Lévy Processes

Seiji HIRABA

Contents

In this text we describe “additive processes”, which are in a basic group of stochastic processes, and which have independent increments. Especially, we investigate “Lévy processes” in detail, which are time homogeneous additive processes, continuous in probability, and have first-order discontinuous sample paths, that is, they are right-continuous and have left-hand limits.

We will show the following: For every Lévy process, it has infinitely divisible distributions and their characteristic functions are given by the **Lévy-Khintchine representations**.

Also its sample path has the **Lévy-Ito decomposition**, which is a sum of Gaussian process with drift and a jump process.

Reference. SATO, Ken-ichi; “Kahou Katei” (in Japanese) as “Additive Processes”, Kinokuniya (1990). This was rewritten in English and revised to “Lévy Processes and Infinitely Divisible Distributions”, Cambridge (1999, 2002).

This text is based on the above text and the proofs are almost the same. However, the author tried to simplify and to make refinements in order to understand easily.

1 Overview of Lévy Processes and Infinitely Divisible Distributions

Let $(\Omega, \mathcal{F}, P)$ be a probability space, that is, Ω is a non-empty set, $\mathcal{F}$ is a σ-additive class on Ω and $P = P(d\omega)$ is a probability measure on a measurable space $(\Omega, \mathcal{F})$.

A *stochastic process* $(X_t)_{t\geq 0}$ is random variables $X_t = X_t(\omega)$ parameterized by time $t \geq 0$.

In the text we only consider the $\mathbf{R}^d$-valued processes. So we denote $X_t = (X_t^j)_{j\leq d}$ and a vector as $x = (x^j)_{j\leq d} = (x^1, \ldots, x^d) \in \mathbf{R}^d$. Also we denote inner product as $\langle x, y\rangle \equiv x \cdot y = \sum_{j\leq d} x^j y^j$.

A **Lévy process** $(X_t)_{t\geq 0}$ is an $\mathbf{R}^d$-valued stochastic process starting from the origin 0, which is continuous in probability, has time homogeneous independent increments and right-continuous sample paths with left-hand limits.

This is equivalent to that $^\forall t > 0$, the distribution of X_t; $\mu_t = P \circ X_t^{-1}$, i.e., $\mu_t(dx) = P(X_t indx)$ is an **infinitely divisible distribution**. This is also equivalent to that if we let $\mu = \mu_1$, then for any $t > 0$, $\mu_t = \mu^{t*}$, where the right-hand side denotes t-convolution of μ.

A convolution of measures μ, ν is defined as

$$\mu * \nu(dx) := \int \mu(dx - y)\nu(dy) = \int \nu(dx - y)\mu(dy) = \int\int 1_{dx}(y + z)\mu(dy)\nu(dz).$$

If $\nu = \mu$, then $\mu^{2*} = \mu * \mu$. In general, for $n \in \mathbf{N}$, we define $\mu^{n+1*} = \mu^{n*} * \mu$. Moreover, if μ is an infinitely divisible distribution, then for any $t \geq 0$, μ^{t*} can be defined.

In this case, the above is equivalent to that the characteristic function of X_t; $\widehat{\mu_t}(z) := E[e^{i\langle z, X_t\rangle}]$ $(i = \sqrt{-1})$ has the **Lévy-Khintchine representation (LK-representation)**, i.e., $\widehat{\mu_t}(z) = e^{t\psi(z)}$ with

$$\psi(z) = -\frac{1}{2}\langle Az, z\rangle + \int_{(|x|\geq 1)} (e^{i\langle z,x\rangle} - 1)\nu(dx) + \int_{(|x|<1)} (e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle)\nu(dx) + i\langle \gamma, z\rangle,$$

where

- $A = (a_{jk})$: a non-negative definite $d \times d$-matrix, i.e., $a_{jk} = \sum_{\ell\leq m} \sigma_\ell^j \sigma_\ell^k$ with $\sigma = (\sigma_\ell^j)_{\ell\leq m, j\leq d}$: a diffusion coefficient.
- $\nu = \nu(dx)$ is a **Lévy measure** on $\mathbf{R}^d$ such that $\nu(\{0\}) = 0$ and that $\displaystyle\int_{\mathbf{R}^d} 1 \wedge |x|^2 \nu(dx) < \infty$.
- $\gamma = (\gamma^j)_{j\leq d} \in \mathbf{R}^d$,

Furthermore, this is equivalent to the following: **Lévy-Ito decomposition (LI-decomposition)**

$$dX_t(\omega) = \gamma dt + \sigma dB_t(\omega) + \int_{(|x|\geq 1)} xN(\omega; dt, dx) + \int_{(|x|<1)} x\widetilde{N}(\omega; dt, dx), X_0 = 0.$$

More precisely,

$$X_t(\omega) = \gamma t + \sigma B_t(\omega) + \int_0^t \int_{(|x|\geq 1)} xN(\omega; ds, dx) + \int_0^t \int_{(|x|<1)} x\widetilde{N}(\omega; ds, dx).$$

If $X_t = (X_t^j)_{j\leq d} = (X_t^1, \ldots, X_t^d)$, then

$$X_t^j = \gamma^j t + \sum_{\ell\leq m} \sigma_\ell^j B_t^\ell + \int_0^t \int_{(|x|\geq 1)} x^j N(\omega; ds, dx) + \int_0^t \int_{(|x|<1)} x^j \widetilde{N}(\omega; ds, dx),$$

where $B_t = (B_t^\ell)$ is an m-dimensional Brownian motion, $N(\omega; dt, dx)$ is a $dt\nu(dx)$-Poisson random measure on $[0, \infty) \times \mathbf{R}^d$, $\widetilde{N} = N - \widehat{N}$ is a compensated Poisson random measure with $\widehat{N} = E[N]$, i.e, $\widehat{N}(dt, dx) = dt\nu(dx)$ is a mean measure of N.

Let $\Delta X_t := X_t - X_{t-}$ be a jump of X_t at time t and let $N(dt, dx) := \sharp\{(t, \Delta X_t) \in dt \times dx; \Delta X_t \neq 0\}$ be a measure of jump-times and jumps on a time space. Then, it can be shown that N is a Poisson random measure by the property of time homogeneous independent increments of the Lévy process (X_t).

The above decomposition theorem means that if we remove large jumps in order from X_t, then the remaining as the limit is a continuous process and it is a Gaussian process. (Kiyoshi ITO showed the way.) That is, let

$$X_t^n = X_t - \sum_{s \le t; |\Delta X_s| \ge 1/n} \Delta X_s = X_t - \int_0^t \int_{(|x| \ge 1/n)} x N(ds, dx).$$

if $n \to \infty$, then $X_t^n \to {}^\exists X_t^c$ in some sense, and X_t^c is a continuous Lévy process, i.e., a Gaussian process.

If X_t have a LI-decomposition, then it can be shown easily by using **Ito formula** that the characteristic function has the above representation as follows;

For $f(x) = e^{ix \cdot z} \in C^2(\mathbf{R}^d)$,

$$\begin{aligned} df(X_t) &= \gamma \cdot Df(X_t)dt + \sigma \cdot Df(X_t)dB_t + \frac{1}{2}\sigma^2 \cdot D^2 f(X_t)dt \\ &\quad + \int_{(|x|\ge 1)} [f(X_{t-} + x) - f(X_{t-})] N(dt, dx) \\ &\quad + \int_{(|x|<1)} [f(X_{t-} + x) - f(X_{t-})] \widetilde{N}(dt, dx) \\ &\quad + \int_{(|x|<1)} [f(X_{t-} + x) - f(X_{t-}) - x \cdot Df(X_{t-})] \nu(dx) dt, \end{aligned}$$

where $\gamma \cdot D = \gamma^j \partial_j$, $\sigma \cdot D = \sigma_\ell^j \partial_j$, $\sigma^2 \cdot D^2 = \sum_{\ell \le m} \sigma_\ell^j \sigma_\ell^k \partial_{jk}^2$ (we use the rule of summing on the same index of supper and lower.) and $\partial_j = \partial/\partial x_j, \partial_{jk}^2 = \partial^2/\partial x_j \partial x_k$.

If we take the expectation, then by $EB_t = E\widetilde{N} = 0$, we have

$$\begin{aligned} d\varphi_t(z) &:= dE[f(X_t)] = E[df(X_t)] \\ &= i\gamma \cdot z \varphi_t(z)dt - \frac{1}{2} \sum_{\ell \le m} \sigma_\ell^j \sigma_\ell^k z_j z_k \varphi_t(z) dt \\ &\quad + \int_{(|x|\ge 1)} \varphi_t(z)[e^{ix\cdot z} - 1] dt \nu(dx) + \int_{(|x|<1)} \varphi_t(z)[e^{ix\cdot z} - 1 - ix \cdot z] dt \nu(dx) \\ &= \varphi_t(z) \Big\{ i\gamma \cdot z - \frac{1}{2} a_{jk} z^j z^k \\ &\quad + \int_{(|x|\ge 1)} [e^{ix\cdot z} - 1] \nu(dx) + \int_{(|x|<1)} [e^{ix\cdot z} - 1 - ix \cdot z] \nu(dx) \Big\} dt. \end{aligned}$$

That is, $d\varphi_t(z) = \varphi_t(z)\psi(z)$. Hence by the initial condition $\varphi_0(z) = E[e^{iz\cdot X_0}] = 1$, we get the desired representation $\varphi_t(z) = e^{t\psi(z)}$.

On the other equivalences, if X_t has the LI-decomposition, then by the properties of the stochastic integral (X_t) has time homogeneous independent increments, thus, it is a Lévy process.

If the characteristic function has LK-representation, then the distribution of X_t is a infinitely divisible distribution and it is corresponding to the Lévy process in the sense of law, one to one (except the law equivalence). A Lévy process in the sense of law is equivalent to a Lévy process, hence they are essentially the same.

It remains to show that a Lévy process has a Lévy-Ito decomposition. This can be shown directly mentioned as above. However, by using above results we can show if X_t is a Lévy process, then the characteristic function has the LK-representation. On the other hand, if Y_t has the LI-decomposition, then the characteristic function has the same representation. Hence, they are equivalent in law and their paths are right-continuous and have left-hand limits. Thus, both have the same distributions on $D([0,\infty) \to \mathbf{R}^d)$. Therefore, X_t can have the same decomposition.

If readers want to know about Ito integrals (stochastic integrals) and Ito formula and so on, please see the text of stochastic analysis; **"Ito integrals and Stochastic Differential Equations with Jumps"**.

2 Definition of Lévy Processes and Basic Examples

In this section, we give a definition of Lévy processes and describe Poisson processes, compound Poisson processes and Brownian motions as basic examples.

2.1 Definition of Lévy processes

Definition 2.1 *An* $\mathbf{R}^d$*-valued stochastic process* $(X_t)_{t\geq 0}$ *is a* **Lévy process** *if it satisfies that*

(1) $X_0 = 0$ *a.s.*

(2) (X_t) *has independent increments, i.e., for* $0 \leq t_0 < t_1 < \cdots < t_n$, $\{X_{t_k} - X_{t_{k-1}}\}_{k\leq n}$ *are independent.*

(3) For $s, t > 0$, $X_{t+s} - X_s \stackrel{(\mathrm{d})}{=} X_t$*, i.e., it is time homogeneous.*

(4) It is continuous in probability, i.e., $^\forall t \geq 0, \varepsilon > 0$, $P(|X_s - X_t| < \varepsilon) \to 1\ (s \to t)$.

(5) With probability one, each sample paths is right-continuous and has left-hand-limits, i.e., $^\exists\Omega_0 \in \mathcal{F}; P(\Omega_0) = 1$, $^\forall\omega \in \Omega_0, (X_t(\omega))_{t\geq 0}$ *is right-continuous and has left-hand-limits as a function of* t.

On the other hand, if it satisfies the conditions except the last one, then it is called a **Lévy process in law**.

In §5.1, we give the result that a Lévy process in law is equivalent to a Lévy process, so the sample path property is not essential. That is, if (Y_t) is a Lévy process in law, then there exists a Lévy process (X_t) such that for $^\forall t > 0, P(X_t = Y_t) = 1$.

The condition of continuity in prob. is equivalent to that at $t = 0$ by starting from 0 and by the time homogeneity, that is,

$$^\forall\varepsilon > 0,\ \lim_{t\downarrow 0} P(|X_t| < \varepsilon) = 1.$$

2.2 Exponential times and Poisson processes

For a constant $\alpha > 0$, a random variable $\tau = \tau(\omega)$ is distributed by an **exponential distribution with parameter** α is that

$$P(\tau > t) = \int_t^\infty \alpha e^{-\alpha s} ds = e^{-\alpha t}$$

That is, τ has a distribution with a density function $f(s) = \alpha e^{-\alpha s}$. In this text, we call τ as α**-exponential time** or simply, **exponential time**.

Its means and variance are the following:

$$E[\tau] = \int_0^\infty \alpha s e^{-\alpha s} ds = \frac{1}{\alpha}, \quad V(\tau) = E[\tau^2] - (E[\tau])^2 = \frac{1}{\alpha^2}.$$

Question 2.1 Make sure the above calculation of variance.

Proposition 2.1 *If* τ *is an exponential time, then it has the following* **memoryless property**. *For* $t, s \geq 0$,

$$P(\tau > t + s|\ \tau > s) = P(\tau > t).$$

Proof.

$$P(\tau > t + s|\ \tau > s) = \frac{P(\tau > t + s)}{P(\tau > s)} = \frac{e^{-(t+s)}}{e^{-s}} = e^{-t} = P(\tau > t).$$

■

Proposition 2.2 *If* $\tau_1, \tau_2, \ldots \tau_n$ *are independent* $\alpha_1, \alpha_2, \ldots, \alpha_n$*-exponential times, respectively, then* $\min\{\tau_1, \tau_2, \ldots \tau_n\}$ *is* $(\alpha_1 + \alpha_2 + \cdots + \alpha_n)$*-exponential time. Moreover,*

$$P(\min\{\tau_1, \tau_2, \ldots \tau_n\} = \tau_k) = \frac{\alpha_k}{\alpha_1 + \alpha_2 + \cdots + \alpha_n}.$$

Proof. For simplicity, we only show the case of $n = 2, k = 1$.

$$P(\tau_1 \wedge \tau_2\} > t) = P(\tau_1 > t, \tau_2 > t) = P(\tau_1 > t)P(\tau_2 > t) = e^{-(\alpha_1+\alpha_2)t}.$$

Moreover, since the joint distribution of τ_1, τ_2 is the product of each ones by their independence, we have

$$\begin{aligned} P(\min\{\tau_1, \tau_2\} = \tau_1) &= P(\tau_1 < \tau_2) \\ &= \int_0^\infty ds\alpha_1 e^{-\alpha_1 s} P(s < \tau_2) \\ &= \int_0^\infty ds\alpha_1 e^{-\alpha_1 s} e^{-\alpha_2 s} \\ &= \frac{\alpha_1}{\alpha_1 + \alpha_2}. \end{aligned}$$

The other cases are the same. ■

Example 2.1 There is a system of two devices A and B. The time to failure of A is an 1-exp. time and the time to failure of B is an 2-exp. time. These are failure independent and the system is failure if at least one is failure. Find the mean time to failure of the system.

By the previous proposition, the time to failure of the system is 3-exp. time, and hence, the mean is 1/3.

For $\lambda > 0$, a stochastic process $(X_t)_{t\geq 0}$ is a **Poisson process with a parameter** λ is a Lévy process such that X_1 has a λ-Poisson distribution (it is simply called a **λ-Poisson process**), that is, it satisfies the following:

(1) $X_0 = 0$,

(2) For $0 \leq s < t$, $X_t - X_s$ has a Poisson distribution with a parameter $\lambda(t-s)$, i.e.,

$$P(X_t - X_s = k) = e^{-\lambda(t-s)} \frac{\lambda^k (t-s)^k}{k!} \quad (k = 0, 1, 2, \dots).$$

(3) X_t has independent increments. That is, for $0 < t_1 < t_2 < \cdots < t_n$, $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ are independent.

Theorem 2.1 (Construction of a Poisson process) *Let $\sigma_1, \sigma_2, \dots$ be independent λ-exponential times. Let $\tau_n = \sum_{k=1}^n \sigma_k$ and $\tau_0 = 0$. Define*

$$X_t = n \iff \tau_n \leq t < \tau_{n+1}, \quad \text{that is,} \quad X_t := \sum_{n=0}^\infty n 1_{[\tau_n, \tau_{n+1})}(t) = \max\{n; \tau_n \leq t\}.$$

Then, (X_t) is a λ-Poisson process.

Note The inverse of the above result holds, that is, if $(X_t)_{t\geq 0}$ is a λ-Poisson process and let $\tau_1, \tau_2, \dots$ be jump times of it, then $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \dots$ are i.i.d. and each of them is a λ-exponential time.

In order to show the above result, we use the following result.

Proposition 2.3 *The sum of independent n-number of λ-exponential times σ_k; $\tau = \sum_{k=1}^n \sigma_k$ is distributed by the gamma distribution $\Gamma(n, \lambda)$, i.e.,*

$$P(\tau < t) = \int_0^t \frac{1}{(n-1)!} \lambda^n s^{n-1} e^{-\lambda s} ds.$$

Proof. By the independence of (σ_n),

$$P(\sigma_1 + \cdots + \sigma_n < t) = \int_{s_1+\cdots s_n<t} \lambda^n e^{-\lambda(s_1+\cdots s_n)} ds_1 \cdots ds_n.$$

By the change of variables such that $u_k = s_1 + \cdots s_k$ $(k = 1, \ldots, n)$ and $s = u_n$,

$$\begin{aligned}
\int_{s_1+\cdots s_n<t} \lambda^n e^{-\lambda(s_1+\cdots s_n)} ds_1 \cdots ds_n &= \int_0^t du_n \int_0^{u_n} du_{n-1} \cdots \int_0^{u_2} du_1 \lambda^n e^{-\lambda u_n} \\
&= \int_0^t du_n \int_0^{u_n} du_{n-1} \cdots \int_0^{u_3} du_2 u_2 \ \lambda^n e^{-\lambda u_n} \\
&= \int_0^t du_n \frac{1}{(n-1)!} u_n^{n-1} \lambda^n e^{-\lambda u_n} \\
&= \int_0^t ds \frac{1}{(n-1)!} \lambda^n s^{n-1} e^{-\lambda s}
\end{aligned}$$

■

[Proof of Theorem 2.1]. Since τ_n is independent of σ_{n+1} and distributed by $\Gamma(n, \lambda)$, we have

$$\begin{aligned}
P(X_t = n) &= P(\tau_n \le t < \tau_{n+1} = \tau_n + \sigma_{n+1}) \\
&= \int_0^t ds \ \frac{1}{(n-1)!} \lambda^n s^{n-1} e^{-\lambda s} P(t < s + \sigma_{n+1}) \\
&= \int_0^t ds \ \frac{1}{(n-1)!} \lambda^n s^{n-1} e^{-\lambda s} e^{-(t-s)\lambda} \\
&= e^{-\lambda t} \frac{\lambda^n}{(n-1)!} \int_0^t s^{n-1} ds = e^{-\lambda t} \frac{\lambda^n t^n}{n!}.
\end{aligned}$$

By a similar way,

$$\begin{aligned}
P(\tau_{n+1} > t + s, X_t = n) &= P(\tau_{n+1} > t + s, \tau_n \le t < \tau_{n+1}) \\
&= P(\tau_n + \sigma_{n+1} > t + s, \tau_n \le t) \\
&= \int_0^t du \ \frac{1}{(n-1)!} \lambda^n u^{n-1} e^{-\lambda u} P(u + \sigma_{n+1} > t + s) \\
&= \int_0^t du \ \frac{1}{(n-1)!} \lambda^n u^{n-1} e^{-\lambda u} e^{-\lambda(t+s-u)} = e^{-\lambda(t+s)} \frac{\lambda^n t^n}{n!}.
\end{aligned}$$

Hence,

$$P(\tau_{n+1} > t + s | \ X_t = n) = e^{-\lambda s} = P(\sigma_1 = \tau_1 > s). \tag{2.1}$$

Moreover,

(2.2)
under the condition $X_t = n$, $\tau_{n+1} - t, \sigma_{n+2}, \ldots, \sigma_{n+m}$ has the same distribution as $\sigma_1, \sigma_2, \ldots, \sigma_m$.

In fact,

$$\begin{aligned}
&P(\tau_{n+1} - t > s_1, \sigma_{n+2} > s_2, \ldots, \sigma_{n+m} > s_m | \ X_t = n) \\
&\quad = P(\tau_n \le t < \tau_{n+1}, \tau_{n+1} - t > s_1, \sigma_{n+2} > s_2, \ldots, \sigma_{n+m} > s_m)/P(X_t = n) \\
&\quad = P(\tau_n \le t, \tau_{n+1} - t > s_1) P(\sigma_{n+2} > s_2, \ldots, \sigma_{n+m} > s_m)/P(X_t = n) \\
&\quad = P(\tau_{n+1} - t > s_1 | \ X_t = n) P(\sigma_2 > s_2, \ldots, \sigma_m > s_m) \\
&\quad = P(\sigma_1 > s) P(\sigma_2 > s_2, \ldots, \sigma_m > s_m) \\
&\quad = P(\sigma_1 > s, \sigma_2 > s_2, \ldots, \sigma_m > s_m).
\end{aligned}$$

By this and noting that $\tau_{n+m} - t = (\tau_{n+1} - t) + \sigma_{n+2} + \cdots + \tau_{n+m}$, we have in general, for $m \geq 1$, we can get

$$P(\tau_{n+m} > t + s |\ X_t = n) = P(\tau_m > s).$$

By subtracting the above from the above with $m + 1$ instead of m, we have

$$P(\tau_{n+m} \leq t + s < \tau_{n+m+1} |\ X_t = n) = P(\tau_m \leq s < \tau_{m+1}) = P(X_s = m).$$

By using this, for $n \geq 0, m \geq 1$,

$$\begin{aligned} P(X_t = n, X_{t+s} - X_t = m) &= P(X_t = n, X_{t+s} = n + m) \\ &= P(X_t = n)P(X_{t+s} = n + m |\ X_t = n) \\ &= P(X_t = n)P(\tau_{n+m} \leq t + s < \tau_{n+m+1} |\ X_t = n) \\ &= P(X_t = n)P(X_s = m). \end{aligned}$$

By summing on $n \geq 0$,

$$P(X_{t+s} - X_t = m) = P(X_s = m) = e^{-\lambda} \frac{\lambda^m s^m}{m!}.$$

In case of $m = 0$, it can be seen $P(X_{t+s} - X_t = m) = e^{-\lambda s}$, and this is included in the above. In fact, by

$$P(\tau_n > t + s |\ X_t = n) = P(\tau_n > t + s |\ \tau_n \leq t < \tau_{n+1}) = 0,$$

if we subtract this from (2.1), then

$$P(X_{t+s} = n |\ X_t = n) = P(\tau_n \leq t + s < \tau_{n+1} |\ X_t = n) = e^{-\lambda s}.$$

Thus,

$$\begin{aligned} P(X_t = n, X_{t+s} - X_t = 0) &= P(X_t = n, X_{t+s} = n) \\ &= P(X_t = n)P(X_{t+s} = n |\ X_t = n) \\ &= P(X_t = n)e^{-\lambda s}. \end{aligned}$$

Hence, by summing on $n \geq 0$, we have $P(X_{t+s} - X_t = 0) = e^{-\lambda s}$.

Finally on the independence of increments, by using (2.2), we have for $0 \leq t_1 < \cdots < t_k$,

$$\begin{aligned} &P(X_{t_0} = n_0, X_{t_1} - X_{t_0} = n_1, \ldots, X_{t_k} - X_{t_{k-1}} = n_k) \\ &\quad = P(X_{t_0} = n_0, X_{t_1} = n_0 + n_1, \ldots, X_{t_k} = n_0 + \cdots + n_k) \\ &\quad = P(X_{t_0} = n_0)P(X_{t_1 - t_0} = n_1, \ldots, X_{t_k - t_0} = n_1 + \cdots + n_k). \end{aligned}$$

Therefore, by repeating this, we have the following independent increments:

$$\begin{aligned} &P(X_{t_0} = n_0, X_{t_1} - X_{t_0} = n_1, \ldots, X_{t_k} - X_{t_{k-1}} = n_k) \\ &\quad = P(X_{t_0} = n_0)P(X_{t_1 - t_0} = n_1) \cdots P(X_{t_k - t_{k-1}} = n_k) \\ &\quad = P(X_{t_0} = n_0)P(X_{t_1} - X_{t_0} = n_1) \cdots P(X_{t_k} - X_{t_{k-1}} = n_k). \end{aligned}$$

■

2.3 Compound Poisson processes

Definition 2.2 *(X_t) is a compound Poisson process on $\mathbf{R}^d$ if it is a Lévy process and the characteristic function is given by the following: Let μ_t be the distribution of X_t.*

$$\widehat{\mu}_t(z) := E[e^{i\langle z, X_t \rangle}] = \exp[tc(\widehat{\sigma}(z) - 1)],$$

where $c > 0$ and $\sigma = \sigma(dx)$ is a distribution on $\mathbf{R}^d$ such that $\sigma(\{0\}) = 0$.

Moreover, it also holds that $\mu_t = e^{-tc}\sum_{n\geq 0}\frac{(tc)^n}{n!}\sigma^{n*}$. Note that $\sigma^{0*} = \delta_0$. (It is clear because characteristic functions coincide.)

[Construction of a compound Poisson process] Let (N_t) be a c-Poisson process. Let (S_n) be a random walk on $\mathbf{R}^d$ starting from $S_0 = 0$, with a one-step distribution σ independent of (N_t). Then, $X_t := S_{N_t}$ is a compound Poisson process. In fact,

$$E[e^{i\langle z,S_{N_t}\rangle}] = \sum_{n\geq 0} E[e^{i\langle z,S_n\rangle}]P(N_t = n) = \sum_{n\geq 0}\widehat{\sigma}(z)^n e^{-tc}\frac{(tc)^n}{n!} = \exp[tc(\widehat{\sigma}(z) - 1)],$$

where for $E[e^{i\langle z,S_n\rangle}] = \widehat{\sigma}(z)^n$, we use that $S_n = \sum_{k=1}^n (S_k - S_{k-1})$ $(S_0 = 0)$, the distribution of $S_k - S_{k-1}$ is σ and $\{S_k - S_{k-1}\}$ are independent.

2.4 Brownian motions (Wiener processes)

A real-valued stochastic process $(B_t)_{t\geq 0}$ is a **Brownian motion** is a continuous Lévy process (a Lévy process with continuous sample paths) such that X_1 has a normal distribution $N(0,1)$, that is,

(1) $B_0 = 0$ a.s.

(2) (B_t) is continuous, i.e., for a.a.ω, the sample path $B_\cdot(\omega)$ is continuous.

(3) For $0 = t_0 < t_1 < \cdots < t_n$, $\{B_{t_k} - B_{t_{k-1}}\}_{k=1}^n$ are independent and $B_{t_k} - B_{t_{k-1}}$ is distributed by a normal distribution $N(0, t_k - t_{k-1})$.

The above definition is a one-dimensional Brownian motion.

If $B_t = (B_t^1, \ldots, B_t^d)$ has d numbers of independent one-dimensional Brownian motions as components, then it is called a **d-dimensional Brownian motion**. (It is realized as a product probability space of d-numbers of independent one-dimensional Brownian motions.)

In this case (B_t) satisfies the same conditions as above with the following (3)' instead of (3);

(3)' For $0 = t_0 < t_1 < \cdots < t_n$, $\{B_{t_k} - B_{t_{k-1}}\}_{k=1}^n$ are independent and $B_{t_k} - B_{t_{k-1}}$ is distributed by the d-dimensional normal distribution $N(0, (t_k - t_{k-1})I_d)$.

Let $W = C([0,\infty) \to \mathbf{R}^1)$ and let $\mathcal{W}$ be the σ-additive class determined by the local uniform convergence topology.

Moreover let $w = w(t) \in W_0 \overset{\text{def}}{\iff} w \in W; w(0) = 0$. For any finite number of time points $\mathbf{t}_n = (t_1, \ldots, t_n); 0 \leq t_1 < t_2 < \cdots < t_n < \infty$ and for any $A_n \in \mathcal{B}^n$, $C(\mathbf{t}_n, A_n) = \{w \in W_0; (w(t_1), \ldots, w(t_n)) \in A_n\}$ is called a **cylinder set)**. We denote the σ-additive class generated by all cylinder sets as $\mathcal{W}_0$ (it is known that this is the same σ-additive class determined by the relative topology of W).

Theorem 2.2 (Existence and uniqueness of Wiener measure) *There exists a unique probability measure P_B on $(\Omega, \mathcal{F}) = (W_0, \mathcal{W}_0)$ such that under this measure $B_t(w) = w(t)$ is a Brown motion.*

P_B is called the **Wiener measure**. The Brownian motion is also called the **Wiener process**.

We give the outline of the proof at the end of this section.

The distribution of d-dimensional Brownian motion $B_t = (B_t^1, \ldots, B_t^d)$ is a probability measure on $W_0^d \ni w; w \in C([0,\infty) \to \mathbf{R}^d), w(0) = 0$, and this is called the **d-dimensional Wiener measure**.

The distribution of B_t is given as $P(B_t \in dx) = p_t(x)dx$, where

$$p_t(x) := \frac{1}{\sqrt{2\pi t}^d}e^{-|x|} \quad (x = (x_1, \ldots, x_d) \in \mathbf{R}^d,\ |x| = \sqrt{x_1^2 + \cdots + x_d^2}).$$

$g_t(x)$ is a density function of d-dimensional normal distribution $N_d(0,t)$.

The **characteristic function** of this normal distribution if given as

$$\varphi(z) = \varphi_{B_t}(z) := E[e^{iz\cdot B_t}] = e^{-t|z|^2/2} \quad (z \in \mathbf{R}^d),$$

where $z \cdot B_t = z_1 B_t^1 + \cdots + z_d B_t^d$.

In one-dimensional case, let

$$p_t(x,y) := p_t(y-x) = \frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/(2t)}.$$

Then the finite dimensional distribution of the Brownian motion is given as follows: for $0 < t_1 < t_2 < \cdots < t_n$ and $A_k \in \mathcal{B}^1$,

$$P(B_{t_k} \in A_k) = \int_{A_1} dy_1 p_{t_1}(0, y_1) \int_{A_2} dy_2 p_{t_2 - t_1}(y_1, y_2) \cdots \int_{A_n} dy_n p_{t_n - t_{n-1}}(y_{n-1}, y_n).$$

In fact, by the independent increments letting $t_0 = 0$, we have

$$P(B_{t_k} - B_{t_{k-1}} \in A_k, k = 1, 2, \ldots, n) = \prod_{k=1}^n \int_{A_k} p_{t_k - t_{k-1}}(x_k) dx_k$$

and by the change of variables $x_k = y_k - y_{k-1}$ ($y_0 = 0$) we get the above equation. Here note that $\{B_{t_1} \in A_1, B_{t_2} \in A_2\} == \{B_{t_1} \in A_1, B_{t_2} - B_{t_1} \in A_2 - A_1\}$ where $A_2 - A_1$ is a family of differences of elements, and this is not the difference set . $A_2 \setminus A_1$

In the following let $(\mathcal{F}_t)$ is a standard filtration by the Brownian motion (B_t).

[Properties of Brownian motions]

(1) $EB_t^{2n} = (2n-1)!!t^n, EB_t^{2n-1} = 0$ $(n \geq 1)$.

(2) For $0 \leq s < t$, $B_t - B_s$ is independent of $\mathcal{F}_s$.

This is equivalent to independent increments. From this (B_t) is martingale (described latter), i.e., $0 \leq s < t \Rightarrow E[B_t - B_s |\ \mathcal{F}_s] = 0$

(3) The covariance $E[B_t B_s] = t \wedge s$ $(s, t > 0)$.

(4) A continuous process (X_t) is a Brownian motion $\iff$ $^\forall 0 \leq s < t$, $E[e^{iz(X_t - X_s)} |\ \mathcal{F}_s] = e^{-(t-s)z^2/2}$, where $(\mathcal{F}_t)$ is the canonical filtration by (X_t).

(5) The Brownian motion is invariant under the following transforms ($a > 0$ is a fixed):

$$B_t^a = B_{a+t} - B_a,\ \overline{B_t} = -B_t,\ S^a(B)_t = \sqrt{a} B_{t/a},$$

where $S^a(B)_t$ is called a **scale conversion** or **scaling**.

(6) The total variation of Brownian motion in $[T_1, T_2]$ is infinite a.s., i.e., denote a division as $\Delta = \{t_k\}; T_1 = t_0 < t_1 < \cdot < t_n = T_2$, then

$$V = \sup_{\Delta} \sum_{k=1} |B_{t_k} - B_{t_{k-1}}| = \infty \quad \text{a.s.}$$

(7) $^\forall \varepsilon > 0$, (B_t) has $(1/2 - \varepsilon)$-Hödler uniform continuous paths a.s., i.e, for all $\gamma > 0$,

$$\lim_{h \to 0} \sup_{s \neq t; |t-s| \leq h} \frac{|B_t - B_s|}{|t-s|^\gamma} = 0 \text{ or } \infty \text{ a.s. if } \gamma < 1/2 \text{ or } \gamma \geq 1/2.$$

(8) Sample paths of Brownian motion are not differentiable at every time points a.s.

(9) Let (B_t) be a d-dimensional Brownian motion and T be a $d \times d$ orthogonal matrix. Then (TB_t) is also a Brownian motion. Moreover, let $\tau_S := \inf\{t > 0; B_t \in S = S_r^{d-1}\}$ be a hitting time to the sphere 球面 $S = \partial B^d(0, r)$. Then the distribution of $B_{\tau_S} = B_{\tau_S(\omega)}(\omega)$ is the uniform measure on S.

Furthermore, the Brownian motion (B_t) has the following properties:

- $X_t = tB_{1/t}$ is also a Brownian motion with $X_0 = 0$.
-
$$\limsup_{t\downarrow 0} \frac{B_t}{\sqrt{2t\log\log(1/t)}} = 1 \quad \text{a.s.}$$
Moreover, by symmetry, $\liminf_{t\downarrow 0}$ is -1, and by scaling,
$$\limsup_{t\uparrow\infty} \frac{B_t}{\sqrt{2t\log\log t}} = 1 \quad \text{a.s.}$$
- $^\forall\varepsilon > 0$, (B_t) has $(1/2-\varepsilon)$-Hödler uniform continuity a.s. as mentioned, more precisely, it satisfies the following:
$$\lim_{h\to 0} \sup_{s\neq t;|t-s|\leq h} \frac{|B_t - B_s|}{\sqrt{2|t-s|\log(1/|t-s|)}} = 1.$$

[Construction of Brownian motions] It is well-known that there are 3 ways, however, we give the simplest way.

It is enough to show the case of $t \in [0,1]$. Because the case of $[0,T]$ is the same, and by the uniqueness it is possible to extend to $[0,\infty)$. Let $D = \bigcup_{n\geq 1}\{k/2^n; k = 0,1,\ldots,2^n\}$ be the family of all binary rational numbers in $[0,1]$.

First, by using **Kolmogorov's Extension Theorem** to the probability space on $\mathbf{R}^\infty$, a probability P_0 can be constructed on $\mathbf{R}^D$ ($\ni w = w(t) : D] \to \mathbf{R}$ is a function) such that the every finite dimensional distribution of $X_t(w) = w(t)$ is the same as the Brownian motion.

Furthermore, it is possible to show that (X_t) satisfies the conditions of the following **Kolmogorov's Continuity Theorem**. Hence, (X_t) is uniform continuous on D a.s., and $\widetilde{X_t} = \lim_{r\downarrow t; r\in D} X_r$ is continuous. Thus, $B_t = \widetilde{X_t}$ is the desired one.

Theorem 2.3 (Kolmogorov's Continuity Theorem) *(1) In general, a stochastic process $\{X_t\}_{t\in D}$ which is in a Banach space $(B, \|\cdot\|)$ satisfies*
$$^\exists C,\alpha,\beta > 0; E\|X_t - X_s\}^\alpha \leq C|t-s|^{1+\beta},$$
then X_t is uniform continuous on D a.s.

(2) If $\{X_t\}_{t\in[0,1]}$ satisfies the above inequation for $^\forall s,t \in [0,1]$, then there exists a continuous modification $\{\widetilde{X_t}\}_{t\in[0,T]}$ uniquely, and it is γ-Hölder uniform continuous a.s. for $^\forall\gamma < \beta/\alpha$; i.e.,
$$\lim_{h\to 0} \sup_{s\neq t;|t-s|\leq h} \frac{\|X_t - X_s\|^\gamma}{|t-s|} = 0 \quad \textit{a.s.}$$

If readers want to know the proofs of the results in this section, please see the text of stochastic analysis; **"Ito integrals and Stochastic Differential Equations with Jumps"**.

Here, we give a several results with respect to characteristic functions, which are needed from the next section.

Let $\mathcal{P}(\mathbf{R}^d)$ be a family of all probability measures on $\mathbf{R}^d$, i.e., distributions on $\mathbf{R}^d$.

A **characteristic function=c.f.**; $\widehat{\mu}(z) := \displaystyle\int_{\mathbf{R}^d} e^{i\langle z,x\rangle}\mu(dx)$, and a **convolution** of $\mu,\nu \in \mathcal{P}(\mathbf{R}^d)$:
$$\mu * \nu(A) := \int_{\mathbf{R}^d}\int_{\mathbf{R}^d} 1_A(x+y)\mu(dx)\nu(dy) = \int_{\mathbf{R}^d} \mu(A-y)\nu(dy) = \int_{\mathbf{R}^d} \nu(A-x)\mu(dx).$$
Then, it is easy to see $\widehat{\mu * \nu}(z) = \widehat{\mu}(z)\widehat{\nu}(z)$. The distribution of a sum of independent random variables is a convolution, i.e., if RV's X, Y are independent and their distributions are μ,ν, respectively, then the

dist. of $X+Y$ is $\mu*\nu$. Because the c.f. of $X+Y$ is $\widehat{\mu}\widehat{\nu} = \widehat{\mu*\nu}$, i.e., $E[e^{i\langle z,X+Y\rangle}] = E[e^{i\langle z,X\rangle}]E[e^{i\langle z,Y\rangle}] = \widehat{\mu}(z)\widehat{\nu}(z)$.

Note that $\mu \in \mathcal{P}(\mathbf{R}^d)$ can be expressed by $\widehat{\mu}$ (**Lévy's inversion formula**, So μ is determined by $\widehat{\mu}$ uniquely, that is, if $\widehat{\mu} = \widehat{\nu}$, then $\mu = \nu$ (**uniqueness theorem**).

We also describe the results related with the convergence of characteristic functions and distributions (for their proofs, see the text "Basics of Probability Theory").

Theorem 2.4 *Let $\mu_n, \mu \in \mathcal{P}(\mathbf{R}^d)$. If $\mu_n \to \mu$, then $\widehat{\mu}_n \to \widehat{\mu}$ (uniform on compact sets).*

Note that $\mu_n \to \mu \overset{\text{def}}{\iff} {}^\forall f \in C_b(\mathbf{R}^d)$, $\mu_n(f) := \int f d\mu_n \to \mu(f)$.

Theorem 2.5 (Lévy's Continuity Theorem) *Let $\mu_n \in \mathcal{P}(\mathbf{R}^d)$. If ${}^\exists\varphi; \widehat{\mu}_n \to \varphi$ (pointwise) and φ is continuous at the origin, then ${}^\exists\mu \in \mathcal{P}(\mathbf{R}^d)$; $\varphi = \widehat{\mu}$, $\mu_n \to \mu$, Moreover, $\widehat{\mu}_n \to \widehat{\mu}$ (uniform on compact sets).*

Corollary 2.1 (Glivenko' Theorem) *Let $\mu_n, \mu \in \mathcal{P}(\mathbf{R}^d)$. If $\widehat{\mu}_n \to \widehat{\mu}$ (pointwise), then $\mu_n \to \mu$.*

3 Lévy Processes and Infinitely Divisible Distributions

The distribution of a Lévy process has a property of infinitely divisible. By this property as a characterizing the characteristic function, it is possible to give the Lévy-Khintchine formula.

3.1 Infinitely divisible distributions

Let $\mathcal{P}(\mathbf{R}^d)$ be a family of all probability measures on $\mathbf{R}^d$, i.e., distributions on $\mathbf{R}^d$.

Definition 3.1 *$\mu \in \mathcal{P}(\mathbf{R}^d)$ is an* **infinitely divisile distribution** *if $^\forall n \geq 2, ^\exists \mu_n \in \mathcal{P}(\mathbf{R}^d) : \mu = \mu_n^{n*}$, i.e., $\widehat{\mu} = \widehat{\mu_n}^n$. The family of all these distributions is denoted as $I(\mathbf{R}^d)$.*

This is equivalent to that if we denote the characteristic function of μ as $\widehat{\mu}$, then $^\forall n \geq 2$, $\widehat{\mu}^{1/n}$ is a characteristic function, where n-root of $\widehat{\mu}$; $\widehat{\mu}^{1/n}$ is determined by the following.

A uniform distribution and a binary distribution are not infinitely divisible. An infinitely divisible distribution with a bounded support is only a δ distribution.

In the following, we give several properties of infinitely divisible distributions.

· If $\mu \in I(\mathbf{R}^d)$, then $\widehat{\mu} \neq 0$.

(Pr.) By the definition; $\widehat{\mu_n}^n = \widehat{\mu}$,

$$\varphi(z) := \lim_{n\to\infty} |\widehat{\mu_n}(z)|^2 = \lim_{n\to\infty} |\widehat{\mu}(z)|^{2/n} = 1_{\{\widehat{\mu}(z)\neq 0\}}.$$

If we set $\mu_-(dx) := \mu(-dx)$: a **dual** of μ, $\mu_2 := \mu * \mu_-$: a **symmetrization** of μ, then $\widehat{\mu_-} = \widehat{\mu}(-\cdot) = \overline{\widehat{\mu}}$, $\widehat{\mu_2} = |\widehat{\mu}|^2$. By $\widehat{\mu}(0) = 1$ and the continuity of $\widehat{\mu}$, we have $\varphi = 1$ on a neighborhood of $z = 0$. Hence, by Lévy's continuity theorem, φ is also a characteristic function, and thus, it is continuous on $\mathbf{R}^d$. So $\varphi \equiv 1$ and $\widehat{\mu} \neq 0$.

· For each $\mu \in I(\mathbf{R}^d)$, $^{\exists_1} f(z) : \mathbf{R}^d \to \mathbf{C}$: continuous; $f(0) = 0, \widehat{\mu}(z) = e^{f(z)}$, and $^\forall n \geq 2, ^{\exists_1} g_n(z) : \mathbf{R}^d \to \mathbf{C}$: continuous; $g_n(0) = 1, g_n(z)^n = \widehat{\mu}(z)$. From now on, we denote as $f = \log \widehat{\mu}$, $g_n = \widehat{\mu}^{1/n}$ ($g_n = e^{f/n}$). By these we can define $\widehat{\mu}^t = \exp[t \log \widehat{\mu}]$ and when this is a characteristic function (it is actually true), denote the distribution as μ^{t*}. Then $\widehat{\mu^{t*}} = \widehat{\mu}^t$ holds.

(Pr.) We can show in more general by changing $\widehat{\mu}$ to $\varphi : \mathbf{R}^d \to \mathbf{C}; \varphi \neq 0, \varphi(0) = 1$. Fix any $z \in \mathbf{R}^d$ and for $t \in [0, 1]$, we choose a branch $h_z(t) = \log |\varphi(tz)| + i \arg \varphi(tz)$ of a complex function $\varphi(tz)$ such that $h_z(t)$ is continuous and $h_z(0) = 0$. $h_z(t)$ is unique and $\arg \varphi(tz)$ is a chosen argument such that it is continuous and 0 if $t = 0$. We define $f(z) = h_z(1) = \log |\varphi(z)| + i \arg \varphi(z)$ and show the continuity of this. Fix z_0, and for $z \neq z_0$, let $w_z(t) : [0, 3] \to \triangle(0, z_0, z)$ be continuous such that $w_z(t) = 0, z_0, z, 0$ if $t = 0, 1, 2, 3$ and having the triangle graph of $0, z_0, z$. Since $\{\varphi(tz_0); t \in [0, 1]\}$ is compact and $\varphi \neq 0$, ii has a positive distance to 0. If $z \to z_0$, then $\max_{0\leq t\leq 1} |\varphi(tz) - \varphi(tz_0)| \to 0$. hence, $^\exists U(z_0)$: a nbd of z_0; $^\forall z \in U(z_0)$, the rotation number of the closed curve $\{\varphi(w_z(t)); t \in [0, 3]\}$ around of the origin is 0. Thus, $\arg \varphi(w_z(3)) = 0$. Therefore, $\mathrm{Im}\, f(z) = \arg \varphi(z) = \arg \varphi(w_z(2))$ ($^\forall z \in U(z_0)$) で and if $z \to z_0$, then $\mathrm{Im}\, f(z) \to \mathrm{Im}\, f(z_0)$. The continuity of $\mathrm{Re}\, f(z)$ is clear. So $f(z)$ is continuous. On the other hand, if $\widetilde{f}(z)$ is continuous; $\widetilde{f}(0) = 0, e^{\widetilde{f}(z)} = \varphi(z)$, then by the uniqueness of h_z, we have $h_z(t) = \widetilde{f}(tz)$ and $\widetilde{f}(z) = h_z(1) = f(z)$. Moreover, for n-root g_n of $\widehat{\mu}$, it is a similar. ■

· For $\mu \in I(\mathbf{R}^d)$, the distribution μ_n such that $\mu = \mu_n^{n*}$ is unique and satisfies $\widehat{\mu_n} = \widehat{\mu}^{1/n}$, That is, $\mu_n = \mu^{1/n*}$.

(Pr.) By $\widehat{\mu} \neq 0$ and the result in the above proof, it is clear. ■

· $\mu_n \in I(\mathbf{R}^d) \to \mu \Longrightarrow \mu \in I(\mathbf{R}^d)$.

(Pr.) For $^\forall k \geq 2$, it is enough to show $\widehat{\mu}^{1/k}$ is also a c.f. . We first show $\widehat{\mu} \neq 0$. $\widehat{\mu}_n \to \widehat{\mu}$ implies $|\widehat{\mu}_n|^{2/k} \to |\widehat{\mu}|^{2/k}$. Since $|\widehat{\mu}_n|^{2/k} = |\widehat{\mu}_n^{1/k}|^2$ is a char.ft and $|\widehat{\mu}|^{2/k}$ is continuous, this is also a char.ft. Hence, the distribution with the c.f. $|\widehat{\mu}|^2$ is in $I(\mathbf{R}^d)$. Thus, $\widehat{\mu} \neq 0$. Therefore, as in the above, $\widehat{\mu}^{1/k}$ exists uniquely and it is continuous. By $\widehat{\mu_n} \to \widehat{\mu}$, $\widehat{\mu_n}^{1/k} \to \widehat{\mu}^{1/k}$. Hence, $\widehat{\mu}^{1/k}$ is also a c.f. ■

· If $\mu_1, \mu_2 \in I(\mathbf{R}^d)$, then $\mu_1 * \mu_2 \in I(\mathbf{R}^d)$.

(Pr.) By $\mu_1 = (\mu_{1,n})^{n*}, \mu_2 = (\mu_{2,n})^{n*}$, we have $\mu_1 * \mu_2 = (\mu_{1,n} * \mu_{2,n})^{n*}$. ■

· If $\mu \in I(\mathbf{R}^d)$, then $^\forall t \geq 0, \mu^{t*}$ is defined and $\mu^{t*} \in I(\mathbf{R}^d)$.

(Pr.) By $\widehat{\mu}^{1/m} = (\widehat{\mu}^{1/(mn)})^n \in I(\mathbf{R}^d)$, $\widehat{\mu}^{n/m} \in I(\mathbf{R}^d)$. Letting $r_n \in \mathbf{Q}_+ \to t > 0$, we have $\widehat{\mu}^{r_n} \to \widehat{\mu}^t$, and since $\widehat{\mu}^t$ is continuous, ${}^{\exists 1}\mu_t \in \mathcal{P}(\mathbf{R}^d); \widehat{\mu_t} = \widehat{\mu}^t$. Therefore, $\mu^{t*} \in I(\mathbf{R}^d)$. ■

Theorem 3.1 *Let (X_t) be a Lévy process in law. The distribution of X_t is $\mu_t = P \circ X_t^{-1} \in I(\mathbf{R}^d)$ and if it is denoted as $\mu_1 = \mu$, then $\mu_t = \mu^{t*}$. On the other hand, if $\mu \in I(\mathbf{R}^d)$, then ${}^{\exists}(X_t)$: a Lévy process in law such that $X_t \overset{(\mathrm{d})}{=} \mu^{t*}$ and this is unique except equivalence in law, that is, if (Y_t) satisfies the same conditions, then it is equivalent to (X_t) in law, i.e., they have the same finite-dimensional distribution; $(X_{t_1}, X_{t_2}, \ldots, X_{t_n}) \overset{(\mathrm{d})}{=} (Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n})$.*

Proof. For $t > 0$, set $t_k^n = kt/n$. $t_0^n = 0$ and by $X_0 = 0$, $X_t = \sum_{k=1}^n (X_{t_k^n} - X_{t_{k-1}^n})$, and independent increments and time homogeneity imply $\mu_t \in I(\mathbf{R}^d)$. By $X_1 \overset{(\mathrm{d})}{=} \mu = \mu_1 \in I(\mathbf{R}^d)$, $X_{1/n} \overset{(\mathrm{d})}{=} \mu_{1/n} = \mu^{1/n*}$ and $X_{m/n} \overset{(\mathrm{d})}{=} \mu^{m/n*}$. Hence, approximating by rational numbers, we have ${}^{\forall}t > 0, X_t \overset{(\mathrm{d})}{=} \mu^{t*}$.

On the inverse, in order to show the existence of Lévy process in law corresponding to $\mu \in I(\mathbf{R}^d)$, we use the following Kolmogorov's Extension Theorem. For $0 \le t_1 < t_2 < \cdots < t_n$, $B_k \in \mathcal{B}^1$, $k = 1, 2, \ldots, n$, we define

$$\begin{aligned}&\mu_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n)\\ &:= \int_{\mathbf{R}} \mu^{t_1*}(dy_1)1_{B_1}(y_1) \int_{\mathbf{R}} \mu^{t_2-t_1*}(dy_2)1_{B_2}(y_1+y_2)\cdots \int_{\mathbf{R}} \mu^{t_n-t_{n-1}*}(dy_n)1_{B_n}(y_1+\cdots+y_n).\end{aligned}$$

This satisfies the consistency condition by $\mu^{s*} * \mu^{t*} = \mu^{s+t*}$. Thus, ${}^{\exists 1}P$: a probability measure on $\Omega = (\mathbf{R}^d)^{[0,\infty)}$; for $X_t(\omega) := \omega(t)$, $X_t \overset{(\mathrm{d})}{=} \mu^{t*}$. Moreover,

$$E\left[e^{i\sum_{k=1}^n \langle z_k, X_{t_k} - X_{t_{k-1}}\rangle}\right] = \prod_{k=1}^n \int_{\mathbf{R}} e^{i\langle z_k, y_k\rangle} \mu^{t_k - t_{k-1}*}(dy_k) = \prod_{k=1}^n E\left[e^{i\langle z_k, X_{t_k} - X_{t_{k-1}}\rangle}\right]$$

and hence, (X_t) has independent increments. (The last equation can be obtained by letting 0 except z_k in the previous equation. Furthermore, the continuity in probability is clear, by the following: as $t \downarrow 0$,

$$P(|X_t| \ge \varepsilon) \to 0 \iff \mu_t \to \delta_0 \iff \widehat{\mu}(z)^t \to 1,$$

and $\mu \in I(\mathbf{R}^d)$ has no zero point. Finally, if (Y_t) satisfies the same conditions, then $X_t - X_s \overset{(\mathrm{d})}{=} Y_t - Y_s \overset{(\mathrm{d})}{=} \mu^{t-s*}$ and $(X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}) \overset{(\mathrm{d})}{=} (Y_{t_0}, Y_{t_1} - Y_{t_0}, \ldots, Y_{t_n} - Y_{t_{n-1}})$. Hence, $(X_{t_0}, X_{t_1}, \ldots, X_{t_n}) \overset{(\mathrm{d})}{=} (Y_{t_0}, Y_{t_1}, \ldots, Y_{t_n})$. ■

Theorem 3.2 (Kolmogorov's Extension Theorem) *Let $\Omega = (\mathbf{R}^d)^{[0,\infty)} \ni \omega$ and $X_t(\omega) := \omega(t)$. Let $\mathcal{F}$ be a Kolmogorov σ-additive class, i.e., σ-additive class generated by all cylinder sets; $C = \{X_{t_k} \in B_k, k =, 1, \ldots, n\}$. For every $0 \le t_1 < t_2 < \cdots < t_n$, distributions $\mu_{t_1,\ldots,t_n}$ on $\mathcal{B}((\mathbf{R}^d)^n)$ are given and satisfy the following* **consistency condition***: For $B_1, \ldots, B_n \in \mathcal{B}^1$, if $B_k = \mathbf{R}^d$ for one $k = 1, 2, \ldots, n$, then*

$$\mu_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n) = \mu_{t_1,\ldots,t_{k-1},t_{k+1},\ldots,t_n}(B_1 \times \cdots \times B_{k-1} \times B_{k+1} \times \cdots \times B_n)$$

Under the above conditions, ${}^{\exists}P$: a probability measure on $(\Omega, \mathcal{F})$; $(X_{t_1}, \ldots, X_{t_n}) \overset{(\mathrm{d})}{=} \mu_{t_1,\ldots,t_n}$.

The proof is that on a total family of cylinder sets $\mathcal{C}$, we define $Q(C) := \mu_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n)$ for $C = \{X_{t_k} \in B_k, k =, 1, \ldots, n\} \in \mathcal{C}$. Then, $Q : \mathcal{C} \to [0, 1]$; $Q((\mathbf{R}^d)^{[0,\infty)}) = 1$ and satisfies finite additivity. Thus, it is sufficient to show the continuity at $\emptyset$, i.e., if $A_n \in \mathcal{C}; A_n \downarrow \emptyset$, then $Q(A_n) \to 0$. (Because, by the extension theorem of measures, there exists a unique probability measure P on $\mathcal{F} = \sigma(\mathcal{C})$ such that $P = Q$ on $\mathcal{C}$.) On the continuity at $\emptyset$, if we assume $Q(A_n) \downarrow \delta > 0$, then by the regularity of $\mu_{t_1,\ldots,t_n}$, we can take a compact set of $B_1 \times \cdots \times B_n$ and we can show that $\bigcap A_n \neq \emptyset$. This contradicts.

For details, see **I. Karatzas & S. E. Shreve, "Brownian motions and Stochastic Integrals", Springer (1988, 1993)**.

3.2 Lévy-Khintchine representations

Theorem 3.3 (LK representation) (X_t) が *is a Lévy process if equivalent to that* $^\forall t \geq 0$, *the characteristic function* $\widehat{\mu}_t(z) := E[e^{i\langle z, X_t\rangle}]$ $(i = \sqrt{-1})$ *of* X_t *has the following* **Lévy-Khintchine (LK) representation**: $\widehat{\mu}_t(z) = e^{t\psi(z)}$;

$$\psi(z) = -\frac{1}{2}\langle Az, z\rangle + + \int_{\mathbf{R}^d} (e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle 1_{\{|x|<1\}}\nu(dx) + i\langle \gamma, z\rangle,$$

where

· $A = (a_{jk})_{j,k\leq d}$ *is a non-negative definite symmetric matrix.*

This is equivalent to that $^\exists\sigma = (\sigma^j_\ell)_{\ell\leq m, j\leq d}$; $a_{jk} = \sum_{\ell\leq m} \sigma^j_\ell \sigma^k_\ell$ *(→ the next Question).*

· $\nu = \nu(dx)$ *is called a* **Lévy measure** *on* $\mathbf{R}^d$ *satisfying that* $\nu(\{0\}) = 0$ *and* $\int_{\mathbf{R}^d} (1 \wedge |x|^2)\nu(dx) < \infty$.

· $\gamma = (\gamma_j)_{j\leq d} \in \mathbf{R}^d$,

This triplet (A, ν, γ) *is determined uniquely.*

If ν satisfies $\int_{|x|\leq 1} |x|\nu(dx) < \infty$, then $\psi(z) = -\frac{1}{2}\langle Az, z\rangle + \int_{\mathbf{R}^d}(e^{i\langle z,x\rangle} - 1)\nu(dx) + i\langle\gamma_0, z\rangle$, where $\gamma_0 = \gamma - \int_{|x|<1} x\nu(dx)$ and γ_0 is called a **drift**.

Question. Show the above expression of A.

Let $U = (u_{jk})$ be an orthogonal matrix diagonalizing A, and its eigen valued be $\lambda_k \geq 0$ $(k \leq d)$. Since ${}^tUAU =$ diag (λ_ℓ), i.e., $A = U$diag $(\lambda_\ell){}^tU$, we have $a_{jk} = \sum_{\ell\leq d} \lambda_\ell u_{j\ell}u_{k\ell}$. If m number of eigen values are positive, i.e., for $\ell \leq m$, let $\lambda_{k_\ell} > 0$ and for each $j \leq d$, set $\sigma^j_\ell = \sqrt{\lambda_{k_\ell}} u_{jk_\ell}$. Then we have $a_{jk} = \sum_{\ell\leq m}\sigma^j_\ell\sigma^k_\ell$.

In other words of μ, the above theorem is equivalent to that

$$\mu \in I(\mathbf{R}^d) \iff \widehat{\mu}(z) = e^{\psi(z)}$$

In the characteristic function of a compound Poisson distribution

$$\psi(z) = \log\widehat{\mu}(z) = c(\widehat{\sigma} - 1) = c\int_{\mathbf{R}^d}(e^{i\langle z,x\rangle} - 1)\sigma(dx)$$

if we set $A = 0$, $\nu = c\sigma$ and $\gamma = c\int_{|x|<1} x\sigma(dx)$, then we have the LK representation.

[Proof of LK representation].

We first show there exists a distribution with this characteristic function $\varphi := e^\psi$ and it is an infinitely divisible distribution.

Let ψ_n be ψ without jumps of $|x| \leq 1/n$. Then $\varphi_n = e^{\psi_n}$ is a characteristic function of a convolution of a Gaussian distribution and a compound Poisson distribution. Hence $^\exists\mu_n \in I(\mathbf{R}^d)$; $\widehat{\mu_n} = \varphi_n \to \varphi$. Since φ is continuous, $^{\exists_1}\mu \in \mathcal{P}(\mathbf{R}^d)$; $\widehat{\mu} = \varphi$. Therefore, $\mu_n \to \mu$ and $\mu \in I(\mathbf{R}^d)$.

Next, on the uniqueness of the representation, let $\psi(z) = \log\varphi(z)$ have the representation by (A, ν, γ). Since

$$\frac{1}{s^2}|e^{i\langle sz,x\rangle} - 1 - i\langle sz, x\rangle| \leq \frac{1}{2}|z|^2|x|^2, \quad \to 0\ (s \to \infty),$$

by Lebesgue's convergence theorem,

$$\lim_{s\to\infty}\frac{1}{s^2}\psi(sz) = -\frac{1}{2}\langle z, Az\rangle.$$

Hence, A is determined by μ and unique. Let $\psi_d(z) = \psi(z) + \langle z, Az\rangle/2$ and set $C = [-1,1]^d$. It can be seen that

$$\int_C (\psi_d(z) - \psi_d(z+w))dw = \int_{\mathbf{R}^d} e^{i\langle z,x\rangle}\rho(dx) \quad \text{with} \quad \rho(dx) = 2^d \left(1 - \prod_{j=1}^d \frac{\sin x_j}{x_j}\right)\nu(dx).$$

By this and by $\rho(dx) \le C(1 \wedge |x|^2)\nu(dx)$ ($\to$ the next question), ρ is a finite measure and its Fourier transform is the above equation. Thus, ρ is determined by ψ_d, that is, ν is determined by μ uniquely. Therefore, γ is also unique. On the above transform, set $D = \{|x| < 1\}$, we have

$$\int_C (\psi_d(z) - \psi_d(z+w))dw = \int_C dw \int_{\mathbf{R}^d} (e^{i\langle z,x\rangle} - e^{i\langle z+w,x\rangle} + i\langle w,x\rangle 1_D(x))\nu(dx)$$

and on D by adding and subtracting $i\langle w,x\rangle e^{i\langle z,x\rangle}$, we have

$$|e^{i\langle z,x\rangle} - e^{i\langle z+w,x\rangle} + i\langle w,x\rangle| \le |1 - e^{i\langle w,x\rangle} + i\langle w,x\rangle)| + |i\langle w,x\rangle(e^{i\langle z,x\rangle} - 1)| \le \frac{1}{2}|w|^2|x|^2 + |w||z||x|^2$$

Hence, it is possible to change integrals on dw and $\nu(dx)$. Furthermore, by

$$\int_C (e^{i\langle z,x\rangle} - e^{i\langle z+w,x\rangle} + i\langle w,x\rangle 1_D(x))dw = e^{i\langle z,x\rangle}\int_C (1 - e^{i\langle w,x\rangle})dw = 2^d e^{i\langle z,x\rangle}\left(1 - \prod_{j=1}^d \frac{\sin x_j}{x_j}\right)$$

we can get the desired equation. ■

Question 3.1 *Show that if* $|x| \le 1$, *then* $1 - \prod_{j=1}^d \frac{\sin x_j}{x_j} \le C|x|^2$.

If $x > 0$, then $\sin x \ge x - x^3/3!$. Hence, it is clear in case of $d = 1$. In genral cases it holds by the following:

$$1 - \prod_{j=1}^d \frac{\sin x_j}{x_j} = \sum_{k=1}^d \left(1 - \frac{\sin x_k}{x_k}\right)\prod_{j=1}^{k-1} \frac{\sin x_j}{x_j}$$

[Proof of the possibility of the representation].
We define a compound Poisson distribution μ_n by

$$\widehat{\mu_n}(z) := \exp[n(\widehat{\mu}(z)^{1/n} - 1)] = \exp\left[n\int_{\mathbf{R}^d\setminus\{0\}} (e^{iz\cdot x} - 1)\mu^{1/n*}(dx)\right]$$

(note that $\mu^{1/n*}(\{0\})$ may not be 0, however, it we restrict this to $\mathbf{R}^d \setminus \{0\}$ and denote as ν_n, then we may change them and it is also a compound Poisson). As $n \to \infty$,

$$\widehat{\mu_n}(z) = \exp[n(e^{n^{-1}\log\widehat{\mu}(z)} - 1)] = \exp[n(n^{-1}\log\widehat{\mu}(z) + o(1/n)] \to \widehat{\mu}(z)$$

implies $\mu_n \to \mu$. Since μ_n has a LK-representation, and by after the next convergence theorem, μ has also a LK representation. ■

From the above proof, the following holds.

Theorem 3.4 *An infinitely divisible distribution are expressed as a limit of a sequence of compound Poisson distributions.*

In order to treat easily, we introduce the following **2nd LK representation** (A,ν,β) : Let $\theta(x)$ be a function on $\mathbf{R}^d$ such that which is 1 for $|x|\le 1$, 0 for $|x|\ge 2$, and has a graph connected by line segments between them.

$$\psi(z) = -\frac{1}{2}\langle Az,z\rangle + \int_{\mathbf{R}^d}(e^{i\langle z,x\rangle}-1-i\langle z,x\rangle\theta(x))\nu(dx) + i\langle \beta,z\rangle.$$

We simply call an (A,ν,β) **representation**.

The 1st representation and the 2nd representation are equivalent and it can be rewritten to each other.

Theorem 3.5 (Convergence theorem of LK representations) *If each $\mu_n \in I(\mathbf{R}^d)$ has an (A_n,ν_n,β_n) representations, then for a distribution μ on $\mathbf{R}^d$, $\mu_n\to\mu$ is equivalent to the following:*

$\mu\in I(\mathbf{R}^d)$ has an (A,ν,β) representation and for all bounded continuous functions f such that $f=0$ on a neighborhood of the origin,

$$\lim_{n\to\infty}\int_{\mathbf{R}^d} f(x)\nu_n(dx) = \int_{\mathbf{R}^d} f(x)\nu(dx).$$

Moreover, for $^\forall\varepsilon>0$, if a non-negative definite symmetric matrix $A_{n,\varepsilon}$ is defined by

$$\langle z, A_{n,\varepsilon}z\rangle = \langle z, A_n z\rangle + \int_{|x|<\varepsilon}\langle x,z\rangle^2\nu_n(dx) \quad (^\forall z\in\mathbf{R}^d),$$

then $\lim\limits_{\varepsilon\downarrow 0}\limsup\limits_{n\to\infty}\langle z,A_{n,\varepsilon}z\rangle = \lim\limits_{\varepsilon\downarrow 0}\liminf\limits_{n\to\infty}\langle z,A_{n,\varepsilon}z\rangle = \langle z,Az\rangle$. $\lim_{n\to\infty}\beta_n=\beta$.

Proof. $(\Rightarrow)$ We show if μ_n has 2nd representation and $\mu_n\to\mu$, then μ has also and each coefficients satisfies the convergence condition. At first, $\mu\in I(\mathbf{R}^d)$, and since $\widehat{\mu}(z)$ has no zero points, $\psi(z)=\log\widehat{\mu}(z)$ exists and by the convergence theorem of characteristic functions, we have $\psi_n(z)=\log\widehat{\mu}_n(z)\to\psi(z)$ (uniform on compact sets). Let $g(z,x):=e^{i\langle z,x\rangle}-1-i\langle z,x\rangle\theta(x)$. Then,

$$\psi_n(z) = -\frac{1}{2}\langle A_n z,z\rangle + \int_{\mathbf{R}^d} g(z,x)\nu_n(dx) + i\langle\beta_n,z\rangle.$$

If we set $\rho_n(dx):=(1\wedge|x|^2)\nu_n(dx)$, then it can be shown that

$$\sup_n \rho_n(\mathbf{R}^d)<\infty, \quad \lim_{L\to\infty}\sup_n\rho_n(|x|>L)=0. \tag{3.1}$$

This means "tightness" of $\{\rho_n\}$ and it is equivalent to be relatively compact in case of probability measures, however, it also holds in case of finite measures. Hence, $^\exists\{n_k\};\rho_{n_k}\to {}^\exists\rho$: a finite measure. Let $\nu(dx):=(1\wedge|x|^2)^{-1}1_{\{x\neq 0\}}\rho(dx)$ and for $\varepsilon>0$, if we set

$$I_{1,n}^\varepsilon(z) := \int_{|x|\ge\varepsilon} g(z,x)(1\wedge|x|^2)^{-1}\rho_n(dx),$$

$$I_{2,n}^\varepsilon(z) := \int_{|x|<\varepsilon}(g(z,x)+\frac{1}{2}\langle z,x\rangle^2)(1\wedge|x|^2)^{-1}\rho_n(dx),$$

then

$$\psi_n(z) = -\frac{1}{2}\langle A_{n,\varepsilon}z,z\rangle + I_{1,n}^\varepsilon(z)+I_{2,n}^\varepsilon(z)+i\langle\beta_n,z\rangle.$$

In the following, n is n_k, and letting $n\to\infty$ (i.e., $k\to\infty$) and for ρ-continuous $\varepsilon>0$, i.e., $\rho(|x|=\varepsilon)=0$ (precisely, this means $\{|x|<\varepsilon\}$ is a ρ-continuous set.) letting $\varepsilon\downarrow 0$, it holds

$$I_{1,n}^\varepsilon(z) \xrightarrow[n\to\infty]{} \int_{|x|\ge\varepsilon} g(z,x)\nu(dx) \xrightarrow[\varepsilon\downarrow 0]{} \int_{\mathbf{R}^d} g(z,x)\nu(dx). \tag{3.2}$$

On the other hand, for $^\forall z$, since $\left|g(z,x)+\langle z,x\rangle^2/2\right|(1\wedge|x|^2)^{-1}\le |z|^3|x|/3!\to 0$ if $|x|<\varepsilon\to 0$, then by $\sup_n\rho_n(\mathbf{R}^d)<\infty$, we have

$$\lim_{\varepsilon\downarrow 0}\sup_n|I_{2,n}^\varepsilon(z)|=0.$$

Hence, by dividing $\psi_n(z)$ to the real part and the imaginary part, we have

$$\lim_{\varepsilon\downarrow 0}\limsup_{k\to\infty}\langle z, A_{n_k,\varepsilon}z\rangle = \lim_{\varepsilon\downarrow 0}\liminf_{k\to\infty}\langle z, A_{n_k,\varepsilon}z\rangle \in \mathbf{R}, \quad \limsup_{k\to\infty}\langle \beta_{n_k}, z\rangle = \liminf_{k\to\infty}\langle \beta_{n_k}, z\rangle \in \mathbf{R}$$

and each can be expressed by ${}^{\exists}A; \langle z, Az\rangle$, ${}^{\exists}\beta; \langle \beta, z\rangle$ ($\to$ the next question). These imply $\psi(z)$ has an (A,ν,β) representation and it is unique. The convergences of coefficients hold for a sub-sequence $\{n_k\}$ and for ρ-continuous ε. However, the restriction of ε can be omitted by monotonicity of integrals, and by the uniqueness of the representation of ψ, it holds that for any sub-sequence of $\{\rho_n\}$, there exists a convergence sub-sequence and the limit is ρ, and hence, this implies $\rho_n \to \rho$. Therefore, every convergence of coefficients holds for n. It remains to show (3.1). Set $C(h) = [-h,h]^d$ and $A_n = (a_{jk}^{(n)})$. We have

$$\begin{aligned} -\int_{C(h)} \psi_n(z)dz &= \frac{1}{2}\sum_{j\le d} a_{jj}^{(n)} \int_{C(h)} z_j^2 dz - \int_{\mathbf{R}^d} \nu_n(dx) \int_{C(h)} g(z,x)dz \\ &= \frac{1}{3}2^{d-1}h^{d+2}\sum_{j\le d} a_{jj}^{(n)} + (2h)^d \int_{\mathbf{R}^d} \left(1 - \prod_{j=1}^d \frac{\sin hx_j}{hx_j}\right) \nu_n(dx) \ge 0. \end{aligned}$$

For a fixed $h>0$, (LHS) $\to -\int_{C(h)} \psi(z)dz$ as $n\to\infty$, and hence, they are bounded. Moreover, by

$$\inf_x \left(1 - \prod_{j=1}^d \frac{\sin hx_j}{hx_j}\right)(1\wedge |x|^2)^{-1} > 0$$

($\to$ the next question), $\{\rho_n\}$ is uniform bounded. By the above equation and Question 3.1, as $h\downarrow 0$,

$$\frac{1}{(2h)^d}\int_{C(h)} \psi_n(z)dz \to 0.$$

Thus, ${}^{\forall}\varepsilon>0, {}^{\exists}n_0, h_0; {}^{\forall}n \ge n_0$,

$$\int_{\mathbf{R}^d} \left(1 - \prod_{j=1}^d \frac{\sin h_0x_j}{h_0x_j}\right) \nu_n(dx) < \varepsilon.$$

If $|x| > 2\sqrt{d}/h_0$, then ${}^{\exists}j_0; |x_{j_0}| > 2/h_0$ and noting that

$$1 - \prod_{j=1}^d \frac{\sin h_0x_j}{h_0x_j} \ge 1 - \left|\frac{\sin h_0x_{j_0}}{h_0x_{j_0}}\right| \ge 1 - \frac{1}{h_0|x_{j_0}|} > \frac{1}{2},$$

we may assume $h_0>0$ is sufficiently small and we have

$$\rho_n\left(|x| > 2\sqrt{d}/h_0\right) = \frac{1}{2}\nu_n\left(|x| > 2\sqrt{d}/h_0\right) < \varepsilon \quad (n \ge n_0).$$

Therefore (3.1) is obtained.

($\Leftarrow$) We show $\mu_n \to \mu$ from the convergences of coefficients. Let ρ_n be the same as above and set $\rho(dx) = (1\wedge|x|^2)\nu(dx)$. Let $\varepsilon>0$ be ρ-continuous and letting $\varepsilon\downarrow 0$, by the assumption of the convergence of ν_n, we have the convergence of $I_{1,n}^{\varepsilon}(z)$; (3.2). Moreover, by the convergences of ν_n and $A_{n,\varepsilon}$, we have uniform boundedness of ρ_n, and hence, $\lim_{\varepsilon\downarrow 0}\sup_n |I_{2,n}^{\varepsilon}(z)| = 0$. Therefore, by considering the real part and the imaginary part of of $\psi_n(z)$, we have $\psi_n(z)\to\psi(z)$, i.e., $\widehat{\mu}_n(z) \to \widehat{\mu}(z)$ and we get the result. ■

Question 3.2 *Show that if a symmetric matrix A_n is non-negative and ${}^{\forall}z, {}^{\exists}\lim\langle z, A_nz\rangle$, then ${}^{\exists}A$: symmetric and non-negative; $\lim\langle z, A_nz\rangle = \langle z, Az\rangle$.*

Question 3.3 *Show*

$$\inf_x \left(1 - \prod_{j=1}^d \frac{\sin hx_j}{hx_j}\right)(1\wedge |x|^2)^{-1} > 0 \ (h>0) \quad \textit{and} \quad \frac{1}{(2h)^d}\int_{C(h)} \psi_n(z)dz \to 0 \ (h\downarrow 0).$$

(Let $d=1$. Use if $x>0$, then $\sin x \le x - x^3/3! + x^5/5!$ and consider $|x|<1, \ge 1$.)

4 Important Examples of Lévy Processes

We gave basic examples in §2. in this section we describe stable processes and L-processes (self-decomposable processes) as important examples.

4.1 Stable processes and stable distributions

A strictly stable process with exponent $0 < \alpha \leq 2$ is considered as an extension of a Brownian motion to a Lévy process. This has the same type scaling property as a Brownian motion That is, $X_t \stackrel{(d)}{=} t^{1/\alpha} X_1$. If $\alpha = 2$, then it is a Gaussian process with mean 0. Moreover, adding a drift, it is simply called a stable process. Their distributions of each time points are called a strictly stable distribution, a stable distribution respectively.

Definition 4.1 *A stochastic process $(X_t)_{t\geq 0}$ on $\mathbf{R}^d$ is called a* **stable process** *if it is a Lévy process and satisfies that $^\forall a > 0, ^\exists b > 0, c \in \mathbf{R}^d$; (X_{at}) and $(bX_t + ct)$ are equivalent in law, i.e., they have the same finite-dimensional distributions. Moreover, if it can be taken as $c = 0$, then it is called a* **strictly stable process**.

Furthermore, the distribution of X_1 is called as a **stable, strictly stable distribution***, respectively.*

If $X_t = \gamma t$ a.s., then it is called as a **trivial Lévy process**. Obviously, this is a strictly stable process. If a stable process is a non-trivial Lévy process, then it is called a **non-trivial stable process**.

Theorem 4.1 *A non-trivial Lévy process $(X_t)_{t\geq 0}$ on $\mathbf{R}^d$ is a stable process $\iff$ $^\forall t > 0, ^{\exists_1} a_t > 0, b_t \in \mathbf{R}^d$; $X_t \stackrel{(d)}{=} a_t X_1 + b_t$, i.e., $\widehat{\mu}(z)^t = \widehat{\mu}(a_t z) e^{ib_t \cdot z}$. Moreover, if it can ge taken as $b_t = 0$ for all $t > 0$, then it is equivalent to that it is a strictly process.*

Proof. Since $^\forall a > 0, ^\exists b > 0, ^\exists c \in \mathbf{R}^d$; (X_{at}) and $(bX_t + ct)$ are equivalent in law, letting $t = 1, a = t$, it is clear that $^\forall t > 0, ^\exists a_t, b_t; X_t \stackrel{(d)}{=} a_t X_1 + b_t$. The uniqueness follows from that for a non constant RV X, letting $aX + b \stackrel{(d)}{=} \widetilde{a}X + \widetilde{b}$, we have $a = \widetilde{a}, b = \widetilde{b}$. In fact, let $aX + b \stackrel{(d)}{=} X$ and it is enough to show $a = 1, b = 0$. (because if $\widetilde{a} \neq 0$, then $\widetilde{a}^{-1}(aX + b - \widetilde{b}) \stackrel{(d)}{=} X$). If X_1, X_2 are indep. and $\stackrel{(d)}{=} X$, then $a(X_1 - X_2) = (aX_1 + b) - (aX_2 + b) \stackrel{(d)}{=} X_1 - X_2$. Thus, $^\forall n \geq 1$, $a^n|X_1 - X_2| \stackrel{(d)}{=} |X_1 - X_2|$. If $a \neq 1$, then $X_1 - X_2 \stackrel{(d)}{=} 0$ and this implies X is a constant a.s. This contradicts. ($\to$ the next question). Hence, $a = 1$. Moreover, we have $X \stackrel{(d)}{=} X + nb$ ($^\forall n$). This implies $b = 0$ ($\to$ the next question).

About the inverse, for $^\forall a > 0$, By $X_a \stackrel{(d)}{=} a_a X_1 + b_a$, letting $b = a_a, c = b_a$, we have $X_a \stackrel{(d)}{=} bX_1 + c$, and $(X_{at}), (bZ_t + ct)$ are both Lévy processes and they have the same distribution at $t = 1$. Hence, they are equivalent in law, and therefore, it is a stable process. For the strictly stable, it is obvious. ■

Question 4.1 *Show that if X_1, X_2 are indep. and $X_1 - X_2 \stackrel{(d)}{=} 0$, then $X_1 = X_2 =$const. a.s.*

Show that $X \stackrel{(d)}{=} X + nb$ ($^\forall n$) implies $b = 0$.

(Ans.) By $P(X_1 - X_2 = 0) = 1$, $X_1 = X_2$ a.s. and they have the same distribution μ. The characteristic function of $X_1 - X_2$ is $|\widehat{\mu}(z)|^2 = 1$ and the following give the desired result.

· $|\widehat{\mu}| = 1$ (on a nbd of the origin), then $^\exists \gamma \in \mathbf{R}^d; \mu = \delta_\gamma$

In fact, it is enough to show the case of $d = 1$. For $z \neq 0$ in a nbd of the origin, $^\exists \gamma_z; \widehat{\mu}(z) = e^{i\gamma_z}$. Thus, the support of μ is in $x = (\gamma_z + 2n\pi)/z$. If the support has two points $x_1 \neq x_2$ at least, then $|x_1 - x_2| \geq 2\pi/|z|$. However, this contradicts.

In case of $X \stackrel{(d)}{=} X + nb$ ($^\forall n$), if we assume $b \neq 0$, then by taking a small set $^\exists A; \delta := P(X \in A) > 0$. Therefore, we have $1 \geq P(X \in \bigcup_{n\geq 1}(A + nb)) = \sum_{n\geq 1} P(X \in A + nb) = \infty \cdot \delta = \infty$. This contradicts. Hence, $b = 0$.

Theorem 4.2 (Existence of Exponent of Stable Process) *If (X_t) is a non-trivial stable process, then $^{\exists_1}\alpha \in (0, 2]$; $^\forall t > 0, ^{\exists_1} b_t \in \mathbf{R}^d$; $X_t \stackrel{(d)}{=} t^{1/\alpha} X_1 + b_t$, i.e., $\widehat{\mu}(z)^t = \widehat{\mu}(t^{1/\alpha} z) e^{iz \cdot b_t}$.*

If (X_t) is a non-zero strictly stable process, then $^{\exists_1}\alpha \in (0,2]$; $^{\forall}t > 0, X_t \overset{(d)}{=} t^{1/\alpha}X_1$, i.e., $\widehat{\mu}(z)^t = \widehat{\mu}(t^{1/\alpha}z)$.

Definition 4.2 *The exponent $0 < \alpha \le 2$ in the above theorem is called an exponent of a non-trivial stable process or a non-zero strictly stable process, respectively.*

An exponent of a stable distribution except δ-distribution, or a strictly stable distribution except δ_0, is defined by the corresponding exponent of a stable process or a strictly stable process.

Note that for a non-zero trivial strictly stable process the exponent is 1, however, the exponent as a stable process is not defined.

a Brownian motion on $\mathbf{R}^d$ is a strictly stable with an exponent 2, and a non constant Gauss process is a stable with an exponent 2.

[Proof of Theorem 4.2].

We first show the existence of an exponent for a strictly stable process (Y_t). Let $\eta = P \circ Y_1^{-1}$. Since for $^{\forall}t > 0, ^{\exists_1}a_t > 0$; $Y_t \overset{(d)}{=} a_tY_1$, we have $\widehat{\eta}(z)^t = \widehat{\eta}(a_tz)$. For $s > 0$,

$$\widehat{\eta}(a_{st}z) = \widehat{\eta}(z)^{st} = (\widehat{\eta}(z)^t)^s = \widehat{\eta}(a_tz)^s = \widehat{\eta}(a_sa_tz).$$

By the uniqueness we have $a_{st} = a_sa_t$ and $a_1 = 1$. moreover, the continuity of a_t in $t > 0$ can be shown. Hence, $^{\exists}\beta; a_t = t^\beta$ (see the next question), Moreover, $\beta > 0$ can be also shown. Hence, we may set $\alpha := 1/\beta$. The uniqueness of a_t implies the uniqueness of α.

On the continuity of a_t, if we let $t_n \to t$, then $\widehat{\eta}(a_{t_n}z) = \widehat{\eta}(z)^{t_n} \to \widehat{\eta}(z)^t = \widehat{\eta}(a_tz)$. If $a_n \to 0$, then $\widehat{\eta}(z)^t = \widehat{\eta}(0) = 1$ and $Y = 0$ a.s. This contradicts $Y \neq 0$ a.s. If $a_n \to \infty$, then $\widehat{\eta}(z) = \widehat{\eta}(a_{t_n}^{-1}z)^{t_n} \to \widehat{\eta}(0)^t = 1$, and this contradicts. If $a_{t_n} \to a \in (0,\infty)$, then we have $\widehat{\eta}(az) = \widehat{\eta}(z)^t = \widehat{\eta}(a_tz)$ and $a = a_t$ by the uniqueness. Therefore, we see the continuity and $0 < a_t < \infty$ (more precisely, if we consider $\limsup, \liminf$ and take sub-sequences with the same limits respectively, then all the above hold for these sub-sequences. Hence, both limits are the same as $a_t \in (0,\infty)$). Moreover, we have $a_t = t^\beta$. If $\beta < 0$, then $a_t \to \infty$ as $t \downarrow 0$ and this contradicts as above. If $\beta = 0$, then $a_t = 1$, $\widehat{\eta}(z)^t = \widehat{\eta}(z)$ and by letting $t \downarrow 0$ we have $\widehat{\eta}(z) \equiv 1$, this contradicts. Thus, $\beta > 0$. Therefore, we can set $\alpha := 1/\beta$.

In case of a stable process (X_t), by considering the symmetrization $Y_t = X_t - \widetilde{X}_t$, where $\widetilde{X}_t \overset{(d)}{=} X_t$ and indep. of X_t, the previous theorem implies $^{\forall}t > 0, ^{\exists_1}a_t > 0, b_t \in \mathbf{R}^d$; $X_t, \widetilde{X}_t \overset{(d)}{=} a_tX_1 + b_t$ and they are non-trivial. Hence, (Y_t) is a non-zero strictly stable process. Therefore, the desired result is obtained as follows; Let $X_1 \overset{(d)}{=} \mu, Y_1 \overset{(d)}{=} \eta$. We have $\widehat{\eta}(z) = |\widehat{\mu}(z)|^2$ and

$$|\widehat{\mu}(z)|^{2t} = \widehat{\eta}(z)^t = \widehat{\eta}(t^{1/\alpha}z) = |\widehat{\mu}(t^{1/\alpha}z)|^2.$$

Hence, $^{\exists}\widetilde{b_t} \in \mathbf{R}^d; \widehat{\mu}(z)^t = e^{iz\cdot\widetilde{b_t}}\widehat{\mu}(t^{1/\alpha}z)$ and the uniqueness of b_t in the previous theorem implies $\widetilde{b_t} = b_t$.

It remains to show $\alpha \le 2$. Let (A,ν,γ) be a triplet of μ and we define ν_t by $\nu_t(dx) := \nu(t^{-1/\alpha}dx)$. Since the c.f.s of $t^{1/\alpha}X_1 + b_t$ and X_t are the same, we have

$$tA = t^{2/\alpha}A, \quad t\nu = \nu_t$$

(we also have $t^{1/\alpha}\gamma + b_t = t\gamma$, i.e., $b_t = (t - t^{1/\alpha})\gamma$). Hence, if $\alpha \neq 2$, then $A = 0$. Moreover, if $\alpha > 2$,then by $1 - 2/\alpha > 0$, $x = t^{-1/\alpha}x'$ and $\nu(t^{-1/\alpha}dx) = \nu_t(dx) = t\nu(dx)$, we have $^{\forall}a > 0$,

$$\int_{|x|<a} |x|^2\nu(dx) = t^{-2/\alpha}\int_{|x|<t^{1/\alpha}a} |x|^2\nu(t^{-1/\alpha}dx) = t^{1-2/\alpha}\int_{|x|<t^{1/\alpha}a} |x|^2\nu(dx) \to 0 \ (t \downarrow 0).$$

Thus, $\nu = 0$. That is, $X_1 = b_1 + \gamma$ and this contradicts X_t is non-trivial. Therefore, $\alpha \le 2$. ■

Question. Show that if $a_t > 0$ is continuous in $t > 0$ and satisfies $a_{st} = a_sa_t$ and $a_1 = 1$, then $^{\exists}\beta; a_t = t^\beta$.

Let $\beta := \log a_e$, i.e., $e^\beta = a_e$. $^{\forall}t > 0$, $a_{t^n} = a_t^n$ and $a_{t^{1/n}} = a_t^{1/n}$ by $a_{t^{1/n}}^n = a_t$. Hence, $^{\forall}r \in \mathbf{Q}, a_{t^r} = a_t^r$. The continuity implies $^{\forall}x \in \mathbf{R}, a_{t^x} = a_t^x$. Set $e^x = t$, then $a_t = a_{e^x} = a_e^x = e^{\beta x} = e^{\beta \log t} = t^\beta$.

The following result can be shown by using the notion of *"type equivalence"*. However in this text, we omit the proof.

Theorem 4.3 *$^\exists(S_n)$: a partial sum of i.i.d. RVs $\{Z_k\}$, i.e., a RW (random walk), If $^\exists a_n > 0, b_n \in \mathbf{R}^d$; $a_nS_n+b_n \to \mu$ in law, then μ is a stable distribution. On the other hand, the inverse holds, i.e., if μ is stable, then it is a limit distribution of the above form, more exactly, let $Z_k \overset{(\mathrm{d})}{=} \mu$, then $^\exists a_n > 0, b_n \in \mathbf{R}^d$; $a_nS_n + b_n \overset{(\mathrm{d})}{=} \mu$.*

In the next, we consider the representations of characteristic functions of stable distributions.

Theorem 4.4 (Representation of Stable Distribution) *Let $\mu \in I(\mathbf{R}^d)$, $\neq \delta$, with a triplet (A, ν, γ).*

(1) μ is a 2-stable distribution $\iff \nu = 0$.

(2) Let $0 < \alpha < 2$. μ is an α-stable distribution $\iff A = 0$, $^{\exists_1}\lambda(d\xi) \neq 0$: a finite measure on $S = S^{d-1}$;

$$\nu(dx) = \int_S \lambda(d\xi) \int_0^\infty 1_{dx}(r\xi) r^{-1-\alpha} dr.$$

That is, μ has the following **1st representation***: $\widehat{\mu}(z) = e^{t\psi(z)}$;*

$$\psi(z) = \int_S \lambda(d\xi) \int_0^\infty \left(e^{i\langle z, r\xi\rangle} - 1 - i\langle z, r\xi\rangle 1_{(0,1)}(r)\right) r^{-1-\alpha} dr + i\langle \gamma, z\rangle.$$

Moreover, it has the following **2nd representation***: For $z = |z|\zeta \in \mathbf{R}^d$,*
in case of $\alpha \neq 1$,

$$\psi(z) = -|z|^\alpha \int_S \left(1 - \tan\frac{\pi\alpha}{2} \mathrm{sgn}\ \langle \zeta, \xi\rangle\right) |\langle \zeta, \xi\rangle|^\alpha \lambda(d\xi) + i\langle \gamma_0, z\rangle,$$

in case of $\alpha = 1$,

$$\psi(z) = -|z| \int_S \left(|\langle \zeta, \xi\rangle| + \frac{2}{\pi}\langle \zeta, \xi\rangle \log|\langle z, \xi\rangle|\right) \lambda(d\xi) + i\langle \gamma_0, z\rangle.$$

In these representations $\lambda, \gamma, \gamma_0$ are unique.

The following is immediately obtained:

Theorem 4.5 (Representation of Strictly Stable Distribution) *Let $\mu \in I(\mathbf{R}^d)$, $\neq \delta_0$ and $0 < \alpha \leq 2$.*
μ is an α-strictly stable distribution $\iff$

(1) If $\alpha = 2$, then μ is a non δ_0 Gaussian distribution with mean 0.

(2) If $0 < \alpha < 2$, then μ has the following **1st representation***: $^{\exists_1}\lambda(d\xi)$: a finite measure on $S = S^{d-1}$; $\lambda \neq 0$ if $\alpha \neq 1$ and satisfying that*

(i) in case of $0 < \alpha < 1$,

$$\widehat{\mu}(z) = \exp\left[\int_S \lambda(d\xi) \int_0^\infty \left(e^{i\langle z, r\xi\rangle} - 1 - i\langle z, r\xi\rangle 1_{(0,1)}(r)\right) r^{-1-\alpha} dr\right],$$

(ii) in case of $1 < \alpha < 2$,

$$\widehat{\mu}(z) = \exp\left[\int_S \lambda(d\xi) \int_0^\infty \left(e^{i\langle z, r\xi\rangle} - 1 - i\langle z, r\xi\rangle\right) r^{-1-\alpha} dr\right],$$

(iii) in case of $\alpha = 1$, $^{\exists_1}\gamma \in \mathbf{R}^d$;

$$\widehat{\mu}(z) = \exp\left[\int_S \lambda(d\xi)\int_0^\infty \left(e^{i\langle z, r\xi\rangle} - 1 - i\langle z, r\xi\rangle 1_{(0,1)}(r)\right) r^{-2}dr + i\langle \gamma, z\rangle\right]$$

and ($\lambda = 0$ *is possible*)

$$\int_S \xi\lambda(d\xi) = 0, \quad \lambda(S) + |\gamma| > 0.$$

Moreover, it also has the **2nd representation** *which is the same as the 2nd representation of a stable distribution and satisfying the following:*

- *In case of* $\alpha \neq 1$, $\gamma_0 = 0$ ($\lambda \neq 0$).
- *In case of* $\alpha = 1$, $\int_S \xi\lambda(d\xi) = 0$, $|\gamma_0| + \lambda(S) > 0$.

[Proof of Theorem 4.4]. Let μ be an α-stable dist,, X_t be a stable process corresponding to it. As in the proof of the existence of an exponent, We know that $tA = t^{2/\alpha}A$, $t\nu = \nu_t$ ($\nu_t(dx) = \nu(t^{-1/\alpha}dx)$), and if $\alpha = 2$, then $\nu = 0$, if $\alpha < 2$, then $A = 0$. Let

$$\lambda(d\xi) := \alpha\nu((1,\infty)d\xi)$$

on $S = \mathbf{S}^{d-1}$. It is a finite measure and let ν' be given by the RHS of (2) in the theorem by using the above λ, i.e.,

$$\nu'(dx) = \int_S \lambda(d\xi)\int_0^\infty 1_{dx}(r\xi)r^{-1-\alpha}dr.$$

Then $\nu' = \nu$ holds. In fact, for $^\forall a > 0, C \in \mathcal{B}(S)$, by $a^{-\alpha}\nu(dx) = \nu_{a^{-\alpha}}(dx) = \nu(adx)$, we have

$$\nu'((a,\infty)C) = \lambda(C)\int_a^\infty r^{-1-\alpha}dr = \frac{1}{\alpha}a^{-\alpha}\lambda(C) = a^{-\alpha}\nu((1,\infty)C) = \nu((a,\infty)C).$$

Since λ is determined by ν, it is unique, and γ, γ_0 are so, too. The inverse is also clear.

On the 2nd representation, it is possible to get by using the following:

$$\int_0^\infty (e^{ir} - 1)r^{-1-\alpha}dr = \Gamma(-\alpha)e^{-i\pi\alpha/2} \quad (0 < \alpha < 1).$$

$$\int_0^\infty (e^{ir} - 1 - ir)r^{-1-\alpha}dr = \Gamma(-\alpha)e^{-i\pi\alpha/2} \quad (1 < \alpha < 2).$$

$$\int_0^\infty (e^{izr} - 1 - izr1_{(0,1)}(r))r^{-2}dr = -\frac{\pi}{2}z - iz\log z + icz \ (z > 0),$$

where

$$c = \int_1^\infty \sin r\ \frac{dr}{r^2} + \int_0^1 (\sin r - r)\frac{dr}{r^2}.$$

■

On the last equations of the above proof if $0 < \alpha < 1$, then by

$$\int_0^\infty (e^{-ur} - 1)r^{-1-\alpha}dr = \int_0^\infty drr^{-1-\alpha}\int_0^u (-re^{-tr})dt = -\int_0^u dt\ t^{\alpha-1}\int_0^\infty s^{(1-\alpha)-1}e^{-s}ds$$

$= -\alpha^{-1}\Gamma(1-\alpha)u^\alpha = \Gamma(-\alpha)u^\alpha$, for $w \in \mathbf{C}; \neq 0, \mathrm{Re}\, w \leq 0$, we have

$$\int_0^\infty (e^{wr} - 1)r^{-1-\alpha}dr = \Gamma(-\alpha)(-w)^\alpha$$

with branching $(-w)^\alpha = |w|^\alpha e^{i\alpha\arg(-w)}$; $\arg(-w) \in (-\pi, \pi)$. In fact, both sides are regular on $\mathrm{Re}\, w < 0$, continuous on $\mathrm{Re}\, w \leq 0, w \neq 0$, equal for negative numbers, and hence, on $\mathrm{Re}\, w \leq 0, w \neq 0$. Thus, we have the 1st equation. On the 2nd equation, by integration by parts it is reduced to the 1st one. On the last equation, by $\int_0^\infty r^{-2}(1 - \cos r)dr = \pi/2$, it can be calculated directly.

Theorem 4.6 (X_t) *is a rotation invariant* α*-stable process* $(0 < \alpha \leq 2)$ $\iff$ $^\exists c > 0; E[e^{i\langle z, X_t\rangle}] = e^{-tc|z|^\lambda}$. *Moreover, if* $\alpha < 2$, *then* λ *is an uniform measure on* S.

4.2 L- processes (self-decomposable processes) and L-distributions

As an extension of a stable process, it is called a self-decomposable process, or L-process.

Definition 4.3 *(X_t) is a* **self-decomposable process**), *or an* L-**process** $\overset{\text{def}}{\iff}$ *(X_t) is a d-dimensional Lévy process and $^\forall c \in (0,1)$, $^\exists(Y_t), (Z_t)$: d-dimensional Lévy processes on $^\exists(\Omega', \mathcal{F}', P')$: a probability space; $(Y_t) \perp\!\!\!\perp (Z_t)$, $(Y_t) = (cX_t)$ in law, $(Y_t + Z_t) = (X_t)$ in law.*

In this case the distribution of X_1 is called a **self-decomposable distribution** *or an* L-**distribution**. *It is equivalent to the above with $t = 1$. That is, $^\forall c \in (0,1)$, $^\exists \rho_c, \eta_c \in I(\mathbf{R}^d)$; $\rho_c \perp\!\!\!\perp \eta$, $\widehat{\rho_c}(z) = \widehat{\mu}(cz)$, $\mu = \rho_c * \eta_c$.*

Note. If μ is an L-dist., then $^\forall t > 0, \mu^{t*}$ is so, too.

Lemma 4.1 *(X_t) is an L-process, i.e., $X_1 \overset{\text{(d)}}{=} \mu$ is an L-distribution $\iff$ $^\forall c \in (0,1), ^\exists \eta_c \in I(\mathbf{R}^d); \widehat{\mu}(z)/\widehat{\mu}(cz) = \widehat{\eta}_c(z)$. $\iff$ Let $\mu \leftrightarrow (A, \nu, \gamma)$. For $r > 0$, set $N(r, d\xi) := \nu((r,\infty)d\xi)$. $^\forall B \in \mathcal{B}(S)$, $n_B(s) := N(e^{-s}, B)$ is convex in $s \in \mathbf{R}$.*

Proof. The first equivalence is clear by $Z_1 \overset{\text{(d)}}{=} \eta_c$, and on the inverse, let (Z_t) be a Lévy process corresponding to $\eta_c \in I(\mathbf{R}^d)$. Note that (Y_t) is determined by $\widehat{\mu}(cz)$.

On the later half equivalence, let μ be an L-distribution and let $\psi(z) = \log \widehat{\mu}(z)$. By $X_1 \overset{\text{(d)}}{=} Y_1 + Z_1$, $Y \overset{\text{(d)}}{=} cX_1$、$Y_1 \perp\!\!\!\perp Z_1$, the log-characteristic function of Z_1 is $\psi_c(z) = \psi(z) - \psi(cz)$. Therefore, It is enough to show that μ is an L-dist. $\iff$ $\eta_c \in I(\mathbf{R}^d)$, and in order to it we may show $\widehat{\eta_c} = e^{\psi}_c$ has an LK-representation. Let $A_c = (1 - c^2)A, \nu_c(dx) : \nu(dx) - \nu(c^{-1}dx)$. Then $^\exists \gamma_c \in \mathbf{R}^d$; $\psi_c \leftrightarrow (A_c, \nu_c, \gamma_c)$. Hence, $\eta_c \in I(\mathbf{R}^d)$ is equivalent to $\nu_c \geq 0$, i.e., $\nu(E) - \nu(c^{-1}E) \geq 0$ ($^\forall E \in \mathcal{B}(\mathbf{R}^d \setminus \{0\})$. Moreover, this is equivalent to the following: for any fixed $B \in \mathcal{B}(S)$, let $n(s) = n_B(s)$ and for $^\forall u > 0$, $n(s+u) - n(s) \geq n(s + u + \log c) - n(s + \log c)$. Furthermore, this is equivalent to the condition given in the theorem ($\to$ the next question. Note that $\log c < 0$ for $c \in (0,1)$). ∎

Question. Show the equivalences in the above proof, that is, $\nu(E) - \nu(c^{-1}E) \geq 0$ ($^\forall E \in \mathcal{B}(\mathbf{R}^d \setminus \{0\})$ $\iff$ For any fixed $B \in \mathcal{B}(S)$, letting $n(s) = n_B(s)$, for $^\forall s \in \mathbf{R}, ^\forall u > 0, c \in (0,1)$, $n(s+u) - n(s) \geq n(s + u + \log c) - n(s + \log c)$. $\iff$ $^\forall B \in \mathcal{B}(S)$, $n_B(s) := N(e^{-s}, B)$ is convex in $s \in \mathbf{R}$.

Theorem 4.7 (Representation of L-process) *(X_t) is an L-process $\iff$ For a Lévy measure ν of X_1, $^\exists \lambda(d\xi)$: a finite measure on S, $^\exists k_\xi(r) \geq 0$: measurable in $\xi \in S$, non-increasing right-continuous in $r > 0$, $k_\xi(0+) > 0$;*

$$\nu(dx) = \int_S \lambda(d\xi) \int_0^\infty 1_{dx}(r\xi) \frac{k_\xi(r)}{r} dr.$$

Proof. Let (X_t) be an L-process. Then by the above lemma, $^\forall B \in \mathcal{B}(S)$, $N(e^{-s}, B)$ is convex in $s \in \mathbf{R}$. Since $N(r, B) = \nu((r,\infty)B)$ is non-increasing in $r > 0$, if we let

$$\lambda(B) := -\int_0^\infty (1 \wedge r^2) dN(r, B) = \int_{(0,\infty)B} (1 \wedge |x|^2) \nu(dx),$$

then λ is a finite measure on S, and for every $r > 0$, $\lambda(d\xi) \ll N(r, d\xi)$. Thus, for each $s \in \mathbf{R}$, $^\exists H_\xi(s)$: non-negative measurable ft of ξ and $N(e^{-s}, d\xi) = H_\xi(s)\lambda(d\xi)$. Since the LHS is non-decreasing and convex in s, for any $s_1 < s_2, p \in (0,1)$ and for λ-a.a.ξ,

$$H_\xi(s_1) \leq H_\xi(s_2), \quad H_\xi(ps_1 + (1-p)s_2) \leq pH_\xi(s_1) + (1-p)H_\xi(s_2).$$

Hence, we may assume for λ-a.a.ξ, $H_\xi(s)$ is non-decreasing and convex in s. More exactly, we can take a such version. In fact, $^\exists C_1 \in \mathcal{B}(S); \lambda(C_1^c) = 0$, and we may assume that for $^\forall \xi \in C_1$, for all rational numbers $s_1 < s_2, p \in (0,1)$, $H_\xi(s)$ satisfies the above inequality. Thus, let

$$\widetilde{H_\xi}(s) := \sup_{r \in (-\infty, s) \cap \mathbf{Q}} H_\xi(r).$$

Then this is measurable in ξ, and satisfies the above conditions and $N(e^{-s}, d\xi) = \widetilde{H_\xi}(s)\lambda(d\xi)$. Hence, ${}^\exists C_2 \subset C_1; C_2 \in \mathcal{B}(S)$ and ${}^\forall \xi \in C_2, \widetilde{H_\xi}(-\infty) = 0$. Let

$$h_\xi(u) := \lim_{n\to\infty} n(\widetilde{H_\xi}(u) - \widetilde{H_\xi}(u - 1/n)).$$

Then this is left-continuous, measurable in ξ and

$$\widetilde{H_\xi}(s) = \int_{-\infty}^{s} h_\xi(u)du.$$

Moreover, set $C = \{\xi; h_\xi \equiv 0\}, C_3 = C_2 \setminus C$, then for $\xi \in C_3$, $h_\xi(\infty) > 0$ and

$$\nu((0,\infty)C) = \lim_{s\to\infty} N(e^{-s}, C) = \lim_{s\to\infty} \int_C \widetilde{H_\xi}(s)\lambda(d\xi) = 0$$

Hence,

$$\begin{aligned}\nu((r,\infty)B) &= N(r,B) = \int_{B\cap C_3} \widetilde{H_\xi}(\log r)\lambda(d\xi) \\ &= \int_{B\cap C_3} \lambda(d\xi) \int_{-\infty}^{\log r} h_\xi(u)du = \int_B \lambda(d\xi) \int_r^\infty h_\xi(-\log v)\frac{dv}{v}.\end{aligned}$$

If we define $k_\xi(v) := h_\xi(-\log v)$ if $\xi \in C_3$, then this is measurable in (ξ, v), non-increasing right-continuous in v, and $k_\xi(0+) = h_\xi(\infty) > 0$. Moreover, set $k_\xi(v) \equiv 1$ outside of C_3. Then this satisfies the desired result.

The inverse is clear. ∎

5 Lévy Processes and Distributions

In this section, we first show that a Lévy process in law is equivalent to a Lévy process. Moreover, we give some sufficient conditions for absolute continuity of the distributions.

5.1 Lévy Processes in law

The following result holds for a general Markov process which is continuous in probability, however, we arrange ti to a Lévy process. (On a case of a Markov process, we describe at the end of this section.)

Theorem 5.1 *Let (X_t) be a Lévy process and $X_1 \overset{(d)}{=} \mu$. For $\varepsilon > 0$, let*

$$\alpha_\varepsilon(t) := P(|X_t| \geq \varepsilon) = P(|X_{t+s} - X_s| \geq \varepsilon) \ ({}^\forall s \geq 0).$$

(1) By the continuity in probability of (X_t), ${}^\forall\varepsilon > 0, \lim_{t\downarrow 0} \alpha_\varepsilon(t) = 0$ and by this (X_t) has a D version, i.e., ${}^\exists(Y_t)$ sis a D process and equivalent to (X_t). Moreover, ${}^\forall t > 0, P(Y_{t-} = Y_t) = 1$ holds. (This immediately follows from the continuity in probability of (X_t), and of (Y_t)).

(2) If (X_t) is a Gaussian process, then ${}^\forall\varepsilon > 0, \lim_{t\downarrow 0} t^{-1}\alpha_\varepsilon(t) = 0$ holds. By this (X_t) has a C version.

Proof. (1) Let $\widetilde{\alpha_\varepsilon}(t) := \sup_{s\in[0,t]} \alpha_\varepsilon(s)$ and $I \subset [a, b] \subset [0, \infty)$. Moreover, set

$$B(p, \varepsilon, I) = \{X_t \text{ has } p \text{ number of } \varepsilon\text{-oscillations (at least) in } I\}$$

That is, this is an event of that there exist $p+1$ number of increasing time points $t_j \in I$ ($j = 1, \ldots, p+1$) such that $|X_{t_{j+1}} - X_{t_j}| \geq \varepsilon$.

(Outline of Proof) The essentials of the proof are that if at a time point, X_t does not have a right-hand-limit or a left-hand-limit, then there exist an $\varepsilon_0 > 0$ such that on a nbd of the time point, it has infinitely many ε_0- oscillations (①), and an inequality obtained by the independent increments (②).

① Let $A_{N,k}$ be an event of that X_t has finite number of $1/k$-oscillations in $t \in [0, N] \cap \mathbf{Q}$. Then it holds that

$$\bigcap_{N,k\geq 1} A_{N,k} \subset \{{}^\forall t \geq 0, {}^\exists X_{t+} \in \mathbf{R}^d, {}^\forall t > 0, {}^\exists X_{t-} \in \mathbf{R}^d\} =: \Omega_1$$

② For $n > p \geq 1$ and let $a \leq t_1 < \cdots < t_n \leq b$, $I = \{t_1, \ldots, t_n\}$, the independent increments implies

$$P(B(p, 4\varepsilon, I)) \leq (2\widetilde{\alpha_\varepsilon}(b - a))^p.$$

By this and the assumption of $\alpha_\varepsilon(t) \to 0$ ($t \downarrow 0$), we have

③ ${}^\forall N, k \geq 1, P(A^c_{N,k}) = 0$, and hence, $P(\Omega_1) = 1$. By the continuity in probability of (X_t), it can be shown $Y_t := X_{t+}$ is a D version of (X_t).

(Detailed Proof)

① We consider the complements. If there is no point $t \geq 0; X_{t+} \in \mathbf{R}^d$, then there is no $t_n \downarrow t; \lim X_{t_n}$, that is,

$${}^\exists k_0 \geq 1; {}^\forall j, {}^\exists n_j, m_j \geq j; |X_{t_{n_j}} - X_{t_{m_j}}| \geq 1/k_0.$$

Moreover, a sub-sequence $\{t_{n_j}\}$ can be taken such that

$$|X_{t_{n_{j+1}}} - X_{t_{n_j}}| \geq 1/k_0.$$

Clearly, this implies (X_t) has infinitely many $1/k_0$-oscillations in $\{t_{n_j}\}$.

② can be shown by the induction on p. Recall $I = \{t_1, \ldots, t_n\} \subset [a, b]$, $1 \leq p < n$. When $p = 1$, let C_k be an event of that $|X_{t_j} - X_a|$ is larger than or equal to 2ε first at $j = k$, i.e.,

$$C_k = \{|X_{t_k} - X_a| \geq 2\varepsilon, |X_{t_j} - X_a| < 2\varepsilon, j = 1, 2, \ldots, k - 1\}$$

and let $D_k = \{|X_b - X_{t_k}| \geq \varepsilon\}$. Then C_k are mutually disjoint and it holds that

$$B(1, 4\varepsilon, I) \subset \bigcup_{k=1}^{n} \{|X_{t_k} - X_a| \geq 2\varepsilon\} = \bigcup_{k=1}^{n} C_k \subset \{|X_b - X_a| \geq \varepsilon\} \cup \bigcup_{k=1}^{n} (C_k \cap D_k).$$

The first inclusion is clear by complements, and the last one comes from

$$C_k \cap D_k^c \subset \{|X_{t_k} - X_a| \geq 2\varepsilon, |X_b - X_{t_k}| < \varepsilon\} \subset \{|X_b - X_a| \geq |X_{t_k} - X_a| - |X_b - X_{t_k}| > \varepsilon\}.$$

Independent increments implies

$$\begin{aligned} P(B(p, 4\varepsilon, I)) &\leq P(|X_b - X_a| \geq \varepsilon) + \sum_{k=1}^{n} P(C_k)P(D_k) \\ &\leq P(|X_{b-a} - X_0| \geq \varepsilon)) + \sum_{k=1}^{n} P(C_k)P(|X_{b-t_k} - X_0| \geq \varepsilon) \\ &\leq \alpha_\varepsilon(b-a) + P(\bigcup_{k=1}^{n} C_k)\widetilde{\alpha_\varepsilon}(b-a) \leq 2\widetilde{\alpha_\varepsilon}(b-a) \end{aligned}$$

Next we assume the desired inequality holds for p (≥ 1).

· Let E_k be an event of that (X_t) has p number of 4ε-oscillations in $\{t_1, \ldots, t_k\}$ and does not have p number of 4ε-oscillations in $\{t_1, \ldots, t_{k-1}\}$.

· Let F_k be an event of that (X_t) has at least 1 number of 4ε-oscillations in $\{t_k, \ldots, t_n\}$. Then we have

$$B(p, 4\varepsilon, I) = \bigcup_{k=1}^{n} E_k, \quad B(p+1, 4\varepsilon, I) \subset \bigcup_{k=1}^{n} (E_k \cap F_k).$$

By $P(F_k) \leq 2\widetilde{\alpha_\varepsilon}(b-a)$ and by using the assumption of the induction and by independent increments,

$$\begin{aligned} P(B(p+1, 4\varepsilon, I)) &\leq \sum_{k=1}^{n} P(E_k)P(F_k) \leq 2\widetilde{\alpha_\varepsilon}(b-a) \sum_{k=1}^{n} P(E_k) \\ &= 2\widetilde{\alpha_\varepsilon}(b-a)P(B(p, 4\varepsilon, I)) \leq (2\widetilde{\alpha_\varepsilon}(b-a))^{p+1}. \end{aligned}$$

Therefore, the desired inequality is obtained.

③ Fix $^\forall N, k \geq 1$. Let $\varepsilon = 1/(4k)$ and by the assumption we have $^\exists \ell \geq 1; \widetilde{\alpha_\varepsilon}(N/\ell) < 1/2$. Let $t_{\ell,j} := jN/\ell$. Then it holds that

$$\begin{aligned} P(A_{N,k}^c) &= P(X_t \text{ has infinitely many number of } 1/k\text{-oscillations in } [0, N] \cap \mathbf{Q}) \\ &= \sum_{j=1}^{\ell} P(X_t \text{ has infinitely many number of } 1/k\text{-oscillations in } [t_{\ell,j-1}, t_{\ell,j}] \cap \mathbf{Q}\ ) \\ &= \sum_{j=1}^{\ell} \lim_{p \to \infty} P(B(p, 1/k, [t_{\ell,j-1}, t_{\ell,j}] \cap \mathbf{Q}) = 0. \end{aligned}$$

In fact, denote $[t_{\ell,j-1}, t_{\ell,j}] \cap \mathbf{Q} = \{t_1, t_2, \ldots\}$. Then $^\forall n \geq 1$,

$$P(B(p, 1/k, \{t_1, \ldots, t_n\}) \leq (2\widetilde{\alpha_\varepsilon}(N/\ell))^p$$

and by letting $n \to \infty$, $p \to \infty$, we have the above. Therefore, $P(\Omega_1) = 1$ and set $Y_t := X_{t+}1_{\Omega_1}$, then it is right-continuous and has left-hand-limits. Moreover, for $^\forall t \geq 0$, take $r_n \in \mathbf{Q}_+, \downarrow t$, then $X_{r_n} \to Y_t$ a.s. and the continuity in probability implies $X_{r_n} \to X_t$ in pr. Hence, we have $P(X_t = Y_t) = 1$.

(2) We first assume $t\alpha_\varepsilon(t) \to 0$ $(t \downarrow 0)$ and show (X_t) has a C-version. By (1) there exists a D-version (Y_t). Thus, it is enough to show $^\forall N \geq 1, P(^\forall t \in (0, N], Y_t = Y_{t-}) = 1$.

Fix $^\forall\ell \geq 1$ and for each $j = 0, 1, \dots, \ell$, set $t_{\ell,j} := jN/\ell$. Fix $^\forall\varepsilon > 0$ and let $M_{\varepsilon,\ell}$ be a number of $j = 1, \dots, \ell$ such that $|Y_{t_{\ell,j}} - Y_{t_{\ell,j-1}}| \geq \varepsilon$. Let M_ε be a number of $t \in (0, N]$ such that $|Y_t - Y_{t-}| \geq \varepsilon$. Then $M_{\varepsilon,\ell}$ is $\mathcal{F}$-measurable and it holds that ($\to$ the next question)

$$M_{2\varepsilon} \leq \liminf_{\ell\to\infty} M_{\varepsilon,\ell}.$$

Moreover, by

$$M_{\varepsilon,\ell} = \sum_{j=1}^{\ell} I(|Y_{t_{\ell,j}} - Y_{t_{\ell,j-1}}| \geq \varepsilon)$$

and by the assumption on $\alpha_\varepsilon(t)$, we have

$$EM_{\varepsilon,\ell} = \sum_{j=1}^{\ell} P(|Y_{t_{\ell,j}} - Y_{t_{\ell,j-1}}| \geq \varepsilon) \leq \ell\alpha_\varepsilon(N/\ell) \to 0 \ (\ell \to \infty).$$

Hence, by Fatou's lemma,

$$EM_{2\varepsilon} \leq E[\liminf_{\ell\to\infty} M_{\varepsilon,\ell}] \leq \liminf_{\ell\to\infty} EM_{\varepsilon,\ell} = 0.$$

Therefore, we have $P(\bigcap_{\varepsilon>0}\{M_\varepsilon = 0\}) = 1$ and get the desired result. (More exactly, let $\Omega_N := \bigcap_{k\geq 1}\{\liminf_{\ell\to\infty} M_{1/k,\ell} = 0\}$, then it is contained the above event and $\Omega_N \in \mathcal{F}$, $P(\Omega_N) = 1$. Thus, we may let $\mathcal{F}$ be the completion.)

It remains to show that a Gaussian distribution satisfies $\alpha_\varepsilon(t)$. In general,

$$\widehat{\mu}(z) = \exp\left[-\frac{1}{2}\langle Az, z\rangle + i\langle\gamma, z\rangle\right]$$

by change of variables we may set $A = \text{diag}\ (\lambda_1, \dots, \lambda_p, 0, \dots, 0)$ $(\lambda_j > 0)$, $\gamma = 0$. Moreover, it is enough to show that for $^\forall\varepsilon > 0$,

$$\lim_{t\downarrow 0} \frac{1}{t}\mu^{t*}(C_\varepsilon^c) = 0 \quad (C_\varepsilon = (-\varepsilon, \varepsilon)^d).$$

By $X_t^j = 0$ if $j > p$,

$$\begin{aligned}
\mu^{t*}(C_\varepsilon^c) &= P(X_t \notin C_\varepsilon) = \sum_{j=1}^{p} P(|X_t^j| \geq \varepsilon) = 2\sum_{j=1}^{p} \frac{1}{\sqrt{2\pi\lambda_j t}} \int_\varepsilon^\infty e^{-x^2/(2\lambda_j t)}dx \\
&= 2\sum_{j=1}^{p} \frac{1}{\sqrt{2\pi}} \int_{\varepsilon/\sqrt{\lambda_j t}}^\infty e^{-x^2/2}dx \\
&\leq \frac{2\sqrt{t}}{\varepsilon} \sum_{j=1}^{p} \sqrt{\frac{\lambda_j}{2\pi}} e^{-\varepsilon^2/(2\lambda_j t)} = o(t) \quad (t \downarrow 0),
\end{aligned}$$

where the last estimation comes from

$$\int_c^\infty e^{-x^2/2}dx \leq \int_c^\infty \frac{x}{c}e^{-x^2/2} = \frac{1}{c}e^{-c^2/2} \quad \text{by } x/c \geq 1$$

or

$$\int_c^\infty e^{-x^2/2}dx \leq \int_c^\infty e^{-x^2/2}\left(1 + \frac{1}{x^2}\right)dx = \frac{1}{c}e^{-c^2/2}.$$

■

Question. In the above proof, show $M_{2\varepsilon} \leq \liminf_{\ell\to\infty} M_{\varepsilon,\ell}$.

If at a $t > 0$, Y_t has a jump with size larger than or equal to 2ε, then by using right-continuity, $^\exists\ell_0; ^\forall\ell \geq \ell_0, ^\exists t_{\ell,j-1} \leq t < t_{\ell,j}; |Y_{t_{\ell,j}} - Y_{t_{\ell,j-1}}| \geq \varepsilon$. Because $Y_{t_{\ell,j-1}}$ can be taken as close to Y_{t-}, and $Y_{t_{\ell,j}}$ can be taken as close to Y_t,

5.2 Absolute continuity of distributions of Lévy Processes

In general, a σ-finite measure μ on $\mathbf{R}^d$ has the following **Lebesgue decomposition** with respect to the Lebesgue measure dx:

$$\mu = \mu_c + \mu_d, \quad \mu_c = \mu_{ac} + \mu_{sc}.$$

they are called "a continuous part+ a discrete part", "continuous = absolute continuous + singular continuous" such that ${}^\forall x, \mu_c(\{x\}) = 0$, $\mu_d = \sum a_n \delta_{x_n}; a_n > 0, x_n \in \mathbf{R}^d$. Moreover, $\mu_{ac} \ll dx$, i.e., $|A| = 0 \Rightarrow \mu_{ac}(A) = 0 \iff {}^{\exists_1} f \geq 0; \mu_{ac}(dx) = f(x)dx$, this f is unique a.e.

In this section, we consider the sufficient conditions for that a distribution μ_t of a Lévy process X_t is absolute continuous.

Theorem 5.2 *For a Lévy process (X_t) with a triplet (A, μ, γ), If* rank $A = d$, *then for ${}^\forall t > 0$, μ_t is absolute continuous.*

Since a non-degenerate Gaussian distribution (i.e., rank $A = d$) is clearly absolute continuous and the convolution of it and an arbitrary distribution is also absolute continuous, the above is obvious.

If $r =$ rank $A < d$, then By orthogonal transform, the first r-dimension has a Gaussian density. Thus, in the remaining $(d-r)$-dimension space if it has a density by ν, then the product of them is a density on the whole space. Therefore, in the following we may assume $A = 0$ and investigate the conditions on ν for absolute continuity of μ_t.

If a Lévy measure ν is absolute continuous, then μ is so, too, as in the following. However, in multi-dimensional case, we have an example of that μ is absolute continuous even if ν is not so. The rotation invariant stable distribution is an example of the first half, and the product of 1 -dimensional symmetric stable distributions is the one of the later half.

Let a finite measure $\widetilde{\nu}(dx) = (1 \wedge |x|^2)\nu(dx)$.

Theorem 5.3 (The 1st sufficient condition for absolute continuity) *If $\nu(\mathbf{R}^d) = \infty$ and ${}^\exists \ell \geq 1; \widetilde{\nu}^{\ell *}$ is absolute continuous, then for ${}^\forall t > 0$, the distribution of X_t の is absolute continuous.*

Proof.

The distribution μ of X_1 can be approximated by compound Poisson distributions μ_n by Lévy measures $\nu_n = \nu|_{\{|x| \geq 1/n\}}$:

$$\mu_n = \sum_{k \geq 0} e^{-c_n} \frac{c_n^k}{k!} \nu_n^{k*} = \left(\sum_{k=0}^{\ell-1} + \sum_{k \geq \ell} \right) e^{-c_n} \frac{1}{k!} \nu_n^{k*}$$

(with $c_n = \nu_n(\mathbf{R}^d)$) and μ has μ_n as a convolution element (i.e., $\mu = \mu_n * \mu_n^c$). Moreover, the above 2nd term is absolute continuous, and by $c_n \to \infty$, we have

$$(\mu_{sc} + \mu_d)(\mathbf{R}^d) \leq (\mu_{n,sc} + \mu_{n,d})(\mathbf{R}^d) \leq \sum_{k=0}^{\ell-1} e^{-c_n} \frac{c_n^k}{k!} \to 0.$$

Finally, for X_t $(t > 0)$, we only change c_n to tc_n, and so the desired result holds. ■

A **RV** X **is degenerate** $\overset{\text{def}}{\iff}$ ${}^\exists a \in \mathbf{R}^d, {}^\exists V \subset \mathbf{R}^d$: a subspace; $\dim V < d, P(X \in a + V) = 1$, i.e., supp $\mu_X \subset a + V$.

A **Lévy process** (X_t) **is degenerate** $\overset{\text{def}}{\iff}$ ${}^\forall t > 0, P(X_t \in at + V) = 1$.

If it is not degenerate, then if is called be **non-degenerate**. Moreover, in general, the following are equivalent: (1) ${}^\forall t > 0, P(X_t \in V) = 1$, (2) ${}^\exists t > 0; P(X_t \in V) = 1$, (3) $A(\mathbf{R}^d)$, supp $\nu \subset V, \gamma \in V$

Theorem 5.4 (The 2nd sufficient condition for absolute continuity) *If (X_t) is a non-degenerate Lévy process and if its Lévy measure ν is absolute continuous in radial directions and satisfies* *divergence condition, that is,* $^\exists\lambda(\xi)$*: a finite measure on* $S = \mathbf{S}^{d-1}$, $^\exists g(r,\xi)$*: a measurable function on* $(0,\infty)\times S$*; (note that if we set* $g(0,\xi)=0$ *and we may consider* $r\in[0,\infty)$*.)*

$$\nu(dx) = \int_S \lambda(d\xi)\int_0^\infty g(r,\xi)1_{dx}(r\xi)dr, \quad \int_0^\infty g(r,\xi)dr = \infty \quad \lambda(d\xi)\text{-a.e.},$$

then $^\forall t>0$*, the distribution of* X_t *is absolute continuous.*

Note. The divergence condition contains the case of $\nu = 0$, i.e., $\lambda = 0$, however, in this case we have rank $A = d$.

In order to proof we use the following lemmas:

Lemma 5.1 *If* ν *is absolute continuous in radial directions and for an arbitrary* $(d-1)$*-dimensional subspace* V, $\nu(V) = 0$*, then* ν^{d*} *is absolute continuous, and by the previous theorem,* μ *is absolute continuous.*

Lemma 5.2 *For a linear subsapce* $V; \dim V \le d-1$*, let* T *be the orthogonal projection from* $\mathbf{R}^d$*. If* ν *is absolute continuous in radial directions, then* νT^{-1} *on* V *is so, and if* ν *satisfies the divergence condition, and if* $\nu T^{-1}\neq 0$*, then* νT^{-1} *is so.*

[Proof of Theorem 5.4].

It is enough to show the case of $t = 1$, that is, μ is absolute continuous. As mentioned before, we may set $A - 0$. If $d = 1$, then ν is absolute continuous by Theroem 5.2. We assume the claim holds in lower d-dimension and show in d-dimension. If $\nu(V) = 0$ for any $(d-1)$-dimensional subspaces V, then by Lemma 5.1 μ is absolute continuous. Thus, it is enough to show the case that $^\exists V$: a $(d-1)$-dimensional subspace; $\nu(V) > 0$. Let V_1 be a linear subspace spanned by the support of the restricted ν to V. Then $1\le \dim V_1 \le d-1$, Let V_2 be the orthogonal complement of V_1, denote orthogonal projection to V_1, V_2 as T_1, T_2 and set $x_j = T_j x$. Then $\mathbf{R}^d = V_1\oplus V_2$. We define $\mu_1 \in I(\mathbf{R}^d)$ by the following:

$$\widehat{\mu_1}(z) = \exp\left[\int_{V_1}(e^{i\langle z,x\rangle} - 1 - i\langle z,x\rangle 1_D(x))\nu(dx)\right] \quad (D = \{|x|<1\}).$$

By Lemma 5.2, on V_1, νT_1^{-1} is absolute continuous in radial directions and satisfies the divergence condition, then by the assumption of induction, $^{\exists_1} f_1(x_1)\ge 0; \mu_1(dx_1) = f(x_1)dx_1$. Hence, for $B\in \mathcal{B}(\mathbf{R}^d); |B| = 0$, it is enough to show $\mu(B) = 0$. Define $\mu_2 \in I(\mathbf{R}^d)$ by $\mu = \mu_1 * \mu_2$. We have

$$\mu(B) = \int_{\mathbf{R}^d} h(y_1,y_2)\mu_2(dy), \quad h(y_1,y_2) := \int_{V_1} 1_B(x_1+y_1,y_2)f(x_1)dx_1.$$

By

$$\int_{V_2} dy_2 \int_{V_1} 1_B(x_1,y_2)dx_1 = |B| = 0,$$

$\int_{V_1} 1_B(x_1,y_2)dx_1 = 0$ dy_2-a.e., that is, for $^\forall y_2\notin B_2$ with $^\exists B_2\in\mathcal{B}(V_2); |B_2| = 0$. Hence, for $^\forall y_1\in V_1$ and $^\forall y_2 \notin B_2$, $\int_{V_1} 1_B(x_1+y_1,y_2)dx_1 = 0$. Thus, $h(y_1,y_2) = h(y_1,y_2)1_{B_2}(y_2)$. Define $Y :\overset{(\mathrm{d})}{=} \mu_2$ on $\mathbf{R}^d$, and $Y_j := T_j Y$, let $\rho_2 :\overset{(\mathrm{d})}{=} Y_2$ on V_2, $\rho_1(dy_1|\, y_2) := P(Y_1\in dy_1|\, Y_2 = y_2)$. Then

$$\mu(B) = \int_{\mathbf{R}^d} h(y_1,y_2)1_{B_2}(y_2)\mu_2(dy) = \int_{B_2}\rho_2(dy_2)\int_{V_1} h(y_1,y_2)\rho_1(dy_1|\, y_2),$$

and hence, $\rho_2\in I(V_2)$. Let ν_2 be a Lévy measure of μ_2, Then the Lévy measure of ρ_2 is $\nu_3 := \nu_2 T_2^{-1}|_{V_2}$. This is absolute continuous in radial directions and satisfies the divergence condition on V_2, and ρ_2 is non-degenerate. In fact, if the support of ν_3 is contained in a proper subspace of V_2; $V_2^0\subset V_2$, then the support of ν_2 is in $V_1 + V_2^0$, and hence, ν is so, however it contradicts the non-degeneracy of μ. Thus, the space spanned by the support of ν_3 is V_2, and ρ_2 is non-degenerate on V_2. Therefore, by the assumption of the induction, ρ_2 is absolute continuous on V_2 and we have $\rho_2(B_2) = 0$. Hence, $\mu(B) = 0$. ■

[Proof of Lemma 5.1]. Let $|B| = 0$ and we show $\widetilde{\nu}^{d*}(B) = 0$.

$$\widetilde{\nu}^{d*}(B) = \int_{S^d} \prod_{j=1}^{d} \lambda(d\xi_j) \int_0^\infty \cdots \int_0^\infty 1_B(r_1\xi_1 + \cdot + r_d\xi_d) \prod_{j=1}^{d} g(r_j, \xi_j)(1 \wedge r_j^2) dr_j.$$

By the assumption, ${}^\forall V \subset \mathbf{R}^d$; a subspace ; $\dim V < d$, we may set $\lambda(V \cap S) = 0$ として良い. Moreover, let $V(\xi_1, \dots, \xi_d)$ be a linear subspace spanned by $\xi_1, \dots, \xi_d \in S$ and for $1 \le r \le d$, let $K_r = \{(\xi_1, \dots, \xi_d) \in S^d; \dim V(\xi_1, \dots, \xi_d) = r$ We divide S^d as the following disjoint union:

$$S^d = \bigcup_{r \le d} K_r, \quad K_r = \bigcup_{\{i_1, \dots, i_r\}} K(i_1, \dots, i_r) \quad \text{if } r < d,$$

where $K(i_1, \dots, i_r)$ is a family of all $(\xi_1, \dots, \xi_d) \in K_r$ such that $\xi_{i_1}, \dots, \xi_{i_r}$ are linear independent. On K_d, by $|B| = 0$ and change of variables it is 0. On the other sets they are 0 by the assumption. Hence, we have $\widetilde{\nu}^{d*}(B) = 0$ を得る. In fact, if $\xi_1, \dots, \xi_d$ are linear independent, then by change of variables $(r_j)_{j \le d} \mapsto r_1\xi_1 + \cdot + r_d\xi_d$, we have

$$\int_0^\infty \cdots \int_0^\infty 1_B(r_1\xi_1 + \cdot + r_d\xi_d) \prod_{j=1}^{d} g(r_j, \xi_j)(1 \wedge r_j^2) dr_j = 0.$$

Hence, it is 0 on K_d. Moreover, let $1 \le r \le d - 1$ and fix $i_0 \ne i_1, \dots, i_r$. By the assumption $\lambda(K(i_1, \dots, i_r)) = 0$ and $K(i_1, \dots, i_r) = S \cap V(\xi_{i_1}, \dots, \xi_{i_r})$ implies

$$\int_{K(i_1, \dots, i_r)} \prod_{j=1}^{d} \lambda(d\xi_j) \le \int_{S^{d-1}} \prod_{j \ne i_0} \lambda(d\xi_j) \int_S 1_{V(\xi_{i_1}, \dots, \xi_{i_r})}(\xi_{i_0}) \lambda(d\xi_{i_0}) = 0.$$

Therefore, $\widetilde{\nu}^{d*}(B) = 0$. ■

[Proof of Lemma 5.2]. Let V_2 be an orthogonal complement of V, and let T_2 be the orthogonal projection to V_2. Let $c := \lambda(S \setminus V_2)$. If $c = 0$, then the support of ν is in V_2, the support of νT^{-1} is $\{0\}$, and hence, it is clear. Let $c > 0$. Let $Q := c^{-1}\nu$ be restricted to $S \setminus V_2$ and as a probability measure we define RVs $Y(\xi) = T\xi/|T\xi|, Z(\xi) = T_2\xi$ and define a distribution of Y as $P_Y(d\eta) = Q(Y \in d\eta)$ on $S \cap V$, and define a conditional distribution of Z under the condition $Y = \eta$ as $P_Z^\eta(d\zeta) = Q(Z \in d\zeta |\ Y = \eta)$ on V_2. $P_Z^\eta(d\zeta)$ is a distribution on $\{|\zeta| < 1\} \cap V_2$ and it is determined except η with 0 P_Y measure. Note that $\xi = T\xi + T_2\xi = (1 - |Z|^2)^{1/2} Y + Z$ (since $1 = |\xi|^2 = |T\xi|^2 + |Z|^2$ and by $|T\xi|^2 = 1 - |Z|^2$). Let $\Lambda(d\eta) := cP_Y(d\eta)$ and

$$G(r, \eta) := \int_{V_2} (1 - |\zeta|^2)^{-1/2} g((1 - |\zeta|^2)^{-1/2} r, (1 - |\zeta|^2)^{1/2}\eta + \zeta)) P_z^\eta(d\zeta)$$

Then we have

$$\nu T^{-1}(B) = \int_{S \cap V} \Lambda(d\eta) \int_0^\infty G(r, \eta) 1_B(r\eta) dr$$

In fact, for ${}^\forall B \in \mathcal{B}(V); 0 \notin B$, under the above distribution by $\xi - \zeta = T\xi = (1 - |\zeta|^2)^{1/2}\eta$, we have

$$\begin{aligned}
\nu T^{-1}(B) &= \int_{S \setminus V_2} \lambda(d\xi) \int_0^\infty g(r, \xi) 1_B(rT\xi) dr \\
&= c \int_{S \cap V} P_Y(d\eta) \int_{V_2} P_z^\eta(d\zeta) \int_0^\infty g(r, (1 - |\zeta|^2)^{1/2}\eta + \zeta)) 1_B(r(1 - |\zeta|^2)^{1/2}\eta) dr \\
&= c \int_{S \cap V} P_Y(d\eta) \int_{V_2} (1 - |\zeta|^2)^{-1/2} h_B(\eta, \zeta) P_z^\eta(d\zeta),
\end{aligned}$$

where

$$h_B(\eta,\zeta) = \int_0^\infty g((1-|\zeta|^2)^{-1/2}r, (1-|\zeta|^2)^{1/2}\eta + \zeta))1_B(r\eta)dr.$$

Therefore we have the above equation.

Moreover, on the divergence condition, it is equivalent to that $^\forall C \in \mathcal{B}(S), \nu((0,\infty)C) = 0$ or ∞. If $C \in \mathcal{B}(S \cap V)$, then $x \in T^{-1}((0,\infty)C) \iff Tx \neq 0, Tx/|Tx| \in C$. Thus, let C_1 be a set of unitarized vectors of $(0,\infty)C + V_2$. Then $T^{-1}((0,\infty)C) = (0,\infty)C_1$ and $C_1 \in \mathcal{B}(S)$ implies $\nu T^{-1}((0,\infty)C) = \nu((0,\infty)C_1) = 0$ or ∞. ∎

6 Lévy Processes and Markov Processes

(X_t): a **Markov process** $\overset{\text{def}}{\Longleftrightarrow}$ For arbitrary time $0 \le s < t$ and bounded Borel function f, $E[f(X_t)|\ \mathcal{F}_s] = E[f(X_t)|\ X_s]$ a.s. Moreover, if (the above)$= E[f(X_{t-s}|\ X_0 = x]|_{\ x=X_s}$ a.s., then it is called a **time homogeneous MP**.

When $X_0 = x$ a.s., it is called a Markov process starting from x. In this case it is denoted as $X_t = X_t^x$ or (X_t, P_x).

For example, for a Lévy process (X_t), let $X_t^x = x + X_t$, then it is a Markov process starting from x.

Let (X_t, P_x) be a time homogeneous Markov process on $\mathbf{R}^d$ starting from x. For a bounded Borel ft φ, let

$$P_t(x, dy) := P_x(X_t \in dy), \quad P_t\varphi(x) := E_x[\varphi(X_t)] = \int_{\mathbf{R}^d} \varphi(y)P_t(x, dy)$$

and this is called a **transition probability**.

For a transition probability $(P_t(x, dy))_{t\ge0}$, if $^\exists(P_t(dy))_{t\ge0}$; $P_t(x, dy) = P_t(dy - x)$ ($^\forall t > 0$), then it is **space homogeneous** and (X_t) is called a **time space homogeneous Markov process**.

This process is equivalent to a Lévy process in law and it is given as $P_t(dy) = \mu^{t*}(dy)$.

Theorem 6.1 *Let (X_t) be a time homogeneous Markov process starting from x_0 and let $P_t(x, dy)$ be its transition probability. For $\varepsilon > 0$, set $D_\varepsilon(x) := \{y; |x - y| < \varepsilon\}$ and*

$$\alpha_\varepsilon(t) := \sup_{x\in\mathbf{R}^d} P_t(x, D_\varepsilon(x)^c) = \sup_{x\in\mathbf{R}^d} P_x(|X_t - x| \ge \varepsilon).$$

(1) If $^\forall\varepsilon > 0, \lim_{t\downarrow0} \alpha_\varepsilon(t) = 0$, then (X_t) is continuous in probability and has a D-version, i.e., (Y_t) is a D-process and equivalent to (X_t). Moreover, $^\forall t > 0, P(Y_{t-} = Y_t) = 1$ holds.

(2) If $^\forall\varepsilon > 0, \lim_{t\downarrow0} t^{-1}\alpha_\varepsilon(t) = 0$, then (X_t) has a C-version.

Proof. The proof is almost the same as in the Lévy case. Only ② is changed to the following, and we show it: Let $\widetilde{\alpha_\varepsilon}(t)$, $B(k, \varepsilon, I)$ be defined by the same as before. Let $0 \le s_1 < \cdots < s_m \le a < b, I \subset [a, b]$ and for a bounded Borel function $g(x_1, \ldots, x_m)$, set $Z := g(X_{s_1}, \ldots, X_{s_m})$.

② By Markov property the following holds.

$$E[Z; B(p, 4\varepsilon, I)] \le EZ(2\widetilde{\alpha_\varepsilon}(b - a))^p.$$

This can be shown by the induction on p. If $p = 1$, then let C_k, D_k be the same as in the Lévy case. That is, let C_k be an event of that $|X_{t_j} - X_a|$ is larger than or equal to 2ε first at $j = k$ and set $D_k = \{|X_b - X_{t_k}| \ge \varepsilon\}$. Then C_k are mutually disjoint and it holds that as in the Lévy case

$$B(1, 4\varepsilon, I) \subset \bigcup_{k=1}^{n}\{|X_{t_k} - X_a| \ge 2\varepsilon\} = \bigcup_{k=1}^{n} C_k \subset \{|X_b - X_a| \ge \varepsilon\} \cup \bigcup_{k=1}^{n}(C_k \cap D_k).$$

If we take conditional expectations on $\mathcal{F}_a$ and on $\mathcal{F}_{t_k}$, then by Markov property, we have

$$\begin{aligned}
E[Z; B(1, 4\varepsilon, I)] &\le E[ZP(|X_b - X_a| \ge \varepsilon|\ X_a)] + \sum_{k=1}^{n} E[Z1_{C_k}P(D_k|\ X_{t_k})] \\
&= E[ZP_{X_a}(|X_{b-a} - X_0| \ge \varepsilon)] + \sum_{k=1}^{n} E[Z1_{C_k}P_{X_{t_k}}(|X_{b-t_k} - X_0| \ge \varepsilon)] \\
&\le EZ\alpha_\varepsilon(b - a) + \sum_{k=1}^{n} E[Z1_{C_k}]\alpha_\varepsilon(b - t_k) \le EZ \cdot 2\widetilde{\alpha_\varepsilon}(b - a).
\end{aligned}$$

Next we assume the desired inequality holds for p (≥ 1). Let E_k, F_k be the same as in case of Lévy. By $P(F_k|\ X_a) \le 2\widetilde{\alpha_\varepsilon}(b - a)$, by the assumption of the induction and by Markov property, we have

$$\begin{aligned}
E[Z; B(p + 1, 4\varepsilon, I)] &\le \sum_{k=1}^{n} E[Z1_{E_k}P(F_k|\ X_a)] \le 2\widetilde{\alpha_\varepsilon}(b - a)\sum_{k=1}^{n} E[Z; E_k] \\
&= 2\widetilde{\alpha_\varepsilon}(b - a)E[Z; B(p, 4\varepsilon, I)] \le EZ(2\widetilde{\alpha_\varepsilon}(b - a))^{p+1}.
\end{aligned}$$

Therefore, the desired inequality is obtained. ■