $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ *almost surely.*
(x) *Assume* $X \geq 0$ *almost surely and random variables* $0 \leq X_n \uparrow X$ *almost surely. then*

$$\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$$

almost surely.

Proof. See Chapter 9.7 of [13] □

2.2. **Stochastic processes.**

Definition 2.4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

(i) A *stochastic process* is a family of random variables $X = (X_t)_{t\geq 0}$ defined on the same probability space and indexed by a continuous time parameter t.
(ii) Two stochastic processes X and Y are said to be *independent* if, for all $m, n \in \mathbb{N}$, all $0 \leq t_1 < \cdots < t_n < \infty$ and all $0 \leq s_1 < \cdots < s_m < \infty$, the σ-algebras $\sigma(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ and $\sigma(Y_{s_1}, Y_{s_2}, \ldots, Y_{s_m})$ are independent.
(iii) Similarly a stochastic process X and a sub-σ-algebra $\mathcal{G}$ are *independent* if $\mathcal{G}$ and $\sigma(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ are independent for all $n \in \mathbb{N}$, $0 \leq t_1 < \cdots < t_n < \infty$.

Next we are going to spend some time to get familiar with the most important concepts related to the stochastic processes.

Definition 2.5 (Stopping times)**.** Let $T : \Omega \to [0, \infty]$. Then T is called *stopping time* if the event $\{T \leq t\} \in \mathcal{F}_t$ for each $t \geq 0$.

Definition 2.6 (Filtration)**.** An increasing family of σ-algebras $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is called a *filtration* on $(\Omega, \mathcal{F}, \mathbb{P})$ if

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}, \quad \text{whenever } 0 \leq s \leq t.$$

A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that comes equipped with such a family $\mathbb{F}$ is said to be *filtered.* A stochastic process X is said to be $\mathbb{F}$ -adapted if the random variable X_t is $\mathcal{F}_t$- measurable for each $t \geq 0$.

Any process X is adapted to its own filtration $\mathbb{F}^X$ given by $\mathcal{F}_t = \sigma\{X_s : 0 \leq s \leq t\}$, and this $\mathbb{F}^X$ is called the *natural filtration* of X.

Definition 2.7 (Martingale)**.** A process M is said to be a *martingale* if M is $\mathbb{F}$-adapted, $\mathbb{E}[|M_t|] < \infty$ for any $t \geq 0$ and for all $s < t$ it holds that

$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s \text{ almost surely.}$$

We can define a *supermartingale* and a *submartingale* by replacing the above equation by

$$\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s \qquad \text{and} \qquad \mathbb{E}[M_t|\mathcal{F}_s] \geq M_s.$$

respectively. A martingale $M = (M_t)_{t\geq 0}$ is said to be *square-integrable* if $\mathbb{E}[|M_t^2|] < \infty$ for each $t \geq 0$ and *continuous* if it has almost surely continuous sample paths.

Definition 2.8. Let M be an adapted process. We call M a *local martingale*, if there exists a sequence of stopping times $\tau_1 \leq \cdots \leq \tau_n \to \infty$, such that each of the processes $(M_{t\wedge\tau_n})_{t\geq 0}$ is a martingale.

A local martingale is a useful generalization of the martingale concept. Note that any martingale is always also a local martingale.

Definition 2.9. A function $f : [0, \infty) \to \mathbb{R}^d$ is said to be *càdlàg* if it is right-continuous with left limits and for each $t \in (0, \infty)$ the limits

$$f(t-) = \lim_{s\nearrow t} f(s), \qquad f(t+) = \lim_{s\searrow t} f(s)$$

exist and $f(t) = f(t+)$. Moreover if t is a discontinuity point we denote the jump of f at t by

$$\Delta f(t) = f(t) - f(t-).$$

A stochastic process X is said to be càdlàg if it almost surely has sample paths which are right continuous, with left limits.

It is easy to see that any continuous function and process is always also càdlàg but càdlàg functions do not have to be continuous. Next we define classes of stochastic processes generated by adapted left and right continuous processes called predictable and optional processes respectively.

Definition 2.10 (Predictable process)**.** Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a filtered probability space.
(i) *The predictable σ-algebra* is the σ-algebra $\mathcal{P}$ generated on $[0, \infty) \times \Omega$ by all adapted left-continuous processes.
(ii) A process $X : [0, \infty) \times \Omega \to \mathbb{R}^d$ which is measurable with respect to $\mathcal{P}$ is called a *predictable process.*

Definition 2.11 (Optional process)**.** Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a filtered probability space.
(i) *The optional σ-algebra* is the σ-algebra $\mathcal{O}$ generated on $[0, \infty) \times \Omega$ by all adapted càdlàg processes.
(ii) A process $X : [0, \infty) \times \Omega \to \mathbb{R}^d$ which is measurable with respect to $\mathcal{O}$ is called *optional process.*

Note that even though every adapted process with left continuous sample paths is a predictable process there exists predictable processes which are not left continuous (see Chapter 2.4 of [4]).

2.3. **Lèvy processes.** Next we are going to define one of the key assumptions of our main result, which is a general class of stochastic processes called Lévy processes.

Definition 2.12 (Lévy process)**.** A process $X = (X_t)_{t \geq 0}$ with $X_t : \Omega \to \mathbb{R}^d$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a *Lèvy process* if it possesses the following properties:
(i) The paths of X are $\mathbb{P}$-almost surely right-continuous with left limits.
(ii) $\mathbb{P}(X_0 = 0) = 1$.
(iii) For $0 \leq s \leq t$, $X_t - X_s$ is equal in distribution to X_{t-s}.
(iv) For $0 \leq s \leq t$, $X_t - X_s$ is independent of $\sigma\{X_u : u \leq s\}$.

In the above definition conditions (iii) and (iv) imply the stationary and the independent increments properties, respectively. Stationary increments means that the distribution of any increments $X_t - X_s$ depends only on the lenght $t - s$. For example for the Poisson process defined in Definition 2.14 the probability distribution of increment $X_t - X_s$ is a Poisson distribution with expected value $\mu(t - s)$. Independency of increments implies that increments of a Lèvy process are always independent whenever their time interval does not overlap.

The most popular and useful examples of the Lévy processes are the so called Brownian motion and the Poisson process which are defined below.

Definition 2.13 (Brownian Motion)**.** A $\mathbb{R}^d$-valued process $B = (B_t)_{t \geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a *d-dimensional Brownian motion* if the following holds:
(i) The paths of B are $\mathbb{P}$-almost surely continuous.
(ii) $\mathbb{P}(B_0 = 0) = 1$.

(iii) For $0 \leq s \leq t$, $B_t - B_s$ is equal in distribution to B_{t-s}.
(iv) For $0 \leq s \leq t$, $B_t - B_s$ is independent of $\sigma\{B_u, u \leq s\}$.
(v) For each $t > 0$, B_t is equal in distribution to a normal random variable with a covariance matrix A and mean vector $\bar{0}$.

Definition 2.14 (Poisson process)**.** A process valued on the non-negative integers $N = (N_t)_{t\geq 0}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is said to be a *Poisson process* with intensity $\mu > 0$ if the following holds:

(i) The paths of N are $\mathbb{P}$-almost surely right-continuous with left limits.
(ii) $\mathbb{P}(N_0 = 0) = 1$.
(iii) For $0 \leq s \leq t$, $N_t - N_s$ is equal in distribution to N_{t-s}.
(iv) For $0 \leq s \leq t$, $N_t - N_s$ is independent of $\sigma\{N_u, u \leq s\}$.
(v) For each $t > 0$, N_t is equal in distribution to a Poisson random variable with parameter μt.

Despite Brownian motion and Poisson process are both Lévy processes they have a lot of differences. For example Brownian motion has continuous paths whereas the Poisson process does not. On the other hand the Poisson process has the finite variation property as we find out later, whereas the total variation (see Definition 2.32) of paths of the Brownian motion is almost surely infinite.

We give one more example of Lévy processes called a compound Poisson process that is a natural extension of the Poisson process, where the jump size Y_i is random instead of constant 1.

Definition 2.15 (Compound Poisson process)**.** Suppose that $(Y_i)_{i\geq 1}$ is a an i.i.d. sequence of random variables with values in $\mathbb{R}^d$ and $N = (N_t)_{t\geq 0}$ is Poisson process with intensity $\mu > 0$, independent of $(Y_i)_{i\geq 1}$. Then the process $X = (X_t)_{t\geq 0}$ defined by

$$X_t = \sum_{i=1}^{N_t} Y_i,$$

for each $t \geq 0$, is said to be a *compound Poisson process* with intensity $\mu > 0$.

Poisson processes can indeed be seen as compound Poisson process on $\mathbb{R}$ with jump size $Y_i \equiv 1$. It is not obvious that compound Poisson processes are Lévy processes. The proof can be found for example in Section 3.2 of [4]. In addition, there exists plenty of other Lévy processes for example hyberbolic Lévy processes and Variance Gamma Lévy processes (see Section 4 of [5]).

Definition 2.16. We define the *compensated Poisson process* $\widetilde{N} = (\widetilde{N}_t)_{t\geq 0}$, where each $\widetilde{N}_t = N_t - \mu t$ and N is a Poisson process with intensity μ.

Lemma 2.17. *Let $\widetilde{N}$ be a compensated Poisson process. Then,*

$$\mathbb{E}[\widetilde{N}_t] = 0 \quad \textit{and} \quad \mathrm{Var}[\widetilde{N}_t] = \mu t,$$

for each $t \geq 0$.

Proof. By (v) of Definition 2.14 for each $t > 0$, the random variable N_t is Poisson distributed with parameter μt. Thus

$$\begin{aligned}\mathbb{E}[\widetilde{N}_t] &= \mathbb{E}[N_t - \mu t] \\ &= \mathbb{E}[N_t] - \mu t \\ &= \mu t - \mu t \\ &= 0.\end{aligned}$$

By using the moment generating function $M_N(\theta) = e^{\mu t(e^\theta - 1)}$ of the Poisson distribution with parameter μt we get

$$\begin{aligned}\mathbb{E}[N_t^2] &= \left.\frac{\partial^2 M_N(\theta)}{\partial \theta^2}\right|_{\theta=0} \\ &= \left.\mu t e^{\theta + \mu t(e^\theta - 1)} + (\mu t)^2 e^{2\theta + \mu t(e^\theta - 1)}\right|_{\theta=0} \\ &= \mu t + (\mu t)^2.\end{aligned}$$

By using the above result and the previously calculated expectation $\mathbb{E}[\widetilde{N}_t] = 0$ we obtain that

$$\begin{aligned}\mathrm{Var}(\widetilde{N}_t) &= \mathbb{E}[\widetilde{N}_t^2] - \mathbb{E}[\widetilde{N}_t]^2 \\ &= \mathbb{E}[(N_t - \mu t)^2] \\ &= \mathbb{E}[N_t^2] - 2\mu t \mathbb{E}[N_t] + (\mu t)^2 \\ &= \mu t + (\mu t)^2 - 2(\mu t)^2 + (\mu t)^2 \\ &= \mu t.\end{aligned}$$

□

Proposition 2.18. *The compensated Poisson process $\widetilde{N}_t$ is a martingale with respect to the filtration $\mathcal{F}_s = \sigma\{\widetilde{N}_u : u \leq s\}$.*

Proof. By definition $\widetilde{N}_t$ is $\mathbb{F}$-adapted and $\mathbb{E}[|\widetilde{N}_t|] \leq \mathbb{E}N_t + \mu t < \infty$. To conclude the martingale property of $\widetilde{N}$ we obtain that linearity of the conditional expectation gives us

$$\begin{aligned}\mathbb{E}\left[\widetilde{N}_t|\mathcal{F}_s\right] &= \mathbb{E}\left[\widetilde{N}_t - \widetilde{N}_s + \widetilde{N}_s|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[\widetilde{N}_t - \widetilde{N}_s|\mathcal{F}_s\right] + \mathbb{E}\left[\widetilde{N}_s|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[N_t - \mu t - N_s + \mu s|\mathcal{F}_s\right] + \mathbb{E}\left[\widetilde{N}_s|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[N_t - N_s|\mathcal{F}_s\right] - \mathbb{E}\left[\mu(t-s)|\mathcal{F}_s\right] + \mathbb{E}\left[\widetilde{N}_s|\mathcal{F}_s\right] \quad \text{a.s.}\end{aligned}$$

By definition we have that $\widetilde{N}_s$ is measurable with respect to $\mathcal{F}_s$ and therefore by (v) of the Proposition 2.3

$$\mathbb{E}\left[\widetilde{N}_s|\mathcal{F}_s\right] = \widetilde{N}_s \quad \text{a.s.}$$

Moreover by (iv) of Definition 2.14 $N_t - N_s$ is independent of $\mathcal{F}_s$ and by (ix) of the Proposition 2.3 and (iii) of Definition 2.14 it holds that

$$\begin{aligned}\mathbb{E}\left[N_t - N_s|\mathcal{F}_s\right] &= \mathbb{E}\left[N_t - N_s\right] \\ &= \mathbb{E}\left[N_{t-s}\right] \\ &= \mu(t-s) \text{ a.s.}\end{aligned}$$

Finally, by the above equations we obtain that

$$\begin{aligned}\mathbb{E}\left[\widetilde{N}_t|\mathcal{F}_s\right] &= \mu(t-s) - \mu(t-s) + \widetilde{N}_s \\ &= \widetilde{N}_s \text{ a.s.}\end{aligned}$$

and therefore $\widetilde{N}_s$ is a martingale with respect to the filtration $\mathcal{F}_s$. □

Next we introduce a couple of general properties of Lévy processes.

Proposition 2.19. *Let X and Y be two independent Lévy processes. Then the sum $X + Y = (X_t + Y_t)_{t\geq 0}$ is again a Lévy process.*

Proof.

(i) Since the paths of the processes X and Y are $\mathbb{P}$-almost surely right-continuous with the left limits then the paths of $X + Y$ are clearly $\mathbb{P}$-almost surely right-continuous with left limits.

(ii) By using independency of X and Y we get

$$\begin{aligned}\mathbb{P}(X_0 + Y_0 = 0) &= \mathbb{P}(X_0 = 0, Y_0 = 0) \\ &= \mathbb{P}(X_0 = 0)\mathbb{P}(Y_0 = 0) \\ &= 1\end{aligned}$$

(iii) By (iii) of Definition 2.12 we obtain that for $0 \leq s \leq t$,

$$\begin{aligned}X_t + Y_t - (X_s + Y_s) &= X_t - X_s + Y_t - Y_s \\ &\overset{d}{=} X_{t-s} + Y_{t-s}\end{aligned}$$

Therefore the process $X_t + Y_t - (X_s + Y_s)$ is equal in distribution to $X_{t-s} + Y_{t-s}$.

(iv) A proof for the last condition can be found in Chapter 1 of [8].

□

Note that if we replace $X + Y$ by the process

$$X = \left(\sum_{i=1}^{n} X_t^{(i)}\right)_{t\geq 0}$$

where the $X^{(i)} = (X_t^{(i)})_{t\geq 0}$ are independent Lévy processes, it is possible to conclude that actually sum of any finite number of independent Lévy processes is also Lévy process.

Definition 2.20. Let ξ be a random variable taking values in $\mathbb{R}^d$ with law μ_ξ. We say that ξ is *infinitely divisible* if, for all $n \in \mathbb{N}$, there exists i.i.d. random variables $Y^{(1,n)}, Y^{(2,n)}, \ldots, Y^{(n,n)}$ such that

$$\xi \overset{d}{=} Y^{(1,n)} + \cdots + Y^{(n,n)}.$$

Proposition 2.21. *If X is a Lévy process, then X_t is infinitely divisible for each $t \geq 0$.*

Proof. For each $n \in \mathbb{N}$ we can write

$$\begin{aligned} X_{\frac{t}{n}} &= X_{\frac{kt-kt+t}{n}} \\ &\overset{d}{=} X_{\frac{kt}{n}} - X_{\frac{(k-1)t}{n}} \\ &= Y_t^{(k,n)}, \end{aligned}$$

$Y_t^{(k,n)}$ are i.i.d by (*ii*) of Definition 2.12. Therefore we can express X_t by using $Y^{(k,n)}$ such that

$$\begin{aligned} X_t &= X_{\frac{t-tn-t}{n}} \\ &= X_{\frac{t}{n}} - X_{\frac{t(n-1)}{n}} \\ &= Y_t^{(1,n)} + \cdots + Y_t^{(n,n)}, \end{aligned}$$

and thus for the Lévy process X each X_t is infinitely divisible for each $t \geq 0$. □

2.4. **Poisson random measure.** The Poisson random measure will later be our way to describe the jump structure of Lévy processes. First we define a general class of measures called random measures and then use this to define Poisson random measure.

Definition 2.22 (Random measure)**.** Let $(E, \mathcal{E})$ be a measurable space and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A *random measure* M on $(E, \mathcal{E})$ is a collection of random variables $\{M(A), A \in \mathcal{E}\}$ such that

(i) $M(\emptyset) = 0$.
(ii) Given any sequence $(A_n)_{n=1}^{\infty}$ of mutually disjoint sets in $\mathcal{E}$

$$M\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} M(A_n),$$

almost surely.
(iii) For each disjoint family $(A_1, \ldots, A_n)$ in $\mathcal{E}$, the random variables $M(A_1), \ldots, M(A_n)$ are independent.

Definition 2.23 (Poisson random measure)**.** Let $(E, \mathcal{E}, \mu)$ be a μ-σ-finite measure space and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random measure N is said to be a *Poisson random measure* on $(E, \mathcal{E}, \mu)$,with σ-finite measure μ if for each measurable set $A \subset E$, with finite intensity measure $\mu(A) < \infty$ it holds that $N(A)$ is a Poisson random variable with parameter $\mu(A)$. In other words for all $k \in \mathbb{N}$ it holds

$$\mathbb{P}(N(A) = k) = e^{-\mu(A)} \frac{(\mu(A))^k}{k!}.$$

Let A be such that $\mu(A) < \infty$. We define the *compensated Poisson random measure* by setting

$$\widetilde{N}(A) = N(A) - \mu(A).$$

It is important to notice that Poisson random measures really exist, which is deduced in the following theorem. The construction itself is irrelevant in our case and can be found in section 2.6.1 of [4].

Theorem 2.24 (Existence of the Poisson random measure)**.** *Let μ be a σ-finite measure on a measurable space $(E, \mathcal{E})$. Then there exists a Poisson random measure N on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mu(A) = \mathbb{E}[N(A)]$ for all $A \in \mathcal{E}$.*

Lemma 2.25. *Let N be a Poisson random measure and $\widetilde{N}$ a compensated Poisson random measure. For disjoint compact sets $A_1, \ldots, A_n \in \mathcal{E}$, the variables $\widetilde{N}(A_1), \ldots \widetilde{N}(A_n)$ are independent and verify*

$$\mathbb{E}[\widetilde{N}(A_i)] = 0, \qquad \mathrm{Var}[\widetilde{N}(A_i)] = \mu(A_i).$$

Proof. Similar to the proof of Lemma 2.17. □

Definition 2.26 (Jump measure of a Lévy process)**.** Assume that $X = (X_t)_{t\geq 0}$ is a Lévy process with values in $\mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$. Then we can define a *Jump measure* of X by setting

$$J_X(\omega, .) = \sum_{t\geq 0}^{\Delta X_t(\omega)\neq 0} \delta_{(t,\Delta X_t(\omega))},$$

where $\delta_{(t,\Delta X_t)}$ is a Dirac measure. Moreover for A it holds that

$$J_X([0,t] \times A) := \#\{(s, X_s - X_{s-}) \in [0,t] \times A\}.$$

Thus for every $A \in \mathcal{B}(\mathbb{R}^d)$, the jump measure $J_X([t_1, t_2] \times A)$ counts the number of jump times of the càdlàg process X between t_1 and t_2 such that their jump sizes are in A. Note that it is possible to derive the jump measure explicitly for certain Lévy processes.

