Convergence de mesures et processus de Lévy

Léo Gayral

Ces notes sont basées sur le cours de Pierre-Loïc Méliot

Table des matières

1 Convergence de variables aléatoires

On considère un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et des variables aléatoires $(X_n)_{n\in\mathbb{N}}$ à valeurs dans un espace $\mathbb{X}$. On s'intéresse aux différentes notions de convergence.

1.1 Convergence presque-sûre

Pour cette notion de convergence, on a besoin que $\mathbb{X}$ soit muni d'une topologie $\mathcal{T}$. On le munit en particulier de la tribu borélienne associée.

Définition 1 (Convergence dans $\mathbb{X}$) **:**
On dit que $(x_n) \in \mathbb{X}^{\mathbb{N}}$ converge vers x lorsque, pour tout ouvert $O \in \mathcal{T}$ contenant $x \in O$, la suite est inclue dans O à partir d'un rang.

Définition 2 (Convergence presque-sûre) **:**
On dit que $X_n \xrightarrow{\text{p.s.}} X$ converge presque-sûrement lorsque $\mathbb{P}\left(\left\{\omega \in \Omega,\, X_n \underset{n\to\infty}{\longrightarrow} X\right\}\right) = 1$, ce qui sous-entend que l'évènement est mesurable.

Lemme 3 :
Lorsque $\mathbb{X}$ est séparable ou métrisable, l'ensemble $\left\{\omega \in \Omega,\, X_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega)\right\}$ est mesurable.

Remarque 4 :
On parle de $\mathbb{X}$ métrisable plutôt que métrique car il existe des espaces métrisables, mais sans aucune notion de métrique canonique, et dont les métriques sont toutes peu pratiques à manipuler par rapport à la topologie.

1.2 Convergence en probabilités

On considère ici $\mathbb{X}$ métrique.

Définition 5 (Convergence en probabilités) **:**
On dit que $X_n \xrightarrow{\mathbb{P}} X$ lorsque, pour tout $\varepsilon > 0$, on a $\lim_{n\to\infty} \mathbb{P}(d(X_n, X) > \varepsilon) = 0$.

Proposition 6 :
Si $X_n \xrightarrow{\text{p.s.}} X$ alors $X_n \xrightarrow{\mathbb{P}} X$.

Théorème 7 :
Les propriétés suivantes sont équivalentes :

1. $X_n \xrightarrow{\mathbb{P}} X$,
2. toute extraction $\big(X_{\sigma(n)}\big)$ admet une sous-suite qui converge presque sûrement vers X.

Démonstration. Montrons $(2 \Rightarrow 1)$ par contraposée. Considérons $\varepsilon > 0$ pour lequel on n'observe pas la bonne limite. On a donc une extraction $\big(X_{\sigma(n)}\big)$ et $\delta > 0$ telles que :

$$\inf\Big\{\mathbb{P}\Big(d\big(X_{\sigma(n)}, X\big) > \varepsilon\Big)\Big\} > \delta\,.$$

En conséquence, aucune sous-extraction ne peut converger presque-sûrement.

Supposons désormais 1. Il suffit de trouver une extraction de (X_n) qui converge p.s. pour montrer l'implication, le cas général suivra le même schéma. On peut ici obtenir une extraction itérativement, telle que pour tout rang $n \in \mathbb{N}$ on a $\mathbb{P}\Big(d\big(X_{\sigma(n)}, X\big) > 2^{-n}\Big) \leq 2^{-n}$, d'où le résultat désiré. □

Remarque 8 :
Ce résultat donne une caractérisation plus générale de la convergence en probabilités, valable pour un espace $\mathbb{X}$ métrisable ou séparable, mais non métrique.

En outre, si d et d' induisent la même topologie sur $\mathbb{X}$, alors elle décrivent la même notion de convergence en probabilités.

1.3 Convergence L^1

On considère ici un espace de Banach $\mathbb{X}$.

Remarque 9 (Construction de l'intégrale) **:**
La construction usuelle de l'intégrale sur $\mathbb{R}$ n'est pas généralisable aux espaces vectoriels normés en dimension infinie.

Dans le cas d'un espace de Banach, on peut considérer la construction alternative où on définit la norme sur les fonctions en escalier :

$$\left\|\sum_{i=1}^{n} \mathbb{1}_{A_i} x_i\right\|_1 = \sum_{i=1}^{n} \mathbb{P}(A_i)\|x_i\|_{\mathbb{X}}\,,$$

où les $A_i \in \mathcal{F}$ sont disjoints. L^1 est alors le complété des fonctions en escalier pour $\|.\|_1$ parmi les fonctions mesurables $\Omega \to \mathbb{X}$.

L'intégrale est alors une application linéaire continue $L^1 \to \mathbb{X}$, en prolongeant l'application :

$$\int_\Omega \sum_{i=1}^n \mathbb{1}_{A_i} x_i \,\mathrm{d}\mathbb{P}(\omega) := \sum_{i=1}^n \mathbb{P}(A_i) x_i \,.$$

Définition 10 (Convergence L^1) **:**
On dit que $X_n \xrightarrow{L^1} X$ lorsque $\mathbb{E}[\|X_n - X\|_{\mathbb{X}}] = \|X_n - X\|_1 \longrightarrow 0$.

Proposition 11 :
Si $X_n \xrightarrow{L^1} X$ alors $X_n \xrightarrow{\mathbb{P}} X$.

Démonstration. Cela découle de l'inégalité de Markov, ici :

$$\mathbb{P}(\|X_n - X\|_{\mathbb{X}} \geq \varepsilon) \leq \frac{\|X_n - X\|_1}{\varepsilon} \underset{n\to\infty}{\longrightarrow} 0 \,.$$

□

Proposition 12 (Convergence dominée) **:**
Si $X_n \xrightarrow{\text{p.s.}} X$ et $(\|X_n\|_{\mathbb{X}})$ est dominée par une variable réelle $R \in L^1$, alors $X_n \xrightarrow{L^1} X$.

1.4 Notion d'uniforme intégrabilité

Définition 13 :
Les variables $X_n \in L^1(\Omega, \mathbb{X})$ sont uniformément intégrables lorsque :

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[\|X_n\|_{\mathbb{X}} \times \mathbb{1}_{\|X_n\|_{\mathbb{X}} \geq M}\right] \underset{M\to\infty}{\longrightarrow} 0 \,.$$

Remarque 14 :
De façon équivalente, les (X_n) sont uniformément intégrables lorsque :

$$\sup_{n\in\mathbb{N}} \sup_{\mathbb{P}(A_\varepsilon)\leq\varepsilon} \mathbb{E}[\|X_n\|_{\mathbb{X}} \times \mathbb{1}_{A_\varepsilon}] \underset{\varepsilon\to 0}{\longrightarrow} 0 \,.$$

Théorème 15 :
Lorsque $X_n \xrightarrow{\mathbb{P}} X$, on a l'équivalence suivante :
— $X_n \xrightarrow{L^1} X$,
— (X_n) est uniformément intégrable.

Démonstration. Le sens direct se fait sans encombres, grâce à la variable $X \in L^1$.

Montrons l'implication réciproque. La suite (X_n) est u.i. donc bornée dans L^1. La suite converge en probabilité vers X donc on peut extraire une sous-suite qui converge presque-sûrement. En conséquence, par le lemme de Fatou, $X \in L^1$. Quitte à translater par la variable X, on préserve donc l'uniforme intégrabilité. En effet :

$$\begin{aligned}\mathbb{E}\big[\|X_n - X\|_{\mathbb{X}} \times \mathbb{1}_{\|X_n - X\|_{\mathbb{X}} \geq M}\big] &\leq \mathbb{E}\Big[\|X_n - X\|_{\mathbb{X}} \times \mathbb{1}_{\|X_n\| \geq \|X\|} \times \mathbb{1}_{\|X_n\| \geq \frac{M}{2}}\Big] \\ &+ \mathbb{E}\Big[\|X_n - X\|_{\mathbb{X}} \times \mathbb{1}_{\|X\| \geq \|X_n\|} \times \mathbb{1}_{\|X\| \geq \frac{M}{2}}\Big] \\ &\leq 2\mathbb{E}\Big[\|X_n\|_{\mathbb{X}} \times \mathbb{1}_{\|X_n\| \geq \frac{M}{2}}\Big] + 2\mathbb{E}\Big[\|X\|_{\mathbb{X}} \times \mathbb{1}_{\|X\| \geq \frac{M}{2}}\Big] ,\end{aligned}$$

or ce terme est uniformément petit lorsque M est assez grand, par uniforme intégrabilité de la suite et intégrabilité de X. On peut alors découper l'espérance $\|X_n - X\|_1$ en trois parties, selon si la norme est plus grande que le M précédent, plus petite que ε, ainsi que le terme :

$$\mathbb{E}\big[\|X_n - X\|_{\mathbb{X}} \times \mathbb{1}_{\varepsilon \leq \|X_n - X\|_{\mathbb{X}} \leq M}\big] \leq M\mathbb{P}(\|X_n - X\|_{\mathbb{X}} \geq \varepsilon) ,$$

qui tend vers 0 par convergence en probabilités. □

1.5 Convergence en loi

$\mathbb{X}$ est de nouveau un espace topologique quelconque.

Définition 16 (Convergence en loi) **:**
On dit que $X_n \rightharpoonup X$ lorsque, pour toute fonction $f \in \mathcal{C}_b(\mathbb{X}, \mathbb{R})$ continue bornée, on a $\mathbb{E}[f(X_n)] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[f(X)]$.

Remarque 17 :
C'est une propriété qui porte sur les lois des variables, les mesures induites dans l'espace $\mathbb{X}$. La convergence en loi correspond à la convergence faible pour les mesures induites, en tant que formes linéaires sur $\mathcal{C}_b(\mathbb{X})$.

Définition 18 (Espace des mesures de probabilité) **:**
On note $\mathcal{M}^1(\mathbb{X})$ l'ensemble des mesures de probabilités sur un espace mesurable $\mathbb{X}$.

Théorème 19 (Théorème de Portemanteau) **:**
Pour une suite de probabilités (μ_n) on a l'équivalence suivante :

1. $\mu_n \rightharpoonup \mu$,
2. $\forall f \in \mathcal{C}_b$ uniformément continue, $\mu_n(f) \xrightarrow{\mathbb{X}} \mu(f)$,
3. pour tout ouvert $O \subset \mathbb{X}$, $\underline{\lim}\, \mu_n(O) \geq \mu(O)$,
4. pour tout fermé $F \subset \mathbb{X}$, $\overline{\lim}\, \mu_n(F) \leq \mu(F)$,
5. pour tout borélien $A \subset \mathbb{X}$ tel que $\mu(\partial A) = 0$, on a $\mu_n(A) \to \mu(A)$.

Démonstration. L'implication $(1 \Rightarrow 2)$ est directe, de même que l'équivalence $(3 \Leftrightarrow 4)$.

Montrons $(2 \Rightarrow 4)$. Pour cela, on approche $\mathbb{1}_F$ par au dessus, via $f_k(x) = (1 - k \times d(x, F))^+$. Cette application est k-lipschitzienne, donc uniformément continue, et bornée. On a l'inégalité $\mu_n(F) \leq \mu_n(f_k)$, d'où $\overline{\lim}\, \mu_n(F) \leq \overline{\lim}\, \mu_n(f_k) = \mu(f_k)$. Ceci étant vrai pour tout $k \in \mathbb{N}$, on peut passer à la limite par convergence dominée et enfin obtenir $\overline{\lim}\, \mu_n(F) \leq \mu(F)$.

Montrons $(3, 4 \Rightarrow 5)$. Ceci découle du fait que, pour un tel borélien, on a :

$$\mu\left(\mathring{A}\right) \leq \underline{\lim}\, \mu_n\left(\mathring{A}\right) \leq \overline{\lim}\, \mu_n(\overline{A}) \leq \mu(\overline{A}) = \mu\left(\mathring{A}\right) + \mu(\partial A) = \mu\left(\mathring{A}\right).$$

Montrons finalement $(5 \Rightarrow 1)$. Considérons donc $f \in \mathcal{C}_b(\mathbb{X}, \mathbb{R})$. À transformation affine près, on se ramène à $0 \leq f \leq 1$. Ainsi $\mu_n(f) = \int\limits_{\mathbb{X}} f(x) d\mu_n(x) = \int\limits_0^1 \int\limits_{\mathbb{X}} \mathbb{1}_{y \leq f(x)} d\mu_n(x) dy = \int\limits_0^1 \mu_n(A_y) dy$ par Fubini positif, avec $A_y = \{x \in \mathbb{X},\ f(x) \geq y\}$. L'application $F : y \mapsto \mu(A_y)$ est décroissante, donc elle admet une quantité dénombrable de discontinuités. En tout point de continuité y, l'ouvert $B_y = \{x \in \mathbb{X},\ f(x) > y\}$ est de même mesure, donc $\mu(\partial A_y) = 0$, et $\mu_n(A_y) \to \mu(A_y)$. Comme f est dominée par 1, par convergence dominée, on obtient finalement :

$$\mu_n(f) = \int\limits_0^1 \mu_n(A_y) dy \underset{n \to \infty}{\longrightarrow} \int\limits_0^1 \mu(A_y) dy = \mu(f).$$

□

Corollaire 20 :
Si $X_n \xrightarrow{\mathbb{P}} X$ alors $X_n \rightharpoonup X$.

Démonstration. On utilise la deuxième caractérisation du théorème précédent, par les fonctions uniformément continues. □

Corollaire 21 (Théorème de transfert) **:**
Soit $f \in \mathcal{C}^0(\mathbb{X}, \mathbb{Y})$. Si $X_n \xrightarrow{\text{p.s.}} X$ (resp. $\mathbb{P}$, en loi) alors $f(X_n) \xrightarrow{\text{p.s.}} f(X)$ (resp. $\mathbb{P}$, en loi).

2 Convergence en loi et théorème de Prohorov

L'objectif est de construire des objets aléatoires X compliquées, par approximation en loi par des variables (X_n) plus simples. Ça sera par exemple le cas du mouvement brownien, approché par les marches aléatoires X_n, avec des pas en espace $\frac{1}{\sqrt{n}}$ et des pas en temps $\frac{1}{n}$, qu'on peut identifier à leurs trajectoires dans $\mathcal{C}^0(\mathbb{R}^+, \mathbb{R})$.

Lemme 22 :
Soit (y_n) une suite dans un espace métrique (Y, d). Si toute sous-suite de (y_n) admet une extraction convergente, et que toutes ces sous-suites ont la même limite y, alors $y_n \to y$.

Démonstration. Si $y_n \not\to y$, alors on a une sous-suite $\left(y_{\sigma(n)}\right)$ qui reste toujours à distance au moins δ de y. En conséquence, aucune extraction de cette sous-suite ne converge vers y. $\square$

Pour appliquer ce lemme, il faut alors donner un sens plus topologique à la notion de convergence en loi, identifier une métrique, des parties compactes, etc.

2.1 Topologie de la convergence en loi

Définition 23 (Espace polonais) **:**
Soit $(\mathbb{X}, \mathcal{O})$ un espace topologique. On dit que $\mathbb{X}$ est un espace polonais si :

1. la topologie $\mathcal{O}$ est métrisable par une distance d,
2. il existe une distance d qui rend l'espace $\mathbb{X}$ complet,
3. $\mathbb{X}$ est séparable, pour une suite dense (x_n).

Idéalement, on étudiera $\mathcal{M}^1(\mathbb{X})$ dans le cadre d'un espace polonais $\mathbb{X}$, en essayant de transporter le plus possible de bonnes propriétés.

Définition 24 (Topologie de la convergence en loi) **:**
On veut définir les ouverts de $\mathcal{M}^1(\mathbb{X})$ pour la convergence en loi. Une base d'ouverts est donnée par les voisinages ouverts des $\mu \in \mathcal{M}^1$ suivants :

$$\forall f \in \mathcal{C}_b^0(\mathbb{X}),\, \forall \varepsilon > 0,\, V_{\mu,\varepsilon,f} = \left\{\nu \in \mathcal{M}^1,\, |\mu(f) - \nu(f)| < \varepsilon\right\}.$$

Définition 25 (Distance de Lévy-Prohorov) **:**
Soit $(\mathbb{X}, d)$ muni de sa tribu borélienne $\mathcal{B}$. Pour $\mu, \nu \in \mathcal{M}^1(\mathbb{X})$ on introduit la distance :

$$d_{LP}(\mu, \nu) = \inf\{\varepsilon > 0,\, \forall A \in \mathcal{B},\, \mu(A) \le \nu(A^\varepsilon) + \varepsilon,\, \nu(A) \le \mu(A^\varepsilon) + \varepsilon\}\,,$$

où $A^\varepsilon = \left\{ y \in \mathbb{X},\, d(A,y) = \inf_{x\in A} d(x,y) < \varepsilon \right\}$.

Remarque 26 :

Dans la définition de d_{LP}, on peut se contenter d'une seule des deux inégalités.

En effet, supposons qu'on vérifie toutes les inégalités $\mu(A) \leq \nu(A^\varepsilon) + \varepsilon$. Remarquons alors l'inclusion $\left((A^\varepsilon)^C\right)^\varepsilon \subset A^C$. Dans ce cas :

$$\nu(A) = 1 - \nu\big(A^C\big) \leq 1 - \nu\Big(\big((A^\varepsilon)^C\big)^\varepsilon\Big) \leq 1 - \mu\Big((A^\varepsilon)^C\Big) + \varepsilon = \mu(A^\varepsilon) + \varepsilon\,.$$

Lemme 27 :

L'application d_{LP} est bien une distance sur $\mathcal{M}^1(\mathbb{X})$.

Démonstration. L'application est clairement symétrique par définition.

Pour l'inégalité triangulaire, constatons que, si $d_{LP}(\mu,\nu) < \varepsilon$ et $d_{LP}(\nu,\rho) < \eta$ alors on a :

$$\mu(A) \leq \nu(A^\varepsilon) + \varepsilon \leq \rho((A^\varepsilon)^\eta) + \varepsilon + \eta \leq \rho\big(A^{\varepsilon+\eta}\big) + \varepsilon + \eta\,.$$

En revenant à la définition. $d_{LP}(\mu,\rho) \leq \varepsilon + \eta$. En passant à la limite inférieure, on obtient l'inégalité triangulaire.

Pour la séparation, supposons $d_{LP}(\mu,\nu) = 0$. Pour tout fermé F, on a $F = \bigcap_{\varepsilon>0} F^\varepsilon$, or $\mu(F) \leq \nu(F^\varepsilon) + \varepsilon$ donc $\mu(F) \leq \nu(F)$ pour $\varepsilon \to 0$. Par symétrie, on obtient l'inégalité inverse, donc μ et ν coincident sur les fermés, et donc $\mu = \nu$. □

Lemme 28 :

Si $\mathbb{X}$ est un espace métrisable séparable, alors la distance d_{LP} correspond à la topologie de la convergence en loi.

Démonstration. Il suffira de montrer l'équivalence entre $d_{LP}(\mu_n,\mu) \to 0$ et $\mu_n \rightharpoonup \mu$.

Si $d(\mu_n,\mu) \to 0$, alors pour tout fermé F et $\varepsilon > 0$, à partir d'un rang, $\mu_n(F) \leq \mu(F^\varepsilon) + \varepsilon$. En conséquence, $\overline{\lim}\, \mu_n(F) \leq \mu(F^\varepsilon) + \varepsilon$ et donc pour $\varepsilon \to 0$ on a enfin $\overline{\lim}\, \mu_n(F) \leq \mu(F)$. Par le théorème de Portemanteau, on a établi la convergence en loi.

Réciproquement, supposons $\mu_n \rightharpoonup \mu$. On fixe $\varepsilon > 0$ et une suite (x_n) dense dans $\mathbb{X}$. On a un rang $M < \infty$ tel que $\mu\left(\bigcup_{n=0}^{M} B(x_n,\varepsilon)\right) \geq 1 - \varepsilon$. Par le théorème de Portemanteau, à partir d'un rang N, pour toute famille $F \subset [\![0,M]\!]$, $\mu_n\left(\bigcup_{n\in F} B(x_n,\varepsilon)\right) \geq \mu\left(\bigcup_{n\in F} B(x_n,\varepsilon)\right) - \varepsilon$.

On pose $U(A) = \bigcup_{n\leq M,\, B(x_n,\varepsilon)\cap A\neq\emptyset} B(x_n,\varepsilon)$, qui contient A à une partie de masse ε près, et tel que $U(A) \subset A^{2\varepsilon}$. On a alors $\mu(A) \leq \mu(U(A)) + \varepsilon \leq \mu_n(U(A)) + 2\varepsilon \leq \mu_n(A^{2\varepsilon}) + 2\varepsilon$. On a donc

$d_{LP}(\mu_n, \mu) \leq 2\varepsilon$ à partir d'un rang. □

Proposition 29 :

Si $\mathbb{X}$ est un espace métrisable séparable, alors $\mathcal{M}^1(\mathbb{X})$ aussi.

Démonstration. Conservons les notations de la démonstration précédente. Il ne reste ici plus qu'à montrer que $\mathcal{M}^1(\mathbb{X})$ est séparable. Soit la famille dénombrable suivante :

$$\left\{ \sum_{n=1}^{M} r_n \delta_{x_n},\ M \in \mathbb{N},\ (r_n) \in \left(\mathbb{Q}^+\right)^M,\ \sum_{n=1}^{M} r_i = 1 \right\}.$$

Soient $\varepsilon > 0$, et M le rang du lemme. Partons de $\mu \in \mathcal{M}^1(\mathbb{X})$. On considère les $B_n \subset B(x_n, \varepsilon)$ tels que $\bigcup_{n=0}^{M} B(x_n, \varepsilon) = \bigsqcup_{n=0}^{M} B_n$. On choisit alors des $r_n \in \mathbb{Q}^+$ tels que $|r_n - \mu(B_n)| \leq \frac{\varepsilon}{M+1}$. On pose enfin $\nu = \sum_{n=0}^{M} r_n \delta_{x_n} + \left(1 - \sum_{n \leq M} r_n\right) \delta_{x_{M+1}}$. Quitte à prendre des $r_n \leq \mu(B_n)$, on peut garantir que ν est bien une probabilité.

Adaptons alors la définition de $U(A) = \bigsqcup_{n \leq M,\, B_n \cap A \neq \emptyset} B_n$ qui reste une ε-approximation de A sous μ. Si n est un des indices utilisés pour $U(A)$, alors on a $x_n \in A^{2\varepsilon}$, et donc $\delta_{x_n}(A^{2\varepsilon}) = 1$. On a donc :

$$\mu(A) \leq \mu(U(A)) + \varepsilon \leq \sum_{n=0}^{M} \mu(B_n)\delta_{x_n}\left(A^{2\varepsilon}\right) + \varepsilon \leq \sum_{n=0}^{M} r_n \delta_{x_n}\left(A^{2\varepsilon}\right) + 2\varepsilon \leq \nu\left(A^{2\varepsilon}\right) + 2\varepsilon\,,$$

d'où $d_{LP}(\mu, \nu) \leq 2\varepsilon$. □

2.2 Cas d'un espace $\mathbb{X}$ compact

Définition 30 (Espace compact) **:**

On dit que $\mathbb{X}$ est compact s'il est séparé (propriété de Hausdorff) et si, de tout recouvrement d'ouverts, on peut extraire un sous-recouvrement fini.

Théorème 31 :

Si $(\mathbb{X}, d)$ est métrique, alors $\mathbb{X}$ est compact ssi toute suite admet une sous-suite convergente.

Théorème 32 :

Soit $\mathbb{X}$ métrisable. Alors $\mathbb{X}$ est compact ssi $\mathcal{M}^1(\mathbb{X})$ l'est.

Démonstration. On utilisera les théorèmes de Riesz-Markov et de Banach-Alaoglu.

Supposons $\mathbb{X}$ compact. Le dual topologique de $\mathcal{C}^0(\mathbb{X})$ est l'espace des mesures signées finies, m, par le théorème de Riesz-Markov. Dans cet espace, $\mathcal{M}^1(\mathbb{X})$ est inclus dans la boule unité fermée, compacte pour la topologie faible-$*$, par le théorème de Banach-Alaoglu.

Montrons donc que $\mathcal{M}^1(\mathbb{X})$ est fermé pour la topologie faible-$*$. Cela vient du fait que c'est une intersection d'images réciproques de fermés par des applications faiblement-$*$ continues.

Réciproquement, supposons $\mathcal{M}^1(\mathbb{X})$ compact. On a un plongement de $\mathbb{X}$ dans $\mathcal{M}^1(\mathbb{X})$ via l'injection $x \mapsto \delta_x$. Or l'ensemble des mesures de Dirac est fermé, car on peut le caractériser par une propriété de variance nulle, autrement dit $\{\mu \in \mathcal{M}^1, \forall f \in \mathcal{C}_b, \mu(f^2) = \mu(f)^2\}$. Toute suite (x_n) admet une sous-suite telle que les Diracs $\left(\delta_{x_{\sigma(n)}}\right)$ convergent vers une mesure $\mu = \delta_x$. Pour conclure que $x_n \to x$, il suffit de prendre les fonctions $f_\varepsilon(y) = \max\left(0, 1 - \frac{d(x,y)}{\varepsilon}\right)$, continues bornées, pour lesquelles $\delta_{x_{\sigma(n)}}(f_\varepsilon) > 0$ à partir d'un rang, d'où $d\left(x_{\sigma(n)}, x\right) < \varepsilon$. □

Remarque 33 (Cas d'un espace $\mathbb{X}$ non compact) **:**
Si on perd la compacité de $\mathbb{X}$, l'espace $(\mathcal{C}_b(\mathbb{X}))^*$ est plus gros que l'espace des mesures signées, et en particulier $\mathcal{M}^1(\mathbb{X})$ n'est pas fermé pour la topologie faible-$*$ dans cet espace.

Théorème 34 (Dunford-Schwartz) **:**
Si $\mathbb{X}$ polonais, alors $(\mathcal{C}_b(\mathbb{X}))^*$ est l'espace de mesures finiment additives régulières.

Pour une telle mesure μ, on n'a que les propriétés suivantes :

1. Additivité finie : $\mu\left(\bigsqcup_{i=1}^{r} A_i\right) = \sum_{i=1}^{r} \mu(A_i)$,
2. Pour tout borélien B, $\|\mu\|(B) := \sup_{B=\bigsqcup_{i=1}^{r} A_i} |\mu(A_i)|$,
3. Régularité : pour tout borélien B, on peut l'intercaler entre un ouvert U et un fermé F de sorte que $\|\mu\|(F \backslash U) < \varepsilon$ pour tout $\varepsilon > 0$.

Dans ce contexte, la notion d'intégrale précédemment présentée permet de donner sens à $\int_{\mathbb{X}} f \,\mathrm{d}\mu$ pour $f \in \mathcal{C}_b$.

2.3 Théorème de Prohorov

On veut caractériser les parties de $\mathcal{M}^1(\mathbb{X})$ relativement compactes, incluses dans un compact.

Définition 35 (Famille tendue) **:**
Une famille $N \subset \mathcal{M}^1(\mathbb{X})$ est tendue si, pour tout $\varepsilon > 0$, il existe un ensemble (relativement) compact $K_\varepsilon \subset \mathbb{X}$ tel que $\inf_{\mu \in \mathbb{N}} \mu(K_\varepsilon) \geq 1 - \varepsilon$.

Lemme 36 :
Sur un espace polonais $\mathbb{X}$, toute mesure $\mu \in \mathcal{M}^1(\mathbb{X})$ est tendue. Le résultat se propage alors naturellement à toute famille finie.

Démonstration. On considère une suite (x_n) dense dans $\mathbb{X}$. Pour $\varepsilon > 0$ et $k \in \mathbb{N}^*$, il existe $N(k,\varepsilon)$ tel que $\mu\left(\bigcup_{n \leq N} B\left(x_n, \frac{\varepsilon}{2^k}\right)\right) \geq 1 - \frac{\varepsilon}{2^k}$. On pose $K_\varepsilon = \bigcap_{k \in \mathbb{N}^*} \bigcup_{n \leq N(k)} B\left(x_n, \frac{\varepsilon}{2^k}\right)$. On a alors $\mu(K_\varepsilon) \geq 1 - \sum_{k \in \mathbb{N}^*} \frac{\varepsilon}{2^k} = 1 - \varepsilon$. En outre, K_ε est précompact (inclus dans une union finie de boules, pour tout rayon), et inclus dans un espace complet, donc relativement compact. □

Lemme 37 :

Si N est une famille relativement compacte, alors pour tous $k \in \mathbb{N}^*$ et $\varepsilon > 0$, il existe $M(k,\varepsilon)$ tel que, pour toute mesure $\mu \in N$, on a $\mu\left(\bigcup_{n \leq N} B\left(x_n, \frac{\varepsilon}{2^k}\right)\right) \geq 1 - \frac{\varepsilon}{2^k}$.

Démonstration. Si ce n'était pas le cas, on aurait un couple (k,ε) tel que, pour tout $m \in \mathbb{N}$, il existe $\mu_n \in N$ telle que $\mu_m\left(\bigcup_{n \leq m} B\left(x_n, \frac{\varepsilon}{2^k}\right)\right) < 1 - \frac{\varepsilon}{2^k}$. Comme N est relativement compact, à extraction près, on a $\mu_m \to \mu$. Or $\mu\left(\bigcup_{n \leq n_0} B\left(x_n, \frac{\varepsilon}{2^k}\right)\right) \leq \underline{\lim}\, \mu_m\left(\bigcup_{n \leq m} B\left(x_n, \frac{\varepsilon}{2^k}\right)\right) \leq 1 - \frac{\varepsilon}{2^k}$. Ceci étant vrai pour tout rang n_0, en passant à la limite $n_0 \to \infty$ on en déduit $\mu(\mathbb{X}) < 1$, donc ce n'est pas une mesure de probabilités. □

Corollaire 38 :

Si N est relativement compacte, alors elle est tendue.

Démonstration. On construit l'ensemble relativement compact K_ε comme au premier lemme, à partir des rangs $M(k,\varepsilon)$ communs à toute la famille N donnés par le second lemme. □

Lemme 39 :

Il existe un espace métrique compact W dans lequel on peut plonger tout espace polonais $\mathbb{X}$ de façon homéomorphe.

Démonstration. Soit $W = [0,1]^{\mathbb{N}}$ muni de la topologie de la convergence simple, qu'on métrise par $d_W(a,b) = \sum_{n \in \mathbb{N}} \frac{|a_n - b_n|}{2^n}$.

Soit $\mathbb{X}$ un espace polonais, avec d une distance adaptée et (x_n) une suite dense. On vérifie alors que $\Phi(x) = \left(\frac{d(x_n,x)}{1+d(x_n,x)}\right)$ définit un homéomorphisme de $\mathbb{X}$ sur son image. □

Théorème 40 (Théorème de Prohorov) **:**

Soit $\mathbb{X}$ un espace polonais. Une famille $N \subset \mathcal{M}^1$ est relativement compacte ssi elle est tendue.

Démonstration. On a déjà montré le sens direct.

Pour l'implication réciproque, considérons une famille tendue N et une suite (μ_n) à valeurs dans cette famille. On pose $\nu_n = \Phi_* \mu_n$ la mesure image sur W issue du plongement précédent.

Comme W est un espace polonais compact, on a une extraction qui converge vers $\nu \in \mathcal{M}^1(W)$.

Vérifions que $\nu(\Phi(\mathbb{X})) = 1$. Soit $\varepsilon > 0$, et K_ε un compact associé à la famille tendue. On a $\Phi(K_\varepsilon)$ compact en tant qu'image d'un compact, donc fermé. En conséquence, on peut minorer $\nu_{\sigma(n)}(\Phi(K_\varepsilon)) = \mu_{\sigma(n)}(K_\varepsilon) \geq 1 - \varepsilon$. Par le théorème de Portemanteau, on a alors l'inégalité $\nu(\Phi(K_\varepsilon)) \geq \overline{\lim}\, \nu_{\sigma(n)}(\Phi(K_\varepsilon)) \geq 1 - \varepsilon$. En passant à la limite $\varepsilon \to 0$ on a enfin $\nu(\Phi(\mathbb{X})) = 1$. On peut donc poser $\mu = \Phi_*^{-1}\nu$. Comme $\nu \in \mathcal{M}^1(\Phi(\mathbb{X}))$, par continuité de Φ^{-1}, la convergence en loi de ν se propage à $\mu_{\sigma(n)} \rightharpoonup \mu$. L'ensemble N est donc bien relativement compact. □

Remarque 41 :
Si $A \subset \mathbb{R} \times \mathbb{R}^+$ et N est l'ensemble des gaussiennes de paramètres $(m, \sigma^2) \in A$, Alors N est relativement compacte ssi N est tendu ssi A est borné.

Remarque 42 (Glivenko-Cantelli pour les espaces polonais) **:**
Soient $\mathbb{X}$ polonais et $X_n \overset{\text{iid}}{\sim} \mu \in \mathcal{M}^1(\mathbb{X})$. Alors $\mu_n = \frac{1}{n}\sum\limits_{i=1}^{n} \delta_{X_i} \rightharpoonup \mu$ presque-sûrement.

Théorème 43 :

Si $\mathbb{X}$ est un espace polonais, alors $\mathcal{M}^1(\mathbb{X})$, pour la topologie de la convergence en loi, est polonais.

Démonstration. On a déjà établi que l'espace est séparable, métrique. Il ne reste qu'à montrer la complétude de l'espace. D'après les résultats précédents, il suffit de montrer qu'une suite de Cauchy $(\mu_n)_{n\in\mathbb{N}}$ sur cet espace est tendue.

Soient $\varepsilon > 0$ et $k \in \mathbb{N}$. On a un rang n_0 tel que, pour $n \geq n_0$, $d_{LP}(\mu_n, \mu_{n_0}) \leq \frac{\varepsilon}{2^{k+1}}$. Pour la famille dense (x_k), comme la famille $\{\mu_0, \dots, \mu_{n_0}\}$ finie est tendue, on a n_k tel que $\mu\left(\bigcup_{n=1}^{n_k} B\big(x_n, \frac{1}{2^{k+1}}\big)\right) \geq 1 - \frac{\varepsilon}{2^{k+1}}$.

Pour $n \geq n_0$, on a alors :

$$\begin{aligned}\mu_n\left(\bigcup_{n=1}^{n_k} B\big(x_n, \tfrac{1}{2^{k+1}}\big)\right) &\geq \mu_n\left(\left(\bigcup_{n=1}^{n_k} B\big(x_n, \tfrac{1}{2^{k+1}}\big)\right)^{\frac{\varepsilon}{2^{k+1}}}\right) \\ &\geq \mu_{n_0}\left(\bigcup_{n=1}^{n_k} B\big(x_n, \tfrac{1}{2^{k+1}}\big)\right) - \tfrac{\varepsilon}{2^{k+1}} \\ &\geq 1 - \tfrac{\varepsilon}{2^k}\,.\end{aligned}$$

En passant à l'intersection $K_\varepsilon = \bigcap_{k\in\mathbb{N}} \bigcup_{n=1}^{n_k} B\big(x_n, \frac{1}{2^{k+1}}\big)$, on obtient donc un ensemble relativement compact, pour lequel $\mu_n(K_\varepsilon) \geq 1 - \varepsilon$. On a bien une famille tendue. □

2.4 Théorème de représentation de Skorokhod

Théorème 44 (Théorème) **:**

Si $\mu_n \rightharpoonup \mu$, alors il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et des variables $X_n \sim \mu_n$, $X \sim \mu$ telles que $X_n \xrightarrow{\text{p.s.}} X$.

Corollaire 45 :

Si (X_n) est une suite de variables u.i., qui converge en loi vers X, alors $X_n \xrightarrow{L^1} X$.

3 Convergence en loi sur $\mathbb{R}$, transformée de Fourier

3.1 Fonctions de répartition

Définition 46 (Fonction de répartition) **:**
La fonction de répartition de X réelle est $F_X : t \mapsto \mathbb{P}(X \leq t)$. C'est une fonction croissante, qui tend vers 0 en $-\infty$ et 1 en $+\infty$, continue à droite.

Lemme 47 :
Pour tout $\varepsilon > 0$, $\{t,\ F_X(t) - F_X(t^-) \geq \varepsilon\}$ est fini, de cardinal $\leq \frac{1}{\varepsilon}$.

En passant à l'union pour les $\varepsilon_n = \frac{1}{n}$, l'ensemble des points de discontinuité est dénombrable.

Théorème 48 :
On a l'équivalence entre :

1. $X_n \rightharpoonup X$,
2. $F_{X_n}(t) \to F_X(t)$ en tout point t où F_X est continue.

Démonstration. L'implication $(1 \Rightarrow 2)$ découle du théorème de Portemanteau. Si F_X est continue en t, alors :

$$\mathbb{P}_X(\partial\,]-\infty, t]) = \mathbb{P}_X(\{t\}) = F_X(t) - F_X\left(t^-\right) = 0\,,$$

donc $F_{X_n}(t) = \mathbb{P}_{X_n}(]-\infty, t]) \to F_X(t)$.

Pour l'implication réciproque, on pose $\mu_n = \mathbb{P}_{X_n}$ et $\mu = \mathbb{P}_X$. Pour tout $\varepsilon > 0$, on a $M > 0$ tel que $\mu([-M, M]) \geq 1 - \varepsilon$. On peut en outre supposer que F_X continue en M et $-M$.

A partir d'un rang n_0, on a donc $|F_{X_n}(M) - F_X(M)| \leq \varepsilon$ et $|F_{X_n}(-M) - F_X(-M)| \leq \varepsilon$. On en déduit que $\mu_n([-M, M]) \geq 1 - 3\varepsilon$. Quitte alors à augmenter la valeur de M, comme la famille finie $\{\mu_0, \ldots, \mu_{n_0-1}\}$ est tendue, on étend la minoration par $1 - 3\varepsilon$ à tout rang n. La famille (μ_n) toute entière est donc tendue.

Soit ν une valeur d'adhérence de la suite. En réinjectant le sens direct, En tout point de continuité de F_X et F_ν – donc sur $\mathbb{R}$ privé d'un ensemble dénombrable – on a $F_\nu(t) = F_X(t)$, or les fonctions sont càdlàg, d'où $F_\nu = F_X$.

Comme μ est l'unique valeur d'adhérence de la suite, on a bien $\mu_n \rightharpoonup \mu$. $\square$

Corollaire 49 :
Si la variable X a une loi sans atomes, à densité par exemple, alors F_X est continue, et $X_n \rightharpoonup X$ ssi F_{X_n} converge simplement vers F_X.

Lemme 50 (Théorème de Dini) **:**
Si (F_{X_n}) converge simplement vers $F_X \in \mathcal{C}^0$ alors la suite converge uniformément.

Autrement dit, si on a convergence en loi vers une variable à densité X, alors on a :

$$d_{\text{Kolmogorov}}(X_n, X) = \|F_{X_n} - F_X\|_\infty \to 0\,.$$

3.2 Transformées de Fourier

Définition 51 (Transformée de Fourier d'une mesure) **:**
La transformée de Fourier de μ est $\widehat{\mu} : \xi \in \mathbb{R} \mapsto \int\limits_{\mathbb{R}} e^{i\xi x}\,\mathrm{d}\mu(x) \in \mathbb{C}$. C'est une fonction continue, et $|\widehat{\mu}| \leq 1$. Lorsque X est une variable aléatoire, on notera généralement $\Phi_X(\xi) = \mathbb{E}\left[e^{i\xi X}\right]$ sa transformée de Fourier.

Lemme 52 (Transformée de Fourier et convolution) **:**
Pour deux lois μ et ν, la convolution $\mu * \nu$ correspond à la loi de la somme de variables indépendantes $X \sim \mu$ et $Y \sim \nu$. On a $\widehat{\mu * \nu} = \widehat{\mu} \times \widehat{\nu}$.

Proposition 53 :
La transformée de Fourier est injective.

Démonstration. Soit $\mu \sim \mathcal{N}(0, \sigma^2)$. Par théorème de dérivation sous l'intégrale, $\widehat{\mu}$ est dérivable, et $\partial_\xi \widehat{\mu}(\xi) = \int\limits_{\mathbb{R}} \frac{ixe^{i\xi x - \frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\,\mathrm{d}x = -\xi\sigma^2\widehat{\mu}(\xi)$. En conséquence, $\widehat{\mu}(\xi) = e^{-\frac{\xi^2\sigma^2}{2}}$.

Si on connaît $\widehat{\mu}$, on peut retrouver $\mu_{\sigma^2} = \mu * \mathcal{N}(0, \sigma^2)$. En effet :

$$\begin{aligned}
\mu_{\sigma^2}(A) &= \int\limits_{\mathbb{R}^2} \mathbb{1}_A(t)\frac{e^{-\frac{(t-x)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\,\mathrm{d}t\,\mathrm{d}\mu(x) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{\mathbb{R}^2} \mathbb{1}_A(t)\widehat{N_{\frac{1}{\sigma^2}}}(x-t)\,\mathrm{d}t\,\mathrm{d}\mu(x) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{\mathbb{R}^2} \mathbb{1}_A(t) \int\limits_{\mathbb{R}} e^{i\xi(x-t)}N_{\frac{1}{\sigma^2}}(\xi)\,\mathrm{d}\xi\,\mathrm{d}\mu(x)\,\mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{\mathbb{R}^2} \mathbb{1}_A(t)e^{-i\xi t} \int\limits_{\mathbb{R}} e^{i\xi(x)}\,\mathrm{d}\mu(x)N_{\frac{1}{\sigma^2}}(\xi)\,\mathrm{d}\xi\,\mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{\mathbb{R}^2} \mathbb{1}_A(t)e^{-i\xi t}\widehat{\mu}(\xi)N_{\frac{1}{\sigma^2}}(\xi)\,\mathrm{d}\xi\,\mathrm{d}t\,.
\end{aligned}$$

Finalement, lorsque $\sigma^2 \to 0$, on a la convergence en loi $\mu_{\sigma^2} \rightharpoonup \mu$, avec le lemme de Slutsky par exemple. Autrement dit, si on connaît $\widehat{\mu}$, on peut en déduire μ, d'où l'injectivité. $\square$

Lemme 54 (Lemme de Slutsky) **:**
Si $X_n \rightharpoonup X$ et $Y_n \xrightarrow{\mathbb{P}} c \in \mathbb{R}$, alors $(X_n, Y_n) \rightharpoonup (X, c)$.

Lemme 55 :
Pour tout $\delta > 0$, on a $\mu\big(\{-\frac{2}{\delta}, \frac{2}{\delta}\}^c\big) \leq \frac{1}{\delta}\int_{-\delta}^{\delta}(1-\widehat{\mu}(\xi))\,\mathrm{d}\xi$.

Démonstration. Le terme de gauche est l'intégrale sur $\mathbb{R}$ de la fonction :

$$\mathbb{1}_{|\delta x|>2} \leq 2 \times (1 - \mathrm{sinc}(\delta x)) = \frac{1}{\delta}\int_{-\delta}^{\delta}\big(1 - e^{i\xi x}\big)d\xi\,,$$

d'où la majoration souhaitée par Fubini. □

Théorème 56 (Théorème de Lévy) **:**
On a $\mu_n \rightharpoonup \mu$ ssi $(\widehat{\mu_n})$ converge simplement vers $\widehat{\mu}$.

Dans ce cas, la convergence de $(\widehat{\mu_n})$ est localement uniforme.

Démonstration. L'implication directe découle du fait que $x \mapsto e^{i\xi x}$ est continue bornée, à ξ fixé.

Pour l'implication réciproque, on va utilise le lemme ci-dessus pour montrer la tension. Par continuité de $\widehat{\mu}$, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $|\xi| < \delta$ implique $|1-\widehat{\mu}(\xi)| \leq \varepsilon$. On en déduit $\left|\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\widehat{\mu}(\xi))\,\mathrm{d}\xi\right| \leq 2\varepsilon$. Comme $\widehat{\mu_n}$ converge simplement et est dominée par 1, à partir d'un rang n_0 on a $\left|\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\widehat{\mu_n}(\xi))\,\mathrm{d}\xi\right| \leq 3\varepsilon$. Autrement dit, pour tout $n \geq n_0$, on a $\mu_n\big(\{-\frac{2}{\varepsilon}, \frac{2}{\varepsilon}\}\big) \geq 1 - 3\varepsilon$. La suite est tendue, donc en réutilisant l'injectivité de la transformée de Fourier, et le sens direct du théorème, on en déduit que $\mu_n \rightharpoonup \mu$. □

Corollaire 57 (TCL) **:**
Soient (A_i) iid, centrées, de variance 1. Alors $X_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} A_i \rightharpoonup \mathcal{N}(0,1)$.

Démonstration. On a :

$$\begin{aligned}
\Phi_{X_n}(\xi) &= \Phi_A\Big(\frac{\xi}{\sqrt{n}}\Big)^n \\
&= \Big(1 - \frac{\xi^2}{2n} + o\Big(\frac{1}{n}\Big)\Big)^n \\
&\sim \exp\Big(-\frac{\xi^2}{2}\Big) \\
&\sim \widehat{N}(\xi)\,,
\end{aligned}$$

donc d'après le théorème de Lévy, $X_n \rightharpoonup \mathcal{N}(0,1)$. □

3.3 Bornes de type Berry-Esseen

On a établi des résultats de convergence. L'objectif de cette parenthèse est d'expliciter la vitesse de convergence.

Lemme 58 :

La fonction $k(x) = \frac{1}{\pi}\left(\frac{1-\cos(x)}{x^2}\right)$ est une densité de probabilités, et $\widehat{\mu} = (1-|\xi|)^+$.

Démonstration. On va procéder par transformée de Fourier-Plancherel inverse, bien définie comme isométrie sur L^2. En particulier, pour $\widehat{f} \in L^2 \cap L^1$, l'inverse est donné par la formule :

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(\xi)e^{-ix\xi}\,\mathrm{d}\xi\,.$$

Dans notre cas, on peut vérifier que $\frac{1}{2\pi}\int_{\Re}^{-ix\xi}(1-|\xi|)^+\,\mathrm{d}\xi = k(x)$. □

Théorème 59 :

On considère X et Y réelles, et Φ_X et Φ_Y leurs transformées de Fourier. Si Y est à densité f pour la mesure de Lebesgue, et $\|f\|_\infty \leq M < \infty$, alors pour tout $T > 0$:

$$d_{Kol}(X,Y) \leq \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\Phi_X(\xi)-\Phi_y(\xi)}{\xi}\right|\mathrm{d}\xi + 24\frac{M}{\pi T}\,.$$

Démonstration. Soit $\Delta(s) = F_X(s) - F_Y(s)$. On veut majorer $\delta = \|\Delta\|_\infty$.

À partir de la fonction k définie dans le lemme précédent, on pose $k_T(x) = T \times k(Tx)$. On a alors $\widehat{k_T}(\xi) = \left(1 - \frac{|\xi|}{T}\right)^+$. On pose $\Delta_T = (k_T * \Delta)$ et $\delta_T = \|\Delta_T\|_\infty$.

Montrons $\delta_T \geq \frac{\delta}{2} - 12\frac{M}{\pi T}$. Soit s tel que $\Delta(s) \geq \delta - \varepsilon$ approche $\|\Delta\|_\infty$. Alors pour $t \geq s$ on a $\Delta(t) \geq \delta - \varepsilon - M(t-s)$. Soient $\eta = \frac{\delta-\varepsilon}{2M}$ et $s_0 = s + \eta$. Sur l'intervalle $[s_0 - \eta, s_0 + \eta]$ on a donc $\Delta(s) \geq \frac{\delta-\varepsilon}{2} - M(s - s_0)$, et en dehors de l'intervalle on a $\Delta(s) \geq -\delta$. On a alors :

$$\begin{aligned}
\Delta_T(s_0) &= \int_{\mathbb{R}} k_T(s_0 - s)\Delta(s)\,\mathrm{d}s \\
&= \int_{[s_0\pm\eta]} k_T(s_0-s)\Delta(s)\,\mathrm{d}s + \int_{[s_0\pm\eta]^c} k_T(s_0-s)\Delta(s)\,\mathrm{d}s \\
&\geq (\delta - \varepsilon - M\eta)\int_{[\pm\eta]} k_T(x)\,\mathrm{d}x - \delta\int_{[\pm\eta]^c} k_T(x)\,\mathrm{d}x \\
&\geq \frac{\delta-\varepsilon}{2}\int_{[\pm\eta]} k_T(x)\,\mathrm{d}x - \delta\int_{[\pm\eta]^c} k_T(x)\,\mathrm{d}x\,.
\end{aligned}$$

En passant à la limite $\varepsilon \to 0$ on a finalement :

$$\begin{array}{rcl} \delta_T & \geq & \Delta_T(s_0) \\ & \geq & \frac{\delta}{2} \int\limits_{\mathbb{R}} k_T(x)\,\mathrm{d}x - \frac{3\delta}{2} \int\limits_{\left[\pm\frac{\delta}{2M}\right]^c} k_T(x)\,\mathrm{d}x \\ & = & \frac{\delta}{2} - 3\delta \int\limits_{\frac{\delta T}{2M}}^{\infty} k(x)\,\mathrm{d}x \\ & \geq & \frac{\delta}{2} - \frac{6\delta}{\pi} \int\limits_{\frac{\delta T}{2M}}^{\infty} \frac{1}{x^2}\,\mathrm{d}x \\ & = & \frac{\delta}{2} - \frac{6\delta}{\pi} \times \frac{2M}{\delta T}\,. \end{array}$$

Autrement dit, on a $\delta \leq 2\delta_T + 24\frac{M}{\pi T}$. Il ne reste qu'à majorer δ_T. On peut ici appliquer une transformée de Fourier inverse :

$$\begin{array}{rcl} \delta_T & \leq & \frac{1}{2\pi} \int\limits_{\mathbb{R}} \left|\widehat{\Delta_T}(\xi)\right| \mathrm{d}\xi \\ & \leq & \frac{1}{2\pi} \int\limits_{\mathbb{R}} \left|\widehat{F_X}(\xi) - \widehat{F_Y}(\xi)\right| \widehat{k_T}(\xi)\,\mathrm{d}\xi \\ & \leq & \frac{1}{2\pi} \int\limits_{-T}^{T} \left|\widehat{F_X}(\xi) - \widehat{F_Y}(\xi)\right| \mathrm{d}\xi \\ & \leq & \frac{1}{2\pi} \int\limits_{-T}^{T} \left|\frac{\Phi_X(\xi)-\Phi_Y(\xi)}{\xi}\right| \mathrm{d}\xi\,. \end{array}$$

□

Remarque 60 (Application) **:**

Soient $(A_i) \in L^3$ centrées, de variance 1. On pose $M_3 = \mathbb{E}\left[|A|^3\right] < \infty$, et $X_n = \frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} A_i$. Alors $\Phi_{X_n}(\xi) = \exp\left(-\frac{\xi^2}{2} + O\left(\frac{M_3|\xi|^3}{\sqrt{n}}\right)\right)$ uniformément sur des intervalles $[-\eta\sqrt{n}, \eta\sqrt{n}]$. On a :

$$d_{\mathrm{Kol}}(X_n, N) \leq \frac{1}{\pi} \int\limits_{-T}^{T} \left|\frac{\Phi_{X_n}(\xi) - e^{-\frac{\xi^2}{2}}}{\xi}\right| \mathrm{d}\xi + O\left(\frac{1}{T}\right) \leq \frac{1}{\pi} \int\limits_{-T}^{T} e^{-\frac{\xi^2}{2}} \left|\frac{\exp\left(O\left(M_3\frac{|\xi|^3}{\sqrt{n}}\right)\right) - 1}{\xi}\right| \mathrm{d}\xi + O\left(\frac{1}{T}\right).$$

Comme $|e^z - 1| \leq |z|e^{|z|}$, on majore la distance par :

$$\frac{M_3}{\pi\sqrt{n}} \int\limits_{-T}^{T} \exp\left(-\frac{\xi^2}{2} + O\left(M_3\frac{|\xi|^3}{\sqrt{n}}\right)\right) |\xi|^2\,\mathrm{d}\xi + O\left(\frac{1}{T}\right) = O\left(\frac{M_3}{\sqrt{n}}\right).$$

Il est prouvé que la constante du $O\left(\frac{M_3}{\sqrt{n}}\right)$ est plus petite que 3 (Feller), et plus grande que $\frac{1}{\sqrt{2\pi}}$. Plus récemment, dans les années 2010, on a atteint 0.42 environ.

Lemme 61 (Convergence en loi et convergence des moments) **:**

En général, même si tous les moments de X_n convergent vers ceux de X, on n'a pas pour autant convergence en loi $X_n \rightharpoonup X$.

Cependant, si $\sum\limits_{n\in\mathbb{N}} \frac{\mathbb{E}[X^n]}{n!} z^n$ a un rayon de convergence non nul, la loi de X est caractérisée par ses moments. Dans ce cas précis, la convergence des moments $\mathbb{E}[X_n^k] \underset{n\to\infty}{\longrightarrow} \mathbb{E}\big[X^k\big]$ pour tout $k \in \mathbb{N}$ implique la convergence en loi $X_n \rightharpoonup X$.

4 Variables, processus et mesures aléatoires de Poisson

4.1 Variables de Poisson

Définition 62 (Variable de Poisson) **:**
Une variable de Poisson de paramètre $\lambda > 0$ est une variable aléatoire $X \in \mathbb{N}$ telle que $\mathbb{P}(X = k) = e^{-k}\frac{\lambda^k}{k!}$. On note $\mathcal{P}(\lambda)$ la loi correspondante.

Lemme 63 (Moments factoriels) **:**
On note $X^{\downarrow k} := \prod_{i=0}^{k-1}(X - i)$. Alors on a $\mathbb{E}\left[X^{\downarrow k}\right] = \lambda^k$.

En particulier, $\mathbb{E}[X] = \mathrm{Var}(X) = \lambda$.

La transformée de Laplace de X est $\mathbb{E}\left[e^{zX}\right] = e^{\lambda\left(e^{z}-1\right)}$.

Proposition 64 (Loi des petits nombres) **:**
Soit $X \sim \mathcal{B}\left(n, \frac{\lambda}{n}\right)$. Alors $X_n \rightharpoonup \mathcal{P}(\lambda)$.

Démonstration. Il suffit de constater la convergence des transformées de Fourier :

$$\mathbb{E}\left[e^{i\xi X_n}\right] = \left(1 + \frac{\lambda}{n}(e^{i\xi} - 1)\right)^n \to e^{\lambda\left(e^{i\xi-1}\right)}.$$

□

Proposition 65 :
Soit une famille de paramètres $(p_{i,n})_{i\leq n}$ tels que $\sum_{i=1}^n p_{i,n} \underset{n\to\infty}{\longrightarrow} \lambda$ et $\sum_{i=1}^n (p_{i,n})^2 \underset{n\to\infty}{\longrightarrow} 0$. Alors $X_n = \sum_{i=1}^n \mathcal{B}(p_{i,n}) \rightharpoonup \mathcal{P}(\lambda)$.

Démonstration. On pourrait prouver le résultat en travaillant sur les transformées de Fourier comme précédemment.

Montrons la convergence en variation totale. La somme de lois de Poisson indépendantes est une loi de Poisson. On a donc :

$$\begin{aligned} d_{\mathrm{VT}}(X_n, \mathcal{P}(\lambda)) &\leq d_{\mathrm{VT}}\left(X_n, \sum_{i=1}^n \mathcal{P}(p_i^n)\right) + d_{\mathrm{VT}}\left(\mathcal{P}(\lambda), \mathcal{P}\left(\sum_{i=1}^n p_i^n\right)\right) \\ &\leq \sum_{i=1}^n d_{\mathrm{VT}}(\mathcal{B}(p_i^n), \mathcal{P}(p_i^n)) + d_{\mathrm{VT}}\left(\mathcal{P}(\lambda), \mathcal{P}\left(\sum_{i=1}^n p_i^n\right)\right). \end{aligned}$$

Comme on considère des variables discrètes à valeurs entières, pour tout paramètre p, on a :

$$\begin{aligned} d_{\mathrm{VT}}(\mathcal{B}(p), \mathcal{P}(p)) &= \frac{1}{2}\sum_{n\in\mathbb{N}}|\mathbb{P}(\mathcal{B}(p)=n)-\mathbb{P}(\mathcal{P}(p)=n)| \\ &= \frac{1}{2}(|1-p-\mathbb{P}(\mathcal{P}(p)=0)|+|p-\mathbb{P}(\mathcal{P}(p)=1)|+\mathbb{P}(\mathcal{P}(p)\geq 2)) \\ &= \frac{1}{2}(e^{-p}+p-1)+\frac{p}{2}(1-e^{-p})+\frac{1}{2}(1-e^{-p}-pe^{-p}) \\ &= p(1-e^{-p}) \\ &\leq p^2\,. \end{aligned}$$

On majore ainsi le terme de gauche par $\sum_{i=1}^{n}(p_{i,n})^2 \to 0$. Pour le terme de droite, comme on a $S_n = \sum_{i=1}^{n} p_{i,n} \to \lambda$, les suites $(\mathbb{P}(\mathcal{P}(S_n)=k)-\mathbb{P}(\mathcal{P}(\lambda)=k))_{k\in\mathbb{N}}$ sont dominées dans $l^1(\mathbb{N})$ est la famille converge simplement vers 0, donc on obtient finalement $d_{\mathrm{VT}}(\mathcal{P}(S_n), \mathcal{P}(\lambda)) \to 0$. □

Remarque 66 (Chen, 1970) **:**

On peut en fait montrer que $d_{\mathrm{VT}}\left(X_n, \mathcal{P}\left(\sum_{i=1}^{n} p_{i,n}\right)\right) \leq \frac{\sum_{i=1}^{n}(p_{i,n})^2}{\sum_{i=1}^{n} p_{i,n}}$.

Remarque 67 (Nombre de cycles d'une permutation) **:**

Soit σ_n une permutation sur n éléments, choisie uniformément au hasard. On peut toujours écrire σ_n de manière unique sous la forme $\tau_{1,i_1}\circ\cdots\circ\tau_{n,i_n}$, avec $i_k\in[k]$.

En effet, dans ce cas $i_n = \sigma^{-1}(n)$. La permutation $\sigma_{n-1} = \sigma_n \circ \tau_{n,i_n}$ laisse n stable, donc peut être vue comme une permutation uniforme sur $[n-1]$. Par récurrence décroissante, on a existence et unicité des i_k pour $k<n$. En outre, σ_{n-1} est indépendante de $i_n \sim \mathcal{U}([n])$.

On note C_n son nombre de cycles. D'après ce qui précède, à σ_{n-1} fixé, σ_n récupère un cycle supplémentaire lorsque $i_n = n$, avec probabilité $\frac{1}{n}$, sans quoi n vient s'intercaler au sein d'un cycle préexistant. On a alors $C_n \overset{\mathrm{d}}{=} \sum_{k=1}^{n}\mathcal{B}\left(\frac{1}{k}\right)$. En conséquence, $d_{\mathrm{VT}}(C_n, \mathcal{P}(H_n)) \leq \frac{1}{H_n}\sum_{k=1}^{n}\frac{1}{k^2} = O\left(\frac{1}{\ln(n)}\right)$. Finalement, C_n suit une loi $\mathcal{P}(\ln(n))$ asymptotiquement.

Proposition 68 (Principe de superposition) **:**

Soient $(\lambda_i)_{i\in I}$ une famille dénombrable de paramètres, et $X_i \sim \mathcal{P}(\lambda_i)$ les variables de Poisson indépendantes correspondantes.

Avec la convention $\mathcal{P}(\infty) \overset{\mathrm{d}}{=} \delta_\infty$, on a alors $X := \sum_{i\in I} X_i \sim \mathcal{P}(\lambda)$ où $\lambda = \sum_{i\in I}\lambda_i$.

Démonstration. Si $\lambda = \infty$, alors par Borel-Cantelli, une infinité de X_i seront strictement positifs simultanément, d'où $X = \infty$ p.s.

Sinon, on a X positive et $\mathbb{E}[X] = \lambda < \infty$. En conséquence, on peut calculer sa transformée de Fourier par convergence dominée :

$$\mathbb{E}\left[e^{i\xi X}\right] = \lim \mathbb{E}\left[e^{i\xi \sum\limits_{k=1}^{n} X_i}\right] = \lim \prod_{k=1}^{n} \mathbb{E}\left[e^{i\xi X_k}\right] = \lim \prod_{k=1}^{n} e^{\lambda_k \left(e^{i\xi - 1}\right)} = e^{\lambda\left(e^{i\xi} - 1\right)},$$

donc X suit bien la loi souhaitée. $\square$

Proposition 69 (Principe de raréfaction) **:**

Soient $X \sim \mathcal{P}(\lambda)$ indépendant des $Y_k \overset{\text{iid}}{\sim} p$ pour une probabilité $(p_i)_{i \in \mathbb{N}^*}$ sur $\mathbb{N}^*$. Alors les variables $X_j := \sum\limits_{k=1}^{X} \mathbb{1}_{Y_k = j}$ sont indépendantes dans leur ensemble, de lois $\mathcal{P}(\lambda p_j)$.

Démonstration. Il suffit de vérifier les lois marginales finies de la famille X. On procède par transformée de Fourier :

$$\begin{aligned} \mathbb{E}\left[e^{i\sum\limits_{j=1}^{N} \xi_j X_j}\right] &= \mathbb{E}\left[e^{\sum\limits_{k=1}^{X} i\left(\sum\limits_{j=1}^{N} \xi_j \mathbb{1}_{Y_k = j}\right)}\right] \\ &= \mathbb{E}\left[\left(1 + \sum_{j=1}^{N} p_j \left(e^{i\xi_j} - 1\right)\right)^X\right] \\ &= e^{\lambda\left(\sum\limits_{j=1}^{N} p_j\left(e^{i\xi_j} - 1\right)\right)} \\ &= \prod_{j=1}^{N} e^{\lambda\left(e^{i\xi_j} - 1\right)}. \end{aligned}$$

$\square$

4.2 Processus de Poisson

Remarque 70 (Espace des trajectoires discontinues) **:**

On introduit l'espace de Skorohod des trajectoires discontinues, l'ensemble des fonctions càdlàg sur $\mathbb{R}^+$ qu'on note $\mathcal{D}(\mathbb{R}^+)$. Il existe une topologie sur cet espace qui le rend polonais.

Définition 71 :

Un processus de Poisson d'intensité λ est une famille $(N_t)_{t \geq 0}$ telle que :

1. $N_0 = 0$,
2. pour tous $0 = t_0 \leq t_1 \leq \cdots \leq t_d$, les variables $\left(N_{t_{i+1}} - N_{t_i}\right)_{0 \leq i < d}$ sont indépendantes, et suivent des lois de Poisson de paramètres $\lambda(t_{i+1} - t_i)$,
3. le processus est à trajectoires càdlàg, donc pour tout $\omega \in \Omega$ on a $t \mapsto N_t(\omega) \in \mathcal{D}(\mathbb{R}^+)$.

4.2.1 Construction du processus

Soient $(\xi_k) \overset{\text{iid}}{\sim} \mathcal{E}(\lambda)$ des variables exponentielles. On pose $S_k = \sum\limits_{j=1}^{k} \xi_j$, et $N_t = \max\{k \in \mathbb{N},\, S_k \leq t\}$. Le processus (N_t) est bien à trajectoires càdlàg. Il reste donc à vérifier la loi du processus.

Lemme 72 :
On a $N_t \sim \mathcal{P}(\lambda t)$.

Démonstration. On a $S_k \sim \Gamma(k, \lambda)$ à densité $\frac{\lambda^k}{(k-1)!} e^{-\lambda x} x^{k-1}$ sur $\mathbb{R}^{+*}$, donc :

$$\begin{aligned}
\mathbb{P}(N_t = k) &= \mathbb{P}(S_k \leq t,\, \xi_{k+1} > t - S_k) \\
&= \int_0^t \frac{\lambda^k e^{-\lambda s} s^{k-1}}{(k-1)!} \left(\int_{t-s}^{\infty} \lambda e^{-\lambda x} \,\mathrm{d}x \right) \mathrm{d}s \\
&= \lambda^k e^{-\lambda t} \int_0^t \frac{s^{k-1}}{(k-1)!} \,\mathrm{d}s \\
&= (\lambda t)^k \frac{e^{-\lambda t}}{k!} \,.
\end{aligned}$$

□

Lemme 73 :
Conditionnellement à $\{N_t = k\}$, la famille $(S_1, \ldots, S_k)$ suit une loi uniforme sur le simplexe $\Delta_k(t) = \{0 < t_1 < \cdots < t_k < t\}$.

Démonstration. