MATH 589: Advanced Probability Theory 2
Final Exam: 14 December 2021 18:30-21:30

Contents

1 Central Limit Theorem, Characteristic Functions, and Convergence of Probability Measures

1.1 Review of Sums of Independent Random Variables

Consider $\{X_n \mid n \in \mathbb{N}\}$ iid random variables with $\mathbb{E}\left[X_1\right] = 0$ (WLOG) and $\mathbb{E}\left[X_1^2\right] = 1$. Set $S_n := \sum_{j=1}^{n} X_J$. From the SSLN,

$$\frac{S_n}{n} \to 0$$

almost surely. In other words, $|S_n|$ has sub-linear growth as $n \to \infty$. In fact, given any sequence $\{b_n \mid n \geq 1\} \subseteq]0,\infty[$ such that $b_n \uparrow \infty$, if

$$\sum_{n=1}^{\infty} \frac{1}{b_n^2} < \infty,$$

i.e., b_n grows sufficiently fast, then $\frac{S_n}{b_n} \to 0$ almost surely (by Kronecker's Lemma, c.f. MATH 587). Why?

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\left[X_n^2\right]}{b_n^2} < \infty \Rightarrow \sum_{n=1}^{\infty} \frac{X_n}{b_n} \text{ converges almost surely } \Rightarrow \frac{S_n}{b_n} \to 0 \text{ almost surely.}$$

Such a sequence $\{b_n\}$ includes:

- $\{n^p\}$ for $p > \frac{1}{2}$.
- $\{\sqrt{n}(\ln(n))^p\}$ for any $p > \frac{1}{2}$.

This means that I can do better than what I know about the LLN. For example, we know that $|S_n|$ grows slower than $\sqrt{n}(\ln(p))^{1/2}$ for any $p > \frac{1}{2}$. Since the inequality is strict, this means you can always do better. There is not a critical level. Now suppose we are interested in the asymptotic behaviour? Can we find a lower bound for the growth rate of S_n?

On the other hand, if $\{X_n \mid n \geq 1\}$ is iid $N(0,1)$ standard Gaussian random variables. Then, set:

$$\breve{S}_n := \frac{S_n}{\sqrt{n}}. \tag{1}$$

$\breve{S}_n$ is again $N(0,1)$ for all $n \geq 1$. At least, in this case, $\breve{S}_n$ doesn't converge to any constant almost surely. In fact, it's easy to see that $\limsup_n \frac{S_n}{\sqrt{n}} = +\infty$ and $\liminf_n \frac{S_n}{\sqrt{n}} = -\infty$ almost surely. Why is this? Let's consider the limsup. For all $R > 0$,

$$\begin{aligned}\mathbb{P}\left(\breve{S}_n > R\right) &= \frac{1}{\sqrt{2\pi}} \int_R^{+\infty} e^{-\frac{x^2}{2}} dx \\ &= p_R \\ &> 0.\end{aligned}$$

Since $\limsup_n \breve{S}_n \in m\mathcal{T}$ (tail σ-algebra, we have from the Kolmogorov 0-1 Law that $\limsup_n \breve{S}_n$ is constant almost surely. What is this constant? Write:

$$\breve{S}_n = \frac{S_n}{\sqrt{n}} = \frac{\sum_{j=1}^n X_j + \sum_{j=N+1}^n X_j}{\sqrt{n}}.$$

As $n \to \infty$, $\frac{\sum_{j=1}^n X_j}{\sqrt{n}}$ goes to infinity. Hence, $\limsup_n \breve{S}_n = \infty$ almost surely. One can do a similar analysis for the liminf.

Remark that $\breve{S}_n \sim N(0,1)$ is also seen for a more general sequence of random variables. This phenomenon is called the **Central Limit Phenomenon**.

Q: Can I have a better description of the asymptotics of S_n?

The answer is the **Law of the Iterated Logarithm**.

Theorem 1 (Law of Iterated Logarithm)**.** *Let $\{X_n\}$ be a sequence of iid RVs with $\mathbb{E}[X_1] = 0$ and $\mathbb{E}\left[X_1^2\right] = 1$. For every $n \geq 1$, set $S_n = \sum_{j=1}^n X_j$, and define Λ_n to be the iterated logarithm:*

$$\Lambda_n := \sqrt{2n \ln(\ln(n \vee 3))}.$$

It turns out that Λ_n will give us the accurate oscillation rate of S_n. Recall that the notation $n \vee 3 = \max\{n, 3\}$. Then, we can conclude:

- $\limsup_n \frac{S_n}{\Lambda_n} = 1$ *almost surely.*
- $\liminf_n \frac{S_n}{\Lambda_n} = -1$ *almost surely.*

In fact, for every $c \in [-1,1]$, for almost every sample point $\omega \in \Omega$, there exists a subsequence $\{n_k\}_\omega \subseteq \mathbb{N}$ such that

$$\lim_{k\to\infty} \frac{S_{n_k}(\omega)}{\Lambda_{n_k}} = c. \tag{2}$$

The picture you want to have in mind is the following:

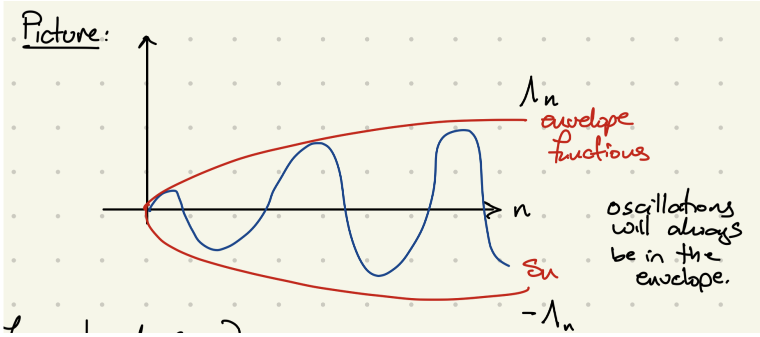


Figure 1: The oscillations of S_n will always be in the envelope given by $\pm\Lambda_n$.

In particular, note that LIL $\Rightarrow$ SLLN. The LIL is a refinement of the SLLN; Λ_n is sub-linear. Another perspective is by looking at it from the Kolmogorov 0-1 Law perspective: the liminf and limsup are constant almost surely.

Task # 1: Prove the Law of Iterated Logarithm.

Q: What can we say about the distribution?

The Central Limit Theorem will answer this question. For now, we will provide a heuristic overview; in the coming sections, we will rigorously do everything.

Idea: in the study of LLN, we consider $\bar{S}_n := \frac{S_n}{n}$, where $\mathbb{E}\left[\bar{S}_n\right] = \mathbb{E}\left[S_1\right] = 0$ for all $n \in \mathbb{N}$. Here, this means that $\bar{S}_n$ preserves the first moment. In **(CLT)** we will consider $\breve{S}_n := \frac{S_n}{\sqrt{n}}$, where $\mathbb{E}\left[\breve{S}_n\right] = 0$ (so, $\breve{S}_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{n}}$, where $\mathbb{E}\left[\breve{S}_n\right] = 0$. Moreover,

$$\mathbb{E}\left[(\breve{S}_n)^2\right] = \frac{n\mathbb{E}\left[X_1^2\right]}{n} = 1.$$

Note that in the CLT, the first and second moments are preserved.

1. The expected value tells us where the mass is centred.
2. The variance measures how the mass is spread out: how random the random variable is.

Heuristically, the CLT studies how the randomness will replace itself under the assumption / condition that the amount of randomness is preserved or fixed. For sure, it will not be going to a constant, and it will resemble a $N(0,1)$ as $n \to \infty$.

We work in the following set-up: $\{X_n\}$ iid random variables with $\mathbb{E}\left[X_1\right] = 0$, $\mathbb{E}\left[X_1^2\right] = 1$, and $S_n = \sum_{j=1}^n X_j$.

Remark: by preserving / stabilizing the second moments, $\breve{S}_n$ stabilizes all the moments. We can see this with the following computation / proof.

Suppose $X_1 \in L^p$ for all $p \geq 1$. We will show this stablization by induction. For some $m \in \mathbb{N}$, define:

$$L_j := \lim_{n\to\infty} \mathbb{E}\left[(\breve{S}_n)^j\right] \text{ exists for } 1 \leq j \leq m. \tag{3}$$

Consider the $(m+1)$st moment of $\breve{S}_n$:

$$
\begin{aligned}
\mathbb{E}\left[S_n^{m+1}\right] &= \mathbb{E}\left[S_n S_n^m\right] \\
&= \sum_{j=1}^{n} \mathbb{E}\left[X_j (X_j + S_{n\backslash j})^m\right] \\
&= \sum_{j=1}^{n} \sum_{k=0}^{m} \binom{m}{k} \mathbb{E}\left[X_j^{k+1}\right] \mathbb{E}\left[S_{n\backslash j}^{m-k}\right] \text{ (by the binomial formula)} \\
&= n \left(\mathbb{E}\left[X_1\right] \mathbb{E}\left[S_{n\backslash 1}^{m}\right] + m \underbrace{\mathbb{E}\left[X_1^2\right]}_{=1} \mathbb{E}\left[S_{n\backslash 1}^{m-1}\right] + \sum_{k=2}^{m} \binom{m}{k} \mathbb{E}\left[X_1^{k+1}\right] \mathbb{E}\left[S_{n\backslash 1}^{m-k}\right] \right),
\end{aligned}
$$

where $\mathbb{E}\left[X_1\right] = 0$ means the first term vanishes. Since $\mathbb{E}\left[X_1^2\right] = 1$, we get, by applying the definition of $\breve{S}_n$:

$$
\begin{aligned}
\mathbb{E}\left[(\breve{S}_n)^{m+1}\right] &= n^{-\frac{m+1}{2}} \mathbb{E}\left[S_n^{m+1}\right] \\
&= n^{-\frac{m+1}{2}} \left(m \mathbb{E}\left[S_{n\backslash 1}^{m-1}\right] + \sum_{k=2}^{m} \binom{m}{k} \mathbb{E}\left[X_1^{k+1}\right] \mathbb{E}\left[S_{n\backslash 1}^{m-k}\right] \right).
\end{aligned}
$$

Substituting in the definition of $\breve{S}_n$, we obtain:

$$
= \left(\frac{n-1}{n}\right)^{\frac{m-1}{2}} m \underbrace{\mathbb{E}\left[(\breve{S}_{n\backslash 1})^{m-1}\right]}_{:=L_{m-1}} + \sum_{k=2}^{m} \underbrace{\frac{(n-1)^{\frac{m-k}{2}}}{n^{\frac{m-1}{2}}}}_{\to 0 \text{ as } n \to \infty} \binom{m}{k} \mathbb{E}\left[X_1^{k+1}\right] \underbrace{\mathbb{E}\left[(\breve{S}_{n-1})^{m-k}\right]}_{:=L_{m-k}}.
$$

So as $n \to \infty$, we obtain:

$$
1 \cdot m \cdot L_{m-1}. \tag{4}
$$

This gives us the following recursive relationship: $L_{m+1} = m L_{m-1}$. Since $L_1 = 0$ and $L_2 = 1$, the *second moment stabilizes all the moments*:

$$
L_{2m+1} = 0 \text{ (all odd indices)} \tag{5}
$$

$$
L_{2m} = 1 \cdot 3 \cdot 4 \cdot \ldots \cdot (2m-1) \text{ (product of all the odd numbers) } = (2m+1)!! \tag{6}
$$

These are the moments of the standard Gaussian. So, the moments of $\breve{S}_n$ converge to the corresponding moments of a $N(0,1)$ random variable as $n \to \infty$. Therefore, intuitively, the distribution of $\breve{S}_n$ "approximates" $N(0,1)$ as $n \to \infty$. As a corollary, if φ is a polynomimal of any degree, then

$$
\lim_{n\to\infty} \mathbb{E}\left[\varphi\left(\breve{S}_n\right)\right] = \frac{1}{\sqrt{2\pi}} \int \varphi(x) e^{-\frac{x^2}{2}} dx = \gamma 0, 1(\varphi)
$$

where $\gamma_{0,1} = N(0,1)$.

1.2 Central Limit Theorems

Theorem 2 (Lindeberg's Central Limit Theorem (CLT))**.** *Assume that* $\{X_n\}$ *is a sequence of independent square-integrable random variables on a probability space,* $\mathbb{E}\left[X_n\right] = 0$*. For every* $n \in \mathbb{N}$*, set:*

$$
\sigma_n := \sqrt{Var(X_n)}
$$

$$
\Sigma_n := \sqrt{Var(S_n)} = \sqrt{\sum_{j=1}^{n} \sigma_j^2},
$$

where the final equality is true only if the X_n *are independent. Set*

$$\breve{S}_n = \frac{S_n}{\Sigma_n}$$

(so $\mathbb{E}\left[\breve{S}_n\right] = 0$ *and* $\mathbb{E}\left[\breve{S}_n\right] = 1$*). For all* $\varepsilon > 0$*, set:*

$$g_n(\varepsilon) := \frac{1}{\Sigma_n^2}\sum_{j=1}^{n}\mathbb{E}\left[X_j^2; |X_j| > \varepsilon\Sigma_n\right] \ \ or$$
$$g_n(\varepsilon) := \sum_{j=1}^{n}\mathbb{E}\left[\left(\frac{X_j}{\Sigma_n}\right)^2; \left|\frac{X_j}{\Sigma_n}\right| > \varepsilon\right].$$

Under this setting, for every $\varphi \in C^3(\mathbb{R})$ *with* φ'' *and* φ''' *being bounded on* $\mathbb{R}$ *and for every* $\varepsilon > 0$*,*

$$\left|\mathbb{E}\left[\varphi(\breve{S}_n)\right] - \gamma_{0,1}(\varphi)\right| \leq \frac{1}{2}(\varepsilon + \sqrt{g_n(\varepsilon)})||\varphi'''||_n + g_n(\varepsilon)||\varphi''||_n. \tag{7}$$

In particular, if for all $\varepsilon > 0$*,*

$$\lim_{n\to\infty} g_n(\varepsilon) = 0, \tag{8}$$

(this is called ***Lindeberg's Condition****), then*

$$\lim_{n\to\infty}\mathbb{E}\left[\varphi(\breve{S}_n)\right] = \gamma_{0,1}(\varphi).$$

Before the proof, we first make a quick remark. In the case when $\{X_n \mid n \geq 1\}$ is iid with $\mathbb{E}[X_1] = 0$, $\mathbb{E}\left[X_1^2\right] = 1$ for all $n \geq 1$, $\sigma_n = 1$, $\Sigma_n = \sqrt{n}$. Hence,

$$\breve{S}_n = \frac{S_n}{\sqrt{n}},$$

and so, for all $\varepsilon > 0$,

$$\begin{aligned} g_n(\varepsilon) &= \frac{1}{\Sigma_n^2}\sum_{j=1}^{n}\mathbb{E}\left[X_j^2; |X_j| > \varepsilon\Sigma_n\right] \\ &= \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left[X_j^2; |X_j| > \varepsilon\sqrt{n}\right] \\ &= \mathbb{E}\left[X_1^2; |X_1| > \varepsilon\sqrt{n}\right] \to 0 \text{ as } n \to \infty. \end{aligned}$$

So, in this case, Lindeberg's Condition is always satisfied.

Proof. Before the proof, the insight is as follows: as $n \to \infty$, the contribution of the X_j's are getting closer and closer to a centered Gaussian $N(0, \sigma_j^2)$ random variable.

Introduce $\{Z_n \mid n \geq 1\}$ iid random variables independent of $\{X_n \mid n \geq 1\}$. For all $n \geq 1$, set $Y_n := \sigma_n Z_n$. Then, as we know Y_n is a $N(0, \sigma_n^2)$ random variable. Further define $\breve{T}_n := \frac{1}{\Sigma_n}\sum_{j=1}^{n} Y_j$. Note that $\breve{T}_n$ is a $N(0, 1)$ random variable. Hence,

$$\gamma_{0,1}(\varphi) = \mathbb{E}\left[\varphi(\breve{T}_n)\right] \Rightarrow \mathbb{E}\left[\varphi(\breve{S}_n)\right] - \gamma_{0,1}(\varphi) = \mathbb{E}\left[\varphi(\breve{S}_n) - \varphi(\breve{T}_n)\right].$$

Hence,

$$\begin{aligned}\varphi(\breve{S}_n) - \varphi(\breve{T}_n) =& \varphi\left(\frac{1}{\Sigma_n}(X_1 + ... + X_n)\right) - \varphi\left(\frac{1}{\Sigma_n}(X_1 + ... + X_{n-1} + Y_n)\right) + \varphi\left(\frac{1}{\Sigma_n}(X_1 + ... + X_{n-1} + Y_n)\right)\\ &- \varphi\left(\frac{1}{\Sigma_n}(X_1 + ... + Y_{n-1} + Y_n)\right) + \varphi\left(\frac{1}{\Sigma_n}(X_1 + ... + Y_{n-1} + Y_n)\right) - ...\\ &- \varphi\left(\frac{1}{\Sigma_n}(X_1 + Y_2 + ... + Y_n)\right) + \varphi\left(\frac{1}{\Sigma_n}(X_1 + Y_2 + ... + Y_n)\right) + \varphi\left(\frac{1}{\Sigma_n}(Y_1 + ... + Y_n)\right).\end{aligned}$$

In light of this representation, for all $1 \leq j \leq n$, set:

$$U_j := \frac{1}{\Sigma_n}(X_1 + ... + X_{j-1} + X_{j+1} + Y_{j+2} + ... + Y_n). \tag{9}$$

Then, we can express the above more compactly as:

$$\varphi(\breve{S}_n) - \varphi(\breve{T}_n) = \sum_{j=1}^{n}\left(\varphi\left(U_j + \frac{X_j}{\Sigma_n}\right) - \varphi\left(U_j + \frac{Y_j}{\Sigma_n}\right)\right)$$

The idea is to now use Taylor expansions: recall that the Taylor Expansion of φ is:

$$\varphi(U_j + \xi) = \varphi(U_j) + \xi\varphi'(U_j) + \frac{\xi^2}{2}\varphi''(U_j). + ...$$

Set $R_j(\xi) = \varphi(U_j + \xi) - \varphi(U_j) - \xi\varphi'(U_j) - \frac{1}{2}\xi^2\varphi'(U_j)$. Then,

$$\mathbb{E}\left[\varphi\left(U_j + \frac{X_j}{\Sigma_n}\right)\right] = \mathbb{E}\left[R_j\left(\frac{X_j}{\Sigma_n}\right)\right] + \mathbb{E}\left[\varphi(U_j)\right] + \mathbb{E}\left[\frac{X_j}{\Sigma_n}\varphi'(U_j)\right] + \frac{1}{2}\mathbb{E}\left[\frac{X_j^2}{\Sigma_n}\varphi''(U_j)\right].$$

Let's simplify all these terms:

- Since X_j is independent of U_j, we can write:

$$\mathbb{E}\left[\frac{X_j}{\Sigma_n}\varphi'(U_j)\right] = \frac{1}{\Sigma_n}\mathbb{E}\left[X_j\right]\mathbb{E}\left[\varphi'(U_j)\right] = 0.$$
$$\frac{1}{2}\mathbb{E}\left[\frac{X_j^2}{\Sigma_n}\varphi''(U_j)\right] = \frac{1}{2}\mathbb{E}\left[\frac{X_j^2}{\Sigma_n^2}\right]\cdot\mathbb{E}\left[\varphi''(U_j)\right] = \frac{\sigma_j^2}{\Sigma_n^2}\mathbb{E}\left[\varphi''(U_j)\right]$$

Similarly,

$$\mathbb{E}\left[\varphi\left(U_j + \frac{Y_j}{\Sigma_n}\right)\right] = \mathbb{E}\left[R_j\left(\frac{Y_j}{\Sigma_n}\right)\right] + \mathbb{E}\left[\varphi(U_j)\right] + 0 + \frac{1}{2}\frac{\sigma_j^2}{\Sigma_n^2}\cdot\mathbb{E}\left[\varphi''(U_j)\right].$$

Therefore,

$$\begin{aligned}\left|\mathbb{E}\left[\varphi(\breve{S}_n) - \varphi(\breve{T}_n)\right]\right| &\leq \sum_{j=1}^{n}\left|\mathbb{E}\left[R_j\left(\frac{X_j}{\Sigma_n}\right)\right] - \mathbb{E}\left[R_j\left(\frac{Y_j}{\Sigma_n}\right)\right]\right|\\ &\leq \sum_{j=1}^{n}\left|\mathbb{E}\left[R_j\left(\frac{X_j}{\Sigma_n}\right)\right]\right| + \left|\mathbb{E}\left[R_j\left(\frac{Y_j}{\Sigma_n}\right)\right]\right|\end{aligned}$$

Moreover, $|R_j(\xi)| \leq \left(\frac{1}{6}\xi^3||\varphi'''||_n\right) \wedge \left(\xi^2||\varphi''||_n\right)$, where the first case happens if ξ is small and the second case happens if ξ is not small. Hence, for all $\varepsilon > 0$, we have:

$$\sum_{j=1}^{n}\left|\mathbb{E}\left[R_j\left(\frac{X_j}{\Sigma_n}\right)\right]\right| \leq \frac{1}{6}||\varphi''||_n\sum_{j=1}^{n}\mathbb{E}\left[\frac{|X_j|^3}{\Sigma_n^3}; |X_j| \leq \varepsilon\Sigma_n\right] + ||\varphi''||_n\sum_{j=1}^{n}\mathbb{E}\left[\frac{|X_j|^2}{\Sigma_n^2}; \frac{|X_j|}{\Sigma_n} > \varepsilon\right],$$

where the first term in the sum comes from the bound for ξ being small and the second term in the sum comes fro the bound for ξ being not so small. Pulling one of the $|X_j|$ out of the fraction in the first term of the sum, and using the bound given, we obtain:

$$\leq \frac{\varepsilon}{6}||\varphi''||_n \sum_{j=1}^{n} \frac{\mathbb{E}\left[X_j^2\right]}{\Sigma_n^2} + ||\varphi''||_n \cdot g_n(\varepsilon),$$

which is good, since we have $\sum_{j=1}^{n} \frac{\sigma_j^2}{\Sigma_n^2} = 1$. Hence,

$$\sum_{j=1}^{n} \left| \mathbb{E}\left[R_j\left(\frac{X_j}{\Sigma_n}\right)\right]\right| \leq \frac{\varepsilon}{6}||\varphi''|||_n + ||\varphi''||_n \cdot g_n(\varepsilon).$$

Similarly,

$$\begin{aligned} \sum_{j=1}^{n} \mathbb{E}\left[\left|R_j\left(\frac{Y_j}{\Sigma_n}\right)\right|\right] &\leq \frac{1}{6}||\varphi'''||_n \mathbb{E}\left[|Z_n|^3\right] \sum_{j=1}^{n} \frac{\sigma_j^3}{\Sigma_n^3} \\ &\leq \frac{1}{3}||\varphi'''||_n \max_{1\leq j\leq n} \frac{\sigma_j}{\Sigma_n} \cdot \underbrace{\sum_{j=1}^{n} \frac{\sigma_j^2}{\Sigma_n^2}}_{=1}. \end{aligned}$$

We have that for all $1 \leq j \leq n$,

$$\begin{aligned} \sigma_j^2 = \mathbb{E}\left[X_j^2\right] &= \mathbb{E}\left[X_j^2; |X_j| \leq \varepsilon\Sigma_n\right] + \mathbb{E}\left[X_j^2; |X_j| > \varepsilon\Sigma_n\right] \\ &= \varepsilon^2\Sigma_n^2 + \sum_{l=1}^{n} \mathbb{E}\left[X_l^2; |X_l| > \varepsilon\Sigma_n\right]. \end{aligned}$$

Hence,

$$\max_{1\leq j\leq n} \frac{\sigma_j^2}{\Sigma_n^2} \leq \varepsilon^2 + g_n(\varepsilon) \Rightarrow \max_{1\leq j\leq n} \frac{\sigma_j}{\Sigma_n} \leq \sqrt{\varepsilon^2 + g_n(\varepsilon)} \leq \varepsilon + \sqrt{g_n(\varepsilon)}.$$

Collecting all the bounds,

$$\begin{aligned} \left|\mathbb{E}\left[\varphi(\breve{S}_n)\right] - \mathbb{E}\left[\varphi(\breve{T}_n)\right]\right| &\leq \frac{\varepsilon}{6}||\varphi'''||_n + g_n(\varepsilon)||\varphi''||_n + \frac{1}{3}||\varphi'''||_n(\varepsilon + \sqrt{g_n(\varepsilon)} \\ &\leq \frac{1}{2}\left(\varepsilon + \sqrt{g_n(\varepsilon)}\right)||\varphi'''||_n + g_n(\varepsilon)||\varphi''||_n \end{aligned}$$

which proves the theorem. $\square$

Corrolary 1. Under the same setting as before, if Lindeberg's condition holds, then for all $\varphi \in C_c^\infty(\mathbb{R})$,

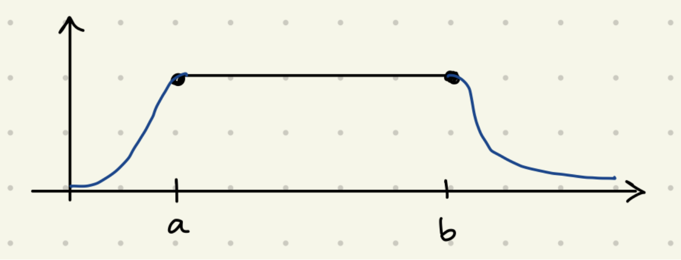
$$\lim_{n\to\infty} \mathbb{E}\left[\varphi(\breve{S}_n)\right] = \gamma_{0,1}(\varphi). \tag{10}$$

In particular, we can show that for all $a, b \in \mathbb{R}$, $a < b$:

$$\mathbb{P}\left(a \leq \breve{S}_n \leq b\right) = \gamma_{0,1}(]a, b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

Proof. The proof only requires a standard fact from analysis, which we will use quite a lot in this course.

Fact. For $[a,b]$ closed, there exists a sequence of functions $\{\varphi_k \mid k \geq 1\} \subseteq C_c^\infty(\mathbb{R})$ such that $0 \leq \varphi_k \leq 1$ for all $k \geq 1$ and $\varphi_k \downarrow \chi_{[a,b]}$. The picture that you want to have in mind is:



Therefore, for all $k \geq 1$, we have

$$\limsup_n \mathbb{P}\left(\breve{S}_n \in [a,b]\right) \leq \lim_n \mathbb{E}\left[\varphi_k(\breve{S}_n)\right] = \gamma_{0,1}(\varphi_k),$$

where the final equality follows from the Lindeberg's CLT. As $k \to \infty$,

$$\gamma_{0,1}(\varphi_k) \to \gamma_{0,1}([a,b]).$$

Hence, $\limsup_n \mathbb{P}\left(a \leq \breve{S}_n \leq b\right) \leq \gamma_{0,1}([a,b])$. Similarly, for $]a,b[$, there exists a sequence of functions $\{\psi_k \mid k \geq 1\}$ such that $0 \leq \psi_k \leq 1$ for all $k \geq 1$, $\psi_k \uparrow \chi_{]a,b[}$ (so, we approach the indicator function from below). Then,

$$\liminf_n \mathbb{P}\left(a < \breve{S}_n < b\right) \geq \lim_n \mathbb{E}\left[\psi_k(\breve{S}_n)\right] = \gamma_{0,1}(\varphi_k) \to \gamma_{0,1}(]a,b[).$$

Since $\gamma_{0,1}(]a,b[) = \gamma_{0,1}([a,b])$ we have the desired limit statement. ☐

So, now we want to look at smooth functions that approximates the indicator function χ of a set we are interested in studying. Let's first do some preparation.

Definition 1 (Convolution)**.** Given μ and ν, two probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ given by: for all $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\mu * \nu := \int_{\mathbb{R}^d} \nu(B-x)\mu(dx), \tag{11}$$

where recall the set $B - x := \{y \in \mathbb{R}^d \mid y + x \in B\}$.

Remarks. It's easy to check with Fubini's Theorem that:

1. $x \mapsto \nu(B-x)$ is a measure with respect to $\mathcal{B}(\mathbb{R}^d)$.
2. $\mu * \nu$ is again a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.
3. $\mu * \nu = \nu * \mu$. If ρ is another probability measure on $\mathbb{R}^d$, then,

$$(\mu * \nu) * \rho = \mu * (\nu * \rho). \tag{12}$$

In the next proposition, we will see how convolution corresponds to taking the sum of two independent random variables.

Proposition 1. Given X and Y two independent random variables, $\mathbb{R}^d$-valued, with $\mathcal{L}_X = \mu$ and $\mathcal{L}_Y = \nu$. If X and Y are independent, then

$$\mathcal{L}_{X+Y} = \mu * \nu. \tag{13}$$

Proof. To see this, we first have that since X and Y are independent,

$$\mathcal{L}_{(X,Y)} = \mathcal{L}_X \cdot \mathcal{L}_Y = \mu \times \nu.$$

So, using Fubini's theorem, we obtain that for all $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned}\mathbb{P}\left(X + Y \in B\right) &= \iint_{\mathbb{R}^d \times \mathbb{R}^d} \chi_B(x+y)(\mu \times \nu) \\ &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \chi_B(x+y)\nu(dy) \right) \mu(dx) \\ &= \mu * \nu(B).\end{aligned}$$

□

Remark. It's also possible to define the convolution of functions. Given f and g two functions on $\mathbb{R}^d$, for all $x \in \mathbb{R}^d$:

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

provided that the integral is defined. Similarly, $f * g = g * f$ and $(f * g) * h = f * (g * h)$.

Corrolary 2. If X and Y are independent, and X has a density f and Y has a density g, then $X + Y$ has a density $f * g$.

Notation. for every $x, \xi \in \mathbb{R}^d$, we will denote by $(\cdot, \cdot)$ the dot product:

$$(x, \xi) := \sum_{j=1}^{d} x_j \xi_j.$$

We will denote by $i := \sqrt{-1}$ the imaginary unit. For $z \in \mathbb{C}$, let $\overline{z}$ be the complex conjugate of z. We consider functions $\varphi : \mathbb{R}^d \to \mathbb{C}$. As we'd expect, φ is Borel $\iff$ the real and imaginary parts of φ are Borel functions in the standard sense. If μ is a probability measure on $\mathbb{R}^d$, then we write that $\varphi \in L^p(\mu)$ if

$$\int_{\mathbb{R}^d} |\varphi(x)|^p \mu(dx) < \infty.$$

(note that $|\varphi(x)|^2 = \mathrm{Re}^2(\varphi) + \mathrm{Im}(\varphi)$). Given two functions ψ and φ, $\mathbb{C}$-valued on $\mathbb{R}^d$, their inner product is given by:

$$\langle \varphi, \psi \rangle = (\varphi, \psi)_{L^2} = \int_{\mathbb{R}^d} \varphi(x)\overline{\psi(x)}dx.$$

Definition 2 (Characteristic Function)**.** Given a probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, the **characteristic function** of μ, denoted by $\hat{\mu}$, is a function on $\mathbb{R}^d$ such that for all $\xi \in \mathbb{R}^d$,

$$\hat{\mu}(\xi) := \int_{\mathbb{R}^d} e^{i(x,\xi)}\mu(dx). \tag{14}$$

$\hat{\mu} : \mathbb{R}^d \to \mathbb{C}$ is well-defined for every ξ, measurable (Fubini) with respect to $\mathcal{B}(\mathbb{R}^d)$, and $|\hat{\mu}| \leq 1$ for all $\xi \in \mathbb{R}^d$.

We can similarly define the characteristic function of a random variable. If X is a random variable on some probability space such that $\mathcal{L}_X = \mu$, then

$$\hat{\mu}(\xi) = \mathbb{E}\left[e^{i(X,\xi)}\right].$$

We now introduce some remarks on characteristic functions.

1. $\hat{\mu} : \mathbb{R}^d \to \mathbb{C}$ is a continuous function. (Can easily verify this by taking a sequence ξ_n, and use **(DOM)** since everything is bounded by 1).
2. If μ is symmetric, i.e., $\forall\ A \in \mathcal{B}(\mathbb{R}^d)$, $\mu(A) = \mu(-A)$. Then, $\hat{\mu}(\xi) \in \mathbb{R}$ for all $\xi \in \mathbb{R}^d$, since by symmetry, the imaginary part will cancel.
3. If μ and ν are two probability measures, then,

$$\mu \mathbin{\hat{*}} \nu(\xi) = \hat{\mu}(\xi) \cdot \hat{\nu}(\xi) = \hat{\mu}(\xi) \cdot \hat{\nu}(\xi),$$

for all $\xi \in \mathbb{R}^d$. To see this, implement with random variables. Take X and Y independent such that $\mathcal{L}_X = \mu$ and $\mathcal{L}_Y = \nu$. Then,

$$\begin{aligned} \mu \mathbin{\hat{*}} \nu(\xi) &= \mathbb{E}\left[e^{i(X+Y,\xi)}\right] \\ &= \mathbb{E}\left[e^{i(X,\xi)} e^{i(Y,\xi)}\right] \\ &= \mathbb{E}\left[e^{i(X,\xi)}\right] \mathbb{E}\left[e^{i(Y,\xi)}\right] \\ &= \hat{\mu}(\xi) \cdot \hat{\nu}(\xi). \end{aligned}$$

4. $\hat{\mu}$ contains information about "moments". To see this, assume that X is a random variable such that $\mathcal{L}_X = \mu$ and $\mathbb{E}\left[|X|^p\right] < \infty$ for some $p \geq 1$. Then, for every multi-index $\alpha = (\alpha_1, \alpha_2, ..., \alpha_d) \in \mathbb{N}^d$ such that $|\alpha| = \alpha_1 + \alpha_2 + ... + \alpha_d \leq p$. Then,

$$\begin{aligned} \partial^\alpha \hat{\mu}(\xi) &:= \partial_1^{\alpha_1} \partial_2^{\alpha_2} ... \partial_d^{\alpha_d} \hat{\mu}(\xi) \\ &= \int_{\mathbb{R}^d} (ix)^\alpha e^{i(x,\xi)} \mu(dx). \end{aligned}$$

This follows from **(DOM)** . The notation $(ix)^\alpha$ means $i^{|\alpha|} x_1^{\alpha_1} \cdot x_2^{\alpha_2} ... x_d^{\alpha_d}$. In particular, we have

$$\left[\partial^\alpha \hat{\mu}(\xi)\right]_{\xi=0} = i^{|\alpha|} \mathbb{E}\left[X^\alpha\right].$$

The $\mathbb{E}\left[X^\alpha\right]$ term is called the **cross-moment**. The notation X^α means $X_1^{\alpha_1} X_2^{\alpha_2} \cdots X_d^{\alpha_d}$.

   (a) In general, $\hat{\mu} \in C^p(\mathbb{R}^d)$ does NOT imply that $\mathbb{E}\left[|X|^p\right] < \infty$. For example, consider that μ is the probability measure on $\mathbb{R}$ with density:

$$f(x) = \begin{cases} 0 & \text{if } |x| < 2 \\ \frac{c}{x^2 \ln(|x|)} & \text{if } |x| \geq 2, \end{cases}$$

   where $c > 0$ is a constant such that $\int_{\mathbb{R}} f(x) dx = 1$. Now let X be a random variable such that $\mathcal{L}_X = \mu$. On one hand,

$$\mathbb{E}\left[|X|\right] = 2 \int_{2\infty} \frac{dx}{x \ln(x)} \cdot c = \infty \Rightarrow X \notin L^1.$$

   On the other hand,

$$\hat{\mu}(\xi) = 2c \int_2^\infty \frac{\cos(x\xi)}{x^2 \ln(x)} dx.$$

   One can verify that $\hat{\mu}$ is differentiable at every $\xi \in \mathbb{R}$ and $\hat{\mu}'(0) = 0$.

Example 1. On $\mathbb{R}$,

$$\hat{\partial}_{m,\sigma^2}(\xi) = e^{im\xi}e^{-\frac{1}{2}\sigma^2\xi^2}.$$

On $\mathbb{R}^d$,

$$\hat{\partial}_{\vec{m},c}(\xi) = e^{i(\vec{m},\xi)} \cdot e^{-\frac{(\xi,c\xi)}{2}},$$

for all $\xi \in \mathbb{R}^d$. Observe that the characteristic functions have super-exponential decay like the densities.

Definition 3. Given a function φ on $\mathbb{R}^d$, the **Fourier Transform** of φ, denoted by $\hat{\varphi}$, is given by: for all $\varphi \in \mathbb{R}^d$:

$$\hat{\varphi}(\xi) := \int_{\mathbb{R}^d} e^{i(x,\xi)}\varphi(x)dx, \tag{15}$$

provided that the integral is defined. In particular, this means that if μ is a probability measure with density φ, then for all $\xi \in \mathbb{R}^d$:

$$\hat{\mu}(\xi) = \hat{\varphi}(\xi). \tag{16}$$

The next theorem tells us that the characteristic function uniquely defines a probability measure.

Theorem 3. *Let μ and ν be two probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. If $\hat{\mu}(\xi) = \hat{\nu}(\xi)$ for all $\xi \in \mathbb{R}^d$, then $\mu = \nu$.*

The proof will follow from these three lemmas.

Lemma 1. Let μ and ν be two probability measures. If $\mu(\varphi) = \nu(\varphi)$ for all $\varphi \in C_b(\mathbb{R}^d; \mathbb{C})$. Then, $\mu = \nu$.

Proof. Since the open sets form a generating π-system, it's sufficient to show that for every open set $B \subseteq \mathbb{R}^d$, $\mu(B) = \nu(B)$. Take B open, and consider $d(x, B^c) := \inf_{y \in B^c} |x - y|$. Then, we know from analysis that $x \mapsto d(x, B^c)$ is continuous. For all $k \geq 1$, set

$$\varphi_k(x) := \left(\frac{d(x, B^c)}{1 + d(x, B^c)}\right)^{1/k}.$$

Ten, $\varphi_k \in [0, 1]$ and for each k, φ_k is continuous. We have that $\varphi_k \uparrow \chi_B$. Why?

$$\lim_{k\to\infty} \varphi_k(x) = \begin{cases} 1 & \text{if } d(x, B^c) > 0 \iff x \in B^c \text{ since closed} \\ 0 & \text{if } d(x, B^c) = 0 \iff x \in B^c. \end{cases}$$

By **(DOM)** or **(MON)** ,

$$\mu(B) = \lim_k \mu(\varphi_k) = \lim_k \nu(\varphi_k) = \nu(B).$$

□

Lemma 2. For all $\varphi \in C_b(\mathbb{R}^d)$ there exists a sequence $\{\varphi_m \mid m \geq 1\} \subseteq C_c^\infty(\mathbb{R}^d)$ such that $||\varphi_m||_n \leq ||\varphi_n||$ for all $m \geq 1$ and $\lim_{m\to\infty} \varphi_m = \varphi$.

(As a result of Lemma 2, $\mu = \nu$ if $\mu(\varphi) = \nu(\varphi)$ for all $C_c^\infty(\mathbb{R}^d)$.

Lemma 3 (A Generalization of Plancherel's Theorem)**.** If $\psi \in C_c^\infty(\mathbb{R}^d)$ and μ is a probability measure on $\mathbb{R}^d$, then

$$\mu(\psi) = \int_{\mathbb{R}^d} \psi(x)\mu(dx) = (2\pi)^{-d}\int_{\mathbb{R}^d} \hat{\psi}(\xi)\overline{\hat{\mu}(\xi)}d\xi, \tag{17}$$

i.e., $\mu(\psi) = \langle \hat{\psi}, \hat{\mu}\rangle$.

As a result of Lemma 3, $\mu = \nu$ if $\hat{\mu} = \hat{\nu}$. We will neatly collect this into a theorem.

Theorem 4. *Let μ and ν be two probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Then,*

$$\begin{aligned}\mu = \nu &\iff \mu(\varphi) = \nu(\varphi)\ \forall\ \varphi \in C_b(\mathbb{R}^d)\\ &\iff \mu(\psi) = \nu(\psi)\ \forall\ \psi \in C_c^\infty(\mathbb{R}^d)\\ &\iff \hat{\mu}(\xi) = \hat{\nu}(\xi)\ \forall \xi \in \mathbb{R}^d.\end{aligned}$$

We can think of $C_b(\mathbb{R}^d)$ and C_c^∞ as classes of test functions that test how measure behaves.

1.3 Weak Convergence of Probability Measures

There are only two types of convergence which will be covered in this course.

Definition 4 (Weak Convergence of Measure)**.** Assume that $\{\mu_n \mid n \geq 1\}$ and μ are probability measures on $\mathcal{B}(\mathbb{R}^d)$. We say that μ_n **converges weakly** to μ, and we write "$\mu_n \Rightarrow \mu$" if for all $\varphi \in C_b(\mathbb{R}^d; \mathbb{C})$:

$$\lim_{n\to\infty} \mu_n(\varphi) = \mu(\varphi). \tag{18}$$

We also have cnvergence in distribution of random variables.

Definition 5. Assume that $\{X_n\}$ and X are $\mathbb{R}^d$-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. X_n converges to X **in distribution**, denoted by "$X_n \to X$ in distribution", if $\mathcal{L}X_n \to \mathcal{L}_X$, i.e., for all $\varphi \in C_b(\mathbb{R}^d)$,

$$\lim_{n\to\infty} \mathbb{E}\left[\varphi(X_n)\right] = \mathbb{E}\left[\varphi(X)\right]. \tag{19}$$

Remark.

1. For two probability measures μ and ν on $\mathbb{R}^d$, the most natural way of putting a metric on the space of probability measures is the **total variation distance** between μ and ν:

$$||\mu - \nu||_{\text{var}} := \sup\{|\mu(A) - \nu(A)| \mid A \in \mathcal{B}(\mathbb{R}^d)\}. \tag{20}$$

2. Given $\{\mu_n\}$ and μ probability measures on $\mathbb{R}^d$, if $\lim_{n\to\infty} ||\mu_n - \mu||_{\text{var}} = 0$, then μ_n converges to μ **in the strong sense**.

Exercise. verify that if $||\mu_n - \mu||_{\text{var}} \to 0$ then $\mu_n \Rightarrow \mu$.

It is often inconvenient to work with strong convergence. For example, we know that $\lim_{n\to\infty} \frac{1}{n} = 0$. If $\mu_n := \delta_{1/n}$ for all $n \geq 0$, then $\mu = \delta_0$. So, naturally, μ_n should be getting closer and closer to μ. However, if you look at the total variation distance, for all $n \geq 1$, $||\mu_n - \mu||_{\text{var}} = 1$. Hence, μ_n does not convergence to μ in the strong sense. However, if we relax our standards, $\mu_n \Rightarrow \mu$ because for all $\varphi \in C_b(\mathbb{R})$, $\mu_n(\varphi) = \varphi\left(\frac{1}{n}\right)\varphi(0) = \mu(\varphi)$ where the convergence follows from continuity.

1. Let $\{X_n\}$ and X be $\mathbb{R}^d$-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.
   (a) If $X_n \to X$ in probability, then $X_n \to X$ in distribution.

   (b) If $X_n \to X$ in distribution and $X \equiv c$ for some constant c, then $X_n \to X$ in probability.

2. Let $\{X_n\}$ and X be $\mathbb{R}$-valued random variables such that X_n has the distribution function F_n for all $n \geq 1$ and X has distribution function F. Then,
   (a) If $X_n \to X$ in distribution, then $\lim_{n\to\infty} F_n(x) = F(x)$at every continuous point x of F.
   (b) If X_n has density f_n for all $n \geq 1$ and X has density f, and $f_n \to f$ a.e. with respect to the Lebesgue measure on $\mathbb{R}$, then $X_n \to X$ in distribution.

Proposition 2. Let $\{\mu_n\}$ and μ be probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. If for every subsequence $\{n_k\} \subseteq \mathbb{N}$, there exists a further subsequence, $\{n_{k_l}\} \subseteq \{n_k\}$ such that $\{\mu_{n_{k_l}}\} \Rightarrow \mu$, then $\mu_n \Rightarrow \mu$.

Proposition 3. If $\{\mu_n \mid n \geq 1\}$ is a sequence of probability measures on $\mathbb{R}^d$ and $\mu_n \Rightarrow \mu$ and $\mu_n \Rightarrow \nu$, then $\mu = \nu$ (the limit of weak convergence is unique).

Proof. We can very briefly sketch the argument: for all $\varphi \in C_b(\mathbb{R}^d)$,

$$\begin{aligned} \mu(\varphi) &= \lim_{n\to\infty} \mu_n(\varphi) = \nu(\varphi) \\ &\Rightarrow \text{ integrals match on all continuous and bounded functions} \\ &\Rightarrow \mu = \nu. \end{aligned}$$

□

The following proposition will be useful for the homework.

Proposition 4. Suppose $\mu_n \Rightarrow \mu$. Then:

1. For all open sets $G \subseteq \mathbb{R}^d$,

$$\mu(G) \leq \liminf_n \mu_n(G). \tag{21}$$

2. For all closed sets $F \subseteq \mathbb{R}^d$,

$$\mu(F) \geq \limsup_n \mu_n(F). \tag{22}$$

Proof. We will only prove (i). Given an open set $G \subseteq \mathbb{R}^d$, there exists a sequence $\{\varphi_k \mid k \geq 1\} \subseteq C_b(\mathbb{R}^d)$ such that $\varphi_k \uparrow \chi_G$. By **(MON)** or **(DOM)** ,

$$\mu(G) = \lim_{k\to\infty} \mu(\varphi_k) = \lim_{k\to\infty} \lim_{n\to\infty} \mu_n(\varphi_k) \leq \liminf_{n\to\infty} \mu_n(G),$$

where the second equality follows from weak convergence, and the final inequality follows from the fact that for all $k \geq 1$, $\varphi_k \leq \chi_G$. □

In fact, if $\mu(G) \leq \liminf_n \mu_n(G)$ for every open set $G \subseteq \mathbb{R}^d$, then $\mu_n \Rightarrow \mu$.

1.4 Tightness of a Family, Class, or Collection of Probability Measures

Definition 6. Let $\{\mu_n \mid n \geq 1\}$ be a sequence of probability measures on $\mathbb{R}^d$. We say that $\{\mu_n\}$ is **tight** if for all $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subseteq \mathbb{R}^d$ such that

$$\sup_n \mu_n(K_\varepsilon^c) < \varepsilon. \tag{23}$$

This is telling us that we can make the whole family uniformly small.

Remark. "tightness" means that the mass is concentrated in a way that is uniform for the μ_n's. For example:

- $\{\mu_n = \gamma_{0,1/n} \mid n \geq 1\}$ is a tight family (the variance goes down as $n \to \infty$.
- $\{\nu_n = \gamma_{0,n} \mid n \geq 1\}$ is *not* tight: as n grows, the variance gets more spread out.

Theorem 5 (Prokhorov's Theorem)**.** *Let $\{\mu_n \mid n \geq 1\}$ be a sequence of probability measures on $\mathbb{R}^d$. Then:*

1. *If there exists a probability measure on $\mathbb{R}^d$ such that $\mu_n \Rightarrow \mu$, then $\{\mu_n \mid n \geq 1\}$ is tight.*
2. *If $\{\mu_n \mid n \geq 1\}$ is tight, then there exists a subsequence $\{n_k \mid n \geq 1\} \subseteq \mathbb{N}$ and a probability measure μ on $\mathbb{R}^d$ such that along the subsequence, $\mu_{n_k} \Rightarrow \mu$ as $k \to \infty$.*

Proof. **(i)**. Assume that $\mu_n \Rightarrow \mu$. For a contradiction, assume that $\{\mu_n \mid n \geq 1\}$ is not tight: there exists an $\eta > 0$ such that for all compact sets $K \subseteq \mathbb{R}^d$,

$$\sup_n \mu_n(K_n^c) > \eta.$$

We will use this statement to extract a subseqeuence: for all $k \geq 1$, there exists an n_k such that $\mu_{n_k}(\overline{B(0,k)}^c) > \eta$. Then, for every $R > 0$ when k is sufficiently large, i.e., $k \geq R$, we get from $\mu_{n_k} \Rightarrow \mu$:

$$\begin{aligned}
\mu(B(0,R)) &\leq \liminf_{k\to\infty} \mu_{n_k}(B(0,R)) \text{ (weak convergence)} \\
&\leq \liminf_{k\to\infty} \mu_{n_k}(B(0,k)) \\
&\leq 1 - \eta.
\end{aligned}$$

Therefore, for all $R > 0$,

$$\mu(B(0,R)) \leq 1 - \eta \Rightarrow \mu(\mathbb{R}^d) < 1 - \eta,$$

where the implication follows from sending $R \to \infty$ and **(MON)** . However, this is not possible, since $\mu(\mathbb{R}^d) = 1$ since μ is a probability measure. ☐

Task: Give a rigorous proof of **(i)** of Prokhorov's Theorem. You may use:

1. Riesz-Representation Theorem
2. Stone-Weierstrass Theorem (separability of space of continuous functions on compact sets).

Theorem 6. *Let $\{\mu_n \mid n \geq 1\}$ and μ be probability measures. If $\mu_n \Rightarrow \mu$, then*

$$\lim_{n\to\infty} \hat{\mu_n}(\xi) = \hat{\mu}(\xi) \ \forall \xi \in \mathbb{R}^d,$$

(so weak convergence of measure gives us convergence of characteristic functions) and this convergence is uniform on compact sets, i.e., for all compact $K \subseteq \mathbb{R}^d$,

$$\lim_{n\to\infty} \sup_{\xi \in K} |\hat{\mu}_n(\xi) - \hat{\mu}(\xi)| = 0. \tag{24}$$

Furthermore, for all $\varphi \in C_b(\mathbb{R}^d)$ if $\{\varphi_n \mid n \geq 1\} \subseteq C_b(\mathbb{R}^d)$ such that $\sup_n ||\varphi_n||_n < \infty$ and $\varphi_n \Rightarrow \varphi$ uniformly on compact sets, then $\lim_{n\to\infty} \mu_n(\varphi_n) = \mu(\varphi)$.

Proof. For all $\xi \in \mathbb{R}^d$, the map $x \in \mathbb{R}^d \mapsto e^{i(x,\xi)} \in \mathbb{C}$ is continuous and bounded. Hence,

$$\mu_n \Rightarrow \mu \Rightarrow \hat{\mu}_n(\xi) \to \hat{\mu}(\xi).$$

We will now prove the last statement. Sicne $\mu_n \Rightarrow \mu$, the sequence $\{\mu_n \mid n \geq 1\}$ is tight for all $\varepsilon > 0$. Hence, there exists a compact set $K_\varepsilon \subseteq \mathbb{R}^d$ such that $\sup_n \mu_n(K_\varepsilon^c) < \varepsilon$ and $\mu(K_\varepsilon^c) < \varepsilon$. Hence,

$$\begin{aligned} |\mu_n(\varphi_n) - \mu(\varphi)| &= |\mu_n(\varphi_n) - \mu_n(\varphi)| + \underbrace{|\mu_n(\varphi) - \mu(\varphi)|}_{\to 0 \text{ as } n \to \infty} \\ &\leq |\mu_n(\chi_{K_\varepsilon} \cdot (\varphi_n - \varphi))| + \left|\mu_n(\chi_{K_\varepsilon^c} \cdot (\varphi_n - \varphi))\right| \\ &\leq \underbrace{\sup_{x \in K_\varepsilon} |\varphi_n(x) - \varphi(x)|}_{\to 0 \text{ as } n \to \infty} \cdot 1 + \underbrace{\sup_n \mu_n(K_\varepsilon^c)}_{<\varepsilon} \cdot \underbrace{(\sup_n ||\varphi_n||_k + ||\varphi||_n)}_{<\infty}, \end{aligned}$$

where the first convergence to zero occurs since $\varphi_n \to \varphi$ uniformly on compact sets K_ε, the second term is assumed to be less than ε and the final term was assumed to be finite. Therefore, $\mu_n(\varphi_n) \to \mu(\varphi)$.

In particular, if $\{\xi_n \mid n \geq 1\} \subseteq \mathbb{R}^d$ such that $\xi_n \to \xi$ as $n \to \infty$, then $\hat{\mu}_n(\xi_n) \to \hat{\mu}(\xi)$ as $n \to \infty$, because we could simply take $\varphi_n = e^{i(\cdot, \xi/n)}$ and $\varphi = e^{i(\cdot, \xi)}$ $(*)$.

We need to now prove that for all compact sets $K \subseteq \mathbb{R}^d$,

$$\sup_{\xi \in K} |\hat{\mu}_n(\xi) - \hat{\mu}(\xi)| \to 0 \text{ as } n \to \infty.$$

For a contradiction, assume otherwise: there exists K compact, $\eta > 0$, a subsequence $\{n_k\} \subseteq \mathbb{N}$ such that

$$\sup_{\xi \in K} |\hat{\mu}_{n_k}(\xi) - \hat{\mu}(\xi)| > \eta.$$

Thus, there exists a subsequence $\xi_{n_k} \in K$ such that

$$|\hat{\mu}_{n_k}(\xi_{n_k}) - \hat{\mu}(\xi_{n_k})| > \eta.$$

Since K is compact, $\{\xi_{n_k} \mid k\} \subseteq K \Rightarrow$ that there exists $\{n_{k_l} \subseteq \{n_k\}$ and there exists a $\xi_0 \in K$ such that $\xi_{n_{k_l}} \to \xi_0$. Hence,

$$\left|\hat{\mu}_{n_{k_l}}(\xi_{n_{k_l}}) - \mu(\xi_{n_{k_l}})\right| \leq \underbrace{\left|\hat{\mu}_{n_{k_l}}(\xi_{n_{k_l}}) - \hat{\mu}(\xi_0)\right|}_{\to 0 \text{ by } (*)} + \underbrace{\left|\hat{\mu}(\xi_0) - \hat{\mu}(\xi_{n_{k_l}})\right|}_{\to 0 \text{ since } \hat{\mu} \text{ is cts}}$$

Contradiction! Therefore, $\hat{\mu}_n \to \hat{\mu}$ uniformly on compact sets. □

Theorem 7. *If $\{\mu_n\}$ is a sequence of probability measures on $\mathbb{R}^d$ and $\{\mu_n\}$ is tight and $\lim_{n\to\infty} \hat{\mu_n}(\xi) = f(\xi)$ for all $\xi \in \mathbb{R}^d$, then there exists a probability measure μ on $\mathbb{R}^d$ such that $\mu_n \Rightarrow \mu$ and $\hat{\mu} = f$.*

Proof. For all subsequences $\{n_k\} \subseteq \mathbb{N}$, since $\{\mu_{n_k}\}$ is tight, there exists a subsequence $\{n_{k_l} \subseteq \{n_k\}$ and a probability measure $\mu^{\{n_k\}}$ (the existence depends on the choice of $\{n_k\}$, hence the superscript) such that $\mu_{n_{k_l}} \Rightarrow \mu^{\{n_k\}}$. According to the assumption on $\hat{\mu_n}$,

$$\widehat{\mu^{\{n_k\}}} = f(\xi) \; \forall \; \xi \in \mathbb{R}^d,$$

i.e., f doesn't depend on the choice of subsequence $\{n_k\}$. Hence, $\mu^{\{n_k\}} = \mu$ is identical for all choices of $\{n_k\}$. Hence, convergence is achieved along the full subsequence:

$$\mu_n \Rightarrow \mu \text{ and } \hat{\mu} = f.$$

□

Remark. the "tightness" condition is necessary in the previous theorem. To see why, consider the following example to see what happens when the tightness assumption is dropped.

- for all $n \geq 1$, set $\mu_n = \gamma_{(0,n)}$. Clearly, $\{\mu_n \mid n \geq 1\}$ is not tight. In the limit, for all $\xi \in \mathbb{R}$:

$$\hat{\mu}_n = \exp\left(\frac{-n\xi}{2}\right) \to \begin{cases} 0 & \text{if } \xi \neq 0 \\ 1 & \text{if } \xi = 0. \end{cases}$$

  So, $\lim_{n\to\infty} \hat{\mu}_n(\xi)$ exists for every $\xi \in \mathbb{R}$. But, we will show that it cannot converge to a probability measure.
- claim: μ_n does not weakly converge to μ for any probability measure μ. For a contradiction, assume otherwise: suppose that there exists a probability measure μ such that $\mu_n \Rightarrow \mu$. Then, for every $L > 0$,

$$\mu(]-L, +L[) \leq \liminf_n \mu_n(]-L, +L[) = 0,$$

  which is not possible.

The following theorem gives us an alternate, easier way to check if there exists a probability measure such that $\mu_n \Rightarrow \mu$, since its condition (continuity at zero) is easier to check than checking if a family of measures is tight.

Theorem 8 (Levy's Continuity Theorem)**.** *Let $\{\mu_n \mid n \geq 1\}$ be a sequence of probability measures on $\mathbb{R}^d$ such that for all $\xi \in \mathbb{R}^d$,*

$$\lim_{n\to\infty} \hat{\mu}_n(\xi) = f(\xi), \tag{25}$$

exists. Further assume that $\underline{f : \mathbb{R}^d \to \mathbb{C} \text{ is continuous at zero}}$ (this is the condition which replaces tightness). Then, there exists a probability measure μ on $\mathbb{R}^d$ such that $\mu_n \Rightarrow \mu$ and $\hat{\mu} = f$.

Before the proof, we first require a lovely technical lemma.

Lemma 4. Given μ a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, for all $r > 0$, $R > 0$, for all unit vectors $\vec{e} \in \mathbb{R}^d$, we have the following two estimates:

1. $|1 - \hat{\mu}(r\vec{e})| \leq rR + 2\mu(\{x \in \mathbb{R}^d \mid |(\vec{x}, \vec{e}) > R\})$
2. set $m(t) := \inf_{|u|>t} \left(1 - \frac{\sin(u)}{u}\right)$ for all $t > 0$. Then,

$$\mu(\{x \in \mathbb{R}^d \mid |(\vec{x}, \vec{e})| > R\}) \leq \left(\frac{1}{r}\int_0^r |1 - \hat{\mu}(s\vec{e})|ds\right) \cdot \frac{1}{m(rR)} \tag{26}$$

Proof. Will include later. □

Proof. **(Proof of Levy's Continuity Theorem)**. We only need to show that if $\hat{\mu}_n(\xi) \to f(\xi)$ for all $\xi \in \mathbb{R}^d$ and f is continuous at zero, then $\{\mu_n \mid n \geq 1\}$ is tight. Obviously,

$$f(0) = \lim_{n\to\infty} \hat{\mu}_n(0) = 1.$$

By continuity, for all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $|\xi| \leq \delta$, $|1 - f(\xi)| < \varepsilon$. For $j = 1, ..., d$, let $\vec{e}_j$ be the jth standard basis vector of $\mathbb{R}^d$. By (ii) of the previous lemma, for all $n \geq 1$,

$$\mu_n\left(\left\{x \in \mathbb{R}^d \mid |(\vec{x}, \vec{e}_j)| > \frac{2}{\delta}\right\}\right) \leq \frac{1}{\delta}\int_0^\delta |1 - \hat{\mu}(s\hat{e}_j)ds \cdot \frac{1}{m(2)} \to \frac{1}{\delta}\int_0^\delta |(1 - f(s\vec{e}_j)|ds\frac{1}{m(t)} \leq 2\varepsilon,$$

where the convergence follows from **(DOM)** . Hence, for all $\varepsilon > 0$, there exists an $R = \frac{2}{\delta}$, there exists an $N \geq 1$, such that for all $n \geq N$,

$$\mu_n\left(\overline{B(0, \sqrt{d}R}^c\right) \leq \sum_{j=1}^{d} \mu_n(\{x \in \mathbb{R}^d \mid |(\vec{x}, \vec{e_j})| > R\}) \leq 2d\varepsilon.$$

So, for $n = 1, ..., N-1$, if necessary, we can make R even larger such that

$$\sup_{1 \leq n \leq N-1} \mu_n\left(\overline{B(0, \sqrt{d}R}^c\right) \leq 2d\varepsilon.$$

Hence, as long as R is sufficiently large, we have $\sup_{n\geq 1} \mu_n\left(\overline{B(0, \sqrt{d}R}^c\right) \leq 2d\varepsilon$. Hence, the family $\{\mu_n \mid n \geq 1\}$ is tight, and the rest follows. ☐

Applications and Examples

1. CLT for i.i.d. sequences:

   Theorem 9. *Let $\{X_n\}$ be a sequence of iid random variables, $\mathbb{R}$-valued, with $\mathbb{E}[X_1] = m$ and $\mathrm{Var}[X_1] = \sigma^2$. If*

   $$\breve{S}_n = \frac{S_n - nm}{\sqrt{n}},$$

   then $\mathcal{L}_{\breve{S}_n} \Rightarrow \gamma_{0,\sigma^2}$ as $n \to \infty$.

   Proof. It suffices to show that if $\mu_n := \mathcal{L}_{\breve{S}_n}$, then $\lim_{n\to\infty} \hat{\mu}_n(\xi) = \exp\left(-\frac{\sigma^2\xi^2}{2}\right)$ for all $\xi \in \mathbb{R}$. W.L.O.G., assume that $m = 0$, otherwise re-centre the random variables. Then, for every $\xi \in \mathbb{R}$,

   $$\begin{aligned}\hat{\mu}_n(\xi) &= \mathbb{E}\left[e^{i\breve{S}_n\cdot\xi}\right] \\ &= \mathbb{E}\left[e^{iS_n\cdot\frac{\xi}{\sqrt{n}}}\right] \\ &= \left(\mathbb{E}\left[e^{iX_1\frac{\xi}{\sqrt{n}}}\right]\right)^n \\ &= \left(\rho\left(\frac{\xi}{\sqrt{n}}\right)\right)^n, \text{ where } \rho(\xi) := \mathbb{E}\left[e^{iX_1\xi}\right].\end{aligned}$$

   Since $\rho'(0) = \mathbb{E}[X_1] = 0$ and $-\rho''(0) = \mathbb{E}\left[X_1^2\right] = 0$, Taylor expand ρ and use this information:

   $$\begin{aligned}\rho\left(\frac{\xi}{\sqrt{n}}\right) &= 1 + \rho'(0)\frac{\xi}{\sqrt{n}} + \frac{\rho''(0)}{2}\frac{\xi^2}{n} + O\left(\frac{1}{n}\right) \\ &= 1 + 0 - \frac{\sigma^2}{2}\frac{\xi^2}{n} + O\left(\frac{1}{n}\right) \\ \Rightarrow \hat{\mu}_n(\xi) &= \left(1 - \frac{\sigma^2\xi^2}{2n} + O\left(\frac{1}{n}\right)\right)^n \to \exp\left(-\frac{\sigma^2\xi^2}{2}\right) = \hat{\gamma_{0,1}}(\xi).\end{aligned}$$

   ☐

2. Characterization of $\gamma_{0,1}$:

   Proposition 5. Let X and Y be i.i.d. random variables with $\mu = \mathcal{L}_X = \mathcal{L}_Y$ such that $\mathbb{E}[X] = 0$ and $\mathbb{E}\left[X^2\right] = 1$. If $X + Y$ and $X - Y$ are also independent, then $\mu = \gamma_{0,1}$.

Proof. Set $Z := X + Y$ and $W := X - Y$. Then, for all $\xi \in \mathbb{R}$,

$$\mathbb{E}\left[e^{iZ\xi}\right] = \mathbb{E}\left[e^{i(X+Y)\xi}\right] = (\hat{\mu}(\xi))^2 \text{ (where equality follows from independence).}$$
$$\mathbb{E}\left[e^{iW\xi}\right] = \hat{\mu}(\xi)\hat{\mu}(-\xi).$$

Note that according to the definition, $2X = Z + W$. Then,

$$\begin{aligned}\hat{\mu}(2\xi) &= \mathbb{E}\left[e^{i2X\xi}\right] \\ &= \mathbb{E}\left[e^{i(Z+W)\xi}\right] \\ &= \mathbb{E}\left[e^{iZ\xi}\right]\mathbb{E}\left[e^{iW\xi}\right] \text{ (by independence)} \\ &= (\hat{\mu}(\xi))^3\hat{\mu}(-\xi).\end{aligned}$$

For all $\xi \in \mathbb{R}$,

$$\hat{\mu}(2\xi) = (\hat{\mu}(\xi))^3\hat{\mu}(-\xi).$$

Similarly,

$$\hat{\mu}(-2\xi) = (\hat{\mu}(-\xi))^3\hat{\mu}(\xi).$$

Taking the ratio of the two expression yields,

$$\frac{\hat{\mu}(2\xi)}{\hat{\mu}(-2\xi)} = \left(\frac{\hat{\mu}(\xi)}{\hat{\mu}(-\xi)}\right)^2 \text{ for all } \xi \in \mathbb{R}.$$

If $g(\xi) := \frac{\hat{\mu}(\xi)}{\hat{\mu}(-\xi)}$, then it satisfies $\underline{g(2\xi) = (g(\xi))^2 \text{ for all } \xi \in \mathbb{R}}$. This relationship is the key. Iterating this identity gives

$$g(\xi) = \left(g\left(\frac{\xi}{2}\right)\right)^2 = \ldots = \left(g\left(\frac{\xi}{2^n}\right)\right)^{2^n} \text{ for all } \xi \in \mathbb{R}, n \geq 1.$$

Taylor expand g near zero:

$$g\left(\frac{\xi}{2^n}\right) = g(0) + g'(0)\frac{\xi}{2^n} + O\left(\frac{1}{2^n}\right).$$

Using the fact that $g(0) = 1$ and $g'(0) = 0$, we recover the special limit for e:

$$g(\xi) = \left(1 + 0 + O\left(\frac{1}{2^n}\right)\right)^{2^n} \to 1 \text{ as } n \to \infty.$$

Hence, $\hat{\mu}(\xi) = \hat{\mu}(-\xi)$ for all ξ which gives that $\hat{\mu}$ is even and hence $\hat{\mu}$ is symmetric. Hence,

$$\begin{aligned}&\Rightarrow \hat{\mu}(2\xi) = (\hat{\mu}(\xi))^4 \text{ for all } \xi \in \mathbb{R} \\ &\Rightarrow \hat{\mu}(\xi) = \left(\hat{\mu}\left(\frac{\xi}{2^n}\right)\right)^{2^{2n}} \text{ for all } \xi \in \mathbb{R}, \text{ for all } n \geq 1.\end{aligned}$$

Taylor expanding $\hat{\mu}$ near zero yields:

$$\hat{\mu}\left(\frac{\xi}{2^n}\right) = 1 + 0 - \frac{1}{2}\frac{\xi^2}{2^{2n}} + O\left(\frac{1}{2^{2n}}\right).$$

Hence,

$$\hat{\mu}(\xi) = \left(1 - \frac{\xi^2}{2}\frac{1}{2^{2n}} + O\left(\frac{1}{2^{2n}}\right)\right)^{2^{2n}} \to e^{-\frac{-\xi^2}{2}} \text{ as } n \to \infty.$$

□

The following theorem is another application of Levy's Continuity Theorem.

Theorem 10 (Levy's Equivalence Theorem)**.** *Let $\{X_n \mid n \geq 1\}$ be independent random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Set $S_n := \sum_{j=1}^{n} X_j$. Then, the following are equivalent:*

1. *S_n converges a.s. to some random variable S, i.e., $\sum_{j=1}^{n} X_j$ converges almost surely.*
2. *S_n converges in probability to some random variable S.*
3. *S_n converges in distribution to S.*

The idea of this theorem is that it's very hard for independent random variables to converge, so if they converge in one sense, they converge in all senses.

Proof. This time, we only need to show that $(iii) \Rightarrow (ii)$. Assume that $\mu_n := \mathcal{L}_{S_n}$ for every $n \in \mathbb{N}$ and $\mu_n \Rightarrow \mu$ for some probability measure μ. We will show that $\{S_n\}$ forms a Cauchy sequence in probability. Recall the definition of that:

$$\forall\, \varepsilon > 0,\ \exists\, N > 1 \ \text{ s.t. } \sup_{m \geq N} \mathbb{P}\left(|S_m - S_n| > \varepsilon\right) \leq \varepsilon.$$

For a contradiction, assume otherwise. Then, there exists an $\eta > 0$ and a subsequence $\{n_k\}$ such that along the subsequence, the Cauchy condition above is violated:

$$\mathbb{P}\left(|S_{n_{k+1}} - S_{n_k}| > \eta\right) \geq \eta. \tag{27}$$

For every k, set $v_k := \mathcal{L}_{S_{n_{k+1}} - S_{n_k}} \Rightarrow \mu_{n_{k+1}} = \mu_{n_k} * v_k$ (since this is a sum of independent random variables) $(*)$. Since $\mu_n \Rightarrow \mu$, the sequence $\{\mu_n \mid n \geq 1\}$ is tight $\Rightarrow$ for all $\varepsilon > 0$, there exists an $M > 0$ such that

$$\sup_n \mu_n(\overline{B(0,M)}^c) \leq \varepsilon.$$

Hence, for all $k \geq 1$,

$$\begin{aligned} v_k(\overline{B(0,2M)}^c) &= \mathbb{P}\left(|S_{n_{k-1}} - S_{n_k}| > 2M\right) \\ &\leq 2 \sup_n \mathbb{P}\left(|S_n| > M\right) \\ &= 2 \sup_n \mu_n(\overline{B(0,M)^c}) \\ &\leq 2\varepsilon. \end{aligned}$$

This shows that the sequence $\{v_k\}$ is tight. Hence, by the second part of the previous theorem, there exists a subsequence $\{k_l \mid l > 1\} \subseteq \mathbb{N}$ and a probability measure v such that $v_{k_l} \Rightarrow v$ as $l \to \infty$ $(**)$. Combining $(*)$ and $(**)$, we obtain that $\mu = \mu * v$.