Example 2.27 (Jump measure of a Poisson process)**.** One can see that the jump measure of a Poisson process N is given by

$$J_N(\omega, .) = \sum_{t\geq 0}^{\Delta N_t\neq 0} \delta_{(t,\Delta N_t)} = \sum_{n\geq 1} \delta_{(T_n,1)},$$

where $T = (T_n)_{n\geq 1}$ is a sequence of adapted random times that describes the jump times of a Poisson process N and jump-sizes are always 1. Thus for measurable $A \subset \mathbb{R}^d$ it holds that

$$J_N([0,t) \times A) = \begin{cases} \#\{i \geq 1, T_i \in [0,t]\}, \text{ if } 1 \in A \\ 0, \text{ if } 1 \notin A. \end{cases}$$

Next we define a measure that counts the average number of jumps of a Lévy process, called Lévy measure.

Definition 2.28 (Lévy measure)**.** Let X be a Lévy process on $\mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$ such that $0 \notin A$. The measure ν on $\mathbb{R}^d$ defined by:

$$\nu(A) = \mathbb{E}[\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in A\}],$$

is called the *Lévy measure* of the process X. Therefore $\nu(A)$ is the expected number, per unit time, of jumps whose size belongs to A.

Definition 2.29. A Lévy process $X \in \mathbb{R}^d$ is *of infinite activity*, if there are infinite number of jumps on any finite time interval i.e. $\nu(\mathbb{R}^d) = \infty$, where ν is a Lévy measure of X.

Next we introduce two essential theorems associated to Lévy processes called Lévy-Khinchin formula and Lévy-Itô decomposition. First we are going to give a characterization for the characteristic function of a Lévy process called Lévy-Khinchin formula.

Theorem 2.30 (Lévy-Khinchin formula)**.** *Let X be a Lévy process on $\mathbb{R}^d$. If there exists a vector $\gamma \in \mathbb{R}^d$, a positive definite symmetric $d \times d$-matrix A and a Lévy measure ν on $\mathbb{R}^d \backslash \{0\}$ given by Definition 2.28 such that*

$$\int_{|x|\leq 1} |x|^2 \, \nu(dx) < \infty \quad \textit{and} \quad \int_{|x|\geq 1} \nu(dx) < \infty.$$

Then the characteristic function of the process X can be written in form

$$\begin{aligned} \phi_{X_t}(u) &= \mathbb{E}\left[e^{i\langle u, X_t\rangle}\right] \\ &= \exp\left(t\left\{i\langle \gamma, u\rangle - \frac{1}{2}\langle u, Au\rangle + \int\limits_{\mathbb{R}^d\backslash\{0\}} \left[e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle, \mathbb{1}_{\{|x|\leq 1\}}\right] \nu(dx)\right\}\right), \end{aligned}$$

for all $u \in \mathbb{R}^d$.

Proof. See Section 3.4 [4]. □

Theorem 2.31 (Lévy-Itô decomposition)**.** *Let X be a Lévy process on $\mathbb{R}^d$ and ν its Lévy measure concentrated on $\mathbb{R}^d \backslash \{0\}$ given by Definition 2.28 such that*

$$\int_{|x|\leq 1} |x|^2 \, \nu(dx) < \infty \quad \textit{and} \quad \int_{|x|\geq 1} \nu(dx) < \infty$$

and let the jump measure of the process X be denoted by J_X be a Poisson random measure on $[0, \infty) \times \mathbb{R}^d$ with intensity measure $\nu(dx)dt$. Then there exists a vector $\gamma \in \mathbb{R}^d$ and a d-dimensional Brownian motion B with a covariance matrix A such that

$$X_t = \gamma t + B_t + \iint\limits_{|x|\geq 1, s\in[0,t]} x \, J_X(ds \times dx) + \lim_{\epsilon \downarrow 0} \iint\limits_{\epsilon \leq |x| < 1, s\in[0,t]} x \, \tilde{J}_X(ds \times dx).$$

Note that terms in the previous equation are independent and the convergence in the last term is almost sure and uniform on $[0, \infty)$.

Proof. See Section 2.4 of [1]. □

We called the triplet (A, ν, γ) from the Theorem 2.30 and 2.31 the *Lévy triplet* of the process X, where A is positive definite matrix, ν a measure and γ a vector. The drift terms γ and covariance matrix A from Lèvy triplet defines a Brownian motion with drift $\gamma t + B_t$, which is a continuous part of the above decomposition.

The meaning of the other two terms of the Lévy-Itô decomposition described by the Lévy measure ν is describe the jumps of the Lévy process X. The condition

$\int_{|x|\geq 1} \nu(dx) < \infty$ implies that every Lévy process has finite number of "large" jumps. By [4] every compound Poisson process can be written as

$$X_t = \iint_{[0,t]\times\mathbb{R}^d} x\, J_X(ds \times dx), \tag{2.1}$$

and therefore the term

$$\iint_{|x|\geq 1, s\in[0,t]} x\, J_X(ds \times dx)$$

is a compound Poisson process. Similary the last term

$$\iint_{\epsilon\leq|x|<1, s\in[0,t]} x\, \tilde{J}_X(ds \times dx)$$

describe the "small" jumps of the Lèvy process X and is as well a compound Poisson process. However in this case there can be infinitely many jumps and their sum not necessarily convergence. Hence we fix this by using the compensated version $\tilde{J}_X$ of the jump measure J_X in the sense of the Definition 2.23.

2.5. **The processes of finite variation.** Next we are going to introduce a class of functions and a class of processes defined by their variation.

Definition 2.32. (i) *The total variation* of a function $f : [a, b] \to \mathbb{R}^d$ is defined by

$$V_{\mathcal{P}}(f) = \sup\left\{\sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|\right\},$$

where the supremum is taken over all finite partitions $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b\}$ of the interval $[a, b]$.

(ii) A stochastic process X is said to be of *finite variation* if the paths $X(\omega) : [0, \infty) \to \mathbb{R}^d$ are of finite variation for almost all $\omega \in \Omega$ on finite interval $[a, b]$.

Functions of finite variation have lot of useful applications in stochastics and analysis for example in integration. The following two theorems give us a examples of processes with finite variation.

Theorem 2.33. *We define a* subordinator $T = (T_t)_{t\geq 0}$ *to be a one-dimensional Lévy process that is non-decreasing almost surely. Then every subordinator is of finite variation.*

Proof. Let T be a subordinator defined on fixed interval $[a, b]$ and $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b\}$ a finite partition of $[a, b]$. The total variation of the paths $T(\omega) : [0, \infty) \to \mathbb{R}^d$ on $[a, b]$ is

$$\begin{aligned} V_{\mathcal{P}}(T_t(\omega)) &= \sum_{j=1}^{n} |T_{t_j} - T_{t_{j-1}}| \\ &= \sum_{j=1}^{n} T_{t_j} - T_{t_{j-1}} \\ &= T_{t_n} - T_{t_0} \\ &\leq T_b \\ &< \infty. \end{aligned}$$

Therefore every subordinator is of finite variation. □

Theorem 2.34. *A continuous martingale is of finite variation if and only if it is constant almost surely.*

Proof. See Chapter 4 of [11]. □

Proposition 2.35. *A necessary and sufficient condition for a Lévy process to be of finite variation is that there is no Brownian part i.e $A = 0$ in the Lévy-Khinchine formula and that $\int_{|x|\leq 1} |x|\, \nu(dx) < \infty$.*

Proof. See Section 3.5 of [4] □

The finite variation assumption simplify the behavior of Lévy processes. Thus we can cut out the continuous Brownian motion part in Lévy-Itô decomposition and Lévy-Khintchin formula that are presented in the case of general Lévy processes in previous section. Moreover the term $\iint_{\epsilon\leq|x|<1, s\in[0,t]} x\, \tilde{J}_X(ds \times dx)$ in Theorem 2.31 that describes the small jumps of the original Lévy process is not anymore needed since sum of them might have been infinity and thus does not satisfy the finite variation assumption.

Corollary 2.36 (Lévy-Itô decomposition and Lévy-Khinchin representation in the finite-variation case)**.** *Let X be a Lévy process of finite variation thus its Lévy triple is now given by $(0, \nu, \gamma)$, where ν is a Lévy measure of process X, $\gamma \in \mathbb{R}^d$ and the covariance matrix of Brownian motion is $A = 0$. Then X can be expressed as the sum of its jumps between 0 and t and linear drift term γt by*

$$X_t = \gamma t + \iint\limits_{[0,t]\times\mathbb{R}^d} x\, J_X(ds \times dx),$$

and its characteristic function can be expressed as

$$\mathbb{E}\left[e^{i\langle u, X_t\rangle}\right] = \exp\left(t\left\{i\langle\gamma, u\rangle + \int_{\mathbb{R}^d} (e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbb{1}_{\{|x|\leq 1\}})\, \nu(dx)\right\}\right),$$

for all $u \in \mathbb{R}^d$.

Theorem 2.37. *Let X be a Lévy process of finite variation. Then it can be written as the difference of two independent subordinators.*

Proof. A Lévy process of finite variation can be written and decomposed in the form

$$\begin{aligned} X_t &= \gamma t + \iint_{[0,t]\times\mathbb{R}} x\, J_X(ds \times dx) \\ &= \left\{(\gamma \vee 0)t + \iint_{[0,t]\times(0,\infty)} x\, J_X(ds \times dx)\right\} \\ &\qquad - \left\{|\gamma \wedge 0|t + \iint_{[0,t]\times(-\infty,0)} |x|\, J_X(ds \times dx)\right\} \\ &= X_t^{(1)} - X_t^{(2)}, \end{aligned}$$

where $t \geq 0$, γ is the drift coefficent and J_X is the jump measure. Both integrals are finite (see Section 2.3 of [8]) and since X is of finite variation by Proposition 2.35 we obtain that

$$\int_{(0,\infty)} (1 \wedge x)\, \nu(dx) < \infty \quad \text{and} \quad \int_{(-\infty,0)} (1 \wedge |x|)\, \nu(x) < \infty.$$

As J_X has independent counts on disjoint domains (see Section 2 of [8]), it follows that the integrals in $X_t^{(1)}$ and $X_t^{(2)}$ are independent. Finally by obtaining that the processes

$$X_t^{(1)} = (\gamma \vee 0)t + \iint\limits_{[0,t]\times(0,\infty)} x\, J_X(ds \times dx),$$

and

$$X_t^{(2)} = |\gamma \wedge 0|t + \iint\limits_{[0,t]\times(-\infty,0)} |x|\, J_X(ds \times dx),$$

have monotone paths, the desired result follows. □

More properties of the finite variation functions are later introduced in Section 5.2.

2.6. **Modification of a process.** There are multiple approaches to the concept of equality of stochastic processes.

Definition 2.38. For $0 \leq t_1 < \cdots < t_n$ the distribution of $X_{t_1}, \ldots X_{t_n}$ is *a finite-dimensional distribution* of a process $(X_t)_{t\geq 0}$. If for the processes X and Y all finite-dimensional distributions conincide, they have the *same distribution*.

Definition 2.39. Let X and Y be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The process Y is said to be a *modifications* of X if $\mathbb{P}(X_t = Y_t) = 1$, for each $t \geq 0$. Two processes X and Y are *indistinguishable* if $\mathbb{P}(X_t = Y_t, t \geq 0) = 1$.

Note that if the process X is a modification of Y then their has the same finite-dimensional distributions. Moreover indistinguishable processes are also modification of each other. Next theorem gives us another relation between modification and indistinguishable process.

Theorem 2.40. *Let X and Y be two stochastic processes, such that X is a modification of Y. If X and Y have right continuous paths almost surely, then X and Y are indistinguishable.*

Proof. Let A and B be two null sets where X and Y are not right-continuous respectively. Let $N_t = \{\omega : X_t(\omega) \neq Y_t(\omega)\}$, and $N = \bigcup_{t\in\mathbb{Q}\cap[0,\infty)} N_t$. We define a set $M = A \cup B \cup N$ and it follows that $\mathbb{P}(M) \leq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(N) = 0$.

Assume $t \geq 0$ and let $(t_n)_{n=1}^{\infty}$ be a sequence such that t_n decreases to t through $\mathbb{Q}$. Since X and Y are modifications, it holds for all $\omega \notin M$ that

$$X_{t_n}(\omega) = Y_{t_n}(\omega).$$

Therefore also for all irrationals we have

$$X_t(\omega) = \lim_{n\to\infty} X_{t_n}(\omega) = \lim_{n\to\infty} Y_{t_n}(\omega) = Y_t(\omega).$$

Since $\mathbb{P}(M) = 0$, the processes X and Y are indistinguishable. □

3. Stochastic integration

3.1. **$\mathcal{H}_2$-space.** In this section our aim is to define the stochastic integral for $T \geq 0$ fixed denoted by

$$I_T(F) = \int_0^T \int_E F(t,x)\, M(dt \times dx).$$

We start our construction by defining the measure M, which we are going to use as a integrator.

Definition 3.1. Let M be a random measure on $[0,\infty] \times E$, where E is a topological space and let $\mathcal{B}(E)$ be a Borel σ-algebra of E. For each $A \in \mathcal{B}(E)$, we define a process $M^A = (M_t^A)_{t\geq 0}$ by setting $M_t^A = M([0,t] \times A)$. We say that M is *martingale-valued measure* if there exist $V \in \mathcal{B}(E)$ such that M^A is a martingale whenever $\overline{A} \cap V = \emptyset$. Furthermore we will denote

$$M((s,t] \times A) = M((0,t] \times A) - M((0,s] \times A),$$

where $s < t$.

Definition 3.2. A martingale-valued measure M is said to be *type* $(2,\rho)$ if it satisfies the following conditions

(i) $M(\{0\} \times A) = 0$ almost surely for all $A \in \mathcal{B}(E)$.
(ii) $M((s,t] \times A)$ is independent of $\mathcal{F}_s$.
(iii) If $E \in \mathcal{B}(\mathbb{R}^d)$ there exists a σ-finite measure ρ on $[0,\infty] \times E$ for which

$$\mathbb{E}[M((0,t] \times A)^2] = \rho((0,t] \times A),$$

for all $0 \leq s < t < \infty$ and $A \in \mathcal{B}(E)$.

Definition 3.3. Let M be a $(2,\rho)$-type martingale-valued measure. For fixed $T > 0$, we define $\mathcal{H}_2(T,E)$ to be the *linear space* of all mappings $F : [0,T] \times E \times \Omega \longrightarrow \mathbb{R}$ which coincide almost everywhere with respect to $\rho\times\mathbb{P}$ and which satisfy the following conditions:

(i) F is $\mathcal{P} \otimes \mathcal{B}(E)$-measurable, where $\mathcal{P}$ is the predictable sigma-algebra in the sense of Definition 2.10.
(ii) F is left-continuous.
(iii) $\int_0^T \int_E \mathbb{E}[|F(t,x)|^2\, \rho(dt \times dx) < \infty$.

We may now also define the inner product on $\mathcal{H}_2(T,E)$ by

$$\langle F, G\rangle = \iint\limits_{[0,T]\times E} \mathbb{E}[F(t,x)G(t,x)]\, \rho(dt \times dx),$$

for each $F, G \in \mathcal{H}_2(T,E)$, and we obtain a norm $||\cdot||_{T,\rho}$ defined by

$$(3.1) \quad ||F||^2_{T,\rho} = \iint\limits_{[0,T]\times E} \mathbb{E}[|F(t,x)|^2]\,\rho(dt\times dx) = \mathbb{E}\left[\iint_{[0,T]\times E} |F(t,x)|^2\,\rho(dt\times dx)\right].$$

Note that the latter equality in the above equation is possible by the Fubini-Tonelli theorem (see Theorem 4.9).

Definition 3.4. A process $F : [0,T]\times E\times\Omega \to \mathbb{R}$ is *simple* if, for some $m,n\in\mathbb{N}$, there exists $0\le t_1\le t_2\le\cdots\le t_{m+1}=T$ and disjoint Borel subsets $A_1,A_2,\ldots,A_n$ of E with each $\mu(A_i)<\infty$ such that

$$F=\sum_{j=1}^{m}\sum_{k=1}^{n} c_k F_{t_j}\mathbb{1}_{((t_j,t_{j+1}]\times A_k)} := \sum_{j,k=1}^{m,n} F^{(k)}_{t_j}\mathbb{1}_{((t_j,t_{j+1}]\times A_k)}.$$

In the above equation each $c_k\in\mathbb{R}$ and each F_{t_j} is a bounded $\mathcal{F}_{t_j}$-measurable random variable. Moreover μ is some σ-finite measure on E. We also define $S(T,E)$ to be *the linear space of all simple predictable processes.*

We notice that $\mathcal{H}_2(T,E)$ is a real Hilbert space (see Lemma 4.1.3 of [1]) and that simple processes $S(T,E)$ are dense in $\mathcal{H}_2(T,E)$ (see Lemma 4.1.4 of [1]).

3.2. **Construction of the stochastic integral.** Similarly to the classical integral we begin our construction for simple processes. Let $F\in S(T,E)$, for which we can write

$$F=\sum_{j,k=1}^{m,n} F^{(k)}_{t_j}\mathbb{1}_{(t_j,t_{j+1}]\times A_k)}.$$

We can then define the stochastic integral for the simple process F by setting

$$I_{[0,T]}(F)=\sum_{j,k=1}^{m,n} F^{(k)}_{t_j} M((t_j,t_{j+1}]\times A_k).$$

We deduce for $I_{[0,T]}$ the following linearity and isometry properties.

Lemma 3.5. *Let $F,G\in S(T,E)$ and $\alpha,\beta\in\mathbb{R}$ then*

(i) *$\alpha F+\beta G\in S(T,E)$ and*

$$I_{[0,T]}(\alpha F+\beta G)=\alpha I_{[0,T]}(F)+\beta I_{[0,T]}(G).$$

(ii) *For each $T\ge 0$,*

$$\mathbb{E}[I_{[0,T]}(F)]=0,\quad \mathbb{E}\left[I_{[0,T]}(F)^2\right]=\iint\limits_{[0,T]\times E}\mathbb{E}[|F(t,x)|^2]\,\rho(dt\times dx).$$

Proof. (i) Nevertheless F might have different t_j and A_k than G, by taking intersection we can always find them such that the following holds. For $F,G\in S(T,E)$, there exist $0\le t_1\le t_2\le\cdots\le t_{m+1}=T$ and disjoint Borel subsets $A_1,A_2,\ldots,A_n$ of E with each $\mu(A_i)<\infty$ such that

$$F=\sum_{j,k=1}^{m,n} F^{(k)}_{t_j}\mathbb{1}_{(t_j,t_{j+1}]\times A_k} \quad\text{and}\quad G=\sum_{j,k=1}^{m,n} G^{(k)}_{t_j}\mathbb{1}_{(t_j,t_{j+1}]\times A_k}.$$

Thus

$$\begin{aligned}
\alpha F + \beta G &= \alpha \sum_{j,k=1}^{m,n} \left[F_{t_j}^{(k)} \mathbb{1}_{(t_j,t_{j+1}]\times A_k} \right] + \beta \sum_{j,k=1}^{m,n} \left[G_{t_j}^{(k)} \mathbb{1}_{(t_j,t_{j+1}]\times A_k} \right] \\
&= \sum_{j,k=1}^{m,n} \left[\alpha F_{t_j}^{(k)} \mathbb{1}_{(t_j,t_{j+1}]\times A_k} + \beta G_{t_j}^{(k)} \mathbb{1}_{(t_j,t_{j+1}]\times A_k} \right] \\
&= \sum_{j,k=1}^{m,n} \left[(\alpha F_{t_j}^{(k)} + \beta G_{t_j}^{(k)}) \mathbb{1}_{(t_j,t_{j+1}]\times A_k} \right] \\
&= \sum_{j,k=1}^{m,n} \left[(\alpha F + \beta G)_{t_j}^{(k)} \mathbb{1}_{(t_j,t_{j+1}]\times A_k} \right],
\end{aligned}$$

and therefore $\alpha F + \beta G \in S(T,E)$. Furthermore we can see that

$$\begin{aligned}
I_{[0,T]}(\alpha F + \beta G) &= \sum_{j,k=1}^{m,n} \left[(\alpha F + \beta G)_{t_j}^{(k)} M((t_j,t_{j+1}] \times A_k) \right] \\
&= \sum_{j,k=1}^{m,n} \left[\alpha F_{t_j}^{(k)} M((t_j,t_{j+1}] \times A_k) + \beta G_{t_j}^{(k)} M((t_j,t_{j+1}] \times A_k) \right] \\
&= \alpha \sum_{j,k=1}^{m,n} \left[F_{t_j}^{(k)} M((t_j,t_{j+1}] \times A_k) \right] + \beta \sum_{j,k=1}^{m,n} \left[G_{t_j}^{(k)} M((t_j,t_{j+1}] \times A_k) \right] \\
&= \alpha I_{[0,T]}(F) + \beta I_{[0,T]}(G).
\end{aligned}$$

(ii) Note that for each time interval $[t_j, t_{j+1}]$ the process $F_{t_j}^{(k)}$ is $\mathcal{F}_{t_j}$-adapted. By linearity of the expectation and (vi) of Proposition 2.3 we have

$$\begin{aligned}
\mathbb{E}[I_{[0,T]}(F)] &= \mathbb{E}\left[\sum_{j,k=1}^{m,n} F_{t_j}^{(k)} M((t_j,t_{j+1}] \times A_k) \right] \\
&= \sum_{j,k=1}^{m,n} \mathbb{E}\left[F_{t_j}^{(k)} M((t_j,t_{j+1}] \times A_k) \right] \\
&= \sum_{j,k=1}^{m,n} \mathbb{E}\left[\mathbb{E}\left[F_{t_j}^{(k)} M((t_j,t_{j+1}] \times A_k) | \mathcal{F}_{t_j} \right] \right].
\end{aligned}$$

Since for the each time interval $[t_j, t_{j+1}]$ the process $F_{t_j}^{(k)}$ is $\mathcal{F}_{t_j}$-measurable and $M((t_j,t_{j+1}] \times A_k)$ is independent of $\mathcal{F}_{t_j}$ by (ii) of Definition 3.2, we obtain by (vii) and (ix) of Proposition 2.3 that

$$\mathbb{E}[I_{[0,T]}(F)] = \sum_{j,k=1}^{m,n} \mathbb{E}[F_{t_j}^{(k)}] \mathbb{E}\left[M((t_j,t_{j+1}] \times A_k) \right].$$

Furthermore by the martingale property, for each $1 \leq j \leq m$ and $1 \leq k \leq n$, we have

$$
\begin{aligned}
\mathbb{E}[M((t_j,t_{j+1}]\times A_k)] &= \mathbb{E}[M((0,t_{j+1}]\times A_k)] - \mathbb{E}[M((0,t_j]\times A_k)] \\
&= \mathbb{E}[\mathbb{E}[M((0,t_{j+1}]\times A_k)|\mathcal{F}_{t_j}]] - \mathbb{E}[M((0,t_j]\times A_k)] \\
&= \mathbb{E}[M((0,t_j]\times A_k)] - \mathbb{E}[M((0,t_j]\times A_k)] \\
&= 0
\end{aligned} \tag{3.2}
$$

and therefore it follows that

$$
\mathbb{E}[I_{[0,T]}(F)] = 0. \tag{3.3}
$$

By linearity of the expectation again and splitting the sum into three pieces we find that

$$
\begin{aligned}
\mathbb{E}[I_T(F)^2] &= \sum_{j,k=1}^{m,n}\sum_{l,p=1}^{m,n}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_l}^{(p)}M((t_l,t_{l+1}]\times A_p)\right] \\
&= \sum_{j,k=1}^{m,n}\sum_{l,p=1}^{n}\sum_{l<p}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_l}^{(p)}M((t_l,t_{l+1}]\times A_p)\right] \\
&\quad + \sum_{j,k=1}^{m,n}\sum_{p=1}^{n}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_l}^{(p)}M((t_l,t_{l+1}]\times A_p)\right] \\
&\quad\quad + \sum_{j,k=1}^{m,n}\sum_{l,p=1}^{n}\sum_{l>p}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_l}^{(p)}M((t_l,t_{l+1}]\times A_p)\right]
\end{aligned}
$$

We can apply similar procedure as in case of $\mathbb{E}[I_{[0,T]}(F)]$ to each of the three term above. First we note that when $l < j$ we have that $M((t_j,t_{j+1}]\times A_k)$ is independent of $\mathcal{F}_{t_j}$ by (ii) of Definition 3.2. Moreover $F_{t_j}^{(k)}, F_{t_l}^{(p)}$ and $M((t_l,t_{l+1}]\times A_p)$ are $\mathcal{F}_{t_j}$-measurable. Hence by (vi), (iiv) and (ix) of Proposition 2.3

$$
\begin{aligned}
&\sum_{j,k=1}^{m,n}\sum_{l,p=1}^{n}\sum_{l<p}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_l}^{(p)}M((t_l,t_{l+1}]\times A_k)\right] \\
&= \sum_{j,k=1}^{m,n}\sum_{l,p=1}^{n}\sum_{l<p}\mathbb{E}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_l}^{(p)}M((t_l,t_{l+1}]\times A_k)|\mathcal{F}_{t_j}\right] \\
&= \sum_{j,k=1}^{m,n}\sum_{l,p=1}^{n}\sum_{l<p}\mathbb{E}\left[F_{t_j}^{(k)}M((t_l,t_{l+1}]\times A_k)F_{t_l}^{(p)}\right]\mathbb{E}\left[M((t_j,t_{j+1}]\times A_k)\right] \\
&= 0,
\end{aligned}
$$

where the last equality holds by Equation (3.2). Note that similar argument holds in the case $l > j$.

In the last case where $l = j$ it holds that $M((t_j,t_{j+1}],A_k)$ and $M((t_j,t_{j+1}],A_p)$ are independent of $\mathcal{F}_{t_j}$ by (ii) of Definition 3.2. Moreover since $F_{t_j}^{(k)}$ and $F_{t_j}^{(p)}$ are $\mathcal{F}_{t_j}$-measurable we have again by (vi), (vii) and (ix) of Proposition 2.3

that

$$\sum_{j,k=1}^{m,n}\sum_{p=1}^{n}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_j}^{(p)}M((t_j,t_{j+1}]\times A_p)\right]$$
$$=\sum_{j,k=1}^{m,n}\sum_{p=1}^{n}\mathbb{E}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_j}^{(p)}M((t_j,t_{j+1}]\times A_p)|\mathcal{F}_{t_j}\right]$$
$$=\sum_{j,k=1}^{m,n}\sum_{p=1}^{n}\mathbb{E}\left[F_{t_j}^{(k)}F_{t_j}^{(p)}\right]\mathbb{E}\left[M((t_j,t_{j+1}]\times A_k)M((t_j,t_{j+1}]\times A_p)\right].$$

We can split the sum into two pieces where $k=p$ and $k\neq p$. Hence by (iii) of Definition 2.22 and Equation (3.3) we obtain that

$$\sum_{j,k=1}^{m,n}\sum_{p=1}^{n}\mathbb{E}\left[F_{t_j}^{(k)}M((t_j,t_{j+1}]\times A_k)F_{t_j}^{(p)}M((t_j,t_{j+1}]\times A_p)\right]$$
$$=\sum_{j,k=1}^{m,n}\sum_{k=p}\mathbb{E}\left[F_{t_j}^{(k)}F_{t_j}^{(p)}\right]\mathbb{E}\left[M((t_j,t_{j+1}]\times A_k)M((t_j,t_{j+1}]\times A_k)\right]$$
$$+\sum_{j,k=1}^{m,n}\sum_{k\neq p}\mathbb{E}\left[F_{t_j}^{(k)}M(t_j,t_{j+1}]\times A_k\right]\mathbb{E}\left[F_{t_j}^{(p)}M((t_j,t_{j+1}]\times A_p)\right]$$
$$=\sum_{j,k=1}^{m,n}\mathbb{E}\left[\left(F_{t_j}^{(k)}\right)^2\right]\mathbb{E}\left[M((t_j,t_{j+1}]\times A_k)^2\right].$$

By (iii) of Definition 3.2 there exists a σ-finite measure ρ on $[0,t]\times E$ for which

$$\mathbb{E}\left[M((0,t_{j+1}]\times A_k)^2\right]=\rho((0,t_{j+1}]\times A_k).$$

Therefore by the martingale property and the above equation we obtain that

$$\sum_{j,k=1}^{m,n}\mathbb{E}\left[\left(F_{t_j}^{(k)}\right)^2\right]\mathbb{E}\left[M((t_j,t_{j+1}]\times A_k)^2\right]$$
$$=\sum_{j,k=1}^{m,n}\mathbb{E}\left[\left(F_{t_j}^{(k)}\right)^2\right]\left(\mathbb{E}\left[M((0,t_{j+1}]\times A_k)^2\right]-\mathbb{E}\left[M((0,t_j]\times A_k)^2\right]\right)$$
$$=\sum_{j,k=1}^{m,n}\mathbb{E}\left[\left(F_{t_j}^{(k)}\right)^2\right](\rho((0,t_{j+1}]\times A_k)-\rho((0,t_j]\times A_k))$$
$$=\sum_{j,k=1}^{m,n}\mathbb{E}\left[\left(F_{t_j}^{(k)}\right)^2\right]\rho((t_j,t_{j+1}]\times A_k).$$

Since the sums in cases $l<j$ and $l>j$ were 0 we conclude that

$$\mathbb{E}\left[I_{[0,T]}(F)^2\right]=\sum_{j=1}^{n}\mathbb{E}\left[\left(F_{t_j}^{(k)}\right)^2\rho((t_j,t_{j+1}]\times A_k\right]=\iint_{[0,T]\times E}\mathbb{E}[|F(t,x)|^2]\,\rho(dt\times dx),$$

which is the desired result.

□

Theorem 3.6. *$I_{[0,T]}$ is a linear isometry from $S(T,E)$ into $L^2(\Omega,\mathcal{F},\mathbb{P})$, and it extends to an isometric embedding of the whole of $\mathcal{H}_2(T,E)$ into $L^2(\Omega,\mathcal{F},\mathbb{P})$.*

Proof. By (i) of Lemma 3.5 the integral $I_{[0,T]}$ is linear for simple processes $F \in S(T,E)$ and by *(ii)* of Lemma 3.5 and Equation (3.1)

$$||I_{[0,T]}(F)||_2^2 = \mathbb{E}[|I_{[0,T]}(F)|^2] = \iint\limits_{[0,T]\times E} \mathbb{E}[|F(t,x)|^2]\,\rho(dt\times dx) = ||F||_{T,\rho}^2.$$

Therefore $I_{[0,T]}$ is a linear isometry from $S(T,E)$ into $L^2(\Omega,\mathcal{F},\mathbb{P})$.

Moreover the space $S(T,E)$ is dense in $\mathcal{H}_2(T,E)$ and therefore for any $F \in \mathcal{H}_2(T,E)$ we can find a sequence $(F_n)_{n=1}^{\infty} \in S(T,E)$ such that

$$\lim_{n\to\infty} ||F_n - F||_{T,\rho} = \lim_{n\to\infty} \mathbb{E}\left[\iint_{[0,T]\times E} |F_n(t,x) - F(t,x)|\,\rho(dt\times dx)\right] = 0. \tag{3.4}$$

Hence by isometry formula of Lemma 3.5 and Equation (3.4) it holds that

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}[|I_{[0,T]}(F_n) - I_{[0,T]}(F_m))|^2] &= \lim_{n\to\infty} \mathbb{E}[|I_{[0,T]}(F_n - F_m)|^2] \\ &= \lim_{n\to\infty} \iint\limits_{[0,T]\times E} \mathbb{E}[|F_n(t,x) - F_m(t,x)|^2]\,\rho(dt\times dx) \\ &= 0.\end{aligned}$$

Therefore we define $I_{[0,T]}(F)$ as limit in $L^2(\mathbb{P})$ of $I_{[0,T]}(F_n)$. Furthermore $I_{[0,T]}$ extends to an isometric embedding of the whole of $\mathcal{H}_2(T,E)$ into $L^2(\Omega,\mathcal{F},\mathbb{P})$. □

Definition 3.7. We called extension $I_{[0,T]}(F)$ in Theorem 3.6 the *stochastic integral* of a process $F \in \mathcal{H}_2(T,E)$ and denote

$$I_{[0,T]}(F) = \iint\limits_{[0,T]\times E} F(t,x)\,M(dt\times dx).$$

By Theorem 3.6 we have also the following equality

$$||I_T(F)||_2^2 = \mathbb{E}[|I_T(F)|^2] = ||F||_{T,\rho}^2,$$

for all $F \in \mathcal{H}_2(T,E)$ called *Ito's isometry.*

Finally we can define the stochastic integral for more general sets by replacing a set $[0,T]\times E$ by $(a,b)\times A$, where $0 \le a \le b \le T$ and $A \in \mathcal{B}(E)$. If $F \in \mathcal{H}_2(T,E)$ we can see that $F\mathbb{1}_{(a,b)\times A} \in \mathcal{H}_2(T,E)$ and we may define

$$I_{(a,b)\times A}(F) = I_{[0,T]}(F\mathbb{1}_{(a,b)\times A}) = \iint\limits_{[a,b]\times A} F(t,x)\,M(dt\times dx) = I_T(F\mathbb{1}_{(a,b)\times A}).$$

We will also write $I_{(a,b)} = I_{(a,b)\times E}$. Note that if $||F||_{t,\rho} < \infty$ for all $t \ge 0$ the integral $I(F) = (I_t(F))_{t\ge 0}$ itself is a stochastic process. The following theorem contains some of the basic properties of stochastic integral.

Theorem 3.8. *If $F, G \in \mathcal{H}_2(T,E)$ and $\alpha, \beta \in \mathbb{R}$ then:*

(i) $I_{[0,T]}(\alpha F+\beta G)=\alpha I_{[0,T]}(F)+\beta I_{[0,T]}(G)$.
(ii) $\mathbb{E}[I_{[0,T]}(F)]=0$ *and* $\mathbb{E}[I_{[0,T]}(F)^2]=\iint_{[0,T]\times E}\mathbb{E}[|F(t,x)|^2]\,\rho(dt\times dx)$.
(iii) $I_t(F)$ *is* $\mathcal{F}_t$*-adapted.*
(iv) $I_t(F)$ *is a square-integrable martingale.*

Proof. See Section 4.2 of [1]. □

3.3. **Poisson stochastic integrals.** Previous construction of stochastic integration allows us to integrate with respect to various different integrators. One of the most common case is when we have Poisson random measure N introduced in Chapter 2.4. Until we can properly define the stochastic integral with respect to Poisson random measure we have to extend the set of mapping $\mathcal{H}_2(T,E)$ to more general case.

Similarly to $\mathcal{H}_2(T,E)$ we assume that $\mathcal{P}_2(T,E)$ includes all mappings $F:[0,T]\times E\times\Omega\to\mathbb{R}$ which coincide almost everywhere with respect to the $\rho\times\mathbb{P}$ and which satisfies the following condition.

(i) F is predictable.
(ii) $\mathbb{P}\left(\int_0^T\int_E|F(t,x)|^2\,\rho(dt\times dx)<\infty\right)=1$.

Note that indeed the only difference to definition of $H_2(T,E)$ is the condition (ii). Now by using previous construction one can derive (see section 4.2.2 of [1]) the *extended stochastic integral* denoted by $\hat{I}_{[0,T]}(F)=(\hat{I}_t(F))_{t\geq 0}$, where

$$\hat{I}_{[0,T]}(F)=\int_0^T\int_E F(t,x)\,M(dt\times dx).$$

The above integral is the limit of the sequence that satisfy the following equation

$$\int_0^T\int_E F(t,x)\,M(dt\times dx)=\lim_{n\to\infty}\int_0^T\int_E F_n(t,x)\,M(dt\times dx),$$

in probability. Furthermore we consider $\hat{I}_{[0,T]}$ as a process for all $t\geq 0$ that provided the condition

$$\mathbb{P}\left(\int_0^t\int_E|F(t,x)|^2\,\rho(dt\times dx)<\infty\right)=1.$$

Assume next N to be a Poisson random measure on $[0,T]\times E$ with intensity $\rho(dt\times dx)$. Then by [1] for any predictable random function $F:[0,T]\times E\times\Omega\longrightarrow\mathbb{R}$ verifying

$$\mathbb{P}\left(\int_0^t\int_E|F(t,x)|^2\,\mu(dt\times dx)<\infty\right)=1.$$

we can define a stochastic integral, where integrator is Poisson random measure by setting

$$I_{[0,T]}=\iint\limits_{[0,T]\times E}F(t,x)M(dt\times dx)=\sum_{j,k=1}^{n,m}F_k(t_j)N((t_j,t_{j+1}]\times A_k).$$

Similarly the compensated integral is defined by

$$I_{[0,T]}=\iint\limits_{[0,T]\times E}F(t,x)\widetilde{N}(dt\times dx)=\sum_{j,k}^{n,m}F_k(t_j)N((t_j,t_{j+1}]\times A_k)-\rho((t_j,t_{j+1}]\times A_k).$$

There exists a special case later where the integrator of Poisson stochastic integral is jump measure J_X defined in Section 2.4. We deduce earlier that for Lévy processes X with Lévy measure ν the J_X is a Poisson random measure with intensity $\nu(dx)dt$. Then for predictable random function $F \in \mathcal{P}(T, E)$ the stochastic Poisson integral with respect to jump measure J_X can be written as a sum of terms involving jump times and jump sizes of X. By using properties of Dirac measure $\delta_{(t,\Delta S_t)}$ it holds that

$$\begin{aligned}
\iint\limits_{[0,T]\times\mathbb{R}^d} F(s,x)\, J_X(ds\times dx) &= \iint\limits_{[0,T]\times\mathbb{R}^d} F(s,x) \sum_{t\in[0,T]}^{\Delta X_t\neq 0} \delta_{(t,\Delta S_t)} \\
&= \sum_{t\in[0,T]}^{\Delta X_t\neq 0} \iint\limits_{[0,T]\times\mathbb{R}^d} F(s,x)\, \delta_{(t,\Delta S_t)} \qquad (3.5) \\
&= \sum_{t\in[0,T]}^{\Delta X_t\neq 0} F(t,\Delta X_t).
\end{aligned}$$

3.4. **Itô's formula.** Next we will introduce one of the change of the variable formula called Itô formula which is in key position in our aim to prove the the main result of the thesis. Note that similar Itô formulas can be formulated also in much more general form, for example for Lévy processes without finite variation assumption (see Theorem 4.3 of [8]).

Theorem 3.9. *Assume that X is a Lévy process of finite variation when by Lèvy-Itô-decomposition from Corollary 2.36 the process X can be represented (in case $d = 1$) as*

$$X_t = \gamma t + \int\limits_{[0,t]\times\mathbb{R}} x\, J_X(ds\times dx)$$

where $\gamma \in \mathbb{R}$ is drift coefficient and J_X is a Jump measure of the process X in the sense of Definition 2.26. Let $C^{1,1}([0,\infty)\times\mathbb{R})$ be the space of functions $f : [0,\infty)\times\mathbb{R} \longrightarrow \mathbb{R}$ which are continuously differentiable in each variable (in the case of the derivative in the first variable at the origin a right derivative is understood). If $f \in C^{1,1}([0,\infty)\times\mathbb{R})$ then

$$\begin{aligned}
f(t,X_t) = f(0,X_0) &+ \int_0^t \frac{\partial f}{\partial s}(s,X_s)\, ds + \gamma \int_0^t \frac{\partial f}{\partial x}(s,X_s)\, ds \\
&+ \iint_{[0,t]\times\mathbb{R}} (f(s,X_{s-}+x) - f(s,X_{s-}))\, J_X(ds\times dx),
\end{aligned}$$

Proof. We define for all $\epsilon > 0$, and $t \leq 0$ a random variable

$$X_t^\epsilon = \gamma t + \iint\limits_{[0,t]\times\{|x|\geq\epsilon\}} x\, J_X(ds\times dx),$$

where by equation 2.1 the process $X^\epsilon = (X^\epsilon)_{t\geq 0}$ is a compound Poisson process with a drift term γt. By properties of Levy measure, $\nu(\mathbb{R}\backslash(-\epsilon,\epsilon)) < \infty$ and it follows that

J_X counts an almost surely finite numbers of jumps over $[0,t]\times\{\mathbb{R}\backslash(-\epsilon,\epsilon)\}$. Suppose that the sequence $(T_i,Y_i)_{i=1}^N$ described jumps of the process X up to time t, where $N=J_X([0,t]\times\{\mathbb{R}\backslash(-\epsilon,\epsilon)\})$ and $T_0=0$. Then by using telescopic sum we get

$$\begin{aligned} f(t,X_t^\epsilon) &= f(t,X_t^\epsilon)+f(0,X_0^\epsilon)-f(0,X_0^\epsilon)+f(X_N,X_{T_N}^\epsilon)-f(X_N,X_{T_N}^\epsilon) \\ &= f(0,X_0^\epsilon)+\sum_{i=1}^N(f(T_i,X_{T_i}^\epsilon)-f(T_{i-1},X_{T_{i-1}}^\epsilon))+(f(t,X_t^\epsilon)-f(T_N,X_{T_N}^\epsilon). \end{aligned}$$

Noting that X^ϵ is piece-wise linear and function f is smooth we can now decompose $f(t,X_t^\epsilon)$ by the following way

$$\begin{aligned} f(t,X_t^\epsilon) &= f(0,X_0^\epsilon)+\sum_{i=1}^N(f(T_i,X_{T_i}^\epsilon)-f(T_{i-1},X_{T_{i-1}}^\epsilon))+(f(t,X_t^\epsilon)-f(T_N,X_{T_N}^\epsilon)) \\ &= f(0,X_0^\epsilon) \\ &\quad+\sum_{i=1}^N\left(\int_{T_{i-1}}^{T_i}\left[\frac{\partial f}{\partial s}(s,X_s^\epsilon)+\gamma\frac{\partial f}{\partial s}(s,X_s^\epsilon)\right]ds+(f(T_i,X_{T_i-}^\epsilon+Y_i)-f(T_i,X_{T_i-}^\epsilon))\right) \\ &\quad+\int_{T_N}^t\left[\frac{\partial f}{\partial s}(s,X_s^\epsilon)+\gamma\frac{\partial f}{\partial x}(s,X_s^\epsilon)\right]ds \\ &= f(0,X_0^\epsilon)+\int_0^t\left[\frac{\partial f}{\partial s}(s,X_s^\epsilon)+\gamma\frac{\partial f}{\partial x}(s,X_s^\epsilon)\right]ds \\ &\quad+\iint\limits_{[0,t]\times\{\mathbb{R}\backslash\{0\}\}}(f(s,X_{s-}^\epsilon+x)-f(x,X_{s-}^\epsilon))\mathbb{1}_{(|x|\geq\epsilon)}\,J_X(ds\times dx). \end{aligned}$$

To complete the proof we have to consider limiting behaviour of the terms in above decomposition as $\epsilon\to\infty$. By Theorem 2.37 every finite variation Lévy process may be written by the difference of two subordinator. In that spirit let us decompose the process X in two terms by $X_t=X_t^{(1)}-X_t^{(2)}$, where

$$X_t^{(1)}=(\gamma\vee 0)t+\iint\limits_{[0,t]\times[0,\infty)}x\,J_X(ds\times dx),\quad t\geq 0,$$

and

$$X_t^{(2)}=|\gamma\wedge 0|t-\iint\limits_{[0,t]\times(-\infty,0)}x\,J_X(ds\times dx),\quad t\geq 0.$$

Therefore $X^{(1)}=(X_t^{(1)})_{t\geq 0}$ and $X^{(2)}=(X_t^{(2)})_{t\geq 0}$ are desired subordinators. Similarly we can now define subordinators

$$X_t^{(1,\epsilon)}=(\gamma\vee 0)t+\iint\limits_{[0,t]\times[\epsilon,\infty)}x\,J_X(ds\times dx),\quad t\geq 0,$$

and

$$X_t^{(2,\epsilon)}=|\gamma\wedge 0|t+\iint\limits_{[0,t]\times(-\infty,\epsilon)}x\,J_X(ds\times dx),\quad t\geq 0,$$

such that $X_t^\epsilon = X_t^{(1,\epsilon)} - X_t^{(2,\epsilon)}$. Hence for each fixed $t > 0$ we have that $X_t^{(1,\epsilon)} \uparrow X_t^{(1)}$, $X_t^{(2,\epsilon)} \uparrow X_t^{(2)}$ and by almost sure monotone convergence (see Theorem 4.1) it holds that

$$\begin{aligned}\lim_{\epsilon\downarrow 0} X_t^\epsilon &= \lim_{\epsilon\downarrow 0}[X_t^{(1,\epsilon)} - X_t^{(2,\epsilon)}]\\ &= X_t^{(1)} - X_t^{(2)}\\ &= X_t,\end{aligned}$$

almost surely. Note that in above equation we can replace $[0,t]$ by $[0,t)$ such that for each fixed $t > 0$, we there exists a limit $\lim_{\epsilon\downarrow 0} X_{t-}^\epsilon = X_{t-}$ almost surely.

Next we define a almost surely bounded random region $B = \{0 \le x \le |X_s^\epsilon| : s \le t \text{ and } \epsilon > 0\}$ in $\mathbb{R}$ such that

$$B \subset \{0 \le x \le X_s^{(1)} : s \le t\} \cup \{0 \ge x \ge -X_s^{(2)} : s \le t\}.$$

The latter two sets are almost surely bounded on account of right-continuity of paths of $X_s^{(1)}$ and $X_s^{(2)}$. Since we assumed that $f \in C^{1,1}([0,\infty)\times\mathbb{R})$ both $\frac{\partial f}{\partial s}$ and $\frac{\partial f}{\partial x}$ are uniformly bounded on $[0,t]\times\overline{B}$, where $\overline{B}$ is the closure of a set B. Hence by dominated convergence theorem (see Theorem 4.3) it holds that

$$\lim_{\epsilon\downarrow 0}\int_0^t \frac{\partial f}{\partial s}(s, X_s^\epsilon) + \gamma\frac{\partial f}{\partial x}(s, X_s^\epsilon)\,ds = \int_0^t \frac{\partial f}{\partial s}(s, X_s) + \gamma\frac{\partial f}{\partial x}(s, X_s)\,ds.$$

The derivative $\frac{\partial f}{\partial x}$ is uniformly bounded also in set $[0,t]\times\{x+\overline{B} : |x| \le 1\}$ and therefore by using Mean Value Theorem we can deduce the following upper bound

$$|(f(s, X_{s-}^\epsilon + x) - f(s, X_{s-}^\epsilon))\mathbb{1}_{\epsilon\le|x|<1}| \le C|x|\mathbb{1}_{(|x|<1)},$$

where C is a random variable independent of s, ϵ and x. The process X is of finite variation by assumption when

$$\iint_{[0,t]\times(-1,1)} C|x|\,J_X(ds\times dx) < \infty,$$

and therefore the function $C|x|\mathbb{1}_{(|x|<1)}$ integrates againts J_X on $[0,t]\times(-1,1)$. Similarly we can find a upperbound, when the delimiters of above integrals are replaced by $[0,t]\times\{\mathbb{R}\setminus(-1,1)\}$. We can now again apply almost sure dominated convergence to obtain that

$$\begin{aligned}&\lim_{\epsilon\downarrow 0}\iint_{[0,t]\times\mathbb{R}\setminus\{0\}} (f(s, X_{s-}^\epsilon + x) - f(x, X_{s-}^\epsilon))\mathbb{1}_{(|x|\ge\epsilon)}\,J_X(ds\times dx)\\ &= \iint_{[0,t]\times(-1,1)} \lim_{\epsilon\downarrow 0}(f(s, X_{s-}^\epsilon + x) - f(x, X_{s-}^\epsilon))\mathbb{1}_{(|x|\ge\epsilon)}\,J_X(ds\times dx)\end{aligned}$$

$$+ \iint\limits_{[0,t]\times\{\mathbb{R}\setminus(-1,1)\}} \lim_{\epsilon\downarrow 0} (f(s, X^{\epsilon}_{s-} + x) - f(x, X^{\epsilon}_{s-}))\mathbb{1}_{(|x|\geq\epsilon)}\, J_X(ds \times dx)$$

$$= \iint\limits_{[0,t]\times\mathbb{R}\setminus\{0\}} (f(s, X_{s-} + x) - f(x, X_{s-}))\, J_X(ds \times dx).$$

Finally by combining above results together we have

$$\begin{aligned} f(t, X_t) &= \lim_{\epsilon\downarrow 0} f(t, X^{\epsilon}_t) \\ &= f(0, X_0) + \int_0^t \frac{\partial f}{\partial s}(s, X_s)\, ds + \gamma \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, ds \\ &\quad + \iint\limits_{[0,t]\times\mathbb{R}} (f(s, X_{s-} + x) - f(s, X_{s-}))\, J_X(ds \times dx), \end{aligned}$$

which is the desired result.

□

4. Distribution theory

First we consider three basic theorems dealing with the limit properties of integral. Proofs for monotone convergence theorem and Fatou's Lemma and also some more properties of integral can be found in section 5.3 of [3].

4.1. **Measures and Integration.**

Theorem 4.1 (Monotone convergence theorem)**.** *Let $(X, \mathcal{F}, \mu)$ a measure space. If $(f_n)_{n=1}^{\infty}$ is a sequence of non-negative measurable functions $f_n : X \to \mathbb{R}$ that is monotone increasing and converging pointwise to f almost everywhere, then*

$$\lim_{n\to\infty} \int_X f_n(x)\, d\mu = \int_X f(x)\, d\mu.$$

Corollary 4.2 (Fatou's lemma)**.** *Let $(X, \mathcal{F}, \mu)$ a measure space. If $(f_n)_{n=1}^{\infty}$ is a sequence of non-negative measurable functions $f_n : X \to \mathbb{R}$ such that $f(x) = \liminf_{n\to\infty} f_n(x)$ almost everywhere, then*

$$\int_X f(x)\, d\mu \leq \liminf_{n\to\infty} \int_X f_n(x)\, d\mu.$$

Theorem 4.3 (Lebesgue's dominated convergence theorem)**.** *Let $(X, \mathcal{F}, \mu)$ a measure space. If $(f_n)_{n=1}^{\infty}$ is a sequence of measurable functions $f_n : X \longmapsto \mathbb{R}^d$ converging pointwise to f almost everywhere and there exists a function $g \in L^1$ such that $|f_n(x)| \leq g(x)$ almost everywhere for all $n \in \mathbb{N}$ and $x \in X$, then $f \in L^1$, and*

$$\lim_{n\to\infty} \int_X f_n(x)\, d\mu = \int_X f(x)\, d\mu.$$

Proof. First we notice that $f \in L^1$, since $|f(x)| \leq g(x)$ for almost every $x \in X$ and $g \in L^1$. Applying Fatou's lemma to the nonnegative functions $g - f_n$, we obtain that

$$\begin{aligned}\int_X g(x)\,d\mu - \int_X f(x)\,d\mu &= \int_X (g-f)(x)\,d\mu \\ &\leq \liminf_{n\to\infty} \int_X (g-f_n)(x)\,d\mu \\ &= \int_X g(x)\,d\mu - \limsup_{n\to\infty} \int_X f_n(x)\,d\mu.\end{aligned}$$

It now follows that

$$\int_X f(x)\,d\mu \geq \limsup_{n\to\infty} \int_X f_n(x)\,d\mu. \tag{4.1}$$

Similar way appylying Fatou's lemma to the nonnegative functions $g + f_n$ we obtain

$$\begin{aligned}\int_X g(x)\,d\mu + \int_X f(x)\,d\mu &= \int_X (g+f)(x)\,d\mu \\ &\leq \liminf_{n\to\infty} \int_X (g+f_n)(x)\,d\mu \\ &= \int_X g(x)\,d\mu + \liminf_{n\to\infty} \int_X f_n(x)\,d\mu.\end{aligned}$$

and

$$\int_X f(x)\,d\mu \leq \liminf_{n\to\infty} \int_X f_n(x)\,d\mu. \tag{4.2}$$

It follows from equation 4.1 and 4.2 that

$$\lim_{n\to\infty} \int_X f_n(x)\,d\mu = \int_X f(x)\,d\mu.$$

□

Definition 4.4. Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be measure spaces. If $E \subset X \times Y$, for $x \in X$ and $y \in Y$ we define the *x-section* E_x and the *y-section* E^y of E by

$$E_x = \{y \in Y : (x, y) \in E\} \quad \text{and} \quad E^y = \{x \in X : (x, y) \in E\}.$$

Also, if f is a function on $X \times Y$ we define the *x-section* f_x and the *y-section* f^y of f by

$$f_x(y) = f^y(x) = f(x, y)$$

Lemma 4.5. *Let $(X, \mathcal{M})$ be a measurable space. If $f : X \to [0, \infty]$ is measurable, there is a sequence $(\phi_n)_{n=1}^{\infty}$ of simple functions such that $0 \leq \phi_1 \leq \phi_2 \leq \cdots \leq f$, and $\phi_n \to f$ pointwise. Moreover $\phi_n \to f$ uniformly on any set on which f is bounded.*

Proof. See Section 2.1 of [6]. □

Definition 4.6. Subset $\mathcal{C}$ of the power set $\mathcal{P}(X)$ is called a *monotone class* on a space X if the following conditions holds

(i) If $E_j \in \mathcal{C}$ and $E_1 \subset E_2 \subset \ldots$, then $\bigcup_{j=1}^{\infty} E_j \in \mathcal{C}$.
(ii) If $E_j \in \mathcal{C}$ and $E_1 \supset E_2 \supset \ldots$, then $\bigcap_{j=1}^{\infty} E_j \in \mathcal{C}$.

Let $\mathcal{E} \subset \mathcal{P}(X)$. Then there is a unique smallest monotone class containing $\mathcal{E}$, called the monotone class *generated by* $\mathcal{E}$.

Lemma 4.7. *If $\mathcal{A}$ is an algebra of sets, then the monotone class $\mathcal{C}$ generated by $\mathcal{A}$ coincides with the σ-algebra $\mathcal{M}$ generated by $\mathcal{A}$.*

Proof. See section 2.5 of [6]. □

Theorem 4.8 (Theorem 2.36 of Folland)**.** *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. If $E = A \times B$, where $A \subset X$ and $B \subset Y$ and $E \in \mathcal{M} \otimes \mathcal{N}$, then the functions $f(x) = \nu(E_x)$ and $f(y) = \mu(E^y)$ are measurable on X and Y, respectively, and*

$$(\mu \times \nu)(E) = \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y). \tag{4.3}$$

Proof. Suppose that measures μ and ν are finite. Thus by Lemma 4.7 it is suffice to show that the set $\mathcal{C}$ of all $E \in \mathcal{M} \otimes \mathcal{N}$, for which conclusions of the theorem are true is monotone class generated of some algebra of sets. If $E = A \times B$, then

$$(\mu \times \nu)(E) = \iint_{X \times Y} \mathbb{1}_{A \times B}\, d(\mu \times \nu) = \int_Y \nu(B) \mathbb{1}_A(x)\, d\mu(x) = \int_X \mu(A) \mathbb{1}_B(y)\, d\nu(y),$$

where $\nu(E_x) = \nu(B)\mathbb{1}_A(x)$ and $\mu(E^y) = \mu(A)\mathbb{1}_B(y)$. Therefore $E \in \mathcal{C}$ and by additivity it follows that also the finite union $\bigcup_{n=1}^{k} E_n \in \mathcal{C}$ so $\mathcal{C}$ is a algebra of sets. If $(E_n)_{n=1}^{\infty}$ is an increasing sequence in $\mathcal{C}$ and $E = \bigcup_{n=1}^{\infty}$, then the each of functions $f_n(y) = \mu((E_n)^y)$ and $f_n(x) = \nu((E_n)_x)$ are measurable and $f_n(y) \longrightarrow f(y)$ and $f_n(x) \longrightarrow f(x)$ pointwise. Hence f is measurable and by monotone convergence theorem,

$$(\mu \times \nu)(E) = \lim_{n \to \infty} (\mu \times \nu)(E_n) = \lim_{n \to \infty} \int_X \mu((E_n)^y)\, d\nu(y) = \int_X \mu(E^y)\, d\nu(y).$$

Similary

$$(\mu \times \nu)(E) = \lim_{n \to \infty} (\mu \times \nu)(E_n) = \lim_{n \to \infty} \int_Y \nu((E_n)_x)\, d\mu(x) = \int_Y \nu(E_x)\, d\mu(x),$$

so $E \in \mathcal{C}$.