Proposition / Remark: Let v be a probability measure on $\mathbb{R}^d$. If there exists a probability measure μ on $\mathbb{R}^d$ such that $\mu = \mu * v$, then $v = \delta_0$.

Proof: For every $\xi \in \mathbb{R}^d$,

$$\hat{\mu}(\xi) = \hat{\mu}(\xi) \cdot \hat{v}(\xi) \Rightarrow \hat{v}(\xi) = 1 \text{ if } \hat{\mu}(\xi) \neq 0.$$

Hence, there exists some positive number r such that $\hat{\mu}(\xi) \neq 0$ for $\xi \in B(0,r)$. Hence, for all $\xi \in B(0,r)$,

$$\hat{v}(\xi) = 1.$$

Hence, for all $\xi \in B(0,r)$,

$$\int_{\mathbb{R}^d} \cos(x,\xi) v(dx) = 1 \Rightarrow (x,\xi) = 0 \text{ mod } 2\pi \text{ for } v\text{- a.e } x.$$

Given any unit vector $e \in \mathbb{R}^d$, choose $\xi_1, \xi_2 \in B(0,r)$, ξ_1 and ξ_2 are both along the direction e and $\xi_2 = \rho\xi_1$ where $\rho \notin \mathbb{Q}$. Hence,

$$\begin{aligned}
&\Rightarrow (x, \xi_1) = 0 \text{ mod } 2\pi \text{ for } v \text{ almost every } x. \\
&\Rightarrow (x, \xi_2) = \rho(x, \xi_1) = 0 \text{ mod } 2\pi \text{ for } v \text{ almost every } x. \\
&\Rightarrow (x, \xi_1) = 0 \text{ for } v \text{ almost every } x. \\
&\Rightarrow (x, e) = 0 \ v \text{ almost everywhere for all unit vectors} \\
&\Rightarrow v = \delta_0.
\end{aligned}$$

Going back to the main proof, we have therefore proven that $v_{k_l} \Rightarrow \delta_0$, i.e., $S_{n_{k_l}+1} - S_{n_{k_l}} \to 0$ in distribution. Since this is a constant, we have therefore that

$$S_{n_{k_l+1}} - S_{n_{k_l}} \to 0 \text{ in probability.}$$

But, this contradicts (27). Hence, we have that $\{S_n\}$ forms a Cauchy sequence in probability $\Rightarrow S_n \to S$ in probability for some random variable S. $\square$

2 Infinitely Divisible Laws

Definition 7 (Infinitely Divisible)**.** Let μ be a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Then, we say that μ is **infinitely divisible** if for every $n \geq 1$, there exists a $v_{(n)}$ probability measure on $(R^d, \mathcal{B}(\mathbb{R}^d))$ such that

$$\underbrace{v_{(n)} * v_{(n)} * \dots * v_{(n)}}_{n \text{ times}} = \mu, \tag{28}$$

Equivalently, $(\hat{v}(\xi))^n = \hat{\mu}(\xi)$ for all $\xi \in \mathbb{R}^d$.

Notation-wise, we write $I(\mathbb{R}^d)$ as the collection of all the infinitely divisible laws on $\mathbb{R}^d$. If $\mu \in I(\mathbb{R}^d)$, then we will write $v_{(n)}$ in the definition as $\mu_{1/n}$, i.e., this means that

$$\underbrace{\mu_{1/n} * \mu_{1/n} * \dots * \mu_{1/n}}_{n \text{ copies}} = \mu.$$

We introduce a few remarks:

1. If $\mu \in I(\mathbb{R}^d)$, then for every $n \geq 1$,

   $$(\hat{\mu}_{1/n}(\xi))^n = \hat{\mu}(\xi) \ \forall \xi \in \mathbb{R}^d.$$

   Heuristically, we want to study the "nth" root of $\hat{\mu}$.
2. If $\mu, \nu \in I(\mathbb{R}^d)$, then for every $n \in \mathbb{N}$,

   $$\left(\mu_{1/n} * \hat{\nu}_{1/n}(\xi)\right)^n = \hat{\mu}(\xi) \cdot \hat{\nu}(\xi) = \mu \mathbin{\hat{*}} \nu(\xi).$$

   This implies that I is closed under convolution:

   $$\mu_{1/n} * \nu_{1/n} = (\mu * \nu)_{1/n} \Rightarrow \mu * \nu \in I(\mathbb{R}^d). \tag{29}$$
3. If $\{\mu_k\}$ is a sequence of infinitely divisible laws, $\mu_k \Rightarrow \mu$ as $k \to \infty$ and for all $n \geq 1$, $\mu_{k,1/n} \Rightarrow \nu_{(n)}$ as $k \to \infty$ for some probability measure $v_{(n)}$, then $\mu \in I(\mathbb{R}^d)$ and $\mu_{1/n} = \nu_{(n)}$.

2.1 Examples of Infinitely Divisible Laws

1. Trivial Examples: for $\vec{m} \in \mathbb{R}^d$, $\delta_m \in I(\mathbb{R}^d)$.

$$(\delta_m)_{1/n} = \delta_{m/n}.$$

2. Gaussian measures: for all $m \in \mathbb{R}^d$, for all $C = (C_{ij})_{d\times d} \geq 0$ (non-negative definite) such that

$$\hat{\gamma_{m,C}}(\xi) = e^{i(m,\xi)} e^{-1/2(\xi,C\xi)}$$

Hence, for all $n \in N$,

$$\left(\gamma_{m/n,C/n}\right)^n = \hat{\gamma}_{m,C}(\xi)$$

for all $\xi \in \mathbb{R}^d$. Hence, all Gaussian measures are infinitely divisible with $(\gamma_{m,C})_{1/n} = \gamma_{m/n,C/n}$.

3. Poisson Measures.
   (a) Standard Poisson distribution / measure is supported on $\{0, 1, 2, ...\}$. Given $\alpha > 0$, let π_α be the Poisson distribution with parameter α, i.e., for all $k \geq 0$,

$$\pi_\alpha(\{k\}) = e^{-\alpha} \frac{\alpha^k}{k!}$$

   Equivalently, write,

$$\pi_\alpha = \sum_{k=0}^{\infty} e^{-\alpha} \frac{\alpha^k}{k!} (\delta_k) = \sum_{k=0}^{\infty} e^{-\alpha} \frac{\alpha^k}{k!} (\delta_1)^{*k}$$

   (b) General Poisson Measure on $\mathbb{R}^d$:
      i. Given $\alpha > 0$ and a probability measure ν on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$,

$$\pi_{\alpha,\nu} = \sum_{k=0}^{\infty} e^{-\alpha} \frac{\alpha^k}{k!} \nu^{(*k)}$$

Let's try to understand this from the random variable point of view. Let $\{X_n\}$ be iid random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathcal{L}_{X_1} = \nu$. Let N be a random variable on the same probability space independent of $\{X_n\}$, $\mathcal{L}_N = \pi_\alpha$. Define $S = \sum_{i=1}^{N} X_j$. Hence, for all $\omega \in \Omega$, S is defined point-wise as:

$$S(\omega) = \sum_{j=1}^{N(\omega)} X_j(\omega). \tag{30}$$

Then, for all $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned}
\mathbb{P}\left(S \in B\right) &= \sum_{k=0}^{\infty} \mathbb{P}\left(S \in B, N = k\right) \\
&= \sum_{k=0}^{\infty} \mathbb{P}\left(\sum_{j=1}^{k} X_j \in B, N = k\right) \\
&= \sum_{k=0}^{\infty} \mathbb{P}\left(\sum_{j=1}^{k} X_j \in B\right) \mathbb{P}\left(N = k\right) \\
&= \sum_{k=0}^{\infty} \mathbb{P}\left(\sum_{j=1}^{k} X_j \in B\right) e^{-\alpha} \frac{\alpha^k}{k!} \\
&= \sum_{k=0}^{\infty} e^{-\alpha} \frac{\alpha^k}{k!} v^{*k}(B) \\
&= \pi_{\alpha,\nu}(B).
\end{aligned}$$

Note that we have a taylor expansion of the exponential, and so for every $\xi \in \mathbb{R}^d$,

$$\widehat{\pi_{\alpha,v}}(\xi) = \sum_{k=0}^{\infty} e^{-\alpha} \frac{\alpha^k}{k!} (\hat{v}(\xi))^k = e^{-\alpha} e^{\alpha \hat{v}(\xi)} = e^{\alpha(\hat{v}(\xi)-1)}.$$

This shows that $\pi_{\alpha,\nu} \in I(\mathbb{R}^d)$ and for all $n \geq 1$,

$$(\pi_{a,\nu})_{1/n} = e^{\alpha/n(\hat{v}(\xi)-1)} = \pi_{\alpha/n,v}.$$

Notation. Given $\alpha > 0$, probability measure v on $\mathbb{R}^d$,

$$\hat{\pi}_{\alpha,v}(\xi) = \exp(\alpha(\hat{v}(\xi) - 1)) = \exp\left(\alpha \int_{\mathbb{R}^d} (e^{i(x,\xi)} - 1)\nu(dx)\right)$$

Set $M := \alpha\nu$, $M(\mathbb{R}^d) = \alpha$. Set $\pi_M := \pi_{\alpha,\nu}$, i.e.,

$$\hat{\pi}_M(\xi) = \exp\left(\int_{\mathbb{R}^d} (e^{i(x,\xi)} - 1) M(dx)\right).$$

Further, $M(\{0\})$ does not affect π_M, so WLOG we assume that $M(\{0\}) = 0$. Set

$$\mathcal{M}_0(\mathbb{R}^d) := \{M \mid M \text{ is a finite Borel measure on } (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \text{ such that } M(\{0\}) = 0\}. \tag{31}$$

Remarks.

1. Given $\alpha > 0$, a probability measure ν, set $\tilde{\alpha} = \alpha(1 - \nu(\{0\}))$. Set $\tilde{\mu}$ to be a set function such that for all $B \in \mathcal{B}(\mathbb{R}^d)$:

$$\tilde{v}(B) := \frac{\nu(B \setminus \{0\})}{1 - \nu(\{0\})}$$

Then, $\tilde{v}$ is a probability measure and $\tilde{\mu}(\{0\}) = 0$. For all $\xi \in \mathbb{R}^d$,

$$\begin{aligned}
\hat{\pi}_{\alpha,\nu}(\xi) &= \exp\left(\alpha \int_{\mathbb{R}^d} \left(e^{i(x,\xi)} - 1\nu(dx)\right)\right) \\
&= \exp\left(\tilde{\alpha} \int_{\mathbb{R}^d} \left(e^{i(x,\xi)} - 1\right) \tilde{\nu}(dx)\right) \\
&= \hat{\pi}_{\tilde{\alpha},\tilde{\nu}}(\xi).
\end{aligned}$$

WLOG, we can assume, whenever necessary, that $\nu(\{0\}) = 0$. Set $M := \alpha\nu$ and assume that M is a finite measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $M(\{0\}) = 0$. Define the following two spaces:

$$\begin{aligned}\mathcal{M}_0(\mathbb{R}^d) &:= \{M \mid \text{ finite measure} M(\{0\}) = 0\}\\ \mathcal{P}(\mathbb{R}^d) &:= \{\pi_M \mid M \in \mathcal{M}_0(\mathbb{R}^d)\}\end{aligned}$$

2. Gaussian measures can be obtained through weak convergence from Poisson measures. We can see this in the following proposition.

Proposition 6. Given $\vec{m} \in \mathbb{R}^d$, $C = (C_{ij})_{d\times d} \geq 0$, there exists a sequence $\{\alpha_n \mid n \geq 1\} \subseteq]0, \infty[$ and probability measures $\{\nu_n \mid n \geq 1\}$ on $\mathbb{R}^d$ such that

$$\pi_{\alpha_n,\nu_n} \Rightarrow \gamma_{m,C}. \tag{32}$$

Proof. We will only prove it for $d = 1$. Given $m \in \mathbb{R}$, $\sigma^2 > 0$, set $\alpha_n = 2n$. Then, set the sequence of measures to be:

$$\nu_n = \frac{1}{2}\left[\delta_{\frac{m}{n}} + \frac{1}{2}\left(\delta_{\sigma/\sqrt{n}} + \delta_{\ -\sigma/\sqrt{n}}\right)\right] \tag{33}$$

for all $n \geq 1$. Now write down the characteristic function: for all $\xi \in \mathbb{R}$,

$$\begin{aligned}\hat{\pi}_{\alpha_n,\nu_n}(\xi) &= \exp(2n(\hat{\nu}_n - 1))\\ &= \exp\left(n\left(e^{i\frac{m}{n}\xi} + \cos\left(\frac{\xi}{\sqrt{n}}\sigma\right) - 2\right)\right)\\ &= \exp\left(n\left(e^{i\xi\frac{m}{n}} - 1\right)\right)\cdot\exp\left(n\left(\cos\left(\frac{\xi}{\sqrt{n}} - 1\right)\right)\right)\\ &\to \exp\left(i\xi m\right)\cdot\exp\left(-\frac{1}{2}\xi^2\right) = \hat{\gamma}_{m,\sigma^2}(\xi) \text{ when we send } n\to\infty.\end{aligned}$$

□

Exercise. Prove the statement in $\mathbb{R}^d$.

The next theorem tells us that out of Gaussian measures and exponential measures, the Poisson family is more fundamental.

Theorem 11. $I(\mathbb{R}^d) = \overline{\mathcal{P}(\mathbb{R}^d)}$.

First, we'll need some technical lemmas.

Lemma 5 (Facts about $\mathbb{C}$-valued functions.)**.** Given $R > 0$, let f be a complex-valued function that is continuous on $\overline{B(0,R)}$, i.e., $f \in C(\overline{B(0,R)}; \mathbb{C})$ and further assume that $f(0) = 1$. Further, assume that $f \neq 0$ on $\overline{B(0,R)}$, and $f_n \to f$ uniformly on $\overline{B(0,R)}$ for some function $f \in C(\overline{B(0,R)}; \mathbb{C})$ with $f(0) = 1$ and $f \neq 0$ on $\overline{B(0,R)}$. Then, there exists a unique function $\ell_f \in C(\overline{B(0,R)}; \mathbb{C})$ such that

$$e^{\ell_f} = f \text{ on } \overline{B(0,R)} \text{ and } \ell_f(0) = 0.$$

We refer to ℓ_f as the **principle log** of f. Moreover, if $\{f_n\} \subseteq C(\overline{B(0,R)}; \mathbb{C})$ such that $f_n(0) = 1$ and $f_n \neq 0$ on $\overline{B(0,R)}$. Then, if ℓ_{f_n} $n \geq 1$ and ℓ_f are the principal logs of f_n and f, respectively, then

$$\ell_{f_n} \to \ell_f \text{ uniformly on } \overline{B(0,R)}.$$

Lemma 6. Given $r, T \in]0, \infty[$ with $r < T$, there exists an $N = N_{r,T} \in \mathbb{N}$ such that if $\mu \in I(\mathbb{R}^d)$ such that

$$|1 - \hat{\mu}(\xi)| \leq \frac{1}{2} \ \forall \xi \in \overline{B(0,r)},$$

then $\inf_{\xi \in \overline{B(0,T)}} |\hat{\mu}(\xi)| > 2^{-N}$. Informally, this Lemma is telling us that we can use the value of $\hat{\mu}$ near zero to give a lower bound

Proof. Given $0 < r < T < \infty$, let N be sufficiently large. Assume that $\mu \in I(\mathbb{R}^d)$. We have:

$$\left(\hat{\mu}_{\frac{1}{n}}\right)^n = \hat{\mu}.$$

Further assume that $|1 - \hat{\mu}(\xi)| \leq \frac{1}{2}$ for all $\xi \in \overline{B(0,r)}$.

- Then, $\hat{\mu}(\xi) \neq 0$ and hence $\hat{\mu}_{\frac{1}{N}}(\xi) \neq 0$ for all $\xi \in \overline{B(0,r)}$.
- by Lemma 1, there exists a unique principal log $\ell_{\hat{\mu}}$ and a unique $\ell_{\hat{\mu}_{\frac{1}{N}}}$ on $\overline{B(0,r)}$ and $\ell_{\hat{\mu}} = N\ell_{\hat{\mu}_{\frac{1}{N}}}$.

Observe:

1. On one hand, $\xi \in \overline{B(0,r)}$ and so

$$|1 - \hat{\mu}(\xi)| \leq \frac{1}{2} \Rightarrow |\ell_{\hat{\mu}}(\xi)| \leq 2 \Rightarrow \left|\ell_{\hat{\mu}_{\frac{1}{N}}}(\xi)\right| \leq \frac{2}{N} \ \forall \xi \in \overline{B(0,r)}.$$

2. On the other hand, for all $\xi \in \mathbb{R}^d$, $\text{Re}(\ell_{\hat{\mu}_{\frac{1}{N}}}) = \frac{1}{N}\text{Re}(\ell_{\hat{\mu}}(\xi))$.

$$\Rightarrow \text{Re}(\ell_{\hat{\mu}_{\frac{1}{N}}}(\xi)) = \frac{1}{N} \ln |\hat{\mu}(\xi)| \leq \frac{1}{N} \ln(1) = 0$$
$$\Rightarrow |1 - \ell_{\hat{\mu}_{\frac{1}{N}}}(\xi)| \leq |\hat{\mu}_{\frac{1}{N}}(\xi)| \leq \frac{2}{N} \ \forall \xi \in \overline{B(0,r)}.$$

Where the first inequality in the line above follows from the fact that if $\text{Re}(z) \leq 0$, then $|1 - e^z| \leq |z|$.

Now we use the technical estimates from Lecture 5. For all unit vectors $\vec{e} \in \mathbb{R}^d$, for all $R > 0$,

$$\mu_{\frac{1}{N}}(\{x \in \mathbb{R}^d \mid |(x, \vec{e}) > R\}) \leq \frac{1}{m(rR)} \cdot \frac{1}{r} \int_0^r |1 - \hat{\mu}_{\frac{1}{N}}(s\vec{e})| ds$$
$$\leq \frac{2}{N} \frac{1}{m(rR)}.$$

Next, for all $0 < \tilde{r} < T$,

$$|1 - \hat{\mu}_{\frac{1}{N}}(\tilde{r}\vec{e})| \leq \tilde{r}R + 2\mu_{\frac{1}{N}}(\{x \in \mathbb{R}^d \mid |(x, \vec{e})| > R\})$$
$$\leq TR + \frac{4}{N}\frac{1}{m(rR)} \text{ (by the estimate above)}$$
$$= TR + \frac{4}{Nm(rR)}.$$

We first choose R such that $TR < \frac{1}{4}$, and then N sufficiently large such that $\frac{4}{N}\frac{1}{m(rR)} \leq \frac{1}{4}$. Therefore, we've proven that for every $\xi \in \overline{B(0,T)}$, $|1 - \hat{\mu}_{\frac{1}{N}}| \leq \frac{1}{2}$ whenever N is sufficiently large. Hence, $|\hat{\mu}_{\frac{1}{N}}| \geq \frac{1}{2}$, which yields the desired conclusion:

$$|\hat{\mu}(\xi)| = |\hat{\mu}_{\frac{1}{N}}(\xi)|^N \geq 2^{-N}.$$

□

Corrolary 3. If $\mu \in I(\mathbb{R}^d)$, then $\hat{\mu}(\xi) \neq 0$ for all $\xi \in \mathbb{R}^d$.

Proof. For such a μ, there exists an $r > 0$ such that $|1 - \hat{\mu}(\xi)| \leq \frac{1}{2}$ for all $\xi \in B(0, r)$, and for all $T > r > 0$, there exists an $N \in \mathbb{N}$ such that $|\hat{\mu}(\xi)| \geq 2^{-N}$ for all $\xi \in \overline{B(0,T)}$. Therefore, there exists a unique $\ell_{\hat{\mu}} \in C(\mathbb{R}^d; \mathbb{C})$ such that $\hat{\mu} = e^{\ell_{\hat{\mu}}}$. ☐

Furthermore, for every $n \geq 1$, $\hat{\mu}_{\frac{1}{N}}(\xi) \neq 0$ for all $\xi \in \mathbb{R}^d$. Hence, there exists one $\ell_{\hat{\mu}_{\frac{1}{N}}}$ which is uniquely the principle log of $\hat{\mu}_{\frac{1}{N}}$. This yields two important relations:

1. Due to the uniqueness of the principle log: $\hat{\mu}_{\frac{1}{N}} = \frac{1}{N}\ell_{\hat{\mu}}$.
2. (1) implies

$$\hat{\mu}_{\frac{1}{N}}(\xi) = \exp\left(\frac{1}{N}\ell_{\hat{\mu}}(\xi)\right) \ \forall \xi \in \mathbb{R}^d.$$

Hence,

1. the root $\mu_{\frac{1}{n}}$ is unique for all $n \geq 1$.
2. $\mu_{\frac{1}{n}} \in I(\mathbb{R}^d)$ for all $n \geq 1$. For every $m \geq 1$,

$$\left(\hat{\mu}_{\frac{1}{nm}}\right)^m = \exp\left(\frac{m}{nm}\ell_{\hat{\mu}}\right) = \hat{\mu}_{\frac{1}{n}}$$

Proposition 7. If $\{\mu_k \mid k \geq 1\} \subseteq I(\mathbb{R}^d)$ such that $\{\mu_k\} \Rightarrow \mu$ (as $k \to \infty$) for some probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, then $\mu \in I(\mathbb{R}^d)$ and hence $I(\mathbb{R}^d)$ is closed under taking weak convergence limits. Furthermore, for every $n \geq 1$, $(\mu_k)_{\frac{1}{n}} \Rightarrow \mu_{\frac{1}{n}}$ as $k \to \infty$.

Proof. Since $\mu_k \Rightarrow \mu$, we know that $\hat{\mu}_k \Rightarrow \hat{\mu}$ uniformly on compact sets. Hence, there exists an $r > 0$ positive such that $\forall\ \xi \in \overline{B(0,r)}$,

$$|1 - \hat{\mu}_k(\xi)| \leq \frac{1}{2} \ \forall\ k \geq 1$$
$$|1 - \hat{\mu}(\xi)| \leq \frac{1}{2}.$$

By the 2nd technical lemma, for all $T > r > 0$, there exists an $N = N_{T,r}$ such that for all $\xi \in \overline{B(0,T)}$ such that for all $k \geq 1$,

$$|\hat{\mu}_k(\xi)| > 2^{-N}. \tag{34}$$

Since $\hat{\mu}_k \Rightarrow \hat{\mu}$, we also have that $|\hat{\mu}(\xi)| \geq 2^{-N}$. Since T is chosen arbitrarily,

$$\hat{\mu}(\xi) \neq 0 \ \forall\ \xi \in \mathbb{R}^d.$$

Hence, there exists a unique principal log $\ell_{\hat{\mu}}$. So, for every $k \geq 1$, let $\ell_{\hat{\mu}_k}$ be the principal log of $\hat{\mu}_k$. Then, $\ell_{\hat{\mu}_k} \Rightarrow \ell_{\hat{\mu}}$ uniformly on compact sets (as $k \to \infty$). Therefore, for every $n \geq 1$,

$$\hat{(\mu_k)}_{\frac{1}{n}} = \exp\left(\frac{1}{n}\ell_{\hat{\mu}_k}\right) \to \exp\left(\frac{1}{n}\ell_{\hat{\mu}}\right) \text{ (uniformly on compact sets).}$$

By **(Levy's continuity theorem)** , there exists a probability measure $\nu_{(n)}$ such that $(\mu_k)_{\frac{1}{n}} \Rightarrow \nu_{(n)}$ as $k \to \infty$ and

$$\hat{\nu}_{(n)} = \exp\left(\frac{1}{n}\ell_{\mu}\right).$$

Obviously,

$$(\hat{\nu}_{(n)})^{(n)} = \hat{\mu} \Rightarrow \mu \in I(\mathbb{R}^d) \text{ and } \mu_{\frac{1}{n}} = \nu_{(n)},$$

i.e., $(\mu_k)_{\frac{1}{n}} \Rightarrow \mu_{\frac{1}{n}}$ as $k \to \infty$. ☐

Corrolary 4 (Closure under weak convergence)**.** $\overline{\mathcal{P}(\mathbb{R}^d)} \subseteq I(\mathbb{R}^d)$.

Proposition 8. $I(\mathbb{R}^d) \subseteq \overline{\mathcal{P}(\mathbb{R}^d)}$ i.e., for all $\mu \in I(\mathbb{R}^d)$, there exists an $\{M_n \mid n \geq 1\} \subseteq \mathcal{M}_0(\mathbb{R}^d)$ such that $\mu_{M_n} \Rightarrow \mu$ as $n \to \infty$.

Proof. For every $\mu \in I(\mathbb{R}^d)$, let $\mu_{\frac{1}{n}}$ be the "nth root" of μ for $n \geq 1$. Set M_n to be the finite measure on $\mathbb{R}^d$ such that for all $B \in \mathcal{B}(\mathbb{R}^d)$,

$$M_n(B) = n \cdot \mu_{\frac{1}{n}}(B \setminus \{0\}).$$

Clearly, $M_n \in \mathcal{M}_0(\mathbb{R}^d) \Rightarrow \pi_{M_n} \in \mathbb{R}^d$ for all $n \geq 1$. For all $\xi \in \mathbb{R}^d$,

$$\begin{aligned}
\hat{\pi}_{M_n}(\xi) &= \exp\left(\int_{\mathbb{R}^d} (e^{i(x,\xi)} - 1) M_n(dx)\right) \\
&= \exp\left(n \int_{\mathbb{R}^d} (e^{i(x,\xi)} - 1) \mu_{\frac{1}{n}}(dx)\right) \text{ (by definition of } M_n.) \\
&= \exp\left(n(\hat{\mu}_{\frac{1}{n}}(\xi) - 1)\right) \\
&= \exp\left(n(e^{\frac{1}{n}\ell_{\hat{\mu}}(\xi)} - 1)\right) \to e^{\ell_{\hat{\mu}}(\xi)} = \hat{\mu}(\xi) \text{ (as } n \to \infty).
\end{aligned}$$

Hence, $\pi_{M_n} \Rightarrow \mu$. □

So the next theorem just follows.

Theorem 12. $\overline{\mathcal{P}(\mathbb{R}^d)} = I(\mathbb{R}^d)$.

The next theorem gives us an explicit characterization of members in the family $I(\mathbb{R}^d)$. We will start with the two classical families: Gaussian and Poisson

Theorem 13 (Levy-Khintchine Formula)**.** *Given $\vec{m} \in \mathbb{R}^d$, $C = (C_{ij})_{d\times d} \geq 0$, $M \in \mathcal{M}_0(\mathbb{R}^d)$. If $\mu = \gamma_{m,C}$, $\nu = \pi_M$, then*

$$\mu * v \in I(\mathbb{R}^d) \text{ and } (\mu * \nu)_{\frac{1}{n}} = \mu_{\frac{1}{n}} * \nu_{\frac{1}{n}}.$$

We denote by $\pi^{(1)}_{m,C,M}$ the measure $\mu * \nu$, i.e, for all $\xi \in \mathbb{R}^d$, $\hat{\pi}^{(1)}_{m,C,M} = exp\left(\ell^{(1)}_{m,C,M}(\xi)\right)$, where

$$\ell^{(1)}_{m,C,M}(\xi) = i(m,\xi) - \frac{1}{2}(\xi, C\xi) + \int_{\mathbb{R}^d} (e^{i(x,\xi)} - 1) M(dx), \tag{35}$$

where the first two terms come from the Gaussian, and the final term comes from the Poisson.

Goal: expand the family of all such measures $\hat{\pi}^{(1)}_{m,C,M}$ by weak convergence. The only parameter we can only really expand is M. First, we will set the notation up.

Notation. We denote by $\mathcal{M}_\alpha(\mathbb{R}^d)$ the family of σ-finite measures M on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that:

$$M(\{0\}) = 0 \text{ and } \int_{\mathbb{R}^d} (|y|^\alpha \wedge 1) M(dy) < \infty.$$

This notation in the integrand compactly means

$$\int_{B(0,1)} |y|^\alpha M(dy) < \infty \text{ and } \int_{B(0,1)^c} M(dy) < \infty$$

e.g. if $M(dy) = \left(\frac{1}{|y|}\right)^{\rho} dy$ on $B(0,1) \setminus \{0\}$, then we require $\rho < \alpha + d$ to remain integrable. Also note that if $0 < \alpha_1 < \alpha_2$, then $\mathcal{M}_{\alpha_1}(\mathbb{R}^d) \subseteq \mathcal{M}_{\alpha_2}(\mathbb{R}^d)$. Now, let's expand the family of $\pi^{(1)}_{m,C,M}$:

Step 1: Set $\alpha = 1$. Take $M \in \mathcal{M}_1(\mathbb{R}^d)$. For every $r > 0$, set

$$M_r(dy) = \chi_{(\overline{B(0,r)}^c)}(y) M(dy). \tag{36}$$

We have that $M_r \in \mathcal{M}_0(\mathbb{R}^d)$. Then, given any $\vec{m} \in \mathbb{R}^d$, for all $C = (C_{ij})_{d\times d} \geq 0$, $\pi^{(1)}_{m,C,M_r} \in I(\mathbb{R}^d)$, with for all $\xi \in \mathbb{R}^d$,

$$\pi^{(1)}_{m,C,M_r}(\xi) = \exp\left[i(m,\xi) - \frac{1}{2}(\xi, X\xi) + \int_{\mathbb{R}^d \setminus \overline{B(0,r)}} (e^{i(y,\xi)} - 1) M(dy)\right]. \tag{37}$$

We want to take $r \to 0$; to do so, we want to use **(DOM)** . TO justify this,

$$\left| e^{i(\xi,y)} - 1 \right| \chi_{\overline{B(0,r)}^c} \leq 2\chi_{\overline{B(0,1)}^c} + \chi_{\overline{B(0,1}} |\xi||y| \in L^1(M(dy)),$$

because $M \in \mathcal{M}_1(\mathbb{R}^d)$. By **(DOM)** , as $r \downarrow 0$,

$$\to \exp\left[i(m,\xi) - \frac{1}{2}(\xi, C\xi) + \int_{\mathbb{R}^d} (e^{i(y,\xi)} - 1) M(dy)\right] \tag{38}$$

So set

$$\textbf{(M1)} := \exp\left[i(m,\xi) - \frac{1}{2}(\xi, C\xi) + \int_{\mathbb{R}^d} (e^{i(y,\xi)} - 1) M(dy)\right] \tag{39}$$

By **(Levy's Continuity Theorem)** , $\pi^{(1)}_{m,C,M_r}$ weakly converges as $\to 0$ to a limit measure, which we will denote by $\pi^{(1)}_{m,C,M}$. Hence, $\pi^{(1)}_{m,C,M_r} \in I(\mathbb{R}^d)$ for every $\vec{m} \in \mathbb{R}^d$, $C = (C_{ij}) \geq 0$, $M \in \mathcal{M}_1(\mathbb{R}^d)$ where $\pi^{\hat{(1)}}_{m,C,M_r}$ is given by **(M1)**. Hence, we have expanded the family of $\pi^{(1)}_{m,C,M_r}$ from $M \in \mathcal{M}_0(\mathbb{R}^d)$ to $\mathcal{M}_1(\mathbb{R}^d)$.

Step 2: Set $\alpha = 2$. Take $M \in \mathcal{M}_2(\mathbb{R}^d)$. For $r > 0$, consider $M_r(dy)$ the same as above. For all $\xi \in \mathbb{R}^d$, define:

$$\begin{aligned}
\ell_{m,C,M_r}(\xi) &:= i(m,\xi) - \frac{1}{2}(\xi, C\xi) + \int_{\mathbb{R}^d} \left[e^{i(y,\xi)} - 1 - i(\xi,y)\chi_{B(0,1)}(y)\right] M_r(dy) \\
&= i(m,\xi) - \frac{1}{2}(\xi, C\xi) + \int_{\mathbb{R}^d \setminus B(0,r)} \left[e^{i(y,\xi)} - 1 - i(\xi,y)\chi_{B(0,1)}(y)\right] M(dy) \\
&= i\left(\underbrace{m - \int_{\mathbb{R}^d \setminus B(0,r)} yM(dy)}_{:=m_r}, \xi\right) - \frac{1}{2}(\xi, C\xi) + \int_{\mathbb{R}^d} [e^{i(y,\xi)} - 1] M_r(dy) \\
&= \ell^{(1)}_{m_r,C,M_r}(\xi).
\end{aligned}$$

This implies that $e^{\ell_{m,C,M_r}} = e^{\ell_{r,C,M_r}(\xi)} = \hat{\pi}^{(1)}_{m_r,C,M_r}(\xi)$. Now send $r \to 0$ in the original expression. Then, by **(DOM)** ,

$$\lim_{r\downarrow 0} \int_{\mathbb{R}^d \setminus B(0,r)} [e^{i(\xi,y)} - 1 - i(\xi,y)\chi_{B(0,1)}(y)] M(dy) = \int_{R^d} [e^{i(\xi,y)} - 1 - i(\xi,y)\chi_{B(0,1)}(y)] M(dy). \tag{40}$$

Therefore,

$$\begin{aligned}\lim_{r\to 0}\hat{\pi}^{(1)}_{m,C,M_r}(\xi) &= \lim_{r\to 0}\exp(\ell^{(1)}_{m_r,C,M_r}(\xi))\\ &= \lim_{r\downarrow 0}\exp(\ell^{(1)}_{m,C,M_r}(\xi))\\ &= \exp\left(\ell_{m,C,M}(\xi)\right),\end{aligned}$$

where

$$\ell_{m,C,M}(\xi) = i(m,\xi) - \frac{1}{2}(\xi, C\xi) - \int_{\mathbb{R}^d}[e^{i(x,\xi)} - 1 - i(\xi,y)\chi_{B(0,1)}(y)]M(dy) := \textbf{(M2)}$$

(Levy's Continuity Theorem) implies that as $r\downarrow 0$, $\pi^{(1)}_{m,C,M_r}$ converges weakly to a limiting measure, denoted by $\pi_{m,C,M}$. Hence, $\pi_{m,C,M}\in I(\mathbb{R}^d)$ and $\hat{\pi}_{m,C,M} = \exp(\ell_{m,C,M}(\xi))$ for all $\xi\in\mathbb{R}^d$.

Definition 8 (Levy System / Canonical Representation)**.** Given $\vec{m}\in\mathbb{R}^d$, $C = (C_{ij})_{d\times d}\geq 0$, $M\in\mathcal{M}_2(\mathbb{R}^d)$, the triple (m,C,M) is called a **Levy System**. $e^{\ell_{m,C,M}}$ where $\ell_{m,C,M}$ is as in **(M2)** is called the **Canonical Representation** of $\hat{\pi}_{m,C,M}$.

Theorem 14 (Levy Khinchine Formula)**.** *Given $\mu\in I(\mathbb{R}^d)$, there exists a Levy System (m_μ, C_μ, M_μ) such that $\mu = \pi_{m_\mu,C_\mu,M_\mu}$.*

Proof. Task. ☐

Example 2 (Cauchy Distribution on $\mathbb{R}$)**.** For the Cauchy Distribution, μ is the probability measure on $\mathbb{R}$ with density $\frac{1}{\pi(1+x^2)}$. One ca check that for every $\xi\in\mathbb{R}$,

$$\hat{\mu}(\xi) = \int_{\mathbb{R}}\frac{e^{ix\xi}}{\pi(1+x^2)}dx = e^{-|\xi|}. \tag{41}$$

For all ≥ 1,

$$\hat{\mu}(\xi) = e^{-|\xi|} = \left(e^{-\left|\frac{\xi}{n}\right|}\right)^n.$$

If $\nu_{(n)}$ is the distribution of $x\in\mathbb{R}\mapsto\frac{1}{n}x$ under distribution μ. In terms of random variables, if X is a Cauchy random variable, i.e., $\mathcal{L}_X = \mu$ and $Y_n := \frac{1}{n}X$, then $\nu_{(n)} = \mathcal{L}_{Y_n}$. For all $\xi\in\mathbb{R}$,

$$\begin{aligned}\nu_{(n)}(\xi) &= \int_{\mathbb{R}} e^{ix\xi}\nu_{(n)}(dx)\\ &= \int_{\mathbb{R}} e^{i\frac{1}{n}x\xi}\mu(dx)\\ &= \hat{\mu}\left(\frac{\xi}{n}\right)\\ &= e^{-\left|\frac{\xi}{n}\right|}\end{aligned}$$

And hence, $(\hat{\nu_{(n)}})^n = \hat{\mu}$ and $\mu_{\frac{1}{n}} = \nu_{(n)}$. So, by **(LK-Formula)** , should be able to find a Levy System. However, from here it's totally not obvious – we will need to fit it into a canonical representation.

Definition 9. Given $\alpha > 0$, a probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is called an **α-stable law** if there exists a unique principle log $\ell_{\hat{\mu}}$ of $\hat{\mu}$ such that for all $\xi\in\mathbb{R}^d$, for all $t>0$,

$$\ell_{\hat{\mu}}(t\xi) = t^\alpha\ell_{\hat{\mu}}(\xi). \tag{42}$$

Remark.

1. Cauchy distribution one-dimensional is a 1-stable law.
   (a) Centred Gaussian measure are 2-stable Laws.
2. If μ is an α-stable law, then $\mu \in I(\mathbb{R}^d)$ and $\mu_{\frac{1}{n}}$ is the distribution of $x \mapsto \left(\frac{1}{n}\right)^{1/\alpha} x$ under μ.
3. Given $\mu \in I(\mathbb{R}^d)$ with Levy System (m_μ, C_μ, M_μ), then for every $t > 0$,

$$t\ell_{\hat{\mu}} = t\ell_{m_\mu, C_\mu, M_\mu} = \ell_{tm_\mu, tC_\mu, tM_\mu}. \tag{43}$$

Hence, there exists a $\mu_t \in I(\mathbb{R}^d)$ such that $\hat{\mu}_t = \exp\left(\ell_{tm_\mu, tC_\mu, tM_\mu}\right)$. Therefore, μ is an α-stable law $\iff$ $\mu \in I(\mathbb{R}^d) \setminus \{\delta_0\}$ such that for all $t > 0$, for all $\varphi \in C_b(\mathbb{R}^d)$,

$$\mu_t(\varphi) = \mu(\varphi_t) \text{ where } \varphi_t(x) = \varphi(t^{\frac{1}{\alpha}} x) \ \forall x \in \mathbb{R}^d,$$

i.e., μ_t is the distribution of $x \mapsto t^{\frac{1}{\alpha}} x$ under μ.

Lemma 7. If $m \in \mathcal{M}_2(\mathbb{R}^d)$, then,

$$\lim_{|\xi| \to 0} \frac{1}{|\xi|^2} \int_{\mathbb{R}^d} [e^{i(y,\xi)} - 1 - i(\xi, y)\chi_{B(0,1)}(y)] M(dy) = 0. \tag{44}$$

In particular, if $\mu \in I(\mathbb{R}^d)$ with Levy System (m_μ, C_μ, M_μ), then for all $\xi \in \mathbb{R}^d$,

$$(\xi, C\xi) = -2 \lim_{t \to \infty} \frac{\ell_{\hat{\mu}}(t\xi)}{t^2}.$$

Proof. We re-write the first integral in the first statement as:

$$\underbrace{\frac{1}{|\xi|^2} \int_{B(0,r)} [...] M(dy)}_{(L)} + \underbrace{\frac{1}{|\xi|^2} \int_{\mathbb{R} \setminus B(0,r)} [...] M(dy)}_{(R)}.$$

We can bound:

$$|(L)| \leq \frac{1}{|\xi|^2} \int_{B(0,r)} |y|^2 |\xi|^2 M(dy) = \int_{B(0,r)} |y|^2 M(dy),$$

by choosing r sufficiently small, which we can do since $M \in \mathcal{M}_2(\mathbb{R}^d)$, we can make **(L)** arbitrarily small.

$$|(R)| \leq \frac{2 + |\xi|}{|\xi|^2} M(\mathbb{R}^d \setminus B(0,r)).$$

For the given r as above, we can choose $|\xi|$ sufficiently large such that **(R)** is arbitrarily small. Finally,

$$\ell_{\frac{\hat{\mu}}{t^2}}(t\xi) = \frac{i(m, t\xi)}{t^2} - \frac{1}{2}(\xi, C\xi) + \frac{|\xi|^2 \int_{\mathbb{R}^d} [e^{i(y,t\xi)} - 1 - i(y, t\xi)\chi_{B(0,1)}(y)] M(dy)}{t^2 |\xi|^2},$$

and so as $t \to \infty$, this term tends to $0 - \frac{1}{2}(\xi, C\xi) + 0$, as desired. $\square$

Proposition 9. (α-stable laws)

1. There exists *no* (non-trivial, i.e., δ_0) α-stable law for $\alpha > 2$.
2. The only (non-trivial) 2-stable law on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is a centre Gaussian $\gamma_{0,C}$ for $C = (C_{ij}) \geq 0$.
3. If μ is an α-stable law with $\alpha \in]0, 2[$, then $C_\mu = 0$.

Proof. For all $\xi \in \mathbb{R}^d$, for all $t \geq 0$,

$$\begin{aligned}\ell_{\hat{\mu}}(\xi) &= t^{-\alpha}\ell_{\hat{\mu}}(t\xi)\\ &= t^{-\alpha}\underbrace{\left[i(m_\mu, t\xi) - \frac{1}{2}(t\xi, C_\mu t\xi) + \int_{\mathbb{R}^d}\left(e^{i(y,t\xi)} - 1 - \chi_{B(0,1)}(y)i(y,\xi)\right) + M_\mu(dy)\right]}_{:=(\triangle)}\end{aligned}$$

When $\alpha > 2$, as $t \to \infty$, $\lim_{t\to\infty} t^{-\alpha}(\triangle) = 0$ for every ξ.

$$\begin{aligned}&\Rightarrow \ell_{\hat{\mu}}(\xi) = 0 \ \forall\ \xi \in \mathbb{R}^d\\ &\Rightarrow \mu = \delta_0 \text{ (trivial case).}\end{aligned}$$

When $\alpha = 2$, then

$$\begin{aligned}\ell_{\hat{\mu}}(\xi) &= \lim_{t\to\infty}\frac{(\triangle)}{t^\alpha} = \frac{1}{2}(\xi, C_\mu\xi)\\ &\Rightarrow \mu = \gamma_{0,C}\end{aligned}$$

When $\alpha \in]0,2[$, for all $t > 0$,

$$\begin{aligned}\lim_{|\xi|\to\infty}\frac{\ell_{\hat{\mu}}(\xi)}{|\xi|^2} &= \lim_{|\xi|\to\infty}\frac{t^{-\alpha}(\triangle)t^2}{t^2|\xi|^2}\\ &= \lim_{|\xi|\to\infty}\frac{t^{2-\alpha}(\triangle)}{t^2|\xi|^2}\\ &= \lim_{|\xi|\to\infty}\frac{t^{2-\alpha}\ell_{\hat{\mu}}(t\xi)}{t^2|\xi|^2}\\ &= \lim_{|\eta|\to\infty} t^{2-\alpha}\frac{\ell_{\hat{\mu}}(\eta)}{|\eta|^2}.\end{aligned}$$

This is only possible if $\lim_{|\xi|\to\infty}\frac{\ell_{\hat{\mu}}(\xi)}{|\xi|^2} = 0$ i.e., $C_\mu = 0$. □

Homework Problem. prove that for every $\alpha \in]0,2[$, $f(\xi) = e^{-|\xi|^\alpha}$ is the characteristic function of an α-stable law, i.e., there exists a μ α-stable law such that $\hat{\mu} = f$. Find out the Levy System (m_μ, C_μ, M_μ) for this $\mu \in I(\mathbb{R}^d)$.

2.2 Levy Processes

We will start discussing continuous-time stochastic processes. We will first set up the notation that we need.

$$\begin{aligned}(\mathbb{R}^d)^{[0,\infty[} &:= \{f : [0,\infty[\to \mathbb{R}^d\}\\ &= \text{ collection of all } \mathbb{R}^d\text{-valued functions defined on } [0,\infty[\end{aligned}$$

For all $t \geq 0$, we define the **coordinate map** or the **projection map**:

$$\begin{aligned}\mathcal{T}_t &: (\mathbb{R}^d)^{[0,\infty[} \to \mathbb{R}^d\\ \mathcal{T}_t(f) &:= f(t).\end{aligned}$$

Then, we can define the **σ-algebra generated by all the projection maps**:

$$\Sigma_{\mathbb{R}^d}^{[0,\infty[} := \sigma(\{\mathcal{T}_t \mid t \geq 0\}).$$

So, $\Sigma_{\mathbb{R}^d}^{[0,\infty[}$ is the smallest σ-algebra such that $\mathcal{T}_t$ is measurable for all $t \geq 0$.

Definition 10 (Stochastic Process)**.** Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a family of $\mathbb{R}^d$-valued functions $\{X_t \mid g \geq 0\}$ is called a **stochastic process** if the mapping

$$X : \omega \in \Omega \mapsto X_\cdot(\omega) \in (\mathbb{R}^d)^{[0,\infty[}, \tag{45}$$

is measurable with respect to $\Sigma_{\mathbb{R}^d}^{[0,\infty[}$. Here, $X_\cdot(\omega)$ denotes a *path*:

$$X_\cdot(\omega) = \{X_t \mid t \geq 0\}. \tag{46}$$

We often say that "X is a stochastic process" or "X is an $(\mathbb{R}^d)^{[0,\infty[}$-valued random variable" or equivalently, for all $t \geq 0$, $\omega \in \Omega \mapsto \mathcal{T}_t \cdot X_\cdot(\omega) = X_t(\omega) \in \mathbb{R}^d$ is an $\mathbb{R}^d$-valued random variable.

Definition 11. Let X be a stochastic processes on $(\Omega, \mathcal{F}, \mathbb{P})$. The distribution of X is a probability measure $\mathcal{L}_X$ on $((\mathbb{R}^d)^{[0,\infty[}, \Sigma_{\mathbb{R}^d}^{[0,\infty[})$, given by: for all $B \in \Sigma_{\mathbb{R}^d}^{[0,\infty[}$,

$$\begin{aligned}\mathcal{L}_X(B) &:= \mathbb{P}\left(X \in B\right) \\ &= \mathbb{P}\left(\{\omega \in \Omega \mid X_\cdot(\omega) \in B\}\right).\end{aligned}$$

For example, one could define:

$$\begin{aligned}B &:= \{f : [0, +\infty] \to \mathbb{R}^d \mid f(1) \in B(0,1)\} \\ B &:= \{f : [0, +\infty] \to \mathbb{R}^d \mid |f(1/2)| > |f(1)|\}.\end{aligned}$$

Proposition 10. Let X and Y be two $(\mathbb{R}^d)^{[0,\infty[}$-valued stochastic processes on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, X and Y have the same distribution if for all $p \in \mathbb{N}$, for all $0 \leq t_1 < t_2 < ... < t_p < \infty$, for all $B_j \in \mathcal{B}(\mathbb{R}^d)$, $j = 1, ..., p$:

$$\mathbb{P}\left(\bigcap_{j=1}^{p} \{X_{t_j} \in B_j\}\right) = \mathbb{P}\left(\bigcap_{j=1}^{p} \{Y_{t_j} \in B_j\}\right). \tag{47}$$

This is true because $\{\bigcap_{j=1}^{p} \mathcal{T}_{t_j}^{-1}(B_j) \mid p \in \mathbb{N},\ 0 \leq t_1 < ... < t_p < \infty, B_1, ..., B_p \in \mathcal{B}(\mathbb{R}^d)\}$ is a generating π-system of $\Sigma_{\mathbb{R}^d}^{[0,\infty[}$.

Definition 12 (Indistinguishible / Modification)**.** Let X and Y be two stochastic processes on $(\Omega, \mathcal{F}, \mathbb{P})$.

- We say that X and Y are **indistinguishible** if

$$\mathbb{P}\left(X = Y\right) = 1. \tag{48}$$

- We say that X is a **modification** of Y if for all $t \geq 0$,

$$\mathbb{P}\left(X_t = Y_t\right) = 1. \tag{49}$$

Note that indistingushible ⇒ modification ⇒ having the same distribution, but the converse doesn't follow in general.

Definition 13 (RCLL Function)**.** We denote by $D([0,\infty[; \mathbb{R}^d) =: D([0,\infty[)$ the collection of all the $\mathbb{R}^d$-valued functions f on $[0,\infty[$ such that:

1. for all $t \geq 0$, f is continuous from the right:

$$f(t^+) = \lim_{s \to t} f(s) = f(t). \tag{50}$$

2. the left-handed limit exists: for all $t \geq 0$,

$$f(t^-) = \lim_{s \to t^-} f(s) \text{ exists in } \mathbb{R}^d.$$

   i.e., f is a **RCLL Function**.

We have a few facts about RCLL Functions.

1. **(Boundedness)**. For all $t \geq 0$,

$$\begin{aligned} ||f||_{U,[0,t]} &:= \sup_{s \in [0,t]} |f(s)| < \infty \\ ||f||_{n,[0,t]} &:= \lim_{n \to \infty} \max_{m \in \{0,1,\ldots,2^n\}} |f(m2^{-m}t)| \end{aligned}$$

2. **(Total Variation).** it's bounded:

$$\begin{aligned} ||f||_{\text{var},[0,t]} &:= \sup \left\{ \sum_{j=1}^{p} |f(t_j) - f(t_{j-1})| \mid p \in \mathbb{N}, 0 = t_0 < t_1 < \ldots < t_p = t \right\} < \infty \\ ||f||_{\text{var},[0,t]} &:= \lim_{n \to \infty} \sum_{m=1}^{2^n} |f(m2^{-n}t) - f((m-1)2^{-n}t)|. \end{aligned}$$

3. **(Discontinuity Points)**. For all $t \geq 0$, for all $r \geq 0$, set:

$$J(t,r,f) := \{s \in]0,t] \mid |f(s) - f(s^-)| \geq r\}. \tag{51}$$

   (This set is collecting all the discontinuity points where the jump size is at least r). Then, $J(t,r,f)$ is a finite set for all $r > 0$, and hence

$$J(t,f) := \{s \in]0,t] \mid f(s) \neq f(s^-)\}, \tag{52}$$

   i.e., the set of all discontinuity points, is at most countable.

Hence, for all $t \geq 0$, we can define a set function on $\mathcal{B}(\mathbb{R}^d)$ as follows: for all Borel sets $B \in \mathcal{B}(\mathbb{R}^d)$,

$$j(t,B,f) := \sum_{s \in J(t,f)} \chi_B(f(s) - f(s^-)). \tag{53}$$

In words, $j(t,B,f)$ is the number of jumps such that the amount of jump is in B. Equivalently,

$$j(t,\cdot,f) := \sum_{s \in J(t,f)} \delta_{f(s)-f(s^-)}.$$

It's easy to check that $j(t,\cdot,f)$ is a measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ (check that it satisfies countable additivity). And, $j(t,\{0\},f) = 0$. Further, for every $r > 0$, $j(t, \mathbb{R}^d \setminus B(0,r), f) = J(t,r,f) < \infty$. Hence, this implies that

$$\begin{aligned} j(t,\cdot,f) \in \mathcal{M}_\infty(\mathbb{R}^d) :=& \{M \mid \sigma - \text{finite measure on } (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \\ & \text{such that } M(\{0\}) = 0 \text{ and } M(\mathbb{R}^d \setminus B(0,r)) < \infty \ \forall r > 0\}. \end{aligned}$$

(Recall the ordering we have on these families: $\mathcal{M}_0(\mathbb{R}^d) \subseteq \mathcal{M}_1(\mathbb{R}^d) \subseteq \mathcal{M}_2(\mathbb{R}^d) \subseteq \mathcal{M}_\infty(\mathbb{R}^d)$, where $\mathcal{M}_i(\mathbb{R}^d)$ allows for ith order singularities). We call $j(t,\cdot,f)$ the **jump measure corresponding to f by time t**.

Given a $\varphi : \mathbb{R}^d \to \mathbb{C}$, φ a Borel function that vanishes near zero $\varphi|_{B(0,r)} \equiv 0$ for some $r > 0$. Then, $\varphi \in L^1(j(t, \cdot, f))$

$$\int_{\mathbb{R}^d} \varphi(y) j(t, dy, f) = \sum_{s \in J(t,r,f)} \varphi(f(s) - f(s^-)).$$

If in addition we assume that $\varphi \in C_b(\mathbb{R}^d; \mathbb{C})$, then we can re-write the above as:

$$\int_{\mathbb{R}^d} \varphi(y) j(t, dy, f) = \lim_{n\to\infty} \sum_{m=1}^{2^n} \varphi(f(m2^{-n}f) - f((m-1)2^{-n}t)).$$

Furthermore, $t \in]0, \infty[\mapsto \int_{\mathbb{R}^d} \varphi(y) j(t, dy, f)$ is right continuous and piece-wise constant. To see this, we notice that for every $t > 0$, since f is right-continuous, there exists a $\varepsilon > 0$ such that for all $s, s' \in [t, t+\varepsilon]$, by the triangle inequality:

$$|f(s) - f(s')| \leq |f(s) - f(t)| + |f(s') - f(t)| < r/2.$$

This implies that there is no discontinuity in $[t, t+\varepsilon]$ that has jump size at least r. Hence, $j(t, r, f) = j(t + \varepsilon, r, f)$ and

$$\int_{\mathbb{R}^d} \varphi(y) j(t, dy, f) = \int_{\mathbb{R}^d} \varphi(y) j(t + \varepsilon, dy, f).$$

Definition 14. Also a Proposition. Set Σ_d to be the σ-algebra of subsets of $D([0, \infty[)$ generated by all the projection maps $\{\mathcal{T}_t \mid g \geq 0\}$. Then, for all $t \geq 0$,

$$||\cdot||_{U,[0,t]} : f \in D([0, \infty[) \mapsto ||f||_{U,[0,t]} \text{ is } \Sigma_D\text{-measurable.}$$
$$||\cdot||_{\text{var},[0,t]} : f \in D([0, \infty[) \mapsto ||f||_{\text{var},[0,t]} \text{ is } \Sigma_D\text{-measurable.}$$

For all $\varphi \in C_0(\mathbb{R}^d)$ such that φ vanishes near zero,

$$f \in D([0, \infty[) \mapsto \int_{\mathbb{R}^d} \varphi(y) j(t, dy, f) \text{ is } \Sigma_D\text{-measurable.}$$

In particular,

$$\{f \in D([0, \infty[\mid f \text{ is continuous on } [0, \infty[\} \in \Sigma_D.$$

Definition 15 (Levy Process)**.** Let $X = \{X_t \mid t \geq 0\}$ be a Stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ such that for all $\omega \in \Omega$, $X_\cdot(\omega) \in D([0, \infty])$ (the sample path of $X_\cdot$ is RCLL). X is called a **Levy Process** if X has independent and homogeneous increments, i.e., for all $0 \leq s < t$, the increment $X_t - X_s$ is *independent* of $\{X_r \mid 0 \leq r \leq s\}$ and $\mathcal{L}_{X_t - X_s}$ only depends on $t - s$ (*homogeneous*).

Remark. Given a Levy process $X = \{X_t \mid t \geq 0\}$ set $\mu = \mathcal{L}_{X_1}$ (by convention, set $X_0 \equiv 0$). *Claim*: $\mu \in I(\mathbb{R}^d)$. To see this, for every $n \geq 1$, we may write:

$$X_1 = \sum_{j=1}^{n} \left(X_{\frac{j}{n}} - X_{\frac{j-1}{n}} \right). \tag{54}$$

We have that the set

$$\left\{ X_{\frac{j}{n}} - X_{\frac{j-1}{n}} \mid j = 1, ..., n \right\} \tag{55}$$

is iid and have distribution $\mathcal{L}_{X_{\frac{1}{n}}}$. Hence, this implies that

$$\left(\hat{\mathcal{L}_{X_{\frac{1}{n}}}}\right)^n = \hat{\mu} \Rightarrow \mu \in I(\mathbb{R}^d) \text{ with } \mu_{\frac{1}{n}} = \mathcal{L}_{X_{\frac{1}{n}}},$$

(by being homogeneous). For every $q \in \mathbb{Q}$, write $q = \frac{m}{n}$ for $m, n \in \mathbb{N}$. Then,

$$\hat{\mathcal{L}_{X_q}} = \left(\hat{\mathcal{L}_{X_{\frac{1}{n}}}}\right)^m = e^{\frac{m}{n}\ell_{\hat{\mu}}} \Rightarrow \mathcal{L}_{X_q} = \mu_q.$$

For $t > 0$, choose $\{q_k \mid k \geq 1\} \subseteq \mathbb{Q}^+$ such that $q_k \to t^+$ as $k \to \infty$. Then, for all $\xi \in \mathbb{R}^d$,

$$\begin{aligned}
\mathbb{E}\left[e^{i(\xi, X_t)}\right] &= \lim_{k\to\infty} \mathbb{E}\left[e^{i(\xi, X_{q_k}}\right] \text{ (by } \textbf{(DOM)}, \textbf{(RCLL)}\text{)}. \\
&= \lim_{k\to\infty} e^{q_k \ell_{\hat{\mu}}(\xi)} \\
&= e^{t\ell_{\hat{\mu}}(\xi)} \\
&= \hat{\mu}_t(\xi).
\end{aligned}$$

Hence, for all $t \geq 0$, $\mathcal{L}_{X_t} = \mu_t$, where $\mu = \mathcal{L}_{X_1}$. And further, for all $0 < s < t$, $\mathcal{L}_{X_t - X_s} = \mu_{t-s}$.

Definition 16 (Levy Process Associated with a Measure)**.** We say that X is a **Levy Process Associated with** $\mu \in I(\mathbb{R}^d)$ if

1. X is a Levy process.
2. for all $0 \leq s < t$,

$$\mathcal{L}_{X_t - X_s} = \mu_{t-s}. \tag{56}$$

The next goal is: we want to, given a μ, construct a Levy Process X associated with μ.

Step # 0: Given $\alpha > 0$, let π_α be the Poisson distribution with parameter $\alpha > 0$. We will construct X for this distribution. Let $\{\theta_n\}$ be iid exponential random variables with parameter $\alpha > 0$, i.e., the density function is $\chi_{]0,\infty[}(x)\alpha e^{-\alpha x}$.WLOG, we assume that $T_0 \equiv 0$ for all $n \geq 1$, and set

$$T_n := \sum_{m=1}^{n} \theta_m. \tag{57}$$

Here, T_n represents the total waiting time by the nth occurence, and θ_m is the waiting time. T_n has a gamma distribution: it's density is

$$\chi_{]0,\infty[}(x)\frac{\alpha^n}{(n-1)!}x^{n-1}e^{-\alpha x}.$$

Set $N_0 \equiv 0$ and for all $t \geq 0$, for all $\omega \in \Omega$, define

$$N_t(\omega) := \max\{n \geq 1 \mid T_n(\omega) \leq t\}. \tag{58}$$

This $N_t(\omega)$ counts how many events occurred by time t. Equivalently,

$$\{N_t = 0\} = \{T_1 > t\} = \{\theta_1 > t\}.$$

For all $n \geq 1$,

$$\{N_t = n\} = \{T_n \leq t, T_{n+1} > t\}. \tag{59}$$

This implies that,

$$\begin{aligned}\mathbb{P}\left(N_t=0\right)=\mathbb{P}\left(\theta_1>t\right)=e^{-\alpha t} \Rightarrow\ \forall n\geq 1, \mathbb{P}\left(N_t=n\right) &= \mathbb{P}\left(T_n\leq t\right)-\mathbb{P}\left(T_{n+1}\leq t\right)\\ &= \int_0^t \frac{\alpha^n}{(n-1)!}x^{n-1}e^{-\alpha x}dx-\int_0^t\frac{\alpha^{n+1}}{n!}x^n e^{-\alpha x}dx\\ &= \frac{(\alpha t)^n}{n!}e^{-\alpha t}.\end{aligned}$$

This shows that $\mathcal{L}_{N_t}=\pi_{\alpha t}$, i.e., N_t is a Poisson random variable with parameter αt.

Remark. for all $\omega\in\Omega$, $t\in[0,\infty[\mapsto N_t(\omega)\in\mathbb{N}$ is non-decreasing, piece-wise constant, and right-continuous:

$$N_0(\omega)=\{N_t(\omega)\mid t\geq 0\}\text{ is RCLL.}$$

To see this, e.g.,

$$\{N_t=n, N_{t+\varepsilon}\geq n+1\}\subseteq\{T_n\leq t,\ t<T_{n+1}\leq t+\varepsilon\}.$$

As $\varepsilon\to 0$, the limit of such a set is $\emptyset$. In addition, for all $t>0$, $N_t-N_{t^-}\in\{0,1\}$.

Proposition 11. $N:\omega\in\Omega\mapsto N(\omega)\in D([0,\infty[;\mathbb{N})$ as a stochastic process is a Levy process associated with π_α.

Proof. It's sufficient to prove that for every $k\geq 1$, $\forall 0=t_0<t_1<...<t_k$, for all $0=n_0\leq n_1\leq n_2\leq ...\leq n_k\in\mathbb{N}$,

$$\begin{aligned}\mathbb{P}\left(\bigcap_{j=1}^k\{N_{t_j}=n_j\}\right) &= \mathbb{P}\left(\bigcap_{j=1}^k\{N_{t_j}-N_{t_{j-1}}=n_j-n_{j-1}\}\right)\\ &= \prod_{j=1}^k\frac{(\alpha(t_j-t_{j-1}))^{n_j-n_{j-1}}}{(n_j-n_{j-1})!}e^{-\alpha(t_j-t_{j-1})}\end{aligned}$$

We write the LHS as:

$$\mathbb{P}\left(\bigcap_{j=1}^k\{T_{n_j}\leq t_j\leq T_{1+n_j}\}\right).$$

The event concerns $\{\theta_1,\theta_2,...,\theta_{1+n_j}\}$ which has the joint density function of

$$\alpha^{1+n_k}e^{-\alpha\sum_{j=1}^{n_k+1}X_j}dx_1dx_2...dx_{n_k+1}.$$

Hence,

$$\text{LHS}=\int_A\alpha^{1+n_k}e^{-\alpha(x_1+...+x_{1+n_k})}dx_1...dx_{1+n_k},$$

where $A=\{\sum_{m=1}^{n_j}x_m\leq t_j<\sum_{m=1}^{1+n_j}x_m\ \forall 1\leq j\leq k\}$. Under the change of variables, $u=\sum x_i$, this yields,

$$=e^{\alpha t_k}\alpha^{n_k}\text{Vol}(B),$$

where $\mathrm{Vol}(B) = \{t_k \geq \sum_{m=1}^{n_k} x_m \geq ... \geq \sum_{m=1}^{1+n_{k-1}} x_m > t_{k-1} \geq ...\}$. Hence, this becomes:

$$
\begin{aligned}
&= e^{-\alpha t_k}\alpha^{n_k}\prod_{j=1}^{k}\frac{(t_j - t_{j-1})^{n_j - n_{j-1}}}{(n_j - n_{j-1})!} \\
&= \prod_{j=1}^{k}\frac{(\alpha(t_j - t_{j-1}))^{n_j - n_{j-1}}}{(n_j - n_{j-1})!}e^{-\alpha(t_j - t_{j-1})},
\end{aligned}
$$

which is the probability function of $\pi_{\alpha(t_j - t_{j-1})}$. □

Definition 17 (Simple Poisson Jump Process)**.** We also call $N = \{N_t \mid t \geq 0\}$ the **simple Poisson jump process** associated with π_α.

Now we are ready to move onto the next step.

Step # 1: given $M \in \mathcal{M}_0(\mathbb{R}^d)$, construct a Levy Process $X = \{X_t \mid t > 0\}$.

Write $M = \alpha\nu$, where $\alpha = M(\mathbb{R}^d)$ and ν a probability measure. Let $\{\theta_m \mid m \geq 1\}$ and $N = \{N_t \mid t \geq 0\}$ be the same as in **(Step 0)** . Take $\{Y_k \mid k \geq 1\}$ to be a sequence of iid random variables with distribution ν defined on the same probability space as $\{\theta_m \mid m \geq 1\}$ but independent of the whole family $\{\theta_m \mid m \geq 1\}$.

Set $X_0 \equiv 0$ for all $t > 0$, for all $\omega \in \Omega$, define:

$$
X_t(\omega) = \sum_{k=1}^{N_t(\omega)} Y_k(\omega). \tag{60}
$$

If $N_t(\omega) = 0$, then by convention $X_t(\omega) = 0$. If $N_t(\omega) = n \geq 1$, then

$$
X_t(\omega) = \sum_{k=1}^{n} Y_k(\omega).
$$

Proposition 12. This process $X = \{X_t \mid t \geq 0\}$ is a Levy Process associated with π_M.

Proof. We will use the fact that if $Z_1, ..., Z_k$ are some $\mathbb{R}^d$-valued random variables, then $Z_1, ..., Z_k$ are mutually independent $\iff$ for all $\xi_1, ..., \xi_k \in \mathbb{R}^d$:

$$
\mathbb{E}\left[e^{i\xi_1 Z_1 + ... + i\xi_k Z_k}\right] = \prod_{j=1}^{k}\mathbb{E}\left[e^{i\xi_j Z_j}\right].
$$

(Note that this last condition is equivalent to saying that the joint distribution of $(Z_1, ..., Z_K)$ is $\mathcal{L}_{Z_1} \cdot$

$\mathcal{L}_{Z_2}\cdots\mathcal{L}_{Z_k}$). So, for all $k \geq 1$, for all $0 = t_0 < t_1 < ... < t_k$, for all $\xi_1, ..., \xi_k \in \mathbb{R}^d$, we need to compute:

$$\begin{aligned}
\mathbb{E}\left[e^{i\sum_{j=1}^k (X_{t_j}-X_{t_{j-1}},\xi_j)}\right] &= \sum_{\text{all choices of } 0=n_0\leq n_1\leq ...\leq n_k} \mathbb{E}\left[e^{i\sum_{j=1}^k (X_{t_j}-X_{t_{j-1}},\xi_j)}; N_{t_1}=n_1, N_{t_2}=n_2, ..., N_{t_k}=n_k\right] \\
&= \sum_{\text{all choices of } 0=n_0\leq n_1\leq ...\leq n_k} \mathbb{E}\left[e^{i\sum_{j=1}^k \left(\sum_{1+n_j}^{n_j} Y_p,\xi_j\right)}; N_{t_j}=n_j \ \forall\ 1\leq j\leq k\right] \\
&= \sum_{\text{all choices of } 0=n_0\leq n_1\leq ...\leq n_k} \left(\prod_{j=1}^k (\hat{v}(\xi_j))^{n_j-n_{j-1}}\right)\cdot \\
&\left(\prod_{j=1}^k \frac{\exp\left(-\alpha(t_j-t_{j-1})\cdot(\alpha(t_j-t_{j-1}))^{n_j-n_{j-1}}\right)}{(n_j-n_{j-1})!}\right) \\
&= \sum_{\text{all choices of } 0=n_0\leq n_1\leq ...\leq n_k} \prod_{j=1}^k \frac{\exp(-\alpha(t_j-t_{j-1}))}{(n_j-n_{j-1})!}\cdot(\alpha(t_j-t_{j-1})\hat{v}(\xi_j))^{n_j-n_{j-1}} \\
&= \prod_{j=1}^k \exp(-\alpha(t_j-t_{j-1}))\cdot\exp\left(\alpha(t_j-t_{j-1})\hat{v}(\xi_j)\right) \\
&= \prod_{j=1}^k \exp(\alpha(t_j-t_{j-1})(\hat{v}(\xi_j)-1) \\
&= \prod_{j=1}^k \hat{\pi}_{(t_j-t_{j-1})M}(\xi_j).
\end{aligned}$$

This proves homogeneity. ☐

Definition 18 (Compound Poisson Process)**.** We call the process above $X = X_t \mid t \geq 0\}$ a **compound Poisson Process** with jump rate α and jump distribution ν.

Remarks. How do we deal with σ-finite measures? We want to consider the jump measure associated with the sample path of X. Recall: given a function that has RCLL path in $\mathbb{R}^d$ $\varphi \in D([0,\infty[;\mathbb{R}^d))$, the jump measure associated with φ is: $\forall\ t > 0$, for all $B \in \mathcal{B}(\mathbb{R}^d)$,

$$j(t,B,\varphi) := \sum_{s\in J(t,\varphi)} \chi_B(\varphi(s)-\varphi(s^-)).$$

We will consider the jump measures $j(t,\cdot,x)$ (the jump measure associated with the process $X = \{X_t \mid t \geq 0\}$.

1. $j(t,\cdot,X) = \sum_{k=1}^{N_t}\delta_{Y_k}$: for all $B \in \mathcal{B}(\mathbb{R}^d)$, $j(t,B,X) = \sum_{k=1}^{N_t}\chi_B(Y_K) \Rightarrow j(t,\cdot,X)$ is a finite measure ($j(t,\cdot,X) \in \mathcal{M}_0(\mathbb{R}^d)$): ($\forall\omega\in\Omega, j(t,\mathbb{R}^d, X.(\omega) = N_t(\omega)$ is finite).
2. $\forall t > 0$, by definition:

$$X_t = \sum_{k=1}^{N_t} Y_k = \int_{\mathbb{R}^d} yj(t,dy,X).$$

$$||X.||_{\text{var},[0,t]} = \sum_{k=1}^{N_t}|Y_k| = \int_{\mathbb{R}^d}|y|j(t,dy,X)$$

Taking the expectation yields:

$$\begin{aligned}\mathbb{E}\left[||X.||_{\text{var},[0,t]}\right] &= \mathbb{E}\left[\sum_{k=1}^{N_t}|Y_k|\right]\\ &= \mathbb{E}\left[N_t\right]\mathbb{E}\left[|Y_1|\right] \text{ (by independence)}\\ &= \alpha t\int_{\mathbb{R}^d}|y|v(dy)\\ &= t\int_{\mathbb{R}^d}|y|M(dy)\end{aligned}$$

Therefore, if $\int_{\mathbb{R}^d}|y|M(dy) < \infty$, then $\mathbb{E}\left[X_t\right] = t\int_{\mathbb{R}^d} yM(dy)$. Similarly, $\int_{\mathbb{R}^d}|y|^2M(dy) < \infty$, then

$$\begin{aligned}\mathbb{E}\left[|X_t|^2\right] &= \mathbb{E}\left[\sum_{j=1}^{N_t}\sum_{k=1}^{N_t}(Y_j, Y_k)\right]\\ &= \mathbb{E}\left[N_t\right]\mathbb{E}\left[|Y_1|^2\right] + \mathbb{E}\left[N_t^2 - N_t\right]\cdot\left|\mathbb{E}\left[Y_1\right]\right|^2\\ &= t\int_{\mathbb{R}^d}|y|^2M(dy) + t^2\left|\int_{\mathbb{R}^d} yM(dy)\right|^2\end{aligned}$$

This proves that $\mathbb{E}\left[|X_t - \mathbb{E}\left[X_t\right]|^2\right] = t\int_{\mathbb{R}^d}|y|^2M(dy)$.

3. Given $\Gamma \in \mathcal{B}(\mathbb{R}^d)$, $t > 0$,

$$j(t, \Gamma, X) = \sum_{k=1}^{N_t}\chi_\Gamma(Y_k).$$

Consider this as a path: $\{j(t, \Gamma, X) \mid t \geq 0\}$ is non-decreasing, $\mathbb{N}$-valued, and RCLL piece-wise constant. Consider $n \geq 0$:

$$\begin{aligned}\mathbb{P}\left(j(t, \Gamma, X) - j(s, \Gamma, X) = n\right) &= \sum_{k=0}^{\infty}\sum_{l=n}^{\infty}\mathbb{P}\left(j(t, \Gamma, X) - j(s, \Gamma, X) = n, N_s = k, N_t - N_s = l\right)\\ &= \sum_{k=0}^{\infty}\sum_{l=n}^{\infty}\mathbb{P}\left(\#\{p \mid k+1 \leq p \leq k+l \mid Y_p \in \Gamma\} = n\right)\mathbb{P}\left(N_s = k\right)\mathbb{P}\left(N_t - N_s = l\right)\\ &= \sum_{k=0}^{\infty}\mathbb{P}\left(N_s = k\right)\cdot\sum_{l=n}^{\infty}\mathbb{P}\left(\#\{p \mid k+1 \leq p \leq k+l \mid Y_p \in \Gamma\} = n\right)\cdot\mathbb{P}\left(N_{t-s} = l\right)\\ &= \sum_{l=n}^{\infty}\mathbb{P}\left(j(t-s, \Gamma, X) = n; N_{t-s} = l\right)\\ &= \mathbb{P}\left(j(t-s, \Gamma, X) = n\right) \text{ (proves independent increments, by rmk 1)}\\ &= \sum_{l=n}^{\infty}\exp(-\alpha(t-s))\frac{(\alpha(t-s))^l}{l!}\binom{l}{n}(\mathbb{P}\left(Y_n \in \Gamma\right))^n(1 - \mathbb{P}\left(Y_1 \in \Gamma\right))^{l-n} \text{ (rmk 2)}\\ &= \sum_{l=n}^{\infty}\exp(-\alpha(t-s))\frac{(\alpha(t-s))^l}{l!}\binom{l}{n}(v(\Gamma))^n(1 - v(\Gamma))^{l-n}\\ &= \exp(-(t-s)M(\Gamma))\frac{((t-s)M(\Gamma))^n}{n!} \to \pi_{(t-s)M(\Gamma)} \text{ (in distribution)}\end{aligned}$$

Still need to prove that the increments are independent. It's an exercise.

Conclusion: the process $\{j(t,\Gamma,X) \mid t \geq 0\}$ is a simple Poisson process with jump rate $M(\Gamma)$.

Remark. Let $M \in \mathcal{M}_0(\mathbb{R}^d)$, $X = \{X_t \mid t \geq 0\}$ be a Levy process associated with π_M and $\{j(t,\cdot,X) \mid t \geq 0\}$ be the jump measure associated with X.

If $\{\Gamma_m \mid m \geq 1\} \subseteq \mathcal{B}(\mathbb{R}^d)$ and $\Gamma_m \cap \Gamma_l = \emptyset$ for all $l \neq m$, then $\{\{j(t,\Gamma_m,X) \mid t \geq 0\} \mid m \geq 1\}$ is an independent family of stochastic process.

Proof. We will only prove independence for $\{j(t,\Gamma_1,X) \mid t \geq 0\}$ and $\{j(s,\Gamma_2,X) \mid t \geq 0\}$. For all $0 \leq s \leq t$, for all $\xi_1, \xi_2 \in \mathbb{R}$:

$$\begin{aligned}
&\mathbb{E}\left[\exp(i(\xi_1, j(t,\Gamma_1,X)) + i(\xi_2, j(s,\Gamma_2,X)))\right] \\
&= \mathbb{E}\left[\exp(i(\xi_1, j(t,\Gamma_1,X)) - i(\xi_1, j(s,\Gamma_1,X)) \cdot \exp(i(\xi_1, j(s,\Gamma_1,X)) + i(\xi_2, j(s,\Gamma_2,X))\right] \\
&= \mathbb{E}\left[\exp(i(\xi_1, j(t,\Gamma_1,X) - j(s,\Gamma_1,X)))\right] \cdot \mathbb{E}\left[...\right] \\
&= \underbrace{\exp((t-s)M(\Gamma_1)(e^{i\xi_1}-1))}_{:=(F)} \cdot \underbrace{\sum_{L=0}^{\infty} \mathbb{E}\left[\exp(i(\xi_1, j(s,\Gamma_1,X)) + i(\xi_2, j(s,\Gamma_2,X))); N_s = L\right]}_{:=(E)}
\end{aligned}$$

When $N_s = L$, $X = \sum_{k=1}^{L} Y_k$ and $j(s,\Gamma,X) = \sum_{k=1}^{L} \chi_{\Gamma_1}(Y_k)$. Hence,

$$\begin{aligned}
(E) &= \sum_{L=0}^{\infty} \frac{e^{-\alpha s}(\alpha s)^L}{L!} \left(\int_{\mathbb{R}^d} (\exp(i\xi_1 \chi_{\Gamma_1}(y) + i\xi_2 \chi_{\Gamma_2}(y))) v(dy) \right)^L \\
&= \sum_{L=0}^{\infty} \frac{e^{-\alpha s}(\alpha s)^L}{L!} \left(e^{i\xi_1} v(\Gamma_1) + e^{i\xi_2} v(\Gamma_2) + 1 - v(\Gamma_1) - v(\Gamma_2) \right)^L \\
&= e^{-\alpha s} \cdot \exp\left(s e^{i\xi_1} M(\Gamma_1) + s e^{i\xi_2} M(\Gamma_2) - \alpha s - SM(\Gamma_1) + SM(\Gamma_2) \right)
\end{aligned}$$

So,

$$\begin{aligned}
(F) \cdot (E) &= \exp(tM(\Gamma_1)(e^{i\xi_1} - 1) \cdot \exp(sM(\Gamma_2)(e^{i\xi_2} - 1)) \\
&= \mathbb{E}\left[\exp(i\xi_1 j(t,\Gamma_1,X)\right] \cdot \mathbb{E}\left[\exp(i\xi_2 j(s,\Gamma_2,X))\right]
\end{aligned}$$

□

Now assume that we have a sequence of mutually singular measures, $\{M_k \mid 1 \leq k \leq K\} \subseteq \mathcal{M}_0(\mathbb{R}^d)$, i.e., there exists $\{\Gamma_k \mid 1 \leq k \leq K\} \subseteq \mathcal{B}(\mathbb{R}^d)$, $\Gamma_k \cap \Gamma_l = \emptyset$ if $k \neq l$ such that $M_k(\Gamma_k^c) = 0$ for all $1 \leq k \leq K$ (the mass of M_k is concentrated in Γ_k). For each $k \in \{1, ..., K\}$, there exists a Levy process

$$X^{(k)} = \{X_t^{(k)} \mid t \geq 0\}$$

associated with π_{M_k}. We can take all these processes $\{X^{(k)} \mid 1 \leq k \leq K\}$ to be on the same probability space and to be independent.

Claim. $X = \sum_{k=1}^{K} X^{(k)}$ is a Levy process associated with π_M, where $M = \sum_{k=1}^{K} M_k \in \mathcal{M}_0(\mathbb{R}^d)$.

Proof. Obviously, $X = \{X_t \mid t \geq 0\}$ is RCLL with independent and homogeneous increments. For all

$t \geq 0$ and for all $\xi \in \mathbb{R}^d$,

$$\begin{aligned}\hat{\mathcal{L}_{X_t}}(\xi) &= \prod_{k=1}^{K} \hat{\mathcal{L}_{X_t^{(k)}}}(\xi) \\ &= \exp\left(\sum_{k=1}^{K} t \int_{\mathbb{R}^d} \left(e^{i(\xi,y)} - 1\right) M_k(dy)\right) \\ &= \exp\left(t \int_{\mathbb{R}^d} \left(e^{i(\xi,y)} - 1\right) M(dy)\right).\end{aligned}$$

This shows that $\mathcal{L}_{X_t} = \pi_{tM}$, which proves the claim. □

In particular, if $j(t, \cdot, X)$ is the jump measure associated with X, then:

$$j(t,\cdot,X) = \sum_{k=1}^{K} j(t,\cdot,X)\big|_{\Gamma_k} = \sum_{k=1}^{K} j(t,\cdot,X^{(k)}).$$

In words: any jump of X must be from one and only one $X^{(k)}$. WLOG, assume that $\{\Gamma_k \mid 1 \leq k \leq K\}$ is a partition of $\mathbb{R}^d$. Now assume that M, $\{M_k \mid k \geq 1\}$, and $\{\Gamma_k \mid k \geq 1\}$ are the same as above. Let $X = \{X_t \mid t \geq 0\}$ be a Levy process associated with π_M. Then, for all $t > 0$, for all $\omega \in \Omega$:

$$\begin{aligned}X_t(\omega) &= \int_{\mathbb{R}^d} y j(t, dy, X.(\omega)) \\ &= \sum_{k=1}^{K} \int_{\Gamma_k} y j(t, dy, X.(\omega)) \\ &= \sum_{k=1}^{K} \int_{\mathbb{R}^d} y j^{(k)}(t, dy, \omega),\end{aligned}$$

where $j^{(k)}(t,\cdot,\omega) = j(t,\cdot,X(\omega))\big|_{\Gamma_k}$. For every $1 \leq k \leq K$, set

$$X_t^{(k)}(\omega) = \int_{\mathbb{R}^d} y j^{(k)}(t, dy, \omega).$$

Claim. $\{X^{(k)} \mid 1 \leq k \leq K\}$ is independent and for every $1 \leq k \leq K$, $\{X_t^{(k)} \mid t \geq 0\}$ is a Levy process associated with π_{M_k}.

Proof. The $X^{(k)}{}'s$ are independent because the $\{j(t,\cdot,X)\big|_{\Gamma_k} \mid 1 \leq k \leq K\}$ are independent. Moreover, for all $1 \leq k \leq K$, $\{X_t^{(k)} \mid t \geq 0\}$ is RCLL with independent and homogeneous increments. For all $t \geq 0$, for all $\xi \in \mathbb{R}^d$,

$$\mathbb{E}\left[e^{i(\xi, X_t^{(k)})}\right] = \sum_{m=0}^{\infty} \mathbb{E}\left[e^{i(\xi, X_t^{(k)})}; N_t = m\right]$$

When $N_t = m$, then $X_t^{(k)} = \sum_{j=1}^{N_t} Y_j \chi_{\Gamma_k}(Y_j)$. Hence,

$$\begin{aligned}&= \sum_{m=0}^{\infty} e^{-\alpha t} \frac{(\alpha t)^m}{m!} \cdot \left(\int_{\mathbb{R}^d} e^{i(\xi,y)\chi_{\Gamma_k}(y)} v(dy)\right)^m \\ &= \sum_{m=0}^{\infty} e^{-\alpha t} \frac{(\alpha t)^m}{m!} \left(\int_{\Gamma_k} e^{i(\xi,y)} v(dy) - 1 - v(\Gamma_k)\right)^m \\ &= \exp\left(t \int_{\mathbb{R}^d} \left(e^{i(\xi,y)} - 1\right) M_k(dy)\right).\end{aligned}$$

Hence, $\mathcal{L}_{X_t^{(k)}} = \pi_{tM_k}$. □

Lemma 8. Given $M \in \mathcal{M}_\infty(\mathbb{R}^d)$ (i.e., M is σ-finite and $M(\{0\}) = 0$ and for all $r > 0$, $M(B(0,r)^c) < \infty$), there exists a mapping $(t,\omega) \in [0,\infty[\times\Omega \mapsto j(t,\cdot,\omega) \in \mathcal{M}_\infty(\mathbb{R}^d)$ referred to as the **jump measure corresponding to** M such that:

1. $\forall\, \Gamma \in \mathcal{B}(\mathbb{R}^d)$ such that $0 \notin \overline{\Gamma}$, $\{j(t,\Gamma) \mid t \geq 0\}$ is a simple Poisson process with rate $M(\Gamma)$.
2. For all $\{\Gamma_m \mid m \geq 1\} \subseteq \mathcal{B}(\mathbb{R}^d)$ such that $\Gamma_m \cap \Gamma_l = \emptyset$ for all $m \neq l$ and $0 \notin \overline{\bigcup_{m=1}^{L} \Gamma_m}$ for all $L \geq 1$. Then,

$$\{\{j(t,\Gamma_m) \mid t \geq 0\} \mid m \geq 1\}$$

is an independent family of stochastic processes.

Proof. Set $A_0 := \mathbb{R}^d \setminus \overline{B(0,1)}$. Set $A_k := \overline{B(0,2^{-(k+1)})} \setminus \overline{B(0,2^{-k})}$ for all $k \in \mathbb{N}$ (we are cutting the space into annuli, since the origin is the problematic part). For every $k \geq 0$, set $M_k := \chi_{A_k} \cdot M \in \mathcal{M}_0(\mathbb{R}^d)$. Then, there exists an $X^{(k)} = \{X^{(k)}_t \mid t \geq 0\}$ a Levy Process associated with π_{M_k} and $\{X^{(k)} \mid k \geq 1\}$ are independent and defined on $(\Omega,\mathcal{F},\mathbb{P})$. For all $(t,\omega) \in [0,\infty[\times\Omega$ and for all $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,

$$j(t,\Gamma,\omega) := \sum_{k=0}^{\infty} j(t,\Gamma,X^{(k)}_\cdot(\omega))$$

is well-defined. It's straightforward to check that this $j(t,\cdot,\omega)$ is a measure on $\mathbb{R}^d$ (monotone convergence). Moreover, for every $r > 0$, when k is sufficiently large such that $2^{-k+1} < r$, then:

$$j(t,\overline{B(0,r)}^c,X^{(k)}) = 0 \Rightarrow j(t,\overline{B(0,r)}^c,\omega) < \infty \Rightarrow j(t,\cdot,\omega) \in \mathcal{M}_0(\mathbb{R}^d).$$

Proof of **(1)**. Given $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ such that $0 \notin \overline{\Gamma}$, there exists a k sufficiently large such that $\overline{B(0,2^{-k})} \cap \overline{\Gamma} = \emptyset$.

$$\Rightarrow j(t,\Gamma,\omega) = \sum_{k=1}^{K} j(t,\Gamma,X^{(k)}_\cdot(\omega)).$$

We have shown that RHS $= j(t,\Gamma,X)$, where $X = \sum_{k=1}^{K} X^{(k)}$, and X is a Levy process associated with $\pi_{M^{(k)}}$, where

$$M^{(k)} = \sum_{k=1}^{K} M_k.$$

This implies that $\{j(t,\Gamma) = j(t,\Gamma,X) \mid t \geq 0\}$ is a simple Poisson process with rate $M^{(k)}(\Gamma) = M(\Gamma)$.

Proof of **(2)**: exercise. □

Step # 2: We construct a Levy Process associated with $\pi^{(1)}_{0,0,M}$ for $M \in \mathcal{M}_1(\mathbb{R}^d)$ and a Levy Process associated with $\pi_{0,0,M}$ for $M \in \mathcal{M}_2(\mathbb{R}^d)$.

Lemma 9. Given $M \in \mathcal{M}_1(\mathbb{R}^d)$, let $(t,\omega) \in [0,\infty[\times\Omega \mapsto j(t,\cdot,\omega)$ be the jump measure corresponding to M. Then, for almost every $\omega \in \Omega$,

$$\int_{\mathbb{R}^d} |y| j(t,dy,\omega) =: V_t(\omega) < \infty \text{ for all } t > 0.$$

Proof. Let $\{A_k \mid k \geq 0\}$, $\{M_k \mid k \geq 0\}$, and $\{X^{(k)} \mid k \geq 0\}$ be the same as above. For every $t \geq 0$, we write:

$$V_t(\omega) = \sum_{k=0}^{\infty} \underbrace{\int_{A_k} |y| j(t, dy, \omega)}_{V_t^{(k)}(\omega)}$$

Next, write:

$$V_t(\omega) = V_t^{(0)}(\omega) + \sum_{k=1}^{\infty} V_t^{(k)}(\omega).$$

By the previous step, we know that $V_t^{(0)}$ is finite for almost every ω. Hence, it suffices to show that $\sum_{k=1}^{\infty} V_t^{(k)}$ converges for almost every $\omega \in \Omega$. For all $k \geq 1$, $X^{(k)}$ is the Levy Process associated with π_{M_k}. Then,

$$\begin{aligned} & V_t^{(k)} = ||X^{(k)}||_{\text{var},[0,t]}. \\ \Rightarrow & \mathbb{E}\left[V_t^{(k)}\right] = t \int_{A_k} |y| M(dy) \\ \Rightarrow & \sum_{k=1}^{\infty} \mathbb{E}\left[V_t^{(k)}\right] = t \int_{\overline{B(0,1)}} |y| M(dy) < \infty \text{ since } M \in \mathcal{M}_1(\mathbb{R}^d). \end{aligned}$$

By **(MON)** ,

$$\mathbb{E}\left[\sum_{k=1}^{\infty} V_t^{(k)}\right] = \sum_{k=1}^{\infty} \mathbb{E}\left[V_t^{(k)}\right] < \infty \Rightarrow \sum_{k=1}^{\infty} V_t^{(k)} \text{ converges almost surely.}$$

This proves the Lemma. □

Step # 2.1:

Theorem 15. *Given $M \in \mathcal{M}_2(\mathbb{R}^d)$, let $j(t, \cdot, \omega)$ be the jump measure corresponding to M. Given $r \in]0, 1[$, define:*

$$\begin{aligned} & X_0^r \equiv 0 \textit{ for all } \omega \in \Omega, \ \forall t > 0 \\ & X_t^{(r)} := \int_{\mathbb{R}^d \setminus \overline{B(0,r)}} y j(t, dy, \omega) - t \int_{\overline{B(0,1)} \setminus \overline{B(0,r)}} y M(dy). \end{aligned}$$

Then, for almost every $\omega \in \Omega$, $X_t(\omega) := \lim_{r \downarrow 0} X_t^{(r)}$ exists for every $t \geq 0$, and $\{X_t \mid t \geq 0\}$ is a Levy Process associated with $\pi_{0,0,M}$.

Proof. For every $r > 0$, $\{X_t^{(r)} \mid t \geq 0\}$ is a Levy process associated with $\pi_{0,0,M^{(r)}}$, where

$$M^{(r)} := M|_{\overline{B(0,1)}^c}.$$

For every $0 < r < r' < 1$, set $Y_t := X_t^{(r)} - X_t^{(r')}$, i.e.:

$$Y_t = \int_{\overline{B(0,r')} \setminus \overline{B(0,r)}} y j(t, dy) - \int_{\overline{B(0,r')} \setminus \overline{B(0,r)}} y M(dy).$$

Then,

$$
\begin{aligned}
||X^{(r)} - X^{(r')}||_{U,[0,t]} &= ||Y||_{U,[0,t]} \\
&= \lim_{n\to\infty} \max_{0\leq m\leq 2^n} |Y_{m2^{-n}t}|.
\end{aligned}
$$

For all $\varepsilon > 0$,

$$
\begin{aligned}
\mathbb{P}\left(||Y||_{U,[0,t]} > \varepsilon\right) &= \lim_{n\to\infty} \mathbb{P}\left(\max_{0\leq m\leq 2^{-n}} |Y_{m2^{-n}t}| \geq \varepsilon\right) \\
&\leq \lim_{n\to\infty} \sum_{i=1}^{d} \mathbb{P}\left(\max_{0\leq m\leq 2^n} |(Y_{m2^{-n}t}, e_j)| \geq \varepsilon\right) \\
&= \lim_{n\to\infty} \sum_{j=1}^{d} \mathbb{P}\left(\max_{0\leq m\leq 2^n} \left|\sum_{l=1}^{m} \left(Y_{\ell 2^{-n}t} - Y_{(\ell-1)2^{-n}t}, e_j\right)\right| > \frac{\varepsilon}{\sqrt{d}}\right).
\end{aligned}
$$

The $Y_{\ell 2^{-n}t} - Y_{(\ell-1)2^{-n}t}$ are iid random variables, and so by Kolmogorov's Inequality,

$$
\begin{aligned}
&\leq \lim_{n\to\infty} \sum_{j=1}^{d} \frac{d}{\varepsilon^2}\mathbb{E}\left[|(Y_t, e_j)|^2\right] \\
&= \frac{d}{\varepsilon^2}\mathbb{E}\left[|Y_t|^2\right] \\
&= \frac{d}{\varepsilon^2}\int_{\overline{B(0,r^2)}\setminus\overline{B(0,r)}} |y|^2 M(dy).
\end{aligned}
$$

Since $M \in \mathcal{M}_2(\mathbb{R}^d)$, we can choose a sequence $\{r_n \mid n \geq 1\}$ such that $r_n \downarrow 0$ and

$$
t\int_{r_n \leq |y| \leq r_{n-1}} |y|^2 M(dy) \leq 2^{-n},
$$

for every $n \in \mathbb{N}$. This implies that

$$
\mathbb{P}\left(||X^{(r_n)} - X^{(r_{n-1})}||_{U,[0,t]} \geq 2^{-\frac{n}{4}}\right) \leq dt2^{\frac{n}{2}}2^{-n} = dt2^{-\frac{n}{2}}.
$$

By **(BC1)** ,

$$
\mathbb{P}\left(||X^{(r_n)} - X^{(r_{n-1})}||_{U,[0,t]} \geq 2^{-\frac{n}{4}} \text{ i.o.}\right) = 0.
$$

Hence, there exists $\Omega' \subseteq \Omega$ such that $\mathbb{P}(\Omega') = 1$ such that for all $\omega \in \Omega'$,

$$
||X^{(r_n)} - X^{(r_{n-1})}||_{U,[0,t]}(\omega) \leq 2^{-\frac{n}{4}} \text{ for } n \text{ sufficiently large.} \tag{61}
$$

Hence, $X_t(\omega) := \lim_{n\to\infty} X_t^{(r_n)}(\omega)$ exists. This implies that $X = \{X_t \mid t \geq 0\}$ is RCLL. It's straightforward to verify that X is a Levy Process associated with $\pi_{0,0,M}$ by **(Convergence of Characteristic Functions)** . □

Corrolary 5. Let $M \in \mathcal{M}_2(\mathbb{R}^d)$. Let $\{X_t \mid t \geq 0\}$ be a Levy process associated with $\pi_{0,0,M}$.

1. If $M \in \mathcal{M}_1(\mathbb{R}^d)$, then for almost every $\omega \in \Omega$, $||X||_{\text{Var},[0,t]} < \infty$ for all $t > 0$.
2. If $M \notin \mathcal{M}_1(\mathbb{R}^d)$, then for almost every $\omega \in \Omega$, $||X||_{\text{Var},[0,t]}(\omega) = \infty$ for all t.

Step # 3: we construct a Levy Process $\{X_t \mid t \geq 0\}$ associated with $\pi_{0,C,0} = \gamma_{0,C}$.

Remark. WLOG, we can assume that $C = I$ (diagonal matrix). If $\{B_t \mid t \geq 0\}$ is a Levy Process associated with the standard d-dimensional Gaussian measure $\gamma_{0,I}$, then for general $C = (C_{ij})_{d\times d} \geq 0$, set $X_t = \sqrt{C}B_t$. $\sqrt{C}$ is the unique non-negative symmetric matrix such that $\sqrt{C}\sqrt{C} = C$. Then, for all $\xi \in \mathbb{R}^d$,

$$\mathbb{E}\left[e^{i(\xi,X_t)}\right] = \mathbb{E}\left[e^{i(\sqrt{C}\xi,B_t)}\right] = e^{-\frac{t}{2}(C,C\xi)},$$

which shows that $\mathcal{L}_{X_t} = \gamma_{0,tC}$, which shows that $\{X_t \mid t \geq 0\}$ is a Levy Process associated with $\gamma_{0,C}$.

Definition 19. If $\{B_t \mid t \geq 0\}$ is a Levy Process associated with the standard Gaussian $\gamma_{0,I}$, then $B = \{B_t \mid t \geq 0\}$ is a **standard Brownian Motion (B.M.)** in $\mathbb{R}^d$. "Standard" means that $B_0 = 0$.

First, some remarks.

1. $\forall 0 \leq s \leq t$, we know that

$$\begin{aligned}\mathcal{L}_{B_t} &= \gamma_{0,tI}\\ \mathcal{L}_{B_t - B_s} &= \gamma_{0,(t-s)I},\end{aligned}$$

   and $B_t - B_s$ is independent of B_r for all $0 \leq r \leq s$.
2. If $B_t = (B_t^{(1)}, B_t^{(2)}, ..., B_t^{(d)})$, then:
   (a) $\{B_t^{(j)} \mid t \geq 0\}$ is a standard one-dimensional Brownian motion for all $1 \leq j \leq d$.
   (b) For all $1 \leq j \leq d$, $\forall t, s \geq 0$,

$$\mathbb{E}\left[B_t^{(j)} B_s^{(j)}\right] = t \wedge s. \tag{62}$$

   (c) For all $1 \leq j, j' \leq d$, $j \neq j'$, $\{B_t^{(j)} \mid t \geq 0\}$ and $\{B_t^{(j')} \mid t \geq 0\}$ are independent.
   (d) For all $t, s \geq 0$,

$$\begin{aligned}\mathbb{E}\left[(B_t, B_s)\right] &= (t \wedge s)d\\ \mathbb{E}\left[(B_t)^2\right] &= td.\end{aligned}$$

3. $\{(B_t, \xi) \mid t \geq 0, \xi \in \mathbb{R}^d\}$ is a **Gaussian Family**, i.e., for all $k \geq 1$, for all $0 \leq t_1 \leq ... \leq t_k$, for all $\xi_1, ..., \xi_k \in \mathbb{R}^d$, the joint distribution of $((B_{t_1}, \xi_1), (B_{t_2}, \xi_2), ..., (B_{t_k}, \xi_k))$ is a multivariate Gaussian distribution with expectation function:

$$\mathbb{E}\left[(B_t, \xi)\right] = 0 \ \forall t \geq 0, \forall \xi \in \mathbb{R}^d,$$

   and covariance function for all $t, s \geq 0$, for all $\xi, \eta \in \mathbb{R}^d$:

$$\mathbb{E}\left[(B_t, \xi)(B_s, \eta)\right] = (t \wedge s)(\xi, \eta).$$

   Equivalently, for all $k \geq 1$, for all $0 \leq t_1 \leq ... \leq t_k$, $(B_{t_1}, B_{t_2}, ..., B_{t_k})$ as an $\mathbb{R}^{dk}$-valued random variable has distribution $\gamma_{0,C_{(t_1,...,t_k)}}$ where

$$C_{(t_1,...,t_k)} = \left(t_{j\wedge j'} I_{d\times d}\right)_{j=1,...,k,\ j'=1,...,k}$$

2.2.1 Levy's Construction of Brownian Motion (Pathwise Construction)

The "building blocks" will be $\{X_{n,m} \mid n \geq 0, m \geq 0\}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ i.i.d. random variables with distribution $\gamma_{0,I}$. We will define a family of processes $\{B^{(n)}_t \mid t \geq 0\}$ for $n \geq 0$ and prove that $B^{(n)}_{\cdot}$ converges to the Brownian motion.

<u>$n = 0$</u>: For $n = 0$, set $B^{(0)}_0 \equiv 0$ and $B^{(0)}_m := \sum_{l=1}^m X_{0,l}$ for $m \geq 1$. Linearly interpolate $[m-1, m]$ for each m, i.e., if $t \in [m-1, m]$, then

$$B^{(0)}_t = B^{(0)}_{m-1} + (t - (m-1))(B^{(0)}_m - B^{(0)}_{m-1}). \tag{63}$$

The first step is shown in red in the figure below. The map $t \in [0, \infty[\mapsto B^{(0)}_t \in \mathbb{R}^d$ is continuous and piece-wise linear.

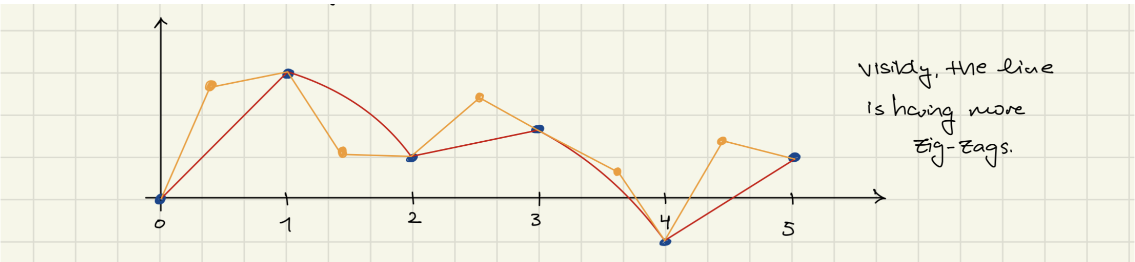


Figure 2: First two steps of construction of the Brownian Motion

<u>$n = 1$</u>: define $B^{(1)}_m = B^{(0)}_m$ for all $m \geq 1$ (i.e., don't change the value at nodes). At the middle point of the interval $[m-1, m]$, define:

$$B_{\frac{2m-1}{2}} := B^{(0)}_{\frac{2m-1}{2}} + \frac{1}{2} X_{1,m} \; \forall m \geq 1. \tag{64}$$

We need to do this, because note that the desired variance is $\frac{2m-1}{2}$, but

$$B^{(0)}_{\frac{2m-1}{2}} = \frac{1}{2}(B^{(0)}_{m-1} + B^{(0)}_m) = \gamma_{0,(m-\frac{3}{4}I)},$$

that is, the current variance is off by $\frac{1}{4}I$, which explains the need for the correction term $\frac{1}{2}X_{1,m}$. Now note that

$$\mathcal{L}_{B^{(1)}_{\frac{2m-1}{2}} = \gamma_{0,\frac{2m-1}{2}I}}.$$

Finally, to complete $B^{(1)}_t$, we use linear interpolation. Hence, $t \in [0, \infty[\mapsto B^{(1)}_t$ is continuous and piece-wise linear. But this next step made it more random with more zig zags then $B^{(0)}_t$.

<u>Induction</u>: Assume that $\{B^{(n)}_t \mid t \geq 0\}$ is built with all the desired properties:

1. $t \in [0, \infty[\mapsto B^{(n)}_t \in \mathbb{R}^d$ is continuous and piece-wise linear.
2. $\mathcal{L}_{B^{(n)}_{m2^{-n}}} = \gamma_{0,m2^{-n}I}$ for all $m \geq 1$.
3. $\forall m, m'$, for all $\xi, \eta \in \mathbb{R}^d$,

$$\mathbb{E}\left[(\xi, B^{(n)}_{m2^{-n}})(\eta, B^{(n)}_{m'2^{-n}})\right] = 2^{-n}(m \wedge m')(\xi, \eta) \tag{65}$$

We will call this property (*).

We call these three properties (A). Then, for $(n+1)$: set $B^{(n+1)}_{m2^{-n}} = B^{(n)}_{m2^{-n}}$ for all $m \geq 1$. As before,

$$B^{(n+1)}_{(2m-1)2^{-n-1}} = B^{(n)}_{(2m-1)2^{-n-1}} + 2^{-\frac{n}{2}-1} X_{n+1,m}.$$

Linearly interpolate to complete $B^{(n+1)}_t$ for all $t \geq 0$. Now we need to check that this inductive construction satisfies (A). We only need to check (*). Given $\xi, \eta \in \mathbb{R}^d$, $k, k' \geq 1$, $k < k'$,

- If both k and k' are even, i.e., if $k = 2m$ and $k' = 2m'$ for some $m < m'$, then (*) is reduced to the assumption given on $\{B^{(n)}_t \mid t \geq 0\}$ (the old nodes).
- If $k = 2m$ and $k' = 2m' - 1$, with $1 \leq m < m'$, then:

$$\begin{aligned}
\mathbb{E}\left[(\xi, B^{(n+1)}_{k2^{-n-1}})(\eta, B^{(n+1)}_{k'2^{-n-1}})\right] &= \mathbb{E}\left[(\xi, B^{(n)}_{m2^{-n}})(\eta, B^{(n)}_{(2m'-1)2^{-n-1}} + 2^{-\frac{n}{2}-1} X_{n+1,m'})\right] \\
&= \frac{1}{2}\left(2^{-n}(m \wedge (m'-1)) + 2^{-n}(m \wedge m')\right)(\xi, \eta) \\
&= 2^{-n} m (\xi, \eta) \\
&= k 2^{-n-1} (\xi, \eta).
\end{aligned}$$

- The other cases are left as an exercise: the computation is exactly identical to what's been done above.
    1. $k = 2m - 1$, $k' = 2m'$ with $m \leq m'$.
    2. $k = 2m - 1$, $k' = 2m' - 1$, with $m \leq m'$.

Next, we prove the convergence of $\{B^{(n)}_t \mid t \geq 0\}$ as $n \to \infty$. For all $L \geq 1$, Consider:

$$\begin{aligned}
||B^{(n+1)}_\cdot - B^{(n)}_\cdot||_{U,[0,2^L]} &= \max_{1 \leq m \leq 2^{n+L}} \left| B^{(n+1)}_{(2m-1)2^{-n+1}} - B^{(n)}_{(2m-1)2^{-n+1}} \right| \\
&= 2^{-\frac{n}{2}-1} \max_{1 \leq m \leq 2^{n+L}} |X_{n+1,m}|
\end{aligned}$$

Taking the expectation of both sides,

$$\begin{aligned}
\mathbb{E}\left[||B^{(n+1)}_\cdot - B^{(n)}_\cdot||_{U,[0,2^L]}\right] &\leq 2^{-\frac{n}{2}-1} \left(\mathbb{E}\left[\max_{1 \leq m \leq 2^{n+L}} |X_{n+1,m}|^4\right]\right)^{\frac{1}{4}} \\
&\leq 2^{-\frac{n}{2}-1} \left(\mathbb{E}\left[\sum_{m=1}^{2^{L+n}} |X_{n+1,m}|^4\right]\right)^{\frac{1}{4}} \\
&\leq C \cdot 2^{-\frac{n}{2}-1} (2^{L+n})^{\frac{1}{4}} \\
&= C 2^{\frac{L}{4}} 2^{-1} 2^{-\frac{n}{4}},
\end{aligned}$$

which decays exponentially fast. Applying **(Markov)** and **(BC)** ,

$$\mathbb{P}\left(||B^{(n+1)}_\cdot - B^{(n)}_\cdot||_{U,[0,2^L]} > 2^{-\frac{n}{8}} \text{ i.o.}\right) = 0.$$

Hence, there exists an $\Omega' \subseteq \Omega$, $\mathbb{P}(\Omega') = 1$, such that for all $\omega \in \Omega'$,

$$B_t(\omega) := \lim_{n \to \infty} B^{(n)}_t(\omega)$$

exists and the convergence is uniform over any compact set. Hence, the map $t \in [0, \infty[\mapsto B_t(\omega) \in \mathbb{R}^d$ is continuous.

Finally, one needs to check that the path $B = \{B_t \mid t \geq 0\}$ is a standard Brownian motion. Note that, for all t,

$$B_t = \lim_{n\to\infty} B_t^{(n)} = \lim_{n\to\infty} B_{\lfloor t\rfloor_n}^{(n)},$$

where $\lfloor t\rfloor_n = (m-1)2^{-n}$ if $(m-1)2^{-n} \leq t < m2^{-n}$. It's straightforward to check that for all $0 \leq r \leq s < t$, for all $\xi_1, \xi_2 \in \mathbb{R}^d$,

$$\begin{aligned}\mathbb{E}\left[e^{i(\xi_1,B_r)+i(\xi_2,B_t-B_s)}\right] &= \lim_{n\to\infty} \mathbb{E}\left[e^{i(\xi_1,B^{(n)}_{\lfloor r\rfloor_n})} e^{i(\xi_2,B^{(n)}_{\lfloor t\rfloor_n}-B^{(s)}_{\lfloor s\rfloor_n})}\right] \\ &= \lim_{n\to\infty} e^{-\frac{1}{2}|\xi_1|^2\lfloor r\rfloor_n} e^{-\frac{|\xi|^2}{2}(\lfloor t\rfloor_n-\lfloor s\rfloor_n)} \textbf{(IH)} \\ &= e^{-\frac{|\xi_1|^2}{2}r} e^{-\frac{|\xi_2|^2}{2}(t-s)}.\end{aligned}$$

Corrolary 6. If $\{B_t \mid t \geq 0\}$ is a Brownian Motion, then $t \in [0,\infty[\mapsto B_t \in \mathbb{R}^d$ is continuous.

Before moving on, we will summarize what we have just done. We constructed $\{B_t\}$ a Levy process associated with $\gamma_{0,I}$. It's standard since $B_0 = 0$. The following three notions are equivalent.

- $\forall\ 0 \leq r \leq s \leq t$, $B_t - B_s$ is independent of B_r, and

$$\mathcal{L}_{B_t-B_s} = \gamma_{0,(t-s)I}. \tag{66}$$

- $\forall\ k \geq 1$, $0 \leq t_1 \leq t_2 \leq \ldots \leq t_k$,

$$\mathcal{L}_{(Bt_1,\ldots,Bt_k)} = \gamma_{0,C(t_1,\ldots,t_k)}, \tag{67}$$

  where $C(t_1,\ldots,t_k) = (t_{j\wedge j'} I_{d\times d})_{j,j'=1,..,k}$.
- $\{B_t \mid t \geq 0\}$ is a **Gaussian Process** such that $\forall\ \eta$ and $\xi \in \mathbb{R}^d$, $\forall\ t, s \geq 0$ with:
  1. $\mathbb{E}\left[(B_t,\xi)_{\mathbb{R}^d}\right] = 0$ (**expectation function**).
  2. $\operatorname{cov}((B_t,\xi),(B_s,\xi) = (t\wedge s)(\xi,\nu)$ (**covariance function**).

$\{X_t\}$ is a Gaussian process if any finite-dimensional distribution is a multivariate Gaussian distribution. From Levy's Construction of Brownian motion, we got:

- $\forall\ n \geq 0$, $\{B_t^{(n)} \mid t \geq 0\}$ is continuous and piece-wise linear a.s., $B_t = \lim_{n\to\infty} B_t^{(n)}$ and the convergence is uniform on compact sets.
- a.s. $t \in [0,\infty[\mapsto B_t \in \mathbb{R}^d$ is **continuous.**

Remark. Levy's Construction of Brownian Motion (B.M.) is an "outward" construction, i.e., we start with "nice" functions (continuous piece-wise linear) and obtain B.M. through convergence (so we go from a small class of functions to something more general). There is also an "inward construction": start with something bigger, then zoom into something smaller. It is done as follows:

1. Start with the construction of the desired probability measures on $(\mathbb{R}^d)^{[0,\infty[}$. Recall that this notation is the collection of all $f : [0,\infty[\to \mathbb{R}$.
2. Confirm that this probability measure is in fact supported on a smaller function space, that is, $C([0,\infty[;\mathbb{R}^d) = C([0,\infty[)$.

This will be done in two tasks.

Task 1: Follow the steps below to construct the distribution of Brownian motion on all functions $(\mathbb{R}^d)^{[0,\infty[}$, i.e., construct a probability measure μ on $((\mathbb{R}^d)^{[0,\infty[}, \Sigma_{\mathbb{R}^d}^{[0,\infty[})$ such that for all $0 \leq r \leq s \leq t$, $\tau_t - \tau_s$ (the projection mappings) and τ_r are independent under μ, and $\mathcal{L}_{\tau_t - \tau_s} = \gamma_{0,(t-s)I}$ under μ. Recall how the projection maps are defined as: given a time t, the projection mapping does:

$$\begin{aligned} \tau_t &: (\mathbb{R}^d)^{[0,\infty[} \to \mathbb{R}^d \\ f &\in (\mathbb{R}^d)^{[0,\infty[} \to \tau_t(f) = f(t) \in \mathbb{R}^d. \end{aligned}$$

is viewed as a random variable on $(\mathbb{R}^d)^{0,\infty[}$ under μ.

Guidelines.

Step 1: Set $\Sigma_0 = \left\{ \bigcup_{l=1}^L S_l \mid L \geq 1, S_l's \text{ are disjoint cylinder sets of } (\mathbb{R}^d)^{[0,\infty[} \right\}$. Note that S is a cylinder set if there exists finitely-many time-stamps $0 \leq t_1 \leq t_2 \leq ... \leq t_k$, and there exist $B_1, ..., B_k \in \mathcal{B}(\mathbb{R}^d)$ Borel sets such that $S = \bigcap_{j=1}^k \tau_{t_j}^{-1}(B_j)$, i.e.,

$$S = \{f \in (\mathbb{R}^d)^{[0,\infty[} \mid f(t_j) \in B_j \ \forall\ 1 \leq j \leq k\}.$$

It's straightforward to check that Σ_0 is an algebra and $\sigma(\Sigma_0) = \Sigma_{\mathbb{R}^d}^{[0,\infty[}$. Define μ_0 on Σ_0 such that if $S_l = \bigcap_{j=1}^{k_l} \tau_{t_j^{(l)}}^{-1}(B_j^{(l)})$ for some $0 \leq t_1^{(l)} \leq ... \leq t_{k_l}^{(l)}$ and $B_1^{(l)}, ..., B_{k_l}^{(l)} \in \mathcal{B}(\mathbb{R}^d)$. Define:

$$\mu_0 \left(\bigcup_{l=1}^L S_l \right) := \sum_{k=l}^L \gamma_{0,C(t_1^{(l)},...,t_{k_l}^{(1)})}(B_1^{(l)} \times B_2^{(l)} \times ... \times B_{k_l}^{(l)}). \tag{68}$$

Recall,

$$C(t_l^{(l)}, ..., t_{k_l}^{(l)}) = (t_{j \wedge j'}^{(l)} I_{d \times d})_{j,j'=1,...,k}.$$

Verify that μ_0 is consistent on Σ_0 and μ_0 is additive.

Step 2: Show that μ_0 is continuous at $\emptyset$.

By Caratheodory's Extension Theorem, we know that μ_0 can be uniquely extended to a measure μ on $\Sigma_{\mathbb{R}^d}^{[0,\infty[}$ such that if $f = \{f(t) \mid t \geq 0\} \in (\mathbb{R}^d)^{[0,\infty[}$ is sampled under μ, then f is a standard B.M.

Task 2. Give a rigorous proof of **Kolmogorov's Continuity Theorem**.

Theorem 16 (Kolmogorov's Continuity Theorem)**.** *Suppose that $T > 0$ and $\{X_t \mid t \in [0,T]\}$ is a stochastic process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X_t \in \mathbb{R}^d$ for all $t \in [0,T]$. If there exists a $p \geq 1$, $C > 0$, and $r > 0$ such that:*

$$(\mathbb{E}[|X_t - X_s|^p])^{1/p} \leq C|t-s|^{\frac{1}{p}+r}, \tag{69}$$

*for all $t, s \in [0,T]$ (this is called the **intrinsic metric**), then there exists a stochastic process $\{\tilde{X}_t \mid t \in [0,T]\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\tilde{X}$ is a modification of X, i.e., for every $t \in [0,T]$, $\tilde{X}_t = X_t$ almost surely, and for all $\alpha \in [0, r[$,*

$$\mathbb{E}\left[\sup_{0 \leq s,t \leq T} \frac{|\tilde{X}_t - \tilde{X}_s|}{|t-s|^\alpha} \right] < \infty. \tag{70}$$

In particular,

$$M_{\alpha,T} := \sup_{0, \leq s,t \leq T} \frac{|\tilde{X}_t - \tilde{X}_s|}{|t-s|^\alpha} < \infty \textit{ almost surely.} \tag{71}$$

In other words, almost surely $\{\tilde{X}_t \mid t \in [0,T]\}$ as a path is α-Holder continuous.

Intuition on α-Holder continuous:

1. If $\alpha = 1$, α-Holder continuity is the same as Lipschitz Continuous.
2. If $\dot{f}$ is bounded, then f is 1-Holder continuous.
   (a) We should view α as a guage which tells us how continuous f is; the higher α, the more continuous f is.

Corrolary 7. If $\{B_t \mid t \geq 0\}$ is a standard Brownian motion, then for every $\alpha \in \left]0, \frac{1}{2}\right[$ a.s. $\{B_t \mid t \geq 0\}$ as a path is α-Holder continuous.

Proof. For all $0 \leq s \leq t$, we have that $\mathbb{E}\left[|B_t - B_s|^2\right] = (t-s)d$. Since this is a Gaussian process, for all $n \in \mathbb{N}$, there exists a $C_{n,d} > 0$ such that

$$\mathbb{E}\left[|B_t - B_s|^{2n}\right] \leq C_{n,d}|t-s|^n.$$

(This follows since it's a fact about the Gaussian distribution). In this expression, apply **(Kolmogorov's Continuity Theorem)** . Set $p = 2n$, $r = \frac{1}{2} - \frac{1}{2n}$. Then, apply **(Kolmogorov's Continuity Theorem)** . For all $\alpha \in \left]0, \frac{1}{2} - \frac{1}{2n}\right[$,

$$\mathbb{E}\left[\sup_{0 \leq t,s \leq T} \frac{|\tilde{B}_t - \tilde{B}_s|}{|t-s|^\alpha}\right] < \infty,$$

where $\tilde{B}$ is a modification of B. Since n is arbitrary, $\frac{1}{2} - \frac{1}{2n}$ can be arbitrarily close to $\frac{1}{2} \Rightarrow \alpha \in \left]0, \frac{1}{2} - \frac{1}{2n}\right[$ can reach arbitrarily close to $\frac{1}{2}$. Hence, for all $\alpha \in \left]0, \frac{1}{2}\right[$, there exists a modification of Brownian motion $\tilde{B}$ of B such that $\tilde{B}$ is α-Holder continuous. WLOG, we assume that $\{B_t \mid t \geq 0\}$ is just that modification. $\square$

Theorem 17. *Let $\{B_t\}$ be a standard B.M. on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, $\forall\ \alpha > \frac{1}{2}$,*

$$\mathbb{P}\left(\exists\ S \geq 0 \ s.t. \ \limsup_{t \downarrow s} \frac{|B_t - B_s|}{|t-s|^\alpha} < \infty\right) = 0. \tag{72}$$

In particular, it implies that almost surely $\{B_t \mid t \geq 0\}$ is nowhere differentiable.

Proof. It's sufficient to focus on $\{B_t \mid t \in [0,1]\}$. Set:

$$E := \left\{\exists\ s \in [0,1] \text{ s.t. } \limsup_{t \downarrow s} \frac{|B_t - B_s|}{|t-s|^\alpha} < \infty\right\} \tag{73}$$

The goal is to show that $\mathbb{P}(E) = 0$. If E "happens", there exists an $\in [0,1]$, there exists a $\delta, A > 0$ such that for all $t \in]s, s+\delta[$,

$$|B_t - B_s| \leq A|t-s|^\alpha \Rightarrow \ \forall t, t' \in]s, s+\delta[, \ |B_t - B_{t'}| \leq A(|t-s|^\alpha + |t'-s|^\alpha)(*).$$

Given $n \geq 1$, divide $[0,1]$ into n segments. Given an $L \geq 1$, there exists an $N \geq 1$ such that for all $n \geq N$, there are at least $L+1$ nodes of $\frac{1}{n}\mathbb{Z}$ inside $]s, s+\delta[$. Choose $m \in \{0, 1, ..., n\}$ such that $\frac{m-1}{n} \leq s$ but $\forall$ $0 \leq l \leq L$, $\frac{m+l}{m} \in]s, s+\delta[$ (i.e., the first one is in the interval). Then, by (*), for all $0 \leq l \leq L-1$,

$$\left|B_{\frac{m+l+1}{n}} - B_{\frac{m+l}{n}}\right| \leq 2A\left(\frac{L+1}{n}\right)^\alpha \tag{74}$$

Therefore,

$$E \subseteq \bigcup_{A=1}^{\infty}\bigcup_{N=1}^{\infty}\bigcap_{n=N}^{\infty}\bigcup_{m=0}^{\infty}\bigcap_{l=0}^{L-1}\left\{\underbrace{\left|B_{\frac{m+l+1}{n}} - B_{\frac{m+l}{n}}\right| \leq 2A\left(\frac{L+1}{n}\right)^{\alpha}}_{:=E_{n,l}}\right\}$$

The $\left|B_{\frac{m+l+1}{n}} - B_{\frac{m+l}{n}}\right|$ has distribution $\gamma_{0,\frac{1}{n}I}$. The increments are independent, and so for all $A \geq 1$, for all $N \geq 1$,

$$\begin{aligned}\mathbb{P}\left(\bigcap_{n=N}^{\infty}\bigcup_{m=0}^{n}\bigcap_{l=0}^{L}E_{n,l}\right) &\leq \lim_{n\to\infty} n\left(\gamma_{0,\frac{1}{n}I}\left(B\left(0,2A\left(\frac{L+1}{n}\right)^{\alpha}\right)\right)\right)^{L} \\ &\leq \lim_{n\to\infty} C_{d,L,A,\alpha}n^{1+(1/2-\alpha)dL}\end{aligned}$$

Since $\alpha > \frac{1}{2}$, $\left(\frac{1}{2}-\alpha\right)$ is negative, and we can choose L large enough such that $(\alpha - 1/2)dL > 1 \Rightarrow$ the power of n is negative. Hence,

$$\lim_{n\to\infty} n^{1+(1/2-\alpha)dL} = 0 \Rightarrow \mathbb{P}(E) = 0.$$

□

Q: What if $\alpha = \frac{1}{2}$?
If $\alpha = \frac{1}{2}$, then in fact, for all $s > 0$,

$$\mathbb{P}\left(\limsup_{t\downarrow s}\frac{|B_t - B_s|}{\sqrt{t-s}} = +\infty\right) = 1,$$

which implies that there is no $\frac{1}{2}$-Holder Continuity.

Step # 3: Construct a Levy process associated with $\gamma_{0,C}$

Such a process is given by $\{\sqrt{C}B_t \mid t \geq 0\}$.

Given $\mu \in I(\mathbb{R}^d)$, assume that $\mu = \pi_{m,C,M}$ for some $m \in \mathbb{R}^d$, $C \geq 0$, $M \in \mathcal{M}_2$. We know how to construct a Levy Process associated with μ. For all $\xi \in \mathbb{R}^d$,

$$\hat{\mu}(\xi) = \exp\left(i(\xi,m) - \frac{1}{2}\underbrace{(\xi,C\xi)}_{C^tB_t} + \underbrace{\int_{\mathbb{R}^d}\left(e^{i(\xi,y)} - 1 - \chi_{B(0,1)}(y)i(\xi,y)\right)M(dy)}_{X_t^J}\right). \tag{75}$$

Given $(\Omega,\mathcal{F},\mathbb{P})$, let $\{B_t \mid t \geq 0\}$ be a standard Brownian Motion on Ω and $\{X_t^J \mid t \geq 0\}$ be a Levy Process associated with $\pi_{0,0,M}$ and $\{B_t \mid t \geq 0\}$ and $\{X_t^J \mid t \geq 0\}$ are independent. Then, for every $t \geq 0$,

$$X_t = \underbrace{mt}_{(1)} + \underbrace{\sqrt{C}B_t}_{(2)} + \underbrace{X_t^J}_{(3)}, \tag{76}$$

is a Levy process associated with $\mu = \pi_{m,C,M}$. We call the term (1) the **Linear Drift**, (2) **Diffusion**, and (3) the **Jump**.

Remarks.

- Let $j(t,\cdot,X)$ be the jump measure of $\{X_t \mid t \geq 0\}$. Then, for all $t \geq 0$,

$$j(t,\cdot,X) = j(t,\cdot,X^J).$$

- If $t \in [0,\infty[\to X_t \in \mathbb{R}^d$ is continuous, then $M = 0$ (i.e., there is no jump component).
- If $t \in [0,\infty[\to X_t \in \mathbb{R}^d$ has locally bounded variation, i.e., $||X||_{\text{var},[0,t]} < \infty$ for all $t \geq 0$ almost surely, then $M \in \mathcal{M}_1(\mathbb{R}^d)$ and $C = 0$) (no diffusion component).

2.3 Classical Wiener Measure

Given $(\Omega, \mathcal{F}, \mathbb{P})$, let B_t be a standard Brownian Motion, $\{B_t \mid t \geq 0\}$ on Ω. Then, for all $\omega \in \Omega$, $t \in [0,\infty[\mapsto B_t(\omega) \in \mathbb{R}^d$ is continuous, i.e., $B.(\omega) \in C([0,\infty[;\mathbb{R}^d) =: C([0,\infty[)$. View the mapping $B : \omega \in \Omega \mapsto B.(\omega) \in C([0,\infty[)$ as a random variable on Ω. then, the distribution of $B.$, $\mathcal{L}_B$, is a probability measure on $(C([0,\infty[), \Sigma_C^{[0,\infty[})$, where

$$\Sigma_C^{[0,\infty[} := \sigma(\{\tau_t \mid t \geq 0\}). \tag{77}$$

This measure is called the **classical Wiener Measure**, denoted by $\mathcal{W}$.

- If $\phi \in C([0,\infty[)$ is sampled under $\mathcal{W}$, then $\{\phi(t) \mid t \geq 0\}$ is a standard Brownian motion.
- If $0 \leq r \leq s \leq t$, then under $\mathcal{W}$, $\tau_t - \tau_s$ is independent of τ_r, and $\tau_t - \tau_s$ has distribution $\gamma_{0,(t-s)I}$.
- Note:

$$\begin{aligned}
&\mathcal{W}\left(\{\phi \in C([0,\infty[\})) \mid \phi(0) = 0\}\right) = 1, \\
&\mathcal{W}\left(\{\phi \in C([0,\infty[) \mid \phi \text{ is nowhere differentiable}\}\right) = 1.
\end{aligned}$$

Proposition 13 (Invariance of Brownian Motion)**.** Let $\{B_t \mid t \geq 0\}$ be a Brownian Motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, $\mathcal{W}$ is invariant under the following transformations.

Brownian Motion Perspective	Wiener Measure Perspective
1. **Scaling**. For all $c \geq 0$, $\left\{\frac{1}{\sqrt{c}} B_{ct} \mid t \geq 0\right\}$ is a standard Brownian motion. (Check it by writing the characteristic function).	1. **Scaling**. For every $C > 0$, $S_c : \varphi \in C([0,\infty[) \to S_c(\varphi) \in C([0,\infty[)$ such that $\forall\ t \geq 0$, $S_c(\varphi)(t) = \frac{1}{\sqrt{c}}\varphi(ct)$.
2. **Time-Inversion**. If $\tilde{B}_0 = 0$ and $\tilde{B}_t = tB_{\frac{1}{t}}$, for all $t \geq 0$, then $\tilde{B}_t$ is a standard Brownian motion	2. **Time-Inversion**. Let $I : \varphi \mapsto I(\varphi)$ such that $I(\varphi)(0) \equiv 0$, then $I(\varphi)(t) = t\varphi(t)$ for all $t \geq 0$.
3. **Time-Reversal**. Given $T > 0$, $\{B_T - B_{T-t} \mid t \in [0,T]\}$ is a standard B.M. on $[0,T]$.	3. **Time-Reversal**
4. **Symmetry**. $\{-B_t \mid t \geq 0\}$ is again a standard B.M.	4. **Symmetry**

Time-inversion is an important property since it helps us link behaviour near time zero to large time behaviour.

3 Continuous Time-Martingales

Definition 20 (Filtration / Right Continuous / Left Continuous)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

- A collection of sigma algebras $\{\mathcal{F}_t \mid t \geq 0\}$ is a **filtration** if for all t, $\mathcal{F}_t$ is a sub-sigma-algebra of the underlying σ-algebra, and for all $0 \leq s \leq t$, $\mathcal{F}_s \subseteq \mathcal{F}_t$.
- Given a filtration $\{\mathcal{F}_t \mid t \geq 0\}$, we set:

  1. $\mathcal{F}_{t^+} := \bigcap_{s>t} \mathcal{F}_s$ for all $t \geq 0$ ("looking into the immediate future")
  2. $\mathcal{F}_{0^-} := \mathcal{F}_0$ and $\mathcal{F}_{t^-} = \sigma\left(\bigcup_{s<t} \mathcal{F}_s\right)$ ("the immediate past"). Recall that the union might not be a sigma-algebra, so we need to generate it.

- $\{\mathcal{F}_t \mid t \geq 0\}$ is **right-continuous** if $\mathcal{F}_t = \mathcal{F}_{t^+}$ if for all t. In words: the immediate future is the same as the present.
- $\{\mathcal{F}_t \mid t \geq 0\}$ is **left-continuous** if $\mathcal{F}_t = \mathcal{F}_{t^-}$ for all t. In words: the immediate past is the same as the present.

Definition 21 (Adapted to a Filtration and Progressively Measurable)**.** Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t \mid t \geq 0\}, \mathbb{P})$ be a filtered space and $\{X_t \mid t \geq 0\}$ a stochastic process on Ω.

- We say that $\{X_t \mid t \geq 0\}$ is **adapted with respect to a filtration** if $X_t \in m\mathcal{F}_t$ for all $t > 0$.
- $\{X_t \mid t \geq 0\}$ is **progressively measurable** with respect to $\{\mathcal{F}_t \mid t \geq 0\}$ if for all $t \geq 0$, the function $(s, \omega) \in [0, T] \times \Omega \mapsto X_t(\omega) \in \mathbb{R}^d$ is measurable with respect to $\mathcal{B}([0, T]) \times \mathcal{F}_t$.

Remark. If $\{X_t \mid t \geq 0\}$ is adapted, then $\int_0^t |X_s| ds$ may NOT be measurable with respect to $\mathcal{F}_t$. If $\{X_t \mid t \geq 0\}$ is progressively measurable, then $\int_0^t |X_s| ds \in m\mathcal{F}_t$ (by Fubini's Theorem).

Proposition 14. Let $\{X_t \mid t \geq 0\}$ be a stochastic process with RCLL paths. Then, if $\{X_t \mid t \geq 0\}$ is adapted with respect to some filtration $\{\mathcal{F}_t \mid t \geq 0\}$, then it is progressively measurable with respect to $\{\mathcal{F}_t \mid t \geq 0\}$.

Proof. For all $t \geq 0$, for all $(s, \omega) \in [0, t] \times \Omega$, $X_s(\omega) = \lim_{n\to\infty} X_{\lceil s \rceil_n}(\omega)$, where $\lceil s \rceil_n = m2^{-n}t$ if $s \in](m-1)2^{-n}t, m2^{-n}t]$. It suffices to show that for every $n \geq 1$ $(s, \omega) \mapsto X_{\lceil s \rceil_n}(\omega)$ is measurable with respect to $\mathcal{B}([0, t[) \times \mathcal{F}_t$. However, for all $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\{(s, \omega) \in [0, t[\times \Omega \mid X_{\lceil s \rceil_n}(\omega) \in B\} = \{0\} \times X_0^{-1}(B) \bigcup \left[\bigcup_{m=1}^{2^n}](m-1)2^{-n}t, m2^{-n}t] \times X_{m2^{-n}t}^{-1}(B)\right]$$
$$\in \mathcal{B}([0, t[) \times \mathcal{F}_t.$$

□

Definition 22 (Natural Filtration)**.** Let $\{X_t \mid t \geq 0\}$ be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, for every $t \geq 0$, the **natural filtration** associated to the process $\{X_t \mid t \geq 0\}$ is:

$$\mathcal{F}_t^s := \sigma(\{X_s \mid s \in [0, t]\}). \tag{78}$$

Hence, $\{\mathcal{F}_t^X \mid t \geq 0\}$ is the natural filtration associated to $\{X_t \mid t \geq 0\}$ and $\{X_t \mid t \geq 0\}$ is adapted to its natural filtration.

Next homework: If $\{X_t \mid t \geq 0\}$ is continuous, then $\{\mathcal{F}_t^X \mid t \geq 0\}$ is left-continuous.

Definition 23 (Stopping and Optional Times)**.** Given a filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_t \mid t \geq 0\}, \mathbb{P})$, let $\tau : \Omega \to [0, \infty]$ be a random variable.

1. τ is a **stopping time** with respect to a filtration if, for all $t \geq 0$, $\{\tau \leq t\} \in \mathcal{F}_t$.
2. τ is an **optional time** with respect to a filtration if, for all $t \geq 0$, $\{\tau < t\} \in \mathcal{F}_t$.

Proposition 15. We have the following two relations between optional times and stopping times.

1. If τ is a stopping time with respect to the filtration $\{\mathcal{F}_t \mid t \geq 0\}$, then τ is also an optional time with respect to $\{\mathcal{F}_t \mid t \geq 0\}$.
2. τ is an optional time with respect to $\{\mathcal{F}_t \mid t \geq 0\}$ $\iff$ τ is a stopping time with respect to $\{\mathcal{F}_{t^+} \mid t \geq 0\}$.

Proof. 1. For all $t \geq 0$,

$$\{\tau < t\} = \bigcup_{n=1}^{\infty} \underbrace{\left\{\tau \leq t - \frac{1}{n}\right\}}_{\in \mathcal{F}_{t-\frac{1}{n}} \subseteq \mathcal{F}_t} \in \mathcal{F}_t, \tag{79}$$

which shows that τ is an optional time.

2. "⇒": Write:

$$\{\tau \leq t\} = \bigcap_{n=1}^{\infty} \underbrace{\{\tau < t + \frac{1}{n}\}}_{\in \mathcal{F}_{t+\frac{1}{n}}} \text{ (τ is an optional time).} \tag{80}$$

For all $s > t$, there exists an N sufficiently large such that for all $n \geq N$, $t + \frac{1}{n} < s$. Hence, we have the following set equality:

$$\bigcap_{n=1}^{\infty} \left\{\tau < t + \frac{1}{n}\right\} = \bigcap_{n=N}^{\infty} \underbrace{\left\{\tau < t + \frac{1}{n}\right\}}_{\in \mathcal{F}_{t+\frac{1}{n}} \subseteq \mathcal{F}_s} \in \mathcal{F}_s.$$

Hence, $\{\tau \leq t\} \in \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_{t^+}$.

"⇐": for all $t > 0$, $\{\tau < t\} = \bigcup_{n=1}^{\infty} \underbrace{\left\{\tau \leq t - \frac{1}{n}\right\}}_{\in \mathcal{F}_{\left(t-\frac{1}{n}\right)^+}}$.

For all $n \geq 1$,

$$\mathcal{F}_{\left(t-\frac{1}{n}\right)^+} = \bigcap_{s > t - \frac{1}{n}} \mathcal{F}_s = \bigcap_{t-\frac{1}{n} < s \leq t} \mathcal{F}_s \subseteq \mathcal{F}_t.$$

This shows that $\{\tau < t\} \in \mathcal{F}_t$. ☐

Proposition 16. Properties about stopping times:

1. If τ_1 and τ_2 are two stopping times with respect to $\{\mathcal{F}_t \mid t \geq 0\}$, then $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$, $\tau_1 + \tau_2$ are all stopping times with respect to the same filtration.
2. If $\{\tau_n \mid n \geq 1\}$ is a sequence of stopping times with respect to $\{\mathcal{F}_t\}$, then so is $\sup_n \tau_n$.
3. Given any $t > 0$ (constant), $\tau \wedge t$ and $\tau \vee t$ are both stopping times.

Proof. Proof that $\tau_1 + \tau_2$ is a stopping time: the proof strategy is to insert a rational number. For all $t \geq 0$, one has:

$$\begin{aligned}
\{\tau_1 + \tau_2 > t\} &= \{\tau_1 > t\} \cup \{\tau_1 \leq t, \tau_1 + \tau_2 > t\} \\
&= \{\tau_1 > t\} \cup \{\tau_2 > t - \tau_1\} \\
&= \underbrace{\{\tau_1 > t\}}_{\in \mathcal{F}_t} \cup \bigcup_{q \in \mathbb{Q} \cap [0,t]} \underbrace{\{q < \tau_1 < t\}}_{\in \mathcal{F}_t} \cap \underbrace{\{\tau_2 > t - q\}}_{\mathcal{F}_{t-q} \subseteq \mathcal{F}_t} \in \mathcal{F}_t.
\end{aligned}$$

Rest: exercise. ☐

Definition 24. Let $\{X_t \mid t \geq 0\}$ be a progressive, measurable stochastic process, and let $B \in \mathcal{B}(\mathbb{R}^d)$. Set $h_b : \Omega \to [0, \infty]$ by the following rule: for all $\omega \in \Omega$,

$$h_b(\omega) := \inf\{t \geq 0 \mid X_t(\omega) \in B\}. \tag{81}$$

h_b is called the **hitting time of the set** B.

Homework problem.

1. If $\{X_t \mid t \geq 0\}$ is RCLL and B is open, then h_b is an optional time with respect to $\{\mathcal{F}_t\}$.
2. If $\{X_t \mid t \geq 0\}$ is continuous and if B is closed, then h_b is a stopping time with respect to $\{\mathcal{F}_t\}$.

Definition 25 (Sigma Algebra Associated with a Stopping Time)**.** Given a stopping time τ, define the following collection of sets:

$$\mathcal{F}_\tau := \{A \in \mathcal{F} \mid A \cap \{\tau \leq t\} \in \mathcal{F}_t \ \forall\ t \geq 0\}. \tag{82}$$

$\mathcal{F}_\tau$ is the σ-algebra **associated** with τ.

Proposition 17. If τ_1 and τ_2 are two stopping times with respect to the same filtration $\{\mathcal{F}_t \mid t \geq 0\}$, then:

1. $\tau_1 \leq \tau_2 \Rightarrow \mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.
2. $\mathcal{F}_{\tau_1 \wedge \tau_2} = \mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2}$.
3. $\{\tau_1 < \tau_2\}, \{\tau_1 > \tau_2\}, \{\tau_1 = \tau_2\} \in \mathcal{F}_{\tau_1 \wedge \tau_2}$.
4. For all $t \geq 0$, $\forall A \in \mathcal{F}_{\tau_1}$, $A \cap \{\tau_1 \leq t\} \in \mathcal{F}_{\tau_1 \wedge t}$.

Definition 26. If $\{X_t \mid t \geq 0\}$ is progressively measurable with respect to $\{\mathcal{F}_t \mid t \geq 0\}$ and τ is a stopping time with respect to $\{\mathcal{F}_t\}$ for all $\omega \in \Omega$, then define:

$$X_\tau(\omega) := \chi_{\{\tau < \infty\}}(\omega) \cdot X_{\tau(\omega)}(\omega) + \chi_{\{\tau = \infty\}}(\omega)\chi_{\{\lim_{\tau \to \infty} X_\tau \text{ exists}\}}(\omega) \lim_{\tau \to \infty} X_\tau(\omega). \tag{83}$$

Proposition 18. We have the following properties about X_τ:

1. X_τ is measurable with respect to $\mathcal{F}_t$.
2. the **stopped process** $\{X_{t \wedge \tau} \mid t \geq 0\}$ is progressively measurable with respect to $\{\mathcal{F}_t\}$.

Proof.

- Proof of (ii): for all $c > 0$, consider:

  $$\{(s, \omega) \in [0, t] \times \Omega \mid \tau(\omega) \wedge s \leq c\} = [0, t] \times \{\omega \in \Omega \mid \tau(\omega) \leq c\} \cup ([0, t \wedge c] \times \Omega) \in \mathcal{B}([0, t]) \times \mathcal{F}_t.$$

  This shows that $(s, \omega) \in [0, t] \times \Omega \mapsto \tau(\omega) \wedge s$ is measurable with respect to $\mathcal{B}([0, t]) \times \mathcal{F}_t$. Hence, $(s, \omega) \in [0, t] \times \Omega \mapsto X_{\tau(\omega) \wedge s}(\omega)$ is measurable with respect to $\mathcal{B}([0, t]) \times \mathcal{F}_t$.

- Proof of (i): for all $B \in \mathcal{B}(\mathbb{R}^d)$, $t \geq 0$,

  $$\underbrace{\{X_\tau \in B\} \cap \{\tau \leq t\}}_{\tau \wedge t = \tau} = \underbrace{\{X_{\tau \wedge t} \in B\}}_{\in \mathcal{F}_t} \cap \underbrace{\{\tau \leq t\}}_{\in \mathcal{F}_t} \in \mathcal{F}_t.$$

  This shows that $\{X_t \in B\} \in \mathcal{F}_t$.

□

Example 3. Let $\{B_t \mid t \geq 0\}$ be a one-dimensional standard B.M. on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\{\mathcal{F}_t \mid t \geq 0\}$ be the natural filtration of $\{B_t \mid t \geq 0\}$. Given $a > 0$, set $\tau_a := \inf\{t \geq 0 \mid B_t \geq a\}$ (the hitting time of the B.M. of $[a, \infty[$. Then, τ_a is a stopping time. We can investigate the distribution of τ_a. Set $M_t := \max_{s \in [0,t]} B_s$ to be the **running maximum of Brownian Motion**. Then, for all $t \geq 0$, we have the following set inclusion:

$$\{\tau_a \leq t\} = \left\{\max_{s \in [0,t]} B_s \geq a\right\} \tag{84}$$

Then, since the B.M. has continuous sample paths,

$$M_t := \lim_{n \to \infty} \max_{0 \leq m \leq 2^n} B_{m2^{-n}t}.$$

For all $\varepsilon > 0$, we have:

$$\mathbb{P}\left(\max_{0 \leq m \leq 2^{-n}} B_{m2^{-n}t} \geq a\right) \leq \mathbb{P}\left(M_t \geq a\right).$$

Define the following set $E_a^{(n)} := \max_{0 \leq m \leq 2^{-n}} B_{m2^{-n}t} \geq a$ and $E_{a-\varepsilon}^{(n)} := \max_{0 \leq m \leq 2^{-n}} B_{m2^{-n}t} \geq a - \varepsilon$. Then, we have the following bounds on $\mathbb{P}\left(M_t \geq a\right)$:

$$\lim_{n \to \infty} \mathbb{P}\left(E_a^{(n)}\right) \leq \mathbb{P}\left(M_t \geq a\right) \leq \lim_{n \to \infty} \mathbb{P}\left(E_{a-\varepsilon}^{(n)}\right). \tag{85}$$

Write:

$$\begin{aligned}
\mathbb{P}\left(B_t \geq a - \varepsilon\right) &= \mathbb{P}\left(\{B_t \geq a - \varepsilon\} \cap E_{a-\varepsilon}^{(n)}\right) \\
&= \sum_{m=1}^{2^n} \mathbb{P}\left(B_t \geq a - \varepsilon, B_{j2^{-n}t} < a - \varepsilon, j = 0, ..., m+1, B_{m2^{-n}t} \geq a - \varepsilon\right) \\
&\geq \sum_{m=0}^{2^n} \mathbb{P}\left(B_t - B_{m2^{-n}t} \geq 0, ..., B_{m2^{-n}t} \geq a - \varepsilon\right) \\
&= \sum_{m=0}^{2^n} \mathbb{P}\left(B_t - B_{m2^{-n}t} \geq 0\right) \cdot \mathbb{P}\left(...\right) \text{ (by Independence)} \\
&= \frac{1}{2} \sum_{m=0}^{2^n} \mathbb{P}\left(E_{a-\varepsilon}^{(n)}\right).
\end{aligned}$$

Hence,

$$\mathbb{P}\left(E_{a-\varepsilon}^{(n)}\right) \leq 2\mathbb{P}\left(B_t > a - \varepsilon\right) \Rightarrow \mathbb{P}\left(M_t \geq a\right) \leq 2\mathbb{P}\left(B_t \geq a - \varepsilon\right).$$

This takes care of the upper bound. To achieve a lower bound, set:

$$A_\varepsilon^{(n)} := \bigcap_{n=1}^{2^n} \{|B_{m2^{-n}t} - B_{(m-1)2^n t}| \leq \varepsilon\}.$$

It's easy to check that $\lim_{n\to\infty} \mathbb{P}\left(A_\varepsilon^{(n)}\right) = 1$ (cut it into sufficiently small pieces). Consider the set:

$$\begin{aligned}
\mathbb{P}\left(\{B_t \geq a+\varepsilon\} \cap A_\varepsilon^{(n)}\right) &= \mathbb{P}\left(\{B_t \geq a+\varepsilon\} \cap A_\varepsilon^{(n)} \cap E_a^{(n)}\right) \\
&= \sum_{m=0}^{2^n} \mathbb{P}\left(B_t \geq a+\varepsilon, A_\varepsilon^{(n)}, B_{j2^{-n}t} \leq a, j=0,\dots,m-1, B_{m2^{-n}t} \geq a\right) \\
&= \sum_{m=0}^{2^n} \mathbb{P}\left(B_t - B_{m2^{-n}t} \geq 0, \dots,\right) \\
&= \frac{1}{2}\mathbb{P}\left(E_a^{(n)}\right)
\end{aligned}$$

Hence, we get that

$$\mathbb{P}\left(E_a^{(n)}\right) \geq 2\left(\mathbb{P}\left(B_t \geq a+\varepsilon\right) - \mathbb{P}\left((A_\varepsilon^{(n)})^c\right)\right) \to 2\mathbb{P}\left(B_t \geq a+\varepsilon\right) \text{ as } n \to \infty.$$

This gives us the following bound:

$$2\mathbb{P}\left(B_t \geq a+\varepsilon\right) \leq \mathbb{P}\left(M_t \geq a\right) \leq 2\mathbb{P}\left(B_t \geq a-\varepsilon\right).$$

Sending $\varepsilon \to 0$, we obtain the desired result, which is called the **reflecting principle**.

$$\mathbb{P}\left(M_t \geq a\right) = \mathbb{P}\left(\tau_a \leq t\right) = 2\mathbb{P}\left(B_t \geq a\right). \tag{86}$$

Task: adapt the rigorous procedure used to determine the distribution of M_t to determine the joint distribution of (M_t, B_t).

Instruction: Consider, for every $t \geq 0$, $\forall\ a > 0$, for all $b < a$, $\varepsilon > 0$ small, consider:

$$\mathbb{P}\left(M_t > a, B_t \geq b+\varepsilon\right) \text{ and } \mathbb{P}\left(M_t \geq a, B_t \geq b-e\right).$$

Further, verify that $M_t - B_t \underbrace{=}_{(d)} M_t \underbrace{=}_{(d)} |B_t|$.

We can now introduce continuous time martingales.

Definition 27 (Martingale)**.** Let $\{X_t \mid t \geq 0\}$ be a stochastic process on some filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be progressively measurable with respect to $\{\mathcal{F}_t\}$ and for all $t \geq 0$, $X_t \in \mathbb{R}$, and $X_t \in L^1(\mathbb{P})$.

1. X_t is called a **martingale** with respect to $\{\mathcal{F}_t\}$ if for all $t \geq s \geq 0$:
$$\mathbb{E}\left[X_t \mid \mathcal{F}_s\right] = X_s, \tag{87}$$
i.e., it's a fair game.
2. X_t is called a **submarginale** with respect to $\{\mathcal{F}_t\}$ if for all $t \geq s \geq 0$, $\mathbb{E}\left[X_t \mid \mathcal{F}_s\right] \geq X_s$.
3. X_t is called a **supermartingale** with respect to $\{\mathcal{F}_t\}$ if $\{-X_t \mid t \geq 0\}$ is a sub-martingale.

Remark. If $X_t \in \mathbb{R}^d$ for all $t \geq 0$, then $\{X_t \mid t \geq 0\}$ is a martingale with respect to some filtration if every component is a martingale: for $X_t = (X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(d)})$, the component $\{X_t^{(j)} \mid t \geq 0\}$ is a martingale for all $1 \leq j \leq d$.

Example 4. (Examples of Continuous-Time Martingales) Let $\mu \in I(\mathbb{R}^d)$, $\mu = \pi_{m,C,M}$ be a Levy System. Let $\{X_t \mid t \geq 0\}$ be a progressively-measurable process on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$. Then, we have the following "**Martingale Characterization of a Levy Process**": $\{X_t \mid t \geq 0\}$ is a Levy process associated with $\pi_{m,C,M} \iff X_0 = 0$ and for all $\xi \in \mathbb{R}^d$, for all $t \geq 0$,

$$\{E_t(\xi) := \exp\left(i\langle \xi, X_t\rangle - t\ell_{\hat{\mu}}(\xi)\right) \mid t \geq 0\},$$

is a martingale with respect to $\{\mathcal{F}_t \mid t \geq 0\}$.

Proof. "$\Rightarrow$": for all $t \geq 0$,

$$\mathbb{E}\left[E_t(\xi)\right] = \mathbb{E}\left[E_0(\xi)\right] = 1 \Rightarrow \mathbb{E}\left[e^{i\langle \xi, X_1\rangle}\right] = e^{i\ell_{\hat{\mu}}(\xi)}$$

Now to show its a Levy process: for all $0 \leq r \leq s \leq t$, for all $\xi_1, \xi_2 \in \mathbb{R}^d$,

$$\begin{aligned}\mathbb{E}\left[e^{i\langle \xi_1, X_r\rangle} e^{i\langle \xi_2, X_t - X_s\rangle}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{i\langle \xi_1, X_r\rangle} e^{i\langle \xi_2, X_t - X_s\rangle} | \mathcal{F}_s\right]\right] \\ &= \mathbb{E}\left[e^{i\langle \xi_1, X_r\rangle} e^{i\langle \xi_2, X_s\rangle} \mathbb{E}\left[e^{i\langle \xi_2, X_t\rangle} | \mathcal{F}_s\right]\right] \\ &= \mathbb{E}\left[e^{i\langle \xi_1, X_r\rangle}\right] e^{(t-s)\ell_{\hat{\mu}}(\xi_2)} \\ &= e^{r\ell_{\hat{\mu}}(\xi_1)} e^{(t-s)\ell_{\hat{\mu}}(\xi_2)}\end{aligned}$$

"$\Rightarrow$": Take $\{X_t \mid t \geq 0\}$ to be a Levy process associated with μ. For all $0 \leq s \leq t$, for all $\xi \in \mathbb{R}^d$:

$$\begin{aligned}\mathbb{E}\left[e^{i\langle X_t, \xi\rangle} | \mathcal{F}_s\right] &= \mathbb{E}\left[e^{i\langle X_t - X_s, \xi\rangle} | \mathcal{F}_s\right] e^{i\langle X_s, \xi\rangle} \text{ (by independent increments)} \\ &= e^{(t-s)\ell_{\hat{\mu}}(\xi)} e^{i\langle X_s, \xi\rangle}.\end{aligned}$$

□

Example 5. Let $\{B_t \mid t \geq 0\}$ be a standard B.M. on $\mathbb{R}^d$. Then, it's easy to check:

1. B.M. itself is a martingale with respect to the natural filtration, $\{\mathcal{F}_t^B \mid t \geq 0\}$.
2. $\forall \xi \in \mathbb{R}^d$, $\{\langle B_t, \xi\rangle^2 - t||\xi||^2 \mid t \geq 0\}$ is a martingale.
3. For all $\xi \in \mathbb{R}^d$, $\{e^{\langle \xi, B_t\rangle + \frac{1}{2}t||\xi||^2} \mid t \geq 0\}$ is a martingale.
4. For all $\xi \in \mathbb{R}^d$, $\{e^{\langle \xi, B_t\rangle - \frac{1}{2}t||\xi||^2} \mid t \geq 0\}$ is a martingale.

Remark. This gives us multiple characterizations of B.M. Assume that $\{B_t \mid t \geq 0\}$ satisfies that $B_0 \equiv 0$, $t \in [0, \infty[\to B_t$ is continuous. Then, B_t is a standard B.M. $\iff$ (3) is true $\iff$ (4) is true $\iff$ (1) and (2) are true.

Proposition 19 (Jensen's Inequality)**.** Let $\{X_t \mid t \geq 0\}$ be an $\mathbb{R}^d$-valued martingale (or a sub-martingale with $d = 1$) on some filtered space and let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a convex function (or a convex non-decreasing function when $d = 1$) such that $\varphi(X_t) \in L^1(\mathbb{P})$ for all $t \geq 0$. Then, $\{\varphi(X_t) \mid t \geq 0\}$ is a sub-martingale with respect to $\{\mathcal{F}_t\}$.

The proof is exactly the same as in the discrete case.

Remarks.

- If $\{X_t \mid t \geq 0\}$ is a martingale, then $\{||X_t||^p \mid t \geq 0\}$ is a martingale for all $p \geq 1$.
- If $\{X_t \mid t \geq 0\}$ is a martingale (or a sub-martingale) and $\{t_n \mid n \geq 0\} \subseteq [0, \infty[$ such that $t_n \uparrow \infty$ as $n \to \infty$, then $\{X_{t_n} \mid n \geq 0\}$ is a martingale/sub-martingale with respect to $\{\mathcal{F}_{t_n} \mid n \geq 0\}$.

We will now generalize the results from the discrete-case from MATH 587 to the continuous case.

Theorem 18 (Martingale Convergence Theorem I)**.** *Let $\{X_t \mid t \geq 0\}$ be a 1D (sub)-martingale with RCLL sample paths and* $\sup_{t \geq 0} \mathbb{E}\left[X_t^+\right] < \infty$*. Then, there exists an $X_\infty \in L^1$ such that $X_t \to X_\infty$ a.s. as $t \to \infty$.*

Proof. We need to adapt Doob's upcrossing inequality. Given $a, b \in \mathbb{R}$, $a < b$, let $U_{[a,b]}$ be the total number of upcrossings from a to b completed by $\{X_t \mid t \geq 0\}$. For every $n \geq 0$, consider $\{Y_m = X_{m2^{-n}} \mid m \geq 0\}$.

Then, $\{Y_m \mid m \geq 0\}$ is a (sub)-martingale. Set $U_{n,[a,b]}$ to be the total number of upcrossings from a to b completed by $\{Y_m \mid m \geq 0\}$. Then, by the RCLL property, $U_{n,[a,b]} \uparrow U_{[a,b]}$ as $n \to \infty$. By **(MON)** :

$$\begin{aligned}\mathbb{E}\left[U_{[a,b]}\right] &= \mathbb{E}\left[U_{n,[a,b]}\right] \\ &\leq \lim_{n\to\infty} \sup_{m\geq 0} \frac{\mathbb{E}\left[(Y_m - a)^+\right]}{b-a} \text{ (By \textbf{(Discrete Doob's Upcrossing Inequality)})} \\ &\leq \sup_{t\geq 0} \frac{\mathbb{E}\left[X_t^+\right] + |a|}{b-a} \\ &< \infty.\end{aligned}$$

This implies that for all $a < b$, $U_{[a,b]} < \infty$ almost surely which implies that $X_\infty = \lim_{t\to\infty} X_t$ exists almost surely. Finally, we see that

$$\mathbb{E}\left[|X_t|\right] = 2\mathbb{E}\left[X^+\right] - \mathbb{E}\left[X_t\right] \leq 2\mathbb{E}\left[X_t^+\right] - \mathbb{E}\left[X_0\right].$$

This implies that $\sup_t \mathbb{E}\left[|X_t|\right] < \infty$ i.e. $\{X_t \mid t \geq 0\}$ is bounded in L^1. Finally, this implies:

$$\mathbb{E}\left[|X_\infty|\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[|X_t|\right] < \infty.$$

□

Theorem 19 (Martingale Convergence II)**.** *If $\{X_t \mid t \geq 0\}$ is a 1D (sub)-martingale with RCLL sample paths and $\{X_t \mid t \geq 0\}$ is **uniformly integrable**, i.e.,*

$$\lim_{A\to\infty} \sup_{t\geq 0} \mathbb{E}\left[|X_t|; |X_t| > A\right] = 0, \tag{88}$$

then there exists an $X_\infty \in L^1(\mathbb{P})$ such that $X_t \to X_\infty$ a.s. and in $L^1(\mathbb{P})$.

Proof. If $\{X_t \mid t \geq 0\}$ is uniformly integrable, then it's easy to see that $\{X_t \mid t \geq 0\}$ is bounded in L^1. Hence, there exists an $X_\infty \in L^1(\mathbb{P})$ such that $X_t \to X_\infty$ almost surely. Before proceeding, an important note of caution:

Remark. *in discrete time settings or for a discrete family, convergence in probability and uniform integrability $\iff$ convergence in L^1. However, this is no longer true in the continuous-time setting, in general. But here it will work.*

For a contradiction, assume that X_t does not converge to X_∞ in L^1. Then, there exists a sequence $\{S_n \mid n \geq 0\} \subseteq \mathbb{N} \uparrow \infty$ as $n \to \infty$ such that X_{S_n} does not converge to X_∞ in L^1 (choose a sub-sequence). $\{X_{S_n} \mid n \geq 0\}$ is uniformly integrable, and $X_{S_n} \to X_\infty$ as $n \to \infty$. However, with the discrete time-result, we know that $X_{S_n} \to X_\infty$ in L^1, which is a contradiction. □

Proposition 20 (Doob's Maximal Inequality)**.** Let $\{X_t \mid t \geq 0\}$ be a 1D (sub)-martingale with RCLL sample paths. Then, for all $\varepsilon > 0$, for all $T > 0$,

$$\mathbb{P}\left(\sup_{t\in[0,T]} X_t > \varepsilon\right) \leq \frac{1}{\varepsilon}\mathbb{E}\left[X_T; \sup_{t\in[0,T]} X_t > \varepsilon\right]. \tag{89}$$

Further, if $X_t \geq 0$ for all $t \geq 0$ and for some $p > 1$, $X_t \in L^p$, then:

$$\mathbb{E}\left[\left(\sup_{t\in[0,T]} X_t\right)^p\right]^{1/p} \leq \frac{p}{p-1}\left(\mathbb{E}\left[X_T^p\right]\right)^{1/p}. \tag{90}$$

Proof. By RCLL, we have that $Y_m = X_{m2^{-n}T}$ for $m = 0, ..., 2^n$. Then,

$$\left\{ \sup_{t\in[0,T]} X_t > \varepsilon \right\} = \bigcup_{n=0}^{\infty} \left\{ \max_{0\leq m\leq 2^n} Y_m > \varepsilon \right\}.$$

Therefore,

$$\begin{aligned}
\mathbb{P}\left(\sup_{[0,T]} X_t > \varepsilon \right) &= \lim_{n\to\infty} \mathbb{P}\left(\max_{0\leq m\leq 2^n} Y_m > \varepsilon \right) \\
\lim_{n\to\infty} &\leq \frac{1}{\varepsilon}\mathbb{E}\left[Y_{2^n}; \max_{0\leq m\leq 2^n} Y_m > \varepsilon \right] \text{ (By } \textbf{(Discrete Doob's Max Ineq)} \text{)} \\
&= \frac{1}{\varepsilon}\mathbb{E}\left[X_T; \sup_{t\in[0,T]} X_t > \varepsilon \right] \text{ (By } \textbf{(MON)} \text{)}
\end{aligned}$$

□

Corrolary 8. Let $\{X_t \mid t \geq 0\}$ be a martingale or a non-negative sub-martingale with RCLL sample paths such that $\{X_t \mid t \geq 0\}$ is uniformly integrable. Then, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\sup_{t\geq 0} |X_t| > \varepsilon \right) \leq \frac{1}{\varepsilon}\mathbb{E}\left[|X_\infty|; \sup_{t\geq 0} |X_t| > \varepsilon \right]. \tag{91}$$

Proof. For all $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\geq 0} |X_t| > \varepsilon \right) &= \lim_{T\to\infty} \mathbb{P}\left(\sup_{t\in[0,T]} |X_t| > \varepsilon \right) \\
&\leq \lim_{T\to\infty} \frac{1}{\varepsilon}\mathbb{E}\left[|X_T|; \underbrace{\sup_{t\in[0,T]} |X_t| > \varepsilon}_{\in\mathcal{F}_t} \right] \\
&\leq \lim_{T\to\infty} \lim_{T<s\to\infty} \frac{1}{\varepsilon}\mathbb{E}\left[|X_s|; \sup_{t\in[0,T]} |X_t| > \varepsilon \right] \text{ (since } |X_t| \text{ is a sub-martingale)} \\
&= \lim_{T\to\infty} \frac{1}{\varepsilon}\mathbb{E}\left[|X_\infty|; \sup_{t\in[0,T]} |X_t| > \varepsilon \right] \text{ (since } |X_t| \to |X_\infty| \text{ in } L^1(\mathbb{P})) \\
&= \frac{1}{\varepsilon}\mathbb{E}\left[|X_\infty|; \sup_{t\geq 0} |X_t| > \varepsilon \right].
\end{aligned}$$

□

Theorem 20 (Martingale Convergence Theorem III)**.** *Let $\{X_t \mid t \geq 0\}$ be a martingale (or a non-negative sub-martingale) on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ with RCLL sample paths. If for some $p > 1$, $\sup_{t\geq 0} \mathbb{E}[|X_t|^p] < \infty$ (i.e., $\{X_t \mid t \geq 0\}$ is bounded in L^p), then there exists an $X_\infty \in L^p$ such that $X_t \to X_\infty$ almost surely and in L^p.*

Proof. By **(Martingale CV Theorem II)** , there exists an $X_\infty \in L^1$ such that $X_t \to X_\infty$ a.s. Hence, by Fatou's Lemma,

$$\mathbb{E}[|X_\infty|^p] \leq \liminf_{t\to\infty} \mathbb{E}[|X_t|^p] < \infty \Rightarrow X_\infty \in L^p.$$

For every $T > 0$, we have:

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^p\right] &\leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|X_T|^p\right]\\ &\leq \left(\frac{p}{p-1}\right)^p \sup_{t\geq 0}\mathbb{E}\left[|X_t|^p\right]\\ &< \infty.\end{aligned}$$

This proves that $\sup_{t\in[0,T]} |X_t|^p \in L^p$. Sending $T \to \infty$, by **(MON)** ,

$$\mathbb{E}\left[(\sup_{t\geq 0} |X_t|)^p\right] = \lim_{T\to\infty} \mathbb{E}\left[(\sup_{[0,T]} |X_t|)^p\right] < \infty.$$

This shows that $\sup_{t\geq 0} |X_t| \in L^p$. Next, by **(DOM)** , we get that $\mathbb{E}\left[|X_t - X_\infty|^p\right] \to 0$ as $t \to \infty$. This follows from the fact that $X_t \to X_\infty$ almost surely and we have the following dominating function:

$$|X_t - X_\infty|^p \leq C_p \left[\sup_{t\geq 0} |X_t|^p + |X_\infty|^p\right].$$

□

Next, we want to establish the stopping time theorems.

Proposition 21. Let $\{X_t \mid t \geq 0\}$ be a martingale (or a non-negative sub-martingale) on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ with RCLL sample paths. (1) Then, for all $T > 0$,

$$\{X_\tau \mid \tau \text{ is a stopping time and } \tau \leq T\},$$

is uniformly integrable. (2) Furthermore, if $\{X_t \mid t \geq 0\}$ is uniformly integrable, then

$$\{X_\tau \mid \tau \text{ is a stopping time such that } \tau < \infty \text{ a.s. }\},$$

is uniformly integrable.

Proof. (1) Given some $n \geq 1$, define $Y_m := X_{m2^{-n}}$ and $\mathcal{G}_m := \mathcal{F}_{m2^{-n}}$ for $m = 0, 1, 2, \ldots$. Then, $\{Y_m \mid m \geq 0\}$ is a martingale (or a non-negative sub-martingale) with respect to $\{\mathcal{G}_m \mid m \geq 0\}$. Fix a $T > 0$. WLOG, assume that $T \in \mathbb{N}$. Let τ be a stopping time such that $\tau \leq T$, and we set $k(\tau) := m + 1$ if $\tau \in [m2^{-n}, (m+1)2^{-n}[$ for $m \geq 0$. Then:

$$\{k(\tau) = m + 1\} = \{\tau \in [m2^{-n}, (m+1)2^{-n}[\} \in \mathcal{F}_{(m+1)2^{-n}} =: \mathcal{G}_{m+1}.$$

Hence, $k(\tau)$ is a discrete stopping time with respect to $\{\mathcal{G}_m \mid m \geq 0\}$. In addition,

$$Y^{(n)}_{k(\tau)} := X_{2^{-n}([2^n\tau]+1)} \to X_\tau \text{ as } n \to \infty,$$

because $2^{-n}([2^n\tau]+1) \to \tau$ from the right. Recall **Discrete Hunt's Theorem**: if $\tau_1 \leq \tau_2 \leq M$ for some $M > 0$, then $\mathbb{E}\left[X_{\tau_2}|\mathcal{F}_{\tau_1}\right] = (\geq) X_{\tau_1}$. Since $k(\tau) \leq (\tau+1)2^n$, by discrete-time Hunt's theorem, we can get:

$$\begin{aligned}|Y^{(n)}_{k(\tau)}| &\leq \mathbb{E}\left[|Y^{(n)}_{2^n(T+1)}| \mid \mathcal{G}_{k(\tau)}\right]\\ &= \mathbb{E}\left[|X_{T+1}| \mid \mathcal{G}_{k(\tau)}\right].\end{aligned}$$

Since $\{Y^{(n)}_{k(\tau)} > A\} \in \mathcal{G}_{k(\tau)}$, for every $A > 0$:

$$\begin{aligned}\mathbb{E}\left[|Y^{(n)}_{k(\tau)}|\ ; |Y^{(n)}_{k(\tau)}| > A\right] &\leq \mathbb{E}\left[|X_{T+1}|; |Y^{(n)}_{k(\tau)}| > A\right] \\ &\leq \mathbb{E}\left[|X_{T+1}|; \sup_{t\in[0,T+1]} |X_t| > A\right]\end{aligned}$$

Sending $n \to \infty$ from the LHS and applying Fatou's Lemma, we obtain:

$$\begin{aligned}\mathbb{E}\left[|X_\tau|; |X_\tau| > A\right] &\leq \liminf_{n\to\infty} \mathbb{E}\left[|Y^{(n)}_{k(\tau)}|\ ;\ |Y^{(n)}_{k(\tau)}| > A\right] \\ &\leq \mathbb{E}\left[|X_{T+1}|; \sup_{t\in[0,T+1]} |X_t| > A\right].\end{aligned}$$

Hence, taking the supremum over all stopping times τ such that $\tau \leq T$ (call this set W), we obtain the following inequality:

$$\sup_W \mathbb{E}\left[|X_\tau|; |X_\tau| > A\right] \leq \mathbb{E}\left[|X_{T+1}|; \sup_{t\in[0,T+1]} |X_t| > A\right].$$

Break the expected value up:

$$\mathbb{E}\left[|X_{T+1}|; \sup_{t\in[0,T+1]} |X_t| > A\right] \leq \mathbb{E}\left[|X_{T+1}|; |X_{T+1}| > \sqrt{A}\right] + \mathbb{E}\left[|X_{T+1}|; |X_{T+1}| \leq \sqrt{A} \text{ and } \sup_{t\in[0,T+1]} |X_t| > A\right]$$

For $\mathbb{E}\left[|X_{T+1}|; |X_{T+1}| > \sqrt{A}\right]$, we have that it goes to zero as $A \to \infty$.
For $\mathbb{E}\left[|X_{T+1}|; |X_{T+1}| \leq \sqrt{A} \text{ and } \sup_{t\in[0,T+1]} |X_t| > A\right]$, we have that:

$$\begin{aligned}\mathbb{E}\left[|X_{T+1}|; |X_{T+1}| \leq \sqrt{A} \text{ and } \sup_{t\in[0,T+1]} |X_t| > A\right] &\leq \sqrt{A}\,\mathbb{P}\left(\sup_{t\in[0,T+1]} |X_t| > A\right) \\ &\leq \sqrt{A}\frac{1}{A}\mathbb{E}\left[|X_{T+1}|\right] \to 0 \text{ as } A \to \infty.\end{aligned}$$

Hence, we have proven that by taking A sufficiently large,

$$\lim_{A\to\infty} \sup_{\tau\in W} \mathbb{E}\left[|X_\tau|; |X_\tau| > A\right] = 0.$$

Next, assume that $\{X_t \mid t \geq 0\}$ is uniformly integrable. Then, there exists an $X_\infty \in L^1$ such that $X_t \to X_\infty$ a.s. and in $L^1(\mathbb{P})$. Then, given any stopping time τ such that $\tau < \infty$ a.s., and any $T > 0$ and $A > 0$,

$$\mathbb{E}\left[|X_{\tau\wedge T}|; |X_{\tau\wedge T}| > A\right] \leq \mathbb{E}\left[|X_{T+1}|\ ; \sup_{t\geq 0} |X_t| > A\right],$$

which follows from the previous statements since $\tau \wedge T \leq T$. Sending $T \to \infty$, by Fatou's Lemma,

$$\sup_{\tau \ \mid\ \tau \text{ is a finite s.t.}} \mathbb{E}\left[|X_\tau|; |X_\tau| > A\right] \leq \mathbb{E}\left[|X_\infty|; \sup_{t\geq 0} |X_t| > A\right].$$

We can follow the same steps as in the above to show that:

$$\lim_{A\to\infty} \mathbb{E}\left[|X_\infty|; \sup_{t\geq 0} |X_t| > A\right] = 0.$$

□

Theorem 21 (Hunt's Theorem)**.** *Let $\{X_t \mid t \geq 0\}$ be a martingale (or a non-negative sub-martingale) with RCLL paths. Let τ_1 and τ_2 be two stopping times such that $\tau_1 \leq \tau_2$.*

1. *If there exists a $T > 0$ such that $\tau_1 \leq \tau_2 \leq T$, then*

$$X_{\tau_1} = (\leq)\mathbb{E}\left[X_{\tau_2} \mid \mathcal{F}_{\tau_1}\right]$$

2. *If $\{X_t \mid t \geq 0\}$ is uniformly integrable, and τ_1 and τ_2 are finite almost surely, then*

$$X_{\tau_1} = (\leq)\mathbb{E}\left[X_{\tau_2} \mid \mathcal{F}_{\tau_1}\right].$$

Proof.

1. For $n \geq 1$, let $Y_m^{(n)}$, $\mathcal{G}_m$, $m \geq 0$, $k(\tau_j)$ for $j = 1, 2, \ldots$ be the same as in the previous proof. Applying **(Discrete Hunt's Theorem)** ,

$$Y_{k(\tau_1)}^{(n)} = (\leq)\mathbb{E}\left[Y_{k(\tau_2)}^{(n)} \mid \mathcal{G}_{k(\tau_1)}\right].$$

i.e.,

$$X_{2^{-n}([2^n\tau_1]+1)} = (\leq)\mathbb{E}\left[X_{2^{-n}([2^n\tau_2]\ +1)} | \mathcal{F}_{2^{-n}([2^n\tau_1]+1)}\right].$$

Hence,

$$\begin{aligned}\mathbb{E}\left[X_{2^{-n}([2^n\tau_2]+1)}|\mathcal{F}_{\tau_1}\right] &= \mathbb{E}\left[\mathbb{E}\left[X_{2^{-n}([2^n\tau_2]+1)}\right]|\mathcal{F}_{\tau_1}\right] \\ &= (\leq)\mathbb{E}\left[X_{2^{-n}([2^n\tau_1]+1)}|\mathcal{F}_{\tau_1}\right] \text{ (by the Tower Property)}\end{aligned}$$

As $n \to \infty$, for $j = 1, 2$, we get that

$$2^n([2^n\tau_j] + 1) \to \tau_j,$$

for $j = 1, 2$ almost surely (this convergence happens pointwise). By the previous proposition, we get that the $\{X_{2^{-n}([2^n\tau_j]+1)} \mid j = 1, 2, n \geq 0\}$ is uniformly integrable. Hence, this implies that

$$X_{2^{-n}([2^n\tau_j]+1)} \to X_{\tau_j} \text{ in } L^1(\mathbb{P}).$$

Thus,

$$\mathbb{E}\left[X_{\tau_2} \mid \mathcal{F}_{\tau_1}\right] = (\geq)X_{\tau_1}.$$

2. Now assume that $\{X_t \mid t \geq 0\}$ is uniformly integrable with $\tau_1 \leq \tau_2 < \infty$ a.s. For any $T > 0$ (WLOG, assume $T \in \mathbb{N}$), by **(i)** we have that:

$$X_{\tau_1 \wedge T} = (\leq)\mathbb{E}\left[X_{\tau_2 \wedge T} | \mathcal{F}_{\tau_1 \wedge T}\right].$$

For all $B \in \mathcal{F}_{\tau_1}$,

$$\mathbb{E}\left[X_{\tau_2 \wedge T}; \underbrace{B \cap \{\tau_1 \leq T\}}_{\in \mathcal{F}_{\tau_1 \wedge T}}\right] = (\geq)\mathbb{E}\left[X_{\tau_1 \wedge T}; B \cap \{\tau_1 \leq T\}\right].$$

As $T \to \infty$, $X_{\tau_j \wedge T} \to X_{\tau_j}$ $(j = 1, 2)$ a.s. and in $L^1(\mathbb{P})$ (due to the uniform integrability and by the previous proposition). For $j = 1, 2$,

$$\left|\mathbb{E}\left[X_{\tau_j \wedge T}; B \cap \{\tau_j \leq T\}\right] - \mathbb{E}\left[X_{\tau_j}; B\right]\right| \leq \mathbb{E}\left[|X_{\tau_j \wedge T} - X_{\tau_j}|\right] + \mathbb{E}\left[|X_{\tau_j}|; B \cap \{\tau_j > T\}\right]$$

□

Theorem 22 (Doob's Stopping Time Theorem)**.** *Let $\{X_t \mid t \geq 0\}$ be a martingale or a non-negative sub-martingale on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ with RCLL sample paths. then, for every stopping time τ (w.r.t. $\{\mathcal{F}_t \mid t \geq 0\}$), the process $\{X_{t\wedge t} \mid t \geq 0\}$ is a martingale or a non-negative sub-martingale w.r.t. $\{\mathcal{F}_t\}$.*

Proof. For all $0 \leq s \leq t < \infty$, for all $B \in \mathcal{F}_s$, $B \cap \{\tau > s\} \in \mathcal{F}_{s\wedge\tau}$. Hence, for τ, $s \wedge \tau \leq t \wedge \tau$. Since we have two stopping times, both of which are bounded by t, **(Hunt's Theorem)** ,

$$\mathbb{E}\left[X_{t\wedge\tau} | \mathcal{F}_{s\wedge\tau}\right] = (\geq) X_{s\wedge\tau}.$$

Hence, we write:

$$\begin{aligned}
\mathbb{E}\left[X_{t\wedge\tau}; B\right] &= \mathbb{E}\left[X_{t\wedge\tau}; B \cap \{\tau \leq s\}\right] + \mathbb{E}\left[X_{t\wedge\tau}; \underbrace{B \cap \{\tau > s\}}_{\in \mathcal{F}_{s\wedge t}}\right] \\
&= (\geq)\mathbb{E}\left[X_\tau; B \cap \{\tau \leq s\}\right] + \mathbb{E}\left[X_{s\wedge\tau}; B \cap \{\tau > s\}\right] \\
&= \mathbb{E}\left[X_{s\wedge\tau}; B\right].
\end{aligned}$$

□

Examples of Hunt's Theorem

Example 6. Let $\{B_t \mid t \geq 0\}$ be a standard one-dimensional B.M. on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$. For $a, b > 0$, set:

- $T_a := \inf\{t \geq 0 \mid B_t \geq a\}$ (the hitting time of the set $[a, \infty[$).
- $T_{-b} := \inf\{t \geq 0 \mid B_t \leq -b\}$ (hitting time of the set $]-\infty, -b]\}$.

Define the set $T := T_a \wedge T_{-b}$. All of these are stopping times. T is the exiting time of the interval $]-b, a[$. Then, we can write Ω as the disjoint union of the following sets:

$$\Omega = \{T = T_a\} \cup \{T = T_{-b}\} = \{T_a < T_{-b}\} \cup \{T_a > T_{-b}\}. \tag{92}$$

By **(Doob's Stopping Time Theorem)** , $\{B_{t\wedge T} \mid t \geq 0\}$ is again a martingale, and for every $t \geq 0$, $|B_{t\wedge T}| \leq a \vee b$ which implies that $\{B_{t\wedge T} \mid t \geq 0\}$ is uniformly integrable.

- If $t < T$, then B_t hasn't left $]-b, a[\Rightarrow |B_t| \leq a \vee b$.
- If $t \geq T$, then $B_{t\wedge T} = B_T = a$ or $-b$.

By **(Hunt's Theorem)** ,

$$\mathbb{E}\left[B_{T\wedge T}\right] = \mathbb{E}\left[B_{0\wedge T}\right].$$

Hence, $\mathbb{E}\left[B_T\right] = 0$, and so

$$\mathbb{E}\left[B_T\right] = a\mathbb{P}\left(T = T_a\right) - b\mathbb{P}\left(T = T_b\right) = 0,$$

and $\mathbb{P}\left(T = T_a\right) + \mathbb{P}\left(T = T_b\right) = 1$. This implies:

$$\begin{aligned}
\mathbb{P}\left(T = T_a\right) &= \frac{b}{a+b} \\
\mathbb{P}\left(T = T_{-b}\right) &= \frac{a}{a+b}.
\end{aligned}$$

Next: next, we try to compute $\mathbb{E}[T]$.

$$\{B_t^2 - t \mid t \geq 0\} \tag{93}$$

is a martingale, which implies that

$$\{B_{t\wedge T}^2 - (t \wedge T) \mid t \geq 0\}$$

is also a martingale. This implies that for all $t \geq 0$,

$$\mathbb{E}[t \wedge T] = \mathbb{E}\left[B_{t\wedge T}^2\right],$$

LHS: $\lim_{t\to\infty} \mathbb{E}[t \wedge T] = \mathbb{E}[T]$ by **(MON)** . On the RHS,

$$\lim_{t\to\infty} \mathbb{E}\left[B_{t\wedge T}^2\right] = \mathbb{E}\left[B_T^2\right] = a^2 \mathbb{P}(T = T_a) + b^2 \mathbb{P}(T = T_{-b}) = ab.$$

The interchanging of the limit and the expected value is justified by **(DOM)** , with dominating function $\sup_{t\geq 0} B_{t\wedge T}^2 = a^2 \vee b^2 < \infty$. Hence,

$$\mathbb{E}[T] = ab.$$

Exercise. Suppose $a = b$. Determine $\mathbb{E}\left[e^{-\alpha T}\right]$ for $\alpha \geq 0$. (This will also determine the distribution of T). **Hint**: Consider / look for a martingale that involves e^{cB_t} for some $c \in \mathbb{R}$.

Prove the following version of the decomposition theorem (**Doob Meyer's Decomposition Theorem**): Let $\{X_t \mid t \geq 0\}$ be a non-negative sub-martingale on some filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ with RCLL sample paths. Then, there exist two stochastic processes $\{M_t \mid t \geq 0\}$ and $\{A_t \mid t \geq 0\}$ such that for all $t \geq 0$,

$$X_t = A_t + M_t. \tag{94}$$

1. $A_0 \equiv 0$, $\{A_t \mid t \geq 0\}$ is increasing (i.e., $A_s \leq A_t$ for all $s \leq t$) and $\{A_t \mid t \geq 0\}$ is predictable (i.e., $A_t \in \mathcal{F}_{t-}$, the equivalent notion of this was "pre-visible" in the discrete case) and $\{A_t \mid t \geq 0\}$ has RCLL sample paths.
2. $\{M_t \mid t \geq 0\}$ is a martingale with respect to the filtration $\{\mathcal{F}_t \mid t \geq 0\}$ with RCLL paths.

Furthermore, the choice of $\{A_t \mid t \geq 0\}$ and $\{M_t \mid t \geq 0\}$ is unique. In addition,

1. If $\{X_t \mid t \geq 0\}$ has continuous sample paths, then so do $\{A_t \mid t \geq 0\}$ and $\{M_t \mid t \geq 0\}$.
2. If $\{X_t \mid t \geq 0\}$ is uniformly integrable, then so is $\{M_t \mid t \geq 0\}$ and $\sup_{t\geq 0} \mathbb{E}[A_t] < \infty$.

Next, we will apply Meyer's Decomposition Theorem to square-integrable martingales.

Definition 28. Let $\{X_t \mid t \geq 0\}$ be a square integrable martingale on some filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ (so, $X_t \in L^2$) with RCLL sample paths. then, by **(Jensen)** 's, $\{X_t^2 \mid t \geq 0\}$ is a non-negative sub-martingale. Applying Meyer's Decomposition Theorem,there exists a process $\{A_t \mid t \geq 0\}$ with $A_0 \equiv 0$ increasing and predictable such that $\{X_t^2 - A_t \mid t \geq 0\}$ is a martinale. We denote by $\langle X \rangle_t = A_t$ and call $\{\langle X \rangle_t \mid t \geq 0\}$ to be the **quadratic variation** of $\{X_t \mid t \geq 0\}$.

Remark 1. We have that $\langle X \rangle_0 \equiv 0$, $t \mapsto \langle X \rangle_t$ is increasing and RCLL, and for all $t \geq 0$, $\mathbb{E}[\langle X \rangle_t] = \mathbb{E}\left[X_t^2\right]$. If $\{X_t \mid t \geq 0\}$ is bounded in L^2, then $\langle X \rangle_\infty = \lim_{t\to\infty} \langle X \rangle_t \in L^1$ and $\langle X \rangle_t \mapsto \langle X \rangle_\infty$ a.s. and in L^1.

Example 7. If $\{B_t \mid t \geq 0\}$ is a standard B.M., then $\langle B\rangle_t = t$ for all $t \geq 0$. If $\{N_t \mid t \geq 0\}$ is a simple Poisson Process with rate 1, then $\{N_t - t \mid t \geq 0\}$ is a square-integrable martingale. Let's find another martingale: all $0 \leq s \leq t$, we have:

$$\begin{aligned}
\mathbb{E}\left[(N_t - t)^2 \mid \mathcal{F}_s\right] &= \mathbb{E}\left[(N_t - N_s - (t-s) + N_s - s)^2 \mid \mathcal{F}_s\right] \\
&= \mathbb{E}\left[(N_t - N_s - (t-s))^2\right] + 2\mathbb{E}\left[(N_t - N_s - (t-s))\right](N_s - s) + (N_s - s)^2 \\
&= (t-s) + (N_s - s)^2.
\end{aligned}$$

Re-arranging shows that $\{(N_t - t)^2 - t \mid t \geq 0\}$ is a martingale. Hence, if $X = N_t - t$, then $\langle X\rangle_t = t$.

Corrolary 9. Let X be a square integrable martingale with continuous sample paths. Then, X is a B.M. $\iff$ $X_0 = 0$ and $\langle X\rangle_t = t$ for all $t \geq 0$.

Theorem 23. *Let $\{X_t \mid t \geq 0\}$ be a square integrable martingale on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ with continuous sample paths. Suppose that for all $\omega \in \Omega$, $t \mapsto \langle X\rangle_t(\omega)$ is strictly increasing, and $\lim_{t\to\infty}\langle X\rangle_t(\omega) = \infty$. Then, there exists a stochastic process $\{B_s \mid s \geq 0\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\{\mathcal{G}_s \mid s \geq 0\}$ with $\mathcal{G}_s \subseteq \mathcal{F}$ for all $s \geq 0$ such that $\{B_s \mid s \geq 0\}$ is a standard Brownian Motion with respect to $\{\mathcal{G}_s \mid s \geq 0\}$ and $X_t = X_0 + B_{\langle X\rangle_t}$ for all $t \geq 0$, i.e., for all $\omega \in \Omega$,*

$$X_t(\omega) = X_0(\omega) + B_{\langle X\rangle_t(\omega)}(\omega). \tag{95}$$

Roughly speaking, this means that there is a "clock" such that the process is a shifted Brownian Motion.

Proof. WLOG, set $X_0 = 0$. By assumption, for all $\omega \in \Omega$, $t \mapsto \langle X\rangle_t(\omega)$ is strictly increasing and continuous $\Rightarrow$ there exists an $s \mapsto S_s(\omega)$ inverse function of $\langle X\rangle(\omega)$ such that $s \mapsto S_s(\omega)$ and $S_s(\omega) \uparrow \infty$ and for all $t \geq 0$, $S_s(\omega) = t \iff \langle X\rangle_t(\omega) = s$. Moreover, for every $t \geq 0$, $s \geq 0$,

$$\{\omega \mid S_s(\omega) \leq t\} = \{\omega \mid \langle X\rangle_t(\omega) \leq s\} \in \mathcal{F}_t.$$

Next we define the filtration: for every $s \geq 0$, set $\mathcal{G}_s := \mathcal{F}_{S_s}$. Since $S_s \leq S_{s'}$ for every $s \leq s'$, we have that $\mathcal{G}_s \subseteq \mathcal{G}_{s'}$ which shows that $\{\mathcal{G}_s \mid s \geq 0\}$ is a filtration. For every $\omega \in \Omega$, for every $s \geq 0$, set:

$$B_s(\omega) := X_{S_s(\omega)}(\omega). \tag{96}$$

Then, for all $s \geq 0$, B_s is measurable with respect to $\mathcal{F}_{S_s} = \mathcal{G}_s$. This implies that $\{B_s \mid s \geq 0\}$ is adapted with respect to $\{\mathcal{G}_s \mid s \geq 0\}$. OTOH, $s \mapsto B_s$ is continuous, so from adaptedness we get that $\{B_s \mid s \geq 0\}$ is progressively measurable with respect to $\{\mathcal{G}_s \mid s \geq 0\}$. By definition, for every $t \geq 0$, for every $\omega \in \Omega$,

$$X_t(\omega) = B_{\langle X\rangle_t(\omega)}(\omega). \tag{97}$$

We only need to check that $\{B_s \mid s \geq 0\}$ is a B.M. now. By continuity of the sample paths, it suffices to show that $\{B_s \mid s \geq 0\}$ and $\{B_s^2 - s \mid s \geq 0\}$ are both martingales. First, observe that

$$\lim_{T\to\infty} X_{T\wedge S_s} = X_{S_s} = B_s,$$

for every $\omega \in \Omega$. Then, by **(MON)** ,

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,S_s]} X_t^2\right] &= \lim_{T\to\infty} \mathbb{E}\left[\sup_{t\in[0,T\wedge S_s]} X_t^2\right] \\
&\leq 4 \lim_{T\to\infty} \mathbb{E}\left[X_{T\wedge S_s}^2\right] \\
&\leq 4 \lim_{T\to\infty} \mathbb{E}\left[X_{T\wedge S_s}^2\right] \\
&= 4 \lim_{T\to\infty} \mathbb{E}\left[\langle X\rangle_{T\wedge S_s}\right] \\
&\leq 4s.
\end{aligned}$$

This shows that $\sup_{t\in[0,S_s]} X_t^2 \in L^1(\mathbb{P})$. This shows that $X_{T\wedge S_s} \to B_s$ and in L^2 by **(DOM)** . By using similar arguments as in the proof of Hunt's Theorem, we can show that for every $s \le t$, for all $A \in \mathcal{F}_s$,

$$\mathbb{E}\left[B_s; A\right] = \mathbb{E}\left[B_t; A\right]$$
$$\mathbb{E}\left[B_s^2 - s; A\right] = \mathbb{E}\left[B_t^2 - t; A\right].$$

□

By applying the LIL to B.M., we obtain the following Corollary:

Corrolary 10. Let $\{X_t \mid t \ge 0\}$ be the same as the above. Then,

$$\limsup_{t\to\infty} \frac{X_t}{\sqrt{2\langle X\rangle_t \ln\ln\langle X\rangle_t}} = 1 = -\liminf_{t\to\infty} \frac{X_t}{\sqrt{2\langle X\rangle_t \ln\ln\langle X\rangle_t}} \text{ a.s.}$$

Definition 29 (Cross-Variation)**.** Let $\{X_t \mid t \ge 0\}$ and $\{Y_t \mid t \ge 0\}$ be two square integrable martingales on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ with RCLL sample paths. Suppose that for all $t \ge 0$, $X_t, Y_t \in \mathbb{R}$. For every $t \ge 0$, set:

$$\langle X, Y\rangle_t := \frac{1}{4}(\langle X+Y\rangle_t - \langle X-Y\rangle_t). \tag{98}$$

This is called the **cross-variation** of the processes $\{X_t\}$ and $\{Y_t\}$.

We have the following remarks about cross-variation.

- $\langle X, X\rangle_t = \langle X\rangle_t$.
- $\langle X, Y\rangle_t = \langle Y, X\rangle_t$.
- If $\{Z_t \mid t \ge 0\}$ is another square integrable martingale with RCLL sample paths, and for all $\alpha, \beta \in \mathbb{R}$, then:

  $$\langle \alpha X + \beta Z, Y\rangle_t = \alpha\langle X, Y\rangle_t + \beta\langle Z, Y\rangle_t,$$

  i.e., $\langle \cdot, *\rangle$ is bilinear.
- $\langle X, Y\rangle_0 \equiv 0$.
- $t \mapsto \langle X, Y\rangle_t$ has locally bounded total variation, i.e., for all $t \ge 0$, $||\langle X, Y\rangle||_{\text{var},[0,t]} < \infty$.

Proposition 22. $\{X_t \cdot Y_t - \langle X, Y\rangle_t \mid t \ge 0\}$ is a martingale with RCLL sample paths.

Proof. For all $0 \le s \le t$, write:

$$\mathbb{E}\left[X_tY_t \mid \mathcal{F}_s\right] = \frac{1}{4}\mathbb{E}\left[(X_t+Y_t)^2 - (X_t-Y_t)^2 \mid \mathcal{F}_s\right]$$

This is true, by the following:

$$\mathbb{E}\left[(X_t+Y_t)^2 - \langle X+Y\rangle_t \mid \mathcal{F}_s\right] = (X_s+Y_s)^2 - \langle X+Y\rangle_s$$
$$\mathbb{E}\left[(X_t-Y_t)^2 - \langle X-Y\rangle_t \mid \mathcal{F}_s\right] = (X_s-Y_s)^2 - \langle X-Y\rangle_s$$

$$= \frac{1}{4}\left((X_s+Y_s)^2 - \langle X+Y\rangle_s - ((X_s-Y_s)^2 - \langle X-Y\rangle_s)\right) + \frac{1}{4}\mathbb{E}\left[\langle X+Y\rangle_t - \langle X-Y\rangle_t | \mathcal{F}_s\right]$$
$$= X_sY_s - \langle X, Y\rangle_s + \mathbb{E}\left[\langle X, Y\rangle_t \mid \mathcal{F}_s\right],$$

which proves the desired statement. □

3.1 Probabilistic Approach to PDEs!

Theorem 24 (Another Martingale Characterization of Brownian Motion)**.** *Let* $\{X_t \mid t \geq 0\}$ *be a progressively measurable stochastic process on some filtered space* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ *with continuous sample paths. Suppose that* $X_t \in \mathbb{R}^d$ *for all* $t \geq 0$ *and that* $X_0 \equiv 0$*. Then,* $\{X_t \mid t \geq 0\}$ *is a standard Brownian Motion* $\iff$ *for all* $\varphi \in C_b^2(\mathbb{R}^d)$*,*

$$\left\{E_t^{\varphi} := \varphi(X_t) - \frac{1}{2}\int_0^t \Delta\varphi(X_s)ds \mid t \geq 0\right\},$$

is a martingale.

Proof. "$\Leftarrow$": We can use the previous martingale characterization of B.M.: for all $j = 1, ..., d$, let $e_j = (0, 0, ..., 0, 1, 0, ..., 0)$, where the 1 is in the jth component. Write $X_t = (X_t^1, ..., X_t^d)$. Consider $\varphi_1(X) := \langle X, e_j \rangle$ for all $X \in \mathbb{R}^d$. Then, $\Delta\varphi_1 \equiv 0$. Then, $\{X_t^j \mid t \geq 0\}$ is a martingale.

Consider $\varphi_2(X) := \langle X, e_j \rangle^2$ for all $X \in \mathbb{R}^d$. We have that $\Delta\varphi_2 = 2$. Then, $\{(X_t^j)^2 - t \mid t \geq 0\}$ is a martingale. By one of the martingale characterizations of a B.M., this shows that $\{X_t^j \mid t \geq 0\}$ is a standard 1D Brownian motion.

"$\Rightarrow$": We will prove this direction for $\varphi \in C_c^{\infty}(\mathbb{R}^d)$. The statement for general $\varphi \in C_b^2(\mathbb{R}^d)$ can be established following an approximation argument. Assume that $\{X_t \mid t \geq 0\}$ is a standard B.M. Then, for every $t > 0$, set φ_t to be the density function of $\gamma_{0,TI}$. By direct computation, we have

$$\frac{d}{dt}\left(\phi * \rho_t\right)(x) = \frac{1}{2}(\Delta\phi * \rho_t)(x). \tag{99}$$

One can see this by writing

$$\phi * \rho_t(x) := \left(\frac{1}{\sqrt{2\pi t}}\right)^d \int_{\mathbb{R}^d} \phi(x-y)\text{exp}\left(-\frac{||y||^2}{2t}\right) dy.$$

Now observe that, for all $0 \leq s \leq t$, one has that for X_t a B.M.,

$$\mathbb{E}\left[\phi(X_t) \mid \mathcal{F}_s\right] = \mathbb{E}\left[\phi(X_t - X_s + X_s) \mid \mathcal{F}_s\right]$$

$X_t - X_s$ is independent of $\mathcal{F}_s$ by independent increments, and X_s is measurable with respect to $\mathcal{F}_s$. Hence,

$$\begin{aligned} &= \int_{\mathbb{R}^d} \phi(y + X_s)\rho_{t-s}(y)dy \\ &= \int_{\mathbb{R}^d} \phi(X_s - y)\rho_{t-s}(y)dy \\ &= (\phi * \rho_{t-s})(X_s). \end{aligned}$$

Therefore, by the Fundamental Theorem of Calculus, we get that:

$$\mathbb{E}\left[\phi(X_t) \mid \mathcal{F}_s\right] = (\phi * \rho_{t-s})(X_s) = \int_s^t \frac{1}{2}\Delta\phi * \rho_{r-s}(X_s)dr + \phi(X_s).$$

Therefore, for every $A \in \mathcal{F}_s$, we may write:

$$\begin{aligned} \mathbb{E}\left[\phi(X_t); A_t\right] &= \mathbb{E}\left[\phi(X_s); A\right] + \int_s^t \frac{1}{2}\mathbb{E}\left[\Delta\phi * \rho_{t-s}(X_s); A\right] dr \\ &= \mathbb{E}\left[\phi(X_s); A\right] + \int_s^t \frac{1}{2}\mathbb{E}\left[\Delta\phi(X_r); A\right] dt \\ &= \mathbb{E}\left[\phi(X_s); A\right] + \mathbb{E}\left[\frac{1}{2}\int_s^t \Delta\phi(X_r)dr; A\right]. \end{aligned}$$

The final equality follows from an application of Fubini's Theorem. This shows that $\left\{\phi(X_t) - \int_0^t \frac{1}{2}\Delta\phi(X_r)dr \mid t \geq 0\right\}$ is a martingale. ☐

Q: Can we do more? Yes.

Task. Prove the following **Martingale Characterization of a Levy Process**.

Theorem 25. *Assume that $\mu = \pi_{m,C,M} \in I(\mathbb{R}^d)$ with some Levy System (m, C, M) and let $\{X_t \mid t \geq 0\}$ be a progressively measurable process on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ with RCLL sample paths. Then, $\{X_t \mid t \geq 0\}$ is a Levy Process associated with $\mu \iff$ for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ if*

$$Z_t^\phi := \phi(X_t) - \int_0^t A^\mu\phi(X_s)ds \tag{100}$$

for all $t \geq 0$, where for all $x \in \mathbb{R}^d$,

$$A^\mu\phi(X) := \frac{1}{2}\sum_{i,j=1}^d C_{ij}\partial_i\partial_j\phi(X) + \sum_{j=1}^d m_j\partial_j\phi(X) +$$

$$\int_{\mathbb{R}^d}\left(\phi(X+y) - \phi(X) - \chi_{B(0,1)}(y)\sum_{j=1}^d y_j\partial_j\phi(X)\right)M(dy).$$

Use the proof of the Levy-Khinchine Formula in the proof, and recall that the **generator of the process** $\{X_t \mid t \geq 0\}$ is given by:

$$A^\mu\phi(0) := \lim_{n\to\infty} n\left(\int \phi(y)\mu_{\frac{1}{n}}(dy) - \phi(0)\right) \tag{101}$$

Corrolary 11. Let $D \subseteq \mathbb{R}^d$ be a bounded, open set, with sufficiently regular (e.g. smooth) boundary ∂D. Let $f : \partial D \to \mathbb{R}$ be continuous and bounded, and let $u_f : D \to \mathbb{R}$ be the **harmonic extension** of f, i.e., $u_f \in C^2(D)$, and

$$\Delta u_f \equiv 0 \text{ on } D \text{ and } u_f = f \text{ on } \partial D. \tag{102}$$

Then, if $\{B_t \mid t \geq 0\}$ is a standard Brownian Motion in $\mathbb{R}^d$ defined on some filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$, and for all $x \in D$, we set:

$$\tau_X := \inf\{t \geq 0 \mid B_t + x \in D^c\}. \tag{103}$$

Then, τ_X is the hitting time of the set D^c, and since that's a closed set, it is a stopping time. Then, $u_f(X) = \mathbb{E}\left[f(B_{\tau_X} + x)\right]$. This is the **probabilistic representation of a harmonic function**, and amounts to solving a Dirichlet Problem.

Proof. We assume that u_f is defined as $u_f \in C_b^2(\mathbb{R}^d)$. Given $x \in D$, define

$$\phi_X(\cdot) = u_f(\cdot + x),$$

on $\mathbb{R}^d$. Then, $\phi_X \in C_b^2(\mathbb{R}^d)$. Therefore, $\left\{\varphi_X(B_t) - \int_0^t \frac{1}{2}\varphi_X(B_s)ds \mid t \geq 0\right\}$ is a martingale. This implies that,

$$\left\{\varphi_X(B_t \wedge \tau_x) - \int_0^{\tau_x \wedge t} \frac{1}{2}\Delta\varphi_X(B_s)ds \mid t \geq 0\right\},$$

is a martingale. For every $s \le t \wedge \tau_X$, $B_s \in D_x := \{y \in \mathbb{R}^d \mid y + x \in D\}$. This implies that

$$\Delta \varphi_X(B_s) = 0,$$

and therefore $\{\varphi_X(B_{t\wedge\tau_X}) \mid t \ge 0\}$ is a martingale. Furthermore, it's uniformly bounded and uniformly integrable. Hence, by **(Hunt's Theorem)** ,

$$\mathbb{E}\left[\varphi_X(B_{\tau_X})\right] = \mathbb{E}\left[\varphi_X(B_0)\right] = \mathbb{E}\left[f(B_{\tau_X} + x)\right] = u_f(X).$$

□

Corrolary 12. Let $u \in C_b^{1,2}([0,\infty[\times\mathbb{R}^d)$ be a solution to the **heat equation**:

$$\partial_t u(t,x) = \frac{1}{2}\Delta u(t,x),$$

for all $t > 0$, for all $x \in \mathbb{R}^d$, with initial condition $u(0,x) = g(x)$. Let $\{B_t \mid t \ge 0\}$ be a standard B.M. on $\mathbb{R}^d$. Then, for all $(t,x) \in]0,\infty[\times\mathbb{R}^d$ almost surely,

$$u(t,x) = \mathbb{E}\left[g(B_t + x)\right]. \tag{104}$$

This is the **probabilistic interpretation of the solution to the heat equation**.

Proof. Given $u(t,x)$ as above, for all $t \ge 0$, $x \in \mathbb{R}^d$, set

$$F : (s,y) \in [0,t] \times \mathbb{R}^d \mapsto F(s,y) = u(t-s, y+x) \in \mathbb{R}. \tag{105}$$

This implies that, for all $(s,y) \in]0,t[\times\mathbb{R}^d$,

$$\left(\partial_s + \frac{1}{2}\Delta\right) F(s,y) = 0. \tag{106}$$

Therefore, $\{F(s,B_s) \mid s \in [0,t]\}$ is a martingale. This implies

$$\begin{aligned} \mathbb{E}\left[F(t,B_t)\right] &= \mathbb{E}\left[F(0,0)\right] \\ \mathbb{E}\left[g(B_t + x)\right] &= u(t,x). \end{aligned}$$

□

4 Stochastic Integrals and Ito's Formula

4.1 Review of Basic Facts of Riemann-Stieljes Integrals

1. Let φ, ψ be two $\mathbb{R}$-valued functions on $[0,\infty[$. If $\varphi \in \mathrm{BV}_{\mathrm{loc}}([0,\infty[)$, i.e., φ has locally-bounded variation ($||\varphi||_{\mathrm{var},[0,t]} < \infty$) for all $t \ge 0$, $\psi \in C([0,\infty[)$. Then, for all $t \ge 0$, ψ is RS-integrable with respect to φ on $[0,t]$, i.e.,

$$\int_0^t \psi(s)d\varphi(s) := \lim_{L\to\infty} \max_{1\le j\le L} |t_j - t_{j-1}| \sum_{j=1}^{L} \psi(r_j^*)(\varphi(t_j) - \varphi(t_{j-1})), \tag{107}$$

where $L \ge 1$, $0 = t_0 < t_1 < \dots < t_L = t$, $r_j^* \in [t_{j-1}, t_j[$ for $1 \le j \le L-1$ and $r_L^* \in [t_{L-1}, t_L]$. The limit does not depend on the choice of L or the partition $\{t_j, r_j^* \mid 1 \le j \le L\}$.

2. Assume that for every $t \geq 0$, ψ is RS-integrable with respect to φ and set:
$$I(\psi, \varphi)(t) := \int_0^t \psi(s) d\varphi(s). \tag{108}$$
Then,
   (a) $t \in [0, \infty[\mapsto I(\psi, \varphi)(t)$ has locally bounded variation and
   $$||I(\psi, \varphi)||_{\text{var},[0,t]} < ||\psi|||_{U,[0,t]}||\varphi||_{\text{var},[0,t]}. \tag{109}$$
   (b) If φ is RCLL, then so is $I(\psi, \varphi)$.
   (c) If φ is continuous, then so is $I(\psi, \varphi)$.
3. If ψ is RS-integrable with respect to φ on $[0, t]$, then φ is also RS-integrable with respect to ψ on $[0, t]$ and the integration-by-parts formula holds:
$$\int_0^t \varphi(s) d\psi(s) = \varphi(t)\psi(t) - \varphi(0)\psi(0) - \int_0^t \psi(s) d\varphi(s). \tag{110}$$

Now, we turn to Brownian Motion. The objective is to gradually build up to integrating a Brownian Motion with respect to a Brownian Motion. Let B_t be a standard B.M. on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$. Assume that $\varphi \in \mathrm{BV}_{\mathrm{loc}}([0, \infty[) \cap C([0, \infty[)$. Then, for every $t \geq 0$, for all $\omega \in \Omega$, define the following:
$$I(\varphi))t(\omega) := \int_0^t \varphi(s) dB_s(\omega). \tag{111}$$
This is a RS-integral, and it's defined path-by-path (the dependence on $\omega \in \Omega$ makes this explicit). Then, we have the following:
$$I(\varphi)_t(\omega) := \varphi(t)B_t(\omega) - \int_0^t B_s(\omega) d\varphi(s). \tag{112}$$
We call $I(\varphi)_t(\omega)$ the **Paley-Wiener Integral** (with respect to Brownian Motion).

Remarks on PW Integrals.

1. *Continuity*: for all $\omega \in \Omega$, $t \mapsto I(\varphi)_t$ is continuous.
2. *Approximation*: we have that $I(\varphi)_0 \equiv 0$ for all $t > 0$, for all $\omega \in \Omega$,
$$I(\varphi)_t(\omega) = \lim_{n\to\infty} \sum_{m=1}^{2^n} \varphi((m-1)2^{-n}t))(B_{m2^{-n}t}(\omega) - B_{(m-1)2^{-n}t}(\omega)). \tag{113}$$
This limit does not depend on the choice of $r_m^* \in [(m-1)2^{-n}t, m2^{-n}t[$.
3. *Progressive Measurability*: $\{I(\varphi)_t \mid t \geq 0\}$ is progressively measurable with respect to $\{\mathcal{F}_t \mid t \geq 0\}$.
4. *Distribution*: Since $\varphi \in C([0, \infty])$, for all $t \geq 0$, $\int_0^t \varphi^2(s)ds < \infty$. For all $\xi \in |R$, we compute the characteristic function:
$$\begin{aligned}\mathbb{E}\left[e^{i\xi I(\varphi)_t}\right] &= \lim_{n\to\infty} \mathbb{E}\left[e^{i\xi \sum_{m=1}^{2^n} \varphi((m-1)2^{-n}t)(B_{m2^{-n}t} - B_{(m-1)2^{-n}t})}\right] \\ &= \lim_{n\to\infty} \prod_{m=1}^{2^n} e^{-\frac{\xi^2}{2}\varphi^2((m-1)2^{-n}t)2^{-n}t} \\ &= \lim_{n\to\infty} \exp\left(\sum_{m=1}^{2^n} \varphi^2((m-1)2^{-n}t)2^{-n}t)\left(-\frac{\xi^2}{2}\right)\right) \\ &= \exp\left(-\frac{\xi^2}{2}\int_0^t \varphi^2(s)ds\right).\end{aligned}$$
Hence, the distribution of $I(\varphi)_t$ is $\gamma_{0,\int_0^t \varphi^2(s)ds}$.

5. *Independent Increments*: for all $0 \le s \le t$,

$$I(\varphi)_t - I(\varphi)_s = \int_s^t \varphi(r) dB_r, \tag{114}$$

is independent of $\mathcal{F}_s$, and with distribution

$$\gamma_{0, \int_s^t \varphi^2(r) dr},$$

so the increments are not necessarily homogeneous !

6. *Gaussian Process*: The family of PW-integrals $\{I(\varphi)_t \mid t \ge 0\}$ is a Gaussian process with

$$\mathbb{E}\left[I(\varphi)_t\right] = 0 \ \forall t \ge 0.$$

For all $0 \le s \le t$,

$$\mathbb{E}\left[I(\varphi)_t I(\varphi)_s\right] = \mathbb{E}\left[I(\varphi)_s^2\right] = \int_0^s \varphi^2(r) dr.$$

Equivalently, $\{I(\varphi)_t \mid t \ge 0\}$ has the same distribution as $\{B_{\int_0^t \varphi^2(r) dr} \mid t \ge 0\}$ where $\{B_s \mid s \ge 0\}$ is a standard B.M.

7. *Martingale*: $\{I(\varphi)_t \mid t \ge 0\}$ is is a square-integrable martingale with continuous sample paths, with quadratic variation given by:

$$\langle I(\varphi) \rangle_t := \int_0^t \varphi^2(r) dr.$$

If φ_1 and φ_2 are two elements in $\mathrm{BV}_{\mathrm{loc}}([0,\infty[) \cap C([0,\infty])$, then for ever $t \ge 0$, $I(\varphi_1 + \varphi_2)_t = I(\varphi_1)_t + I(\varphi_2)_t$. The cross-variation of their PW-integrals is given by:

$$\langle I(\varphi_1) I(\varphi_2) \rangle_t = \int_0^t \varphi_1(r) \varphi_2(r) r.$$

8. *Isometry*: for all $t \ge 0$,

$$I : \varphi \in C([0,t]) \cap L^2([0,t]) \mapsto I(\varphi)_t \in L^2(\Omega)$$

is an isometry between two Hilbert spaces, $L^2([0,t])$ and $L^2(\Omega)$. Since $C([0,t]) \cap BV([0,t])$ is dense in $L^2([0,t])$, I can be uniquely extended to be an isometry as:

$$I : L^2([0,t]) \mapsto L^2(\Omega).$$

For all $\varphi \in L^2([0,t])$, there exists a sequence of functions $\{\varphi_n \mid n \ge 1\} \subseteq C([0,t]) \cap BV([0,t])$ such that $\varphi_n \to \varphi$ in $L^2([0,t])$ and,

$$I(\varphi)_t := \lim_{n \to \infty} I(\varphi_n)_t,$$

in the sense of L^2 convergence. Then, $I(\varphi)_t$ has distribution $\gamma_{0, ||\varphi||^2_{L^2([0,t])}}$. Moreover, $\{I(\varphi)_t \mid \varphi \in L^2([0,t])\}$ is a centred Gaussian family with covariance

$$\mathbb{E}\left[I(\varphi_1)_t I(\varphi_2)_t\right] = \langle \varphi_1, \varphi_2 \rangle_{L^2([0,t])}, \tag{115}$$

for all $\varphi_1, \varphi_2 \in L^2([0,t])$.