Next if $(E_n)_{n=1}^{\infty}$ is decreasing sequence in $\mathcal{C}$ and $E = \bigcup_{n=1}^{\infty} E_n$. By finitness of measures μ and ν we obtain that $\mu((E_1)^y) \leq \mu(X) < \infty$ and $\mu((E_1)_x) \leq \nu(Y) < \infty$, so by the dominated convergence theorem

$$(\mu \times \nu)(E) = \int_X \mu(E^y)\, d\nu(y) \quad \text{and} \quad (\mu \times \nu)(E) = \int_Y \nu(E_x)\, d\mu(x).$$

Thus $E \in \mathcal{C}$ and the theorem holds in finite measures spaces.

Finally we suppose that μ and ν are σ-finite measures, when $X \times Y$ can be written as the union of an increasing sequence $(X_j \times Y_j)_{j=1}^{\infty}$ of rectangles of finite measure. If $E \in \mathcal{M} \otimes \mathcal{N}$, for each j the above result for finite measures gives us

$$\begin{aligned}(\mu\times\nu)(E\cap(X_j\times Y_j)) &= \int \mathbb{1}_{X_j}(x)\nu(E_x\cap Y_j)\,d\mu(x)\\ &= \int \mathbb{1}_{Y_j}(x)\mu(E^y\cap X_j)\,d\nu(x).\end{aligned}$$

By the monotone convergence theorem

$$\begin{aligned}(\mu\times\nu)(E) &= \lim_{j\to\infty}(\mu\times\nu)(E\cap(X_j\times Y_j))\\ &= \lim_{j\to\infty}\int \mathbb{1}_{X_j}(x)\nu(E_x\cap Y_j)\,d\mu(x)\\ &= \int \mathbb{1}_{X}(x)\nu(E_x\cap Y)\,d\mu(x)\\ &= \int_X \nu(E_x)\,d\mu(x).\end{aligned}$$

Similary

$$(\mu\times\nu)(E) = \int_Y \mu(E^y)\,d\nu(y),$$

and the desired result follows. □

Theorem 4.9 (Fubini-Tonelli Theorem; Theorem 2.37 of Folland)**.** *Let L^+ be the space of all measurable functions from X to $[0,\infty]$ and suppose that $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are σ-finite measure spaces.*

(i) *(Tonelli) If $f\in L^+(X\times Y)$, then the functions $g(x)=\int_Y f_x\,d\nu(y)$ and $h(y)=\int_X f^y\,d\mu(x)$ are in $L^+(X)$ and $L^+(Y)$, respectively, and*

$$\begin{aligned}\int_{X\times Y} f(x,y)d(\mu\times\nu) &= \int_X\left[\int_Y f(x,y)d\nu(y)\right]d\mu(x)\\ &= \int_Y\left[\int_X f(x,y)d\mu(x)\right]d\nu(y).\end{aligned}\tag{4.4}$$

(ii) *(Fubini) If $f\in L^1(\mu\times\nu)$, then $f_x\in L^1(\nu)$ for almost every $x\in X$, $f^y\in L^1(\mu)$ for almost every $y\in Y$, the almost everywhere-defined functions $g(x)=\int f_x\,d\nu$ and $h(x)=\int f^y\,d\nu$ are in $L^1(\mu)$ and $L^1(\nu)$, respectively, and equation* 4.4 *holds.*

Proof. (i) In the case f is indicator function, let $E=A\times B$, where $A\subset X$ and $B\subset Y$. Then by Theorem 4.8

$$\begin{aligned}\int_{X\times Y} \mathbb{1}_E \, d(\mu\times\nu) &= \int_E d(\mu\times\nu)\\ &= (\mu\times\nu)(E)\\ &= \int_X \mathbb{1}_A(x)\nu(B)\, d\mu(x)\\ &= \int_Y \mathbb{1}_B(y)\mu(A)\, d\nu(y)),\end{aligned}$$

thus

$$\begin{aligned}\int_{X\times Y} \mathbb{1}_E \, d(\mu\times\nu) &= \int_X \left[\int_Y \mathbb{1}_{A\times B}\, d\nu(y)\right] d\mu(x)\\ &= \int_Y \left[\int_X \mathbb{1}_{A\times B}\, d\mu(x)\right] d\nu(y).\end{aligned}$$

Therefore Tonelli's Theorem holds for nonnegative simple functions by linearity.

If $f \in L^+(X\times Y)$, let $(f_n)_{n=1}^{\infty}$ be a sequence of simple functions which incresase pointwise to f as in Lemma 4.5. The monotone convergence theorem implies first that

$$g_n(x) = \int_{X\times Y} f_n(x,y)\, d\nu(y) \xrightarrow{n\to\infty} \int_{X\times Y} f(x,y)\, d\nu(y) = g(x)$$

and

$$h_n(y) = \int_{X\times Y} f_n(x,y)\, d\mu(x) \xrightarrow{n\to\infty} \int_{X\times Y} f(x,y)\, d\mu(x) = h(y),$$

respectively so that g and h are measurable. By the monotone convergence theorem we can also obtain that

$$\begin{aligned}\int_{X\times Y} f(x,y)\, d(\mu\times\nu) &= \lim_{n\to\infty}\int_{X\times Y} f_n(x,y)\, d(\mu\times\nu)\\ &= \lim_{n\to\infty}\int_X g_n(x)\, d\mu(x)\\ &= \int_X g(x)\, d\mu(x)\\ &= \int_X \left[\int_Y f(x,y)\, d\nu(y)\right] d\mu(x)\end{aligned}$$

and

$$\begin{aligned}
\int_{X\times Y} f(x,y)d(\mu\times\nu) &= \lim_{n\to\infty}\int_{X\times Y} f_n(x,y)\,d(\mu\times\nu)\\
&= \lim_{n\to\infty}\int_Y h_n(y)\,d\nu(y)\\
&= \int_Y h(y)\,d\nu(y)\\
&= \int_Y\left[\int_X f(x,y)\,d\mu(x)\right]\,d\nu(y),
\end{aligned}$$

which proves Tonelli's theorem.

(ii) If $f\in L^+(X\times Y)$ and $\int_{X\times Y} f(x,y)d(\mu\times\nu)<\infty$, then by last two equations above $g<\infty$ and $h<\infty$ almost everywhere. Therefore $f_x\in L^1(\nu)$ for almost every x and $f^y\in L^1(\mu)$ for almost every y. If $f\in L^1(\mu\times\nu)$ then by applying these results to the positive and negative part of real and imaginary parts of f it follows that $g\in L^1(\mu)$ and $h\in L^1(\nu)$ and 4.4 holds.

□

4.2. **Finite variation of function.**

Definition 4.10. Let $U\subset\mathbb{R}^d$ be a open set and f continuous function. The *support* of function f, denoted by $\operatorname{supp}(f)$, is the smallest closed set outside of which f vanishes, that is closure of $\{x : f(x)\neq 0\}$. If $\operatorname{supp}(f)$. is compact, we say that f is *compactly supported*, and define

$$C_c(U)=\{f\in C(U):\operatorname{supp}(f)\text{ is compact}\}.$$

Morover, if $f\in C(U)$, we say that function f *vanishes at infinity* if for every $\epsilon>0$ the set $\{x:|f(x)|\geq\epsilon\}$ is compact, and define

$$C_0(U)=\{f\in C(U): f\text{ vanishes at infinity}\}.$$

Since for $f\in C_0(U)$ the image of the set $\{x:|f(x)|\geq\epsilon\}$ is compact, and $|f|<\epsilon$ on its complement, $C_c(U)\subset C_0(U)$. We also define,

$$C_c^\infty(U)=C^\infty(U)\cap C_c(U)=\cap_{k=1}^\infty C^k(U)\cap C_c(U)$$

where $C^k(U)$ is the space of functions k times continuously differentiable on U.

Theorem 4.11. *The space $C_c(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$ for any $p\in[1,\infty)$.*

Proof. See Section 4.3 of [2] □

Definition 4.12. Let $(X,\mathcal{F},\mu)$ be an arbitrary measure space. A measure ν on $(X,\mathcal{F})$ is said to be *absolutely continuous* with respect to μ if for any set $A\in\mathcal{F}$, $\nu(A)=0$ whenever $\mu(A)=0$. We then write $\nu\ll\mu$. A function $f:\mathbb{R}\to\mathbb{C}$ is called *absolutely continuous* if for every $\epsilon>0$ there exists $\delta>0$ such that for any finite set of disjoint intervals $(a_1,b_1),\ldots,(a_n,b_n)$,

$$\sum_{i=1}^n (b_j-a_j)<\delta \quad\text{implies}\quad \sum_{i=1}^n |f(b_j)-f(a_j)|<\epsilon. \tag{4.5}$$

More generally, f is said to be *absolutely continuous* on $[a, b]$ if this condition holds when the intervals (a_j, b_j) are restricted to lie in $[a, b]$.

Remark 4.13. If $n = 1$ in equation 4.5, it holds that

$$(b_j - a_j) < \delta \quad \text{implies} \quad |f(b_j) - f(a_j)| < \epsilon,$$

and so we obtain that f is uniformly continuous if it is absolutely continuous.

Recall that we call function f to be of finite variation if its total variation denoted by $V(f)$ is finite. Next we consider a few important result, using the finite variation of a function, that we are going to eventually use to proof the fundamental theorem of Lebesgue integral calculus.

Proposition 4.14. *Suppose that total variation of f on $[a, b]$ is finite such that $f(-\infty) = 0$ and f is right-continuous. Then f is absolutely continuous if and only if $\mu_f \ll m$, where μ_f is a unique complex measure such that $F(x) = \mu_f((-\infty, x])$ and $|\mu_f| = \mu_{V(f)}$.*

Corollary 4.15. *If $f \in L^1(m)$, then the total variation of function $F(x) = \int_{-\infty}^{x} f(t)\, dt$ on $[a, b]$ is finite such that $F(-\infty) = 0$, F is absolutely continuous and $f = F'$ almost everywhere. Conversely, if the total variation of function f is finite such that $f(-\infty) = 0$ and and F is absolutely continuous, then $F' \in L^1(m)$ and $F(x) = \int_{-\infty}^{x} F'(t)\, dt$.*

Proof of Proposition 4.14 and Corollary 4.15 can be found in section 3.5 of [6].

Lemma 4.16. *If f is absolutely continuous on $[a, b]$, then its total variation on $[a, b]$ is finite.*

Proof. Let δ be as in the definition 4.12. If $a = x_0 < \cdots < x_n = b$, we collect the intervals (x_{j-1}, x_j) into at most N groups, where $N < \frac{b-a}{\delta} + 1$ and the sum of the lengths in each group is less than δ. By choosing $\epsilon = 1$ we get

$$V(f) \leq \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| < N\epsilon = N,$$

and it follows that the total variation of f on $[a, b]$ is finite. □

Theorem 4.17 (Fundamental theorem of Lebesgue integral calculus)**.** *If $-\infty < a < b < \infty$ and $F : [a, b] \to \mathbb{C}$, the following are equivalent:*

(i) *F is absolutely continuous on $[a, b]$.*
(ii) *$F(x) - F(a) = \int_{a}^{x} f(t)\, dt$ for some $f \in L^1([a, b], m)$.*
(iii) *F is differentiable almost everywhere on $[a, b]$, $F' \in L^1([a, b], m)$, and*

$$F(x) - F(a) = \int_{a}^{x} F'(t)\, dt.$$

Proof. Assume first that F is absolutely continuous on $[a, b]$ and $F(a) = 0$. If we set

$$\widetilde{F} = \begin{cases} 0, & x < a \\ F(x), & a \leq x \leq b \\ F(b), & x > b \end{cases}$$

then $\widetilde{F}$ is of finite variation by Lemma 4.16. Since $\widetilde{F}(-\infty) = 0$ and the function is right-continuous, (iii) holds by Corollary 4.15. By choosing $f(t) = F'(t) \in$

$L^1([a,b],m)$ we get $F(x)-F(a)=\int_a^b F'(t)dt=\int_a^x f(t)dt$ and thus (ii) holds. Next let f be a function such that $F(x)-F(a)=\int_a^x f(x)\,dt$ and $f\in L^1([a,b],m)$. We can define a extension $\tilde{f}$ of function f by

$$\tilde{f}=\begin{cases} f, & \text{if } t\in[a,b]\\ 0, & \text{if } t\notin[a,b].\end{cases}$$

Clearly $\tilde{f}\in L^1$ and

$$F(x)=\int_{-\infty}^{x}\tilde{f}(t)d\,t=\int_a^x f(t)dt$$

is absolutely continuous by Corollary 4.15 and the desired result follows.

□

4.3. **Lebesgue differentiation Theorem.**

Definition 4.18. A point $x\in U\subset\mathbb{R}^d$ is a *Lebesgue point* of a function $f:U\longmapsto\mathbb{R}$ if

$$\lim_{r\to 0^+}\frac{1}{m((B(x,r))}\int_{B(x,r)}|f(y)-f(x)|\,dy=0$$

where $B(x,r)=\{y\in\mathbb{R}^d:|y-x|<r\}$ and the limit is taken for those r small enough to guarantee that $B(x,r)$ is a subset of U.

Definition 4.19. The set of all Lebesgue points of $f:U\longmapsto\mathbb{R}$ is denoted by L_f and it is called the *Lebesgue set*.

Definition 4.20. Let $U\subset\mathbb{R}^d$ be a open set and $1\le p\le\infty$. We say that a function $f:U\to\mathbb{R}$ is *locally integrable* with respect to Lebesgue measure and denoted it $f\in L^p_{loc}(U)$ if $f\mathbb{1}_K\in L^p(U)$ for every compact set K contained in U. Note that if $f\in L^p_{loc}(U)$, then $f\in L^1_{loc}(U)$.

We start the concept of the Lebesgue diferentiation theorem by defining avarage value of function and useful sharper version of the fundamental differentiation theorem. Note that following theorem does not yet utilize Lebesgue point or set of functions.

Theorem 4.21. *Let $f\in L^1_{loc}$, $x\in\mathbb{R}^d$, $r>0$ and let $A_rf(x)$ to be the avarage value of f on $B(x,r)$ such that*

$$A_rf(x)=\frac{1}{m(B(x,r))}\int_{B(x,r)}f(y)\,dy.$$

Then $\lim_{r\to 0}A_rf(x)=f(x)$ for almost every $x\in\mathbb{R}^d$.

Proof. See section 3.4 of [6]. □

We can however be more general by replacing balls with the families $\{E_r\}_{r>0}$ of Borel subset of $\mathbb{R}^d$ which are shrink nicely. This leads to the final version of the differentiation theorem and one that we desire, called Lebesgue differentiation theorem.

Definition 4.22. A family $\{E_r\}_{r>0}$ of a Borel subsets of U is said to *shrink nicely* to $x\in U$ if the following two conditions hold

(i) $E_r \subset B(x,r) \subset U$ for each r.
(ii) There is a constant $\alpha > 0$, independent of r, such that $m(E_r) > \alpha m(B(x,r))$.

Theorem 4.23 (The Lebesgue differentiation theorem)**.** *Suppose that $f \in L^1_{loc}(U)$ and $supp(f) \subset U$. Then we have*

(i) $m(U \setminus L_f) = 0$
(ii) *For every x in the Lebesgue set of f, in particular for almost every x in U, we have*

$$\lim_{r\to 0^+} \frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)|\, dy = 0$$

and

$$\lim_{r\to 0^+} \frac{1}{m(E_r)} \int_{E_r} f(y)\, dy = f(x),$$

where $\{E_r\}_{r>0}$ is a family of Borel subsets of $U \subset \mathbb{R}^d$ that shrinks nicely to x

Proof. We first proof the case $U = \mathbb{R}^d$.

(i) Let $c \in \mathbb{C}$ and $g(x) = |f(x) - c|$. Therefore by Theorem 4.21 we have

$$\lim_{r\to 0} A_r g(x) = \lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - c|\, dy = |f(x) - c|.$$

Next let $R \subset \mathbb{C}$ be a countable dense set. If we define $E = \bigcup_{c\in R} E_c$, where $m(E_c) = 0$, then $m(E) = 0$ and if $x \notin E$ we can choose $c \in R$ such that $|f(x) - c| < \epsilon$. Thus

$$|f(y) - f(x)| \leq |f(y) - c| + |f(x) - c| < |f(y) - c| + \epsilon.$$

Therefore

$$\begin{aligned}
&\limsup_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy \\
&\leq \limsup_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - c| + \epsilon\, dy \\
&= |f(x) - c| + \epsilon \\
&< 2\epsilon,
\end{aligned}$$

and since ϵ is arbitrary it follows that $m((L_f)^c) = 0$.

(ii) By Definition 4.22 $E_r \subset B(x,r)$ and for some $\alpha > 0$ we have $m(E_r) > \alpha m(B(x,r))$. Therefore it holds

$$\begin{aligned}
\frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)|\, dy &\leq \frac{1}{m(E_r)} \int_{B(x,r)} |f(y) - f(x)|\, dy \\
&< \frac{1}{\alpha m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy \\
&< 2\epsilon,
\end{aligned}$$

for some $\alpha > 0$. By (i) of Theorem 4.23, taking limit both sides of above inequality as $r \to 0^+$ the desired result follows.

To proof the latter part of the (ii) we obtain that by Theorem 4.21 for almost every x it holds that

$$\lim_{r\to\infty} \frac{1}{m(B(x,r))} \int_{B(x,r)} [f(x) - f(y)]\, dy = 0.$$

Therefore by using again Definition 4.22 and above equation we have

$$\begin{aligned}\lim_{r\to 0}\left[\frac{1}{m(E_r)}\int_{E_r} f(y)\,dy\right] - f(x) &= \lim_{r\to 0}\left[\frac{1}{m(E_r)}\int_{E_r} f(y)\,dy - \frac{1}{m(E_r)}\int_{E_r} f(x)\,dy\right]\\ &= \lim_{r\to 0}\left[\frac{1}{m(E_r)}\int_{E_r} f(y) - f(x)\,dy\right]\\ &\le \lim_{r\to 0}\left[\frac{1}{\alpha m(B(y,r))}\int_{B(y,r)} f(y) - f(x)\,dy\right]\\ &= 0,\end{aligned}$$

for almost every x. This proves the latter part of the theorem. Next if we suppose that $f \in L^1_{loc}(U)$ and $\mathrm{supp}(f) \subset U$ then $f\mathbb{1}_U \in L^1_{loc}(\mathbb{R}^d)$ and $\mathrm{supp}(f\mathbb{1}_U) \subset \mathbb{R}^d$. Therefore the Lebesgue differentiation theorem in the special case $U = \mathbb{R}^d$ apply and we have desired generalization to open set U. □

Following lemma gives us one possible way to determine the Lebesgue point of a function.

Lemma 4.24. *If $f \in L^1_{loc}(U)$, $U \subset \mathbb{R}^d$, and f is continuous at $x \in U$, then $x \in L_f$.*

Proof. By continuity of f we can find for every given $\epsilon > 0$ a ball with radius $r > 0$ such that $|y - x| < r$ implies $|f(y) - f(x)| < \epsilon$. Therefore

$$\frac{1}{m(B(x,r))}\int_{B(x,r)} |f(y) - f(x)\,|\,dy < \frac{\epsilon}{m(B(x,r))}\int_{B(x,r)} dy = \epsilon.$$

Since ϵ is arbitrary

$$\lim_{r\to 0^+}\frac{1}{m(B(x,r))}\int_{B(x,r)} |f(y) - f(x)|\,dy = 0,$$

and x is a Lebesgue point of function f. □

4.4. **Convolution and mollifier.**

Definition 4.25. If $f : \mathbb{R}^d \longmapsto \mathbb{R}$ and $g : \mathbb{R}^d \longmapsto \mathbb{R}$, are measurable functions, then the *convolution* $f * g : \mathbb{R}^d \longmapsto \mathbb{R}$ is defined by $(f * g)(x) := \int_{\mathbb{R}^d} f(x-y)g(y)\,dy$, provided that for every $x \in \mathbb{R}^d$, the integral is well defined.

The following proposition introduce us some basic properties of convolution.

Proposition 4.26. *Assuming that all integrals in question exist, we have*

(i) $f * g = g * f$
(ii) $(f * g) * h = f * (g * h)$

(iii) $(f * g)(x - z) = f(x - z) * g = f * g(x - z)$

Proof. (i) Let $z = x - y$, then we have

$$\begin{aligned}(f * g)(x) &= \int_{\mathbb{R}^d} f(x-y)g(y)\,dy \\ &= \int_{\mathbb{R}^d} f(z)g(x-z)\,dz \\ &= (g * f)(x).\end{aligned}$$

(ii) By using (i) and Fubini theorem we get

$$\begin{aligned}(f * g) * h &= \int_{\mathbb{R}^d} (g * f)(x-y)h(y)\,dy \\ &= \iint_{\mathbb{R}^d} g(x-y-z)f(z)h(y)\,dz\,dy \\ &= \iint_{\mathbb{R}^d} g(x-z-y)h(y)f(z)\,dy\,dz \\ &= \int_{\mathbb{R}^d} (g * h)(x-z)f(z)\,dz \\ &= f * (g * h).\end{aligned}$$

(iii) We note that

$$(f * g)(x - z) = \int_{\mathbb{R}^d} f(x-z-y)g(y)\,dy = f(x-z) * g$$

and by (i)

$$(f * g)(x - z) = \int_{\mathbb{R}^d} f(y)g(x-z-y)\,dy = f * g(x-z).$$

□

Lemma 4.27. *Let $f \in L^p_{loc}(U)$, $p \geq 1$, and $supp(f) \subset U$. Suppose that $g : \mathbb{R}^d \longmapsto \mathbb{R}$ is bounded and compactly supported. Then $g * f\mathbb{1}_U$, where*

$$f\mathbb{1}_U(x) = \begin{cases} f(x), & x \in U \\ 0, & x \notin U \end{cases}$$

is well defined on $\mathbb{R}$, i.e. the integral $\int_U g(x-y)f(y)\,dy$ is finite for all $x \in \mathbb{R}^d$.

Proof. The function g is bounded so there exits a constant $C < \infty$ such that $|g(x)| \leq C$. By choosing $p = 1$ and $f \in L^1_{loc}(U)$ we obtain that for all $x \in \mathbb{R}^d$

$$\begin{aligned}|(g * f\mathbb{1}_U)(x)| &= \left|\int_{\mathbb{R}^d} g(x-y)f\mathbb{1}_U(y)\,dy\right| \\ &\leq \int_U |g(x-y)||f(y)|\,dy \\ &\leq C\int_U |f(y)|\,dy \\ &< \infty.\end{aligned}$$

Therefore $g * f\mathbb{1}_U$ is well defined. □

Proposition 4.28. *Let $f \in L^1(\mathbb{R})$ and $g \in L^p(\mathbb{R}^d)$ with $1 \le p \le \infty$. Then*

$$\mathrm{supp}(f * g) \subset \overline{\mathrm{supp}(f) + \mathrm{supp}(g)}.$$

Proof. Fix $x \in \mathbb{R}^d$ such that the function $y \longmapsto f(x-y)g(y)$ is integrable. If $x \notin \mathrm{supp}(f) + \mathrm{supp}(g)$, then $(x - \mathrm{supp}(f)) \cap \mathrm{supp}(g) = \emptyset$ and so

$$\begin{aligned}(f * g)(x) &= \int_{\mathbb{R}^d} f(x-y)g(y)\,dy \\ &= \int_{(x-\mathrm{supp}(f))\cap\mathrm{supp}(g)} f(x-y)g(y)\,dy = 0.\end{aligned}$$

Thus $(f * g)(x) = 0$ almost everywhere on $\mathrm{int}[(\mathrm{supp}(f) + \mathrm{supp}(g))^c)]$ and therefore

$$\mathrm{supp}(f * g) \subset \overline{\mathrm{supp}(f) + \mathrm{supp}(g)}.$$

□

Theorem 4.29 (Young)**.** *Let $f \in L^1(\mathbb{R}^d)$ and let $g \in L^p(\mathbb{R}^d)$ with $1 \le p \le \infty$. Then for almost every x, $f * g \in L^p(\mathbb{R}^d)$ and*

$$||f * g||_p \le ||f||_1 ||g||_p$$

Proof. We can consider three cases;

(i) $p = 1$
(ii) $1 < p < \infty$
(iii) $p = \infty$.

Case (*i*): For almost everyone $y \in \mathbb{R}^d$ we have

$$\int_{\mathbb{R}^d} |f(x-y)g(y)|\,dx = |g(y)| \int_{\mathbb{R}^d} |f(x-y)|\,dx = |g(y)|||f||_1 < \infty$$

and

$$\int_{\mathbb{R}^d} dy \int_{\mathbb{R}^d} |f(x-y)g(y)|\,dx = \int_{\mathbb{R}^d} |g(y)| dy \int_{\mathbb{R}^d} |f(x-y)|\,dx = ||g||_1 ||f||_1 < \infty.$$

Therefore by setting $F(x,y) = f(x,y)g(y)$ we deduce from Tonelli's theorem and Fubini's theorem for measurable functions from $X \times Y$ to $\mathbb{R}$, given in section 4.1 of [2] that $F \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$ and

$$\int_{\mathbb{R}^d} |f(x-y)g(y)|\,dy < \infty \quad \text{for a.e. } x \in \mathbb{R}^d$$

respectively. Thus $(f * g)(x) \in L^p(\mathbb{R}^d)$ and moreover

$$\begin{aligned}||(f * g)(x)||_1 &\le \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(x-y)||g(y)|\,dy\,dx \\ &= \int_{\mathbb{R}^d} |g(y)| dy \int_{\mathbb{R}^d} |f(x-y)|\,dx \\ &= ||g||_1 ||f||_1\end{aligned}$$

Case (*ii*) Let $p, q \in \mathbb{R}$ be the conjugate exponent such that $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. By *Case* (*i*) we know that for almost every fixed $x \in \mathbb{R}^d$

$$|f(x-y)|^{1/p}|g(y)| \in L^p_y(\mathbb{R}^d) \quad \text{and} \quad |f(x-y)|^{1/q} \in L^q_y(\mathbb{R}^d).$$

Therefore by Hölder inequality

$$|f(x-y)||g(y)| = |f(x-y)|^{1/q}|f(x-y)|^{1/p}|g(y)| \in L^1_y(\mathbb{R}^d)$$

and

$$\begin{aligned} |(f*g)(x)| &\leq \int_{\mathbb{R}^d} |f(x-y)||g(y)|\,dy \\ &= \int_{\mathbb{R}^d} |f(x-y)|^{1/q}|f(x-y)|^{1/p}|g(y)|\,dy \\ &\leq ||f||_1^{1/q}\left(\int_{\mathbb{R}^d} |f(x-y)||g(y)|^p\,dy\right)^{1/p} \\ &= ||f||_1^{1/q}(|f| * |g|^p)^{1/p}. \end{aligned}$$

By using Young's inequality in case $p = 1$ to the function $|||f| * |g|^p||_1$ we obtain that $f * g \in L^p(\mathbb{R}^d)$ and

$$\begin{aligned} ||f*g||_p &= \left(\int_{\mathbb{R}^d} |(f*g)(x)|^p\,dx\right)^{1/p} \\ &\leq ||f||_1^{p/p}\,||\,|f| * |g|^p||_1 \\ &\leq ||f||_1^{p/q}\,||f||_1^{1/p}\,||g||_p \\ &= ||f||_1||g||_p \end{aligned}$$

Case (*iii*) Let $p = \infty$. Therefore when $g \in L^\infty(\mathbb{R}^d)$ we have

$$|g(y)| \leq ||g||_\infty,$$

almost everywhere in $\mathbb{R}^d$. By using above inequality we obtain that

$$\begin{aligned} |f*g| &\leq \int_{\mathbb{R}^d} |f(x-y)||g(y)|\,dy \\ &\leq \int_{\mathbb{R}^d} |f(x-y)|||g||_\infty\,dy \\ &\leq ||g||_\infty \int_{\mathbb{R}^d} |f(x-y)|\,dy \\ &\leq ||f||_1||g||_\infty, \end{aligned}$$

and therefore

$$||f*g||_\infty \leq ||f||_1||g||_\infty.$$

□

Definition 4.30. (i) Let η be any function $\eta : \mathbb{R}^d \to \mathbb{R} \in C_c^\infty(\mathbb{R}^d)$ such that it satisfies the following conditions

$$\eta \geq 0, \quad \int \eta(x) d\,x = 1, \quad \text{supp}(\eta) = \overline{B(0,1)}.$$

Then η is called *a mollifier.*

(ii) Let x be a point on $\mathbb{R}^d$ and $\epsilon > 0$. We define a *modifed mollifier* η^ϵ by

$$\eta^\epsilon(x) = \frac{1}{\epsilon^d}\eta\left(\frac{x}{\epsilon}\right),$$

when we clearly have

$$\eta^\epsilon \in C_c^\infty(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} \eta^\epsilon(x) d\,x = 1, \quad \text{supp}(\eta^\epsilon) = \overline{B(0,\epsilon)}.$$

(iii) Let

$$\eta(x) = \begin{cases} ce^{\frac{-1}{1-||x||_d^2}}, & ||x||_d < 1 \\ 0, & ||x||_d \geq 1 \end{cases}$$

and take c such that $\int_{\mathbb{R}^d} \eta(x) d\,x = 1$. Then η^ϵ is called the *standard mollifier.* Note that standard mollifier is a example of modifed mollifier.

(iv) Suppose that $f \in L^p_{loc}(U)$, $p \geq 1$ and $\epsilon > 0$. We define a *mollified function* $f^\epsilon : \mathbb{R}^d \to \mathbb{R}$ by

$$f^\epsilon(x) := (\eta^\epsilon * f\mathbb{1}_U)(x) = \int_U \eta^\epsilon(x-y) f(y)\, dy$$

For fixed x and ϵ small enough, $\overline{B(x,\epsilon)} \subset U$ and so $f^\epsilon(x)$ exists. However, if $\text{supp}(f) \subset U$ and since $f \in L^p_{loc}(U)$, $p \geq 1$, by Lemma 4.27, f^ϵ is well defined on $\mathbb{R}^d$ for all $\epsilon > 0$.

Theorem 4.31. *Assume that $f \in L^p_{loc}(U)$, $p \geq 1$, $supp(f) \subset U$ and $\epsilon > 0$. Then*

(i) *$f^\epsilon \in C^\infty(\mathbb{R}^d)$ and $\partial^\alpha f^\epsilon = \partial^\alpha \eta^\epsilon * f\mathbb{1}_U$,*
(ii) *$f^\epsilon \longrightarrow f\mathbb{1}_U$ uniformly as $\epsilon \to 0^+$,*
(iii) *$f^\epsilon \longrightarrow f\mathbb{1}_U$ in $L^p_{loc}(\mathbb{R}^d)$ as $\epsilon \to 0^+$,*
(iv) *$f^\epsilon \longrightarrow f\mathbb{1}_U$ pointwise on L_{f1_U} as $\epsilon \to 0^+$, hence $f^\epsilon \longrightarrow f$ pointwise on L_f as $\epsilon \to 0^+$*

Proof. For simplicity we proof this theorem first in the special case $U = \mathbb{R}^d$ by supposing that $f \in L^p_{loc}(\mathbb{R}^d)$ and $supp(f) \subset U$. After that the generalization to an open set $U \subset \mathbb{R}^d$ follows easily.

(i) Let K be a fixed compact set in $\mathbb{R}^d$ large enough that $x + B(0,1) - supp(\eta^\epsilon) \subset K$. We have for all $y \notin K$ and $h \in B(0,1)$.

$$\eta^\epsilon(x+h-y) - \eta^\epsilon(x-y) - h \cdot \partial\eta^\epsilon(x-y) = 0$$

$$
\begin{aligned}
&|(\eta^\epsilon * f)(x+h) - \eta^\epsilon * f)(x) - h(\partial\eta^\epsilon * f)(x)| \\
&\leq \int_{\mathbb{R}^d} |f(y)||\eta^\epsilon(x+h-y) - \eta^\epsilon(x-y) - h\eta^\epsilon(x-y)|dy \\
&= \int_{\mathbb{R}^d} |f(y)| \left| \int_0^1 h\cdot\partial\eta^\epsilon(x+sh-y) - h\cdot\eta^\epsilon(x-y)ds \right| dy \\
&\leq \int_{\mathbb{R}^d} |f(y)| \int_0^1 |h|| \cdot \partial\eta^\epsilon(x+sh-y) - h\cdot\eta^\epsilon(x-y)|ds\, dy \\
&\leq \int_{\mathbb{R}^d} |f(y)|(|h|\epsilon(|h|)\mathbb{1}_K(y) + |h|\epsilon(|h|)\mathbb{1}_{K^c}(y))dy \\
&\leq |h|\epsilon(|h|) \int_{\mathbb{R}^d} |f(y)|dy
\end{aligned}
$$

It follows that f^ϵ is diffentiable, $\partial f^\epsilon = \partial\eta^\epsilon * f$ and inductively for $\alpha \leq 1$ that

$$
f^\epsilon \in C^\infty(\mathbb{R}^d) \quad \text{and} \quad \partial^\alpha f^\epsilon = \partial^\alpha \eta^\epsilon * f\mathbb{1}_U.
$$

(ii) Let $f_n = (\eta^{\frac{1}{n}} * f)_{n=1}^\infty$ be a sequence of mollified functions converging to f as $n \to \infty$ and let $K \subset \mathbb{R}^d$ be fixed compact set. For $\epsilon > 0$ there exists $\delta > 0$ such that

$$
|f(x-y) - f(x)| < \epsilon
$$

for all $x \in K$ and $y \in B(0,\delta)$. Therefore for $x \in \mathbb{R}^d$, $n > 1/\delta$ and $x \in K$ we have

$$
\begin{aligned}
|f_n - f| &= \left| \int_{\mathbb{R}^d} f(x-y)\eta^{1/n}(y)\, dy - \int_{\mathbb{R}^d} f(x)\,\eta^{1/n}(y)\, dy \right| \\
&= \left| \int_{\mathbb{R}^d} [f(x-y) - f(x)]\,\eta^{1/n}(y)\, dy \right| \\
&\leq \int_{B(0,\frac{1}{n})} |f(x-y) - f(x)|\, |\eta^{1/n}(y)|\, dy \\
&< \epsilon \int\limits_{B(0,\frac{1}{n})} \eta^{1/n}\, dy \\
&= \epsilon,
\end{aligned}
$$

and the desired result follows.

(iii) Let $f_n = (\eta^{\frac{1}{n}} * f)_{n=1}^\infty$ be a sequence of mollified functions converging to f as $n \to \infty$. By densiness of $C_c(\mathbb{R}^d)$ in $L^p(\mathbb{R}^d)$ we can choose a fixed function $g \in C_c(\mathbb{R}^d)$ such that $||f-g||_p < \epsilon$. By Proposition 4.28 we know that

$$
\operatorname{supp}(g_n) = \operatorname{supp}(\eta^{1/n} * g) \subset \overline{B(0,1/n) + \operatorname{supp}(g)} \subset \overline{B(0,1) + \operatorname{supp}(g)},
$$

which is a fixed compact set and therefore by (ii) of Theorem 4.31

$$
||f_n - g||_p \xrightarrow{n\to\infty} 0.
$$

Next by using Young's inequality we conclude that

$$\begin{aligned}||f_n - g||_p &= ||[(\eta^{1/n} * g)] - [(\eta^{1/n} * g) - g] - [g - f]||_p \\ &\le ||(\eta^{1/n} * g)||_p + ||(\eta^{1/n} * g) - g||_p + ||g - f||_p \\ &\le 2||f - g||_p + ||(\eta^{1/n} * g) - g||_p.\end{aligned}$$

Letting $n \to \infty$ we obtain that for $\epsilon > 0$

$$\limsup_{n\to\infty} ||f_n - f||_p \le 2\epsilon.$$

Therefore

$$\lim_{n\to\infty} ||f_n - f||_p = 0.$$

(iv)

$$\lim_{\epsilon\to 0^+} f^\epsilon(x) = f(x)$$

Therefore by the Lebesgue differentiation Theorem $f^\epsilon \to f$, when $r \to 0^+$.

Finally we suppose that $f \in L^p_{loc}(U)$, and $\mathrm{supp}(f) \subset U$, when clearly $f\mathbb{1}_U \in L^p_{loc}(\mathbb{R}^d)$ and $\mathrm{supp}(f\mathbb{1}_U) \subset U$. Thus above results apply and we get desired generalization.

□

4.5. **Weak derivatives.**

Definition 4.32. Suppose that $\alpha \in \mathbb{N}_0^d$ is a multi-index. We say that a function $f \in L^1_{loc}(U)$, $U \subset \mathbb{R}^d$, is *weakly differentiable*; and also its α th-weak derivative denoted by $\partial^\alpha f \in L^1_{loc}(U)$, if

$$\int_U (\partial^\alpha f(x))\phi(x)\,dx = (-1)^{|\alpha|}\int_U f(x)(\partial^\alpha\phi(x))\,dx, \text{ for all } \phi \in C_c^\infty(U),$$

where $|\alpha| = \sum_{i=1}^d \alpha_i$, and the functions $\phi \in C_c^\infty(U)$ are called test functions.

Theorem 4.33. *Let $f \in L^1_{loc}(U)$ and $\mathrm{supp}(f) \subset U$. We further assume that f admits the weak derivative $\partial^\alpha f \in L^1_{loc}(U)$, then*

(i) *$f^\epsilon \in C^\infty(\mathbb{R}^d)$ and $\partial^\alpha(f^\epsilon) = \eta^\epsilon * (\partial^\alpha f)$ on U.*
(ii) *$\partial^\alpha(f^\epsilon) \longrightarrow \partial^\alpha$ in $L^1_{loc}(U)$ as $\epsilon \to 0^+$*
(iii) *$\partial^\alpha(f^\epsilon) \longrightarrow \partial^\alpha$ pointwise on $L_{\partial^\alpha f}$ as $\epsilon \to 0^+$*

Proof. See Section 2 of [9] □

Remark 4.34. Note that part (i) of the Theorems 4.31 and 4.33 still holds if we replace η^ϵ by a test function.

Lemma 4.35. *Assume that $f \in l^1_{loc}(U)$ has the weak derivative $\partial^\alpha f \in L^1_{loc}(U)$. Suppose that $\phi \in C_c^\infty(\mathbb{R}^d)$ is a test function with properties*

$$\mathrm{supp}(\phi) = K, \qquad \phi(x) \ge 0, \qquad \int_{\mathbb{R}^d} \phi(x)\,dx = 1, \tag{4.6}$$

for all $x \in \mathbb{R}^d$. Then every $x \in \mathbb{R}^d$ we have

$$|\partial^\alpha (f * \phi)(x)| \leq \sup_{z \in U \cap \Lambda(x)} |\partial^\alpha f(z)|,$$

where $\Lambda(x) = \{y \in \mathbb{R}^d : x - y \in K\}$

Proof. First we note that since properties of Equation 4.6 holds we can replace η^ϵ by a test function ϕ in the Theorem 4.35. Therefore by (i) of Theorem 4.35 we get a equation

$$\begin{aligned}
\partial^\alpha (f * \phi)(x) &= (\phi * \mathbb{1}_U \partial^\alpha f)(x) \\
&= \int_{\mathbb{R}^d} \phi(x-y) \mathbb{1}_U \partial^\alpha f(y)\, dy \\
&= \int_U \phi(x-y) \partial^\alpha f(y)\, dy
\end{aligned}$$

If $x - y \notin \operatorname{supp}(\phi) = K \subset U$, then $\partial^\alpha(f * \phi) = 0$. Therefore we can define a set $\Lambda(x) = \{y \in \mathbb{R}^d : x - y \in K\}$ and thus it holds that

$$\partial^\alpha (f * \phi) = \int_{U \cap \Lambda(x)} \phi(x-y) \partial^\alpha f(y)\, dy.$$

FInally by using above equation we get and properties of test function from Equation 4.6 we get

$$\begin{aligned}
|\partial^\alpha (f * \phi)(x)| &\leq \int_{U \cap \Lambda(x)} |\phi(x-y)||\partial^\alpha f(y)|\, dy \\
&\leq \sup_{z \in U \cap \Delta(x)} |\partial^\alpha f(z)| \int_{U \cap \Delta(x)} \phi(x-y)\, dy \\
&\leq \sup_{z \in U \cap \Delta(x)} |\partial^\alpha f(z)| \int_{\mathbb{R}^d} \phi(x)\, dx \\
&= \sup_{z \in U \cap \Delta(x)} |\partial^\alpha f(z)|.
\end{aligned}$$

□

Remark 4.36. Note that the value of the right-hand side of the Lemma 4.35 can be infinity.

5. The main result

5.1. **Assumptions of the Main result.** In this section we are going to introduce and prove the Itô formula presented in Theorem 5.7. Though this Itô formula could be extended to a general d, for simplicity we presented it and also all other result in this chapter which are related to the Theorem 5.7 in the case $d = 1$. First we assume a certain class of stochastic processes which are explained in the following assumption.

Assumption 5.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space that means for all $B \in \mathcal{F}$ and $A \subset B$ with $\mathbb{P}(B) = 0$ it holds that $A \in \mathcal{F}$. Suppose moreover that $X : [0, \infty) \times \Omega \to U$ is a càdlàg stochastic process defined on $(\Omega, \mathcal{F}, \mathbb{P})$ such that it satisfies the following condition: If $A \subset U$ is a Borel measurable set such that

$m(A) = 0$, where m is the Lebesgue measure, then for all $s \in [0, \infty)$ it holds that $\mathbb{P}(X_s \in A) = 0$. In other words, for all $s \in [0, \infty)$, the law μ_s of a random variable X_s on U defined by $\mu_s(A) = \mathbb{P}(X_s \in A)$, is absolutely continuous with respect to the Lebesgue measure.

Let us next take a quick look at how the previous assumption restrict appropriate processes. We first consider a few results about the continuity of the laws.

Theorem 5.2. *Let X be a Lèvy process on $\mathbb{R}$ generated by (A, ν, γ) with $\nu(\mathbb{R}) = \infty$. We define a measure $\widetilde{\nu}$ by*

$$\widetilde{\nu}(B) = \int_B (|x|^2 \wedge 1)\, \nu(dx).$$

If $(\widetilde{\nu})^n$ is absolutely continuous for some $n \in \mathbb{N}$, then for every $t > 0$ and $B \in \mathcal{B}(\mathbb{R}^d)$, the law $\mathbb{P}(X_t \in B)$ of the random variable X_t is absolutely continuous with respect to the Lebesgue measure.

Remark 5.3. If $B \in \mathcal{B}(\mathbb{R}^d)$ and N is a compound Poisson process on $\mathbb{R}^d$ with Lèvy measure ν. Then the law of the random variable N_t can be written as

$$\mathbb{P}(N_t \in B) = e^{-t\nu(\mathbb{R}^d)} \sum_{k=0}^{\infty} (k!)^{-1} t^k \nu(B)^k,$$

which is not continuous as $\mathbb{P}(N_t = 0) > 0$.

Proof. See section 27 of [12]. □

By Remark 5.3 we see that Assumption 5.1 is not valid for finite variation Lévy processes in general since the law μ_s defined in Assumption 5.1 is not absolutely continuous with respect to Lebesgue measure in the case of compound Poisson process as $\mathbb{P}(N_t = 0) > 0$ for $t > 0$. However by Theorem 5.2 we notice that Assumption 5.1 is always satisfied for a finite variation Lévy process with infinite activity, if its Lévy measure is absolutely continuous with respect to Lebesgue measure. The following proposition consider a bit more about the behaviour of the processes under Assumption 5.1.

Proposition 5.4. *Assume that the process X satisfies Assumption 5.1. Let $A \subset [0, \infty) \times U$ be any Lebesgue measurable set such that $m(A) = 0$, then for all $t \geq 0$ we have*

$$\mathbb{P}(\{\omega \in \Omega : m(\{s \in [0, t] : (s, X_s(\omega)) \in A\}) = 0\}) = 1.$$

In particular, this implicity implies that for almost all $\omega \in \Omega$, the set $\{s \in [0, t] : (s, X_s(\omega)) \in A\}$ is Lebesgue measurable for all $t \geq 0$.

Proof. First we are going to show that the set $m(\{s \in [0, t] : (s, X_s(\omega)) \in A\})$ is well-defined for almost all $\omega \in \Omega$. Assume that A is a Borel measurable set. We define a process $Y : [0, \infty) \times \Omega \to [0, \infty) \times U$ by

$$Y(s, \omega) = (s, X_s(\omega)).$$

Since X is càdlàg process by Assumption 5.1, the process Y is also càdlàg and $\mathcal{B}_{[0,\infty)} \otimes \mathcal{F}$-measurable by Proposition 5.5, where $\mathcal{B}_{[0,\infty)}$ is the Borel σ-algebra on

$[0,\infty)$ and $\mathcal{F}$ is σ -algebra on Ω. Hence when $A \subset [0,\infty) \times \Omega$ is a Borel set it holds that

$$Y^{-1}(A) \in \mathcal{B}_{[0,\infty)} \otimes \mathcal{F}$$

and so

$$([0,t] \times \Omega) \cap Y^{-1}(A) \in \mathcal{B}_{[0,\infty)} \otimes \mathcal{F} \subset \mathcal{L} \otimes \mathcal{F},$$

where $\mathcal{L}$ is Lebesgue σ-algebra on $[0,\infty)$. Therefore we can define function $f : [0,\infty) \times \to \mathbb{R}$, $f(t,\omega) := \mathbb{1}_{([0,t]\times\Omega)\cap Y^{-1}(A)}$ and hence $f \in L^1(m \times \mathbb{P})$.

We define a function $f_\omega := (.,\omega)$ and by Fubini-Tonelli theorem 4.9 $f_\omega \in L^1(m)$ for almost all ω. Therefore the set $([0,t] \times \Omega) \cap Y^{-1}(A)$ is Lebesgue measurable for a fixed ω and hence $m(\{s \in [0,t] : (s, X_s(\omega)) \in A\})$ is well-defined for almost all $\omega \in \Omega$ that completes the first part of the proof.

Next let

$$Z(\omega) := \int_{\mathbb{R}} f_\omega \, dm = m(\{s \in [0,t] : (s, X_s(\omega)) \in A\}).$$

Again by Fubini-Tonelli theorem Z is a random variable and $Z \in L^1(\mathbb{P})$. Since Z is non-negative random variable it is sufficient to show that $\mathbb{E}[Z] = 0$.

$$\begin{aligned}
\mathbb{E}[Z] &= \int_\Omega \int_{\mathbb{R}} f_\omega \, dm \, d\mathbb{P} \\
&= \int_0^t \int_\Omega f_s \, d\mathbb{P} \, ds \\
&= \int_0^t \mathbb{E}[\mathbb{1}_{\{(s,X_s)\in A\}}] \, ds. \\
&= \int_0^t \mathbb{P}(X_s \in A_s) \, ds,
\end{aligned}$$

where for fixed s, $\mathbb{1}_{\{(s,X_s)\in A\}} = \mathbb{1}_{\{X_s \in A_s\}}$ and $A_s = \{y \in \mathbb{R}^d : (s,y) \in A\}$ is Borel measurable. The A is a Borel measurable by assumption and hence Lebesgue measurable as well. By Theorem 4.8 the function $s \longmapsto m(A_s)$ is Lebesgue measurable and $m(A) = \int_{[0,\infty)} m(A_s) \, ds$. On the other hand we assumed that $m(A) = 0$ and hence $m(A_s) = 0$ for Lebesgue almost all $s \geq 0$. In other words there exist a set $N \subset [0,\infty)$ such that $m(N) = 0$ and for all $s \in N^c$ it holds that $m(A_s) = 0$. Moreover by Assumption 5.1 the measure $\mu_s(A) = \mathbb{P}(X_s \in A)$ is absolutely continuous with respect to the Lebesgue measure and therefore it holds that $\mathbb{P}(X_s \in A) = 0$, when $s \in N^c$. Finally with the help of the above equations we can obtain that

$$\begin{aligned}
\mathbb{E}[Z] &= \int\limits_{[0,t]\cap N^c} \mathbb{P}(X_s \in A_s) \, ds + \int\limits_{[0,t]\cap N} \mathbb{P}(X_s \in A_s) \, ds \\
&= \int\limits_{[0,t]\cap\{s : m(A_s)=0\}} \mathbb{P}(X_s \in A_s) \, ds \\
&= 0.
\end{aligned}$$

The expectation $\mathbb{E}[Z] = 0$ implies that $Z = 0$, $\mathbb{P}$-almost surely and hence $m(\{s \in [0,t] : (s, X_s(\omega)) \in A\}) = 0$ for almost all $\omega \in \Omega$, which completes the proof in the case A is a Borel measurable set.

Finally suppose that A is Lebesgue measurable. We can write $A = A' \cup A''$, $A'' \subset B$, where A' and B are Borel measurable sets such that $m(B) = 0$. Moreover A is a null set by the assumption and therefore also $m(A') = 0$. Since we assume the probability space to be complete it holds that $A'' \in \mathcal{F}$ and $Y^{-1}(A' \cup A'') \in \mathcal{B}_{[0,\infty)} \otimes \mathcal{F}$. By using similar procedure as in above case where A is Borel measurable we can deduce that a set $m(\{s \in [0,t] : (s, X_s(\omega)) \in A' \cup A''\})$ is well defined for almost all $\omega \in \Omega$. Again by using previous part of the proof to the Borel null sets A' and B we obtain that

$$m(\{s \in [0,t] : (s, X_s(\omega)) \in A'\}) = 0 \quad \text{and} \quad m(\{s \in [0,t] : (s, X_s(\omega)) \in B\}) = 0.$$

Therefore

$$m(\{s \in [0,t] : (s, X_s(\omega)) \in A\}) = m(\{s \in [0,t] : (s, X_s(\omega)) \in A' \cup A''\}) \leq 0,$$

for almost all $\omega \in \Omega$ and the desired result follows. □

Intuitively Proposition 5.4 tells about the amount of time a process X spends in set U. For example if $A = [0,t] \times B$, where $B \subset \mathbb{R}^d$ is a Borel set and $\{s \in [0,t] : (s, X_s) \in A\}$ is time X spends in B. Therefore under Assumption 5.1, Proposition 5.4 implies that the Lebesgue measure of the amount of time the process X spends in any Borel measurable set B with $m(B) = 0$ is always zero.

5.2. **Key tools.**

Proposition 5.5. *Let X be an optional process. When considered as a mapping on $[0,\infty) \times \Omega$, it is $\mathcal{B}_{[0,\infty)} \otimes \mathcal{F}$ -measurable, where $\mathcal{B}_{[0,\infty)}$ is Borel σ-algebra on $[0,\infty)$ and $\mathcal{F}$ is σ-algebra on Ω.*

Proof. By [7] it is suffiecient to show that every càdlàg adapted process X satisfies claimed property. We define a new process $X^{(n)}$ for $n, k \in \mathbb{N}$ by setting

$$X^{(n)} = \sum_{k=1}^{2^n} X_{k/2^n} \mathbb{1}_{\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)}(t).$$

Since

$$\{X^{(n)} \in B\} = \bigcup_{k \in \mathbb{N}} \left[\{\omega : X_{k/2^n}(\omega) \in B\} \times \left[\frac{k-1}{2^n}, \frac{k}{2^n} \right) \right],$$

we have $\{X^{(n)} \in B\} \in \mathcal{B}_{[0,\infty)} \otimes \mathcal{F}$ for all Borel sets B, hence $X^{(n)}$ is $\mathcal{B}_{[0,\infty)} \otimes \mathcal{F}$-measurable. Since the Process X is càdlàg by assumption the $X^{(n)} \to X$ pointwise as $n \to \infty$ and therefore X is also $\mathcal{F} \otimes \mathcal{B}_{[0,\infty)}$-measurable.

□

Before we introduce the main result of this section we have to define some elementary tools to help us. For any continuous function $f : [0,\infty) \times U \longrightarrow \mathbb{R}$ on $[0,\infty) \times U$ we can always continuously extend it to a new function $\tilde{f} : \mathbb{R} \times U \longrightarrow \mathbb{R}$:

$$\tilde{f}(t,x) = \begin{cases} f(t,x), & (t,x) \in [0,\infty) \times U; \\ f(-t,x), & (t,x) \in [0,\infty) \times U. \end{cases}$$

Since $f \in L^1_{loc}([0,\infty)\times U)$ then clearly $\tilde{f} \in L^1_{loc}(\mathbb{R}\times U)$ and it is weakly differentiable on the open set $\mathbb{R} \times U$ in the sense of Definition 4.32 and therefore by applying chain rule the weak derivatives of the function $\tilde{f}$ in one-dimensional case are

$$\frac{\partial \tilde{f}}{\partial t}(t,x) = \begin{cases} \frac{\partial f}{\partial t}(t,x), & (t,x) \in [0,\infty) \times U; \\ -\frac{\partial f}{\partial t}(-t,x), & (t,x) \in [0,\infty) \times U, \end{cases}$$

and

$$\frac{\partial \tilde{f}}{\partial x}(t,x) = \begin{cases} \frac{\partial f}{\partial x}(t,x), & (t,x) \in [0,\infty) \times U; \\ \frac{\partial f}{\partial x}(-t,x), & (t,x) \in [0,\infty) \times U, \end{cases}$$

where $\frac{\partial \tilde{f}}{\partial t}(t,x)$ and $\frac{\partial \tilde{f}}{\partial x}(t,x)$ are weak derivatives of f in the sense of Equation 5.1.

By using an extension $\tilde{f}$ of the function $f \longrightarrow \mathbb{R} \times U$ we can now define a sequence of mollified functions by

$$f_n(t,x) = (\phi_n * \tilde{f}\mathbb{1}_{\mathbb{R}\times U})(t,x),$$

where $(t,x) \in \mathbb{R}^2$, $n \geq 1$ and $\phi_n = \eta^{\frac{1}{n}}$ is a standard mollifier in the sense of Definition 4.30. The following proposition summarizes some limiting properties of the sequence f_n.

Proposition 5.6. *Assume that $f : [0,\infty) \times U \longrightarrow \mathbb{R}$ is a continuous function on $[0,\infty) \times U$ such that $f \in L^1_{loc}([0,\infty) \times U)$, $\text{supp}(f) \subset [0,\infty) \times U$ and $U \subset \mathbb{R}$ is an open set. Let the weak derivatives $\frac{\partial f}{\partial s}, \frac{\partial f}{\partial x} \in L^1_{loc}([0,\infty) \times U)$ be locally bounded and defined by equation*

$$\int\limits_{[0,\infty)\times U} (\partial^\alpha f(x))\phi(x)\,dx = (-1)^{|\alpha|} \int\limits_{[0,\infty)\times U} f(x)(\partial^\alpha \phi(x))\,dx, \tag{5.1}$$

for all $\phi \in C_c^\infty([0,\infty) \times U)$. Finally suppose that X is a finite variation Lévy process satisfying Assumption 5.1 such that for all $t \geq 0$, X_t and X_{t-} are in U. Then the following are true:

(i) $\lim\limits_{n\to\infty} f_n(0,X_0) = f(0,X_0)$.

(ii) $\lim\limits_{n\to\infty} \int_0^t \frac{\partial f_n}{\partial s}(s,X_s)\,ds = \int_0^t \frac{\partial f}{\partial s}(s,X_s)\,ds$, *$\mathbb{P}$-almost surely.*

(iii) $\lim\limits_{n\to\infty} \int_0^t \frac{\partial f_n}{\partial x}(s,X_s)\,ds = \int_0^t \frac{\partial f}{\partial x}(s,X_s)\,ds$, *$\mathbb{P}$-almost surely.*

(iv) $\lim\limits_{n\to\infty} \iint_{[0,t]\times\mathbb{R}}(f_n(s,X_{s-}+x) - f_n(s,X_{s-}))\,J_X(ds \times dx)$
$= \iint_{[0,t]\times\mathbb{R}}(f(s,X_{s-}+x) - f(s,X_{s-}))\,J_X(ds \times dx)$.

Proof. (i) Since $\tilde{f}$ is a continuous function, $L_{\tilde{f}} = \mathbb{R} \times U$ by Lemma 4.24. On the other hand for all $t \geq 0$, X_t is in U and so by (iii) of Theorem 4.31

$$f_n(0,X_0) \longrightarrow \tilde{f}(0,X_0),$$

for all $\omega \in \Omega$ and $t \in \mathbb{R}$. It also holds for all $t \geq 0$ that, $\tilde{f}(t,X_t) = f(t,X_t)$ by the definition of $\tilde{f}$.

(ii) From Theorem 4.33, if $(s, X_s) \in L_{\frac{\partial \tilde{f}}{\partial s}}$, then we have

$$\frac{\partial f_n}{\partial s}(s, X_s) \longrightarrow \frac{\partial \tilde{f}}{\partial s}(s, X_s).$$

Let $L_1 = \mathbb{R} \times \left(U \setminus L_{\frac{\partial \tilde{f}}{\partial s}}\right)$ be a Lebesgue set, then

$$\begin{aligned}\int_0^t \frac{\partial f_n}{\partial s}(s, X_s)\, ds = &\int_0^t \frac{\partial f_n}{\partial s}(s, X_s) \mathbb{1}_{\{(s,X_s)\notin L_1\}}\, ds \\ &+ \int_0^t \frac{\partial f_n}{\partial s}(s, X_s) \mathbb{1}_{\{(s,X_s)\in L_1\}}\, ds.\end{aligned}$$

By Lebesgue differentiation theorem $m(L_1) = 0$ and therefore by Proposition 5.4 it holds that $m([0,t] \cap \{s \in [0,t] : (s, X_s) \in L_1\}) = 0$, $\mathbb{P}$-almost surely. Moreover for fixed t we have

$$\int_0^t \frac{\partial f_n}{\partial s}(s, X_s) \mathbb{1}_{\{(s,X_s)\in L_1\}}\, ds = \int\limits_{[0,t]\cap\{s:(s,X_s)\in L_1\}} \frac{\partial f_n}{\partial s}(s, X_s)\, ds = 0,$$

and therefore

$$\int_0^t \frac{\partial f_n}{\partial s}(s, X_s)\, ds = \int_0^t \frac{\partial f_n}{\partial s}(s, X_s) \mathbb{1}_{\{(s,X_s)\notin L_1\}}\, ds, \tag{5.2}$$

$\mathbb{P}$-almost surely.

By Lemma 4.35, for all $(s, x) \in \mathbb{R}^2$

$$\begin{aligned}\left|\frac{\partial f_n}{\partial s}(s, x)\right| &\leq \sup_{z\in(\mathbb{R}\times U)\cap\Lambda(s,x)} \left|\frac{\partial \tilde{f}}{\partial s}(z)\right| \\ &\leq \sup_{z\in\Lambda(s,x)} \left|\frac{\partial \tilde{f}}{\partial s}(z)\right|,\end{aligned} \tag{5.3}$$

where $\Lambda(s, X_s) = \{y \in \mathbb{R}^2 : (s, X_s) - y \in K\}$, and $K = \text{supp}(\phi_n) = \overline{B(0, \frac{1}{n})} \subset \overline{B(0,1)}$. Therefore for $0 \leq s \leq t$, it holds that

$$\left|\frac{\partial f_n}{\partial s}(s, X_s)\right| \leq \sup_{z\in\Lambda(s,X_s)} \left|\frac{\partial \tilde{f}}{\partial s}(z)\right|.$$

Since a càdlàg process is bounded on $[0, t]$, for fixed $\omega \in \Omega$ the set $\Lambda(s, X_s)$ is bounded. Therefore for a fixed $\omega \in \Omega$ and $s \in [0, t]$ we can find a upper bound for $\left|\frac{\partial f_n}{\partial s}(s, X_s)\right|$ that depends only on ω, t and the minimum and maximum of $\frac{\partial \tilde{f}}{\partial s}(s, X_s)$ on $[0, t]$. Since we assumed the weak derivatives of function f be locally bounded, the upper bound of $\left|\frac{\partial f_n}{\partial s}(s, X_s)\right|$ is finite and hence also weak derivatives of $\tilde{f}$ are locally bounded too. Therefore we can apply Lebesgue Dominated convergence theorem, Equation 5.2 and Theorem 4.33 to obtaining

$$\lim_{n\to\infty}\int_0^t \frac{\partial f_n}{\partial s}(s,X_s)\,ds = \int_0^t \lim_{n\to\infty}\frac{\partial f_n}{\partial s}(s,X_s)\mathbb{1}_{\{(s,X_s)\notin L_1\}}\,ds$$
$$= \int_0^t \frac{\partial \tilde f}{\partial s}(s,X_s)\mathbb{1}_{\{(s,X_s)\notin L_1\}}\,ds,$$

$\mathbb{P}$-almost surely. Since $m\{s\in[0,t]:(s,X_s)\in L_1\}=0$, $\mathbb{P}$-almost surely it holds that

$$\int_0^t \frac{\partial \tilde f}{\partial s}(s,X_s)\,ds = \int_0^t \frac{\partial \tilde f}{\partial s}(s,X_s)\mathbb{1}_{\{(s,X_s)\in L_1\}}\,ds,$$

$\mathbb{P}$-almost surely. Finally by definition of extension $\tilde f$, for each $s\in[0,t]$, it holds that $\frac{\partial \tilde f}{\partial s}(s,X_s)=\frac{\partial f}{\partial s}(s,X_s)$ and we obtain

$$\lim_{n\to\infty}\int_0^t \frac{\partial f_n}{\partial s}(s,X_s)\,ds = \int_0^t \frac{\partial \tilde f}{\partial s}(s,X_s)\,ds = \int_0^t \frac{\partial f}{\partial s}(s,X_s)\,ds.$$

(iii) Similar arguments holds if we replace $\frac{\partial f_n}{\partial s}(s,X_s)$ by $\frac{\partial f_n}{\partial x}(s,X_s)$ and therefore we can apply previous procedure from (ii) to conclude that

$$\begin{aligned}
&\lim_{n\to\infty}\int_0^t \frac{\partial f_n}{\partial x}(s,X_s)\,ds\\
&=\lim_{n\to\infty}\left[\int_0^t \frac{\partial f_n}{\partial x}(s,X_s)\mathbb{1}_{\{(s,X_s)\notin L_1\}}\,ds + \int_0^t \frac{\partial f_n}{\partial x}(s,X_s)\mathbb{1}_{\{(s,X_s)\in L_1\}}\,ds\right]\\
&=\lim_{n\to\infty}\int_0^t \frac{\partial f_n}{\partial x}(s,X_s)\mathbb{1}_{\{(s,X_s)\notin L_1\}}\,ds\\
&=\int_0^t \lim_{n\to\infty}\frac{\partial f_n}{\partial x}(s,X_s)\mathbb{1}_{\{(s,X_s)\notin L_1\}}\,ds\\
&=\int_0^t \frac{\partial \tilde f}{\partial x}(s,X_s)\mathbb{1}_{\{(s,X_s)\notin L_1\}}\,ds\\
&=\int_0^t \frac{\partial \tilde f}{\partial x}(s,X_s)\mathbb{1}_{\{(s,X_s)\notin L_1\}}\,ds + \int_0^t \frac{\partial \tilde f}{\partial x}(s,X_s)\mathbb{1}_{\{(s,X_s)\in L_1\}}\,ds\\
&=\lim_{n\to\infty}\int_0^t \frac{\partial f_n}{\partial x}(s,X_s)\,ds,
\end{aligned}$$

$\mathbb{P}$-almost surely.

(iv) Let $I_n=\iint_{[0,t]\times\mathbb{R}} f_n(s,X_{s-}+x)-f_n(s,X_{s-})\,J_X(ds\times dx)$. By using mean-value theorem we have

$$|f_n(s,X_{s-}+x)-f_n(s,X_{s-})| = \left|\frac{\partial f_n}{\partial x}(s,C)\right||x|,$$

where C is a random variable between X_s and $X_{s-}+x$. By Lemma 4.35 and the same procedure in Equation 5.3 we conclude that

$$
\begin{aligned}
|f_n(s, X_{s-} + x) - f_n(s, X_{s-})| &= \left| \frac{\partial f_n}{\partial s}(s, C) \right| |x| \\
&\leq \sup_{z \in (\mathbb{R} \times U) \cap \Lambda(s,C)} \left| \frac{\partial \tilde{f}}{\partial s}(z) \right| |x| \\
&\leq \sup_{z \in \Lambda(s,C)} \left| \frac{\partial \tilde{f}}{\partial s}(z) \right| |x| \\
&\leq C'|x|,
\end{aligned}
$$

where $\Lambda(s, C) = \{y \in \mathbb{R}^2 : (s, C) - y \in K\}$, $K = \text{supp}(\phi_n) = \overline{B(0, \frac{1}{n})} \subset \overline{B(0, 1)}$. Again for fixed $\omega \in \Omega$ the set $\Lambda(s, C)$ is bounded on $[0, t]$, since C is between bounded random variables X_{s-} and $X_{s-} + x$. Hence for fixed $\omega \in \Omega$ and $s \in [0, t]$ there exists a upper bound C' that doesn't depend on s, t or n. Finiteness of C' follows by obtaining that the weak derivatives of f are locally bounded that implies also the boundness of derivatives of $\tilde{f}$.

On the other hand X is finite variation Lévy process when we have by Proposition 2.35 that

$$
\begin{aligned}
I_n &\leq \iint_{[0,t] \times \mathbb{R}} C'|x| \, J_X(ds \times dx) \\
&= C' \iint_{[0,t] \times \mathbb{R}} |x| \, J_X(ds \times dx) \\
&< \infty,
\end{aligned}
$$

and therefore we can apply Lebesgue dominated convergence theorem to interchange the limit and the integral in expression I_n as n goes infinity. Furthermore since $[0, t] \times U \subset \mathbb{R} \times U = L_{\tilde{f}}$ by (iv) of Theorem 4.31 we obtain that $f_n \longrightarrow \tilde{f}$ pointwise as $n \to \infty$ on $[0, t] \times U$ and thus

$$
\begin{aligned}
\lim_{n \to \infty} I_n &= \iint_{[0,t] \times \mathbb{R}} \lim_{n \to \infty} (f_n(s, X_{s-} + x) - f_n(s, X_{s-})) \, J_X(ds \times dx) \\
&= \iint_{[0,t] \times \mathbb{R}} \tilde{f}(s, X_{s-} + x) - \tilde{f}(s, X_{s-}) \, J_X(ds \times dx) \\
&= \iint_{[0,t] \times \mathbb{R}} f(s, X_{s-} + x) - f(s, X_{s-}) \, J_X(ds \times dx).
\end{aligned}
$$

□

Now we are ready to state and prove our main result.

5.3. **Itô's formula for finite variation Lévy processes.**

Theorem 5.7. *Assume that $f : [0,\infty)\times U \longmapsto \mathbb{R}$ is a continuous function on $[0,\infty)\times U$ such that $f \in L^1_{loc}([0,\infty)\times U)$, $supp(f) \subset [0,\infty)\times U$ and U is an open set of $\mathbb{R}$. Let the weak derivatives $\frac{\partial f}{\partial s}, \frac{\partial f}{\partial x} \in L^1_{loc}([0,\infty)\times U)$ be locally bounded and defined by equation 5.1. Suppose that X is a finite variation Lévy process satisfying Assumption 5.1 such that for all $t \geq 0$, X_t and X_{t^-} are in U. Then*

$$f(t,X_t) = f(0,X_0) + \int_0^t \frac{\partial f}{\partial s}(s,X_s)\,ds + \gamma\int_0^t \frac{\partial f}{\partial s}(s,X_s)\,ds + \iint_{[0,t]\times\mathbb{R}} (f(s,X_{s^-}+x) - f(s,X_{s^-}))\,J_X(ds\times dx),$$

where J_X and γ are respectively the Jump measure and the drift coefficient of the process X admitting the following representation from Theorem 2.36 : $X_t = \gamma t + \int_{[0,t]\times\mathbb{R}} x\,J_X(ds\times dx)$.

Proof. Let $\tilde{f}$ and f_n be defined like in Proposition 5.6. Since we assumed $f \in L^1_{loc}([0,\infty)\times U)$ it follows that $\tilde{f} \in L^1_{loc}(\mathbb{R}\times U)$ and by (i) of Theorem 4.33, $f_n \in C^\infty(\mathbb{R}\times\mathbb{R})$ for all $n \geq 1$. Hence by using Ito formula presented in Theorem 3.9, we have

$$f_n(t,X_t) = f_n(0,X_0) + \int_0^t \frac{\partial f_n}{\partial s}(s,X_s)\,ds + \gamma\int_0^t \frac{\partial f_n}{\partial s}(s,X_s)\,ds + \iint_{[0,t]\times\mathbb{R}} (f_n(s,X_{s-}+x) - f_n(s,X_{s-}))\,J_X(ds\times dx),$$

and by Proposition 5.6, for fixed $t \geq 0$ it holds for $\mathbb{P}$-almost surely that;

$$f(t,X_t) = f(0,X_0) + \int_0^t \frac{\partial f}{\partial s}(s,X_s)\,ds + \gamma\int_0^t \frac{\partial f}{\partial s}(s,X_s)\,ds + \iint_{[0,t]\times\mathbb{R}} (f(s,X_{s-}+x) - f(s,X_{s-}))\,J_X(ds\times dx). \tag{5.4}$$

The process X is càdlàg by Assumption 5.1, so the left-hand side and the right-hand side of the above equality are well-defined processes and also modifications of each other in the sense of Definition 2.39. To show that these two sides of Equation 5.4 are actually indistinguishable we obtain that by Theorem 2.40 it is sufficient to show that the processes of the both side of the equation 5.4 are right-continuous.

First we note that since the process $X : [0,\infty)\times\Omega \longrightarrow U$ is càdlàg and the function f is continuous then $(f(t,X_t))_{t\geq 0}$ is also càdlàg. Moreover $(f(0,X_0))_{t\geq 0}$ is càdlàg as well.

The functions $\frac{\partial f}{\partial s}$ and $\frac{\partial f}{\partial x}$ are Borel measurable and for fixed $\omega \in \Omega$, the process $(X_s)_{0\leq s\leq t}$ is also Borel measurable. Hence for a fixed ω, $\frac{\partial f}{\partial s}(s,X_s)$ and $\frac{\partial f}{\partial x}(s,X_s)$ are

Borel measurable and hence Lebesgue measurable. Furthermore by Fundamental theorem of Lebesgue integral calculus it holds that

$$t \longmapsto \int_0^t \frac{\partial f}{\partial s}(s, X_s)\, ds \quad \text{and} \quad t \longmapsto \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, ds,$$

are absolutely continuous and therefore by remark 4.13 also uniformly continuous in t.

Let

$$Z_t := \iint_{[0,t]\times\mathbb{R}} (f(s, X_s) - f(s, X_{s^-}))\, J_X(ds \times dx).$$

Since f is not necessarily smooth, to show the right-continuity of the process $Z = (Z_t)_{t\geq 0}$ we have to deduce that $\lim_{h\to 0^+} |Z_{t+h} - Z_t| = 0$. For all $s \geq 0, X_s$ and X_{s-} are in U, therefore by equation 3.5 it holds that

$$Z_t = \sum_{0\leq s\leq t} (f(s, X_s) - f(s, X_{s-})).$$

By using above equation we obtain that

$$\begin{aligned}
\lim_{h\to 0^+} |Z_{t+h} - Z_t| &= \lim_{h\to 0^+} \left| \sum_{0\leq s\leq t+h} (f(s, X_s) - f(s, X_{s-})) - \sum_{0\leq s\leq t} (f(s, X_s) - f(s, X_{s-})) \right| \\
&= \lim_{h\to 0^+} \left| \sum_{t\leq s\leq t+h} (f(s, X_s) - f(s, X_{s-})) \right| \\
&\leq \lim_{h\to 0^+} \sum_{t\leq s\leq t+h} |(f(s, X_s) - f(s, X_{s-})| \\
&= \lim_{h\to 0^+} \sum_{t\leq s\leq t+h} |\lim_{n\to\infty} [(f_n(s, X_s) - f_n(s, X_{s-})]| \\
&= \lim_{h\to 0^+} \sum_{t\leq s\leq t+h} \lim_{n\to\infty} |((f_n(s, X_{s-} + \Delta X_s) - f_n(s, X_{s-}))|
\end{aligned}$$

Similarly to (iv) of Proposition 5.6 by using mean-value theorem we can find C' such that

$$|((f_n(s, X_{s-} + \Delta X_s) - f_n(s, X_{s-}))| = \sum_{0\leq s\leq t} \left| \frac{\partial f_n}{\partial s}(s, C) \right| |\Delta X_s|,$$

where C is a random variable between X_{s-} and $X_{s-} + \Delta X_s$. By Equation 5.3 we have that

$$|((f_n(s, X_{s-} + \Delta X_s) - f_n(s, X_{s-}))| \leq C''|\Delta X_s|,$$

where C'' is a upper bound for that doesn's depends on s, x or n on $[0, t]$. Furthermore C'' is finite because the weak derivatives of f are locally bounded that implies also boundness of derivatives of $\tilde{f}$. Hence we obtain that

$$\begin{aligned}\lim_{h\to 0^+} |Z_{t+h} - Z_t| &\le \lim_{h\to 0^+} \sum_{t\le s\le t+h} \lim_{n\to\infty} |(f_n(s, X_s) - f_n(s, X_{s-}))| \\ &\le \lim_{h\to 0^+} \sum_{t\le s\le t+h} \lim_{n\to\infty} C''|\Delta X_s| \\ &\le C'' \lim_{h\to 0^+} \sum_{t<s\le +h} \Delta X_s = 0, \quad \mathbb{P}\text{-almost surely.}\end{aligned}$$

This shows that the process Z is right-continuous and thus the conclusion of the theorem holds. □

References

[1] David Applebaum, *Lévy Processes and Stochastic Calculus.* Second edition, D. Applebaum, 2009.

[2] Haim Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations* Universitext, Springer, New York, ISBN 978-0-387-70913-0, 2011.

[3] Andrew M. Bruckner, Judith B. Bruckner, Brian S. Thomson, *Real Analysis* ISBN: 0-13-458886-X

[4] Rama Cont, Peter Tankov, *Financial modelling With Jump Processes* Chapman & Hall/CRC Financ. Math. Ser., Chapman & Hall/CRC, Boca Raton, FL, ISBN 1-5848-8413-4, 2004.

[5] Ernst Eberlein, *Jump-type Lèvy processes* T. G. Andersen, R. A. Davis, J.-P. Kreiss, & T. V. Mikosch (Eds.), Handbook of financial time series (pp. 439–456). Berlin: Springer, 2009.

[6] Gerald B. Folland, *Real Analysis: Modern Techniques and Their Applications* second edition, Pure Appl. Math. (New York), John Wiley , John Wiley & Sons Inc., New York, ISBN 0-471-31716-0. A Wiley–Interscience Publication, 1999.

[7] Jean Jacod, Albert N. Shiryaev, *Limit Theorems for Stochastic Processes* Grundlehren Math. Wiss ciples of Mathematical Sciences), vol. 288, Springer-Verlag, Berlin, ISBN 3-540-17882-1, 1987.

[8] Andreas E. Kyprianou, *Indruductory Lectures on Fluctations of Lévy processes with Applications.* Springer-Verlag Berlin Heidelberg, 2006.

[9] Ramin Okhrati, Uwe Schmock, *Itô's formula for finite variation Lévy proceses: The case of non-smooth functions.* J. Math Anal. Appl 430. 1163-1174, 2015

[10] Philip E. Protter, *Stochastic Integration and Differential Equations* second edition, Stoch. Model. Appl. Probab., vol. 21, Springer-Verlag, Berlin, ISBN 3-540-00313-4, 2004.

[11] Daniel Revuz, Marc Yor, *Continuous Martingales and Brownian Motion* Spinger-Verlag (1990); third edition 1999.

[12] Ken-iti Sato, *Lévy Processes and Infinitely Divisible Distributions* Cambridge Stud. Adv. Math., vol. 68, Cambridge University Press, Cambridge, ISBN 0-521-55302-4. Translated from the 1990 Japanese original. Revised by the author, 1999.

[13] David Williams, *Probability with Martingales* Cambridge University Press, 1991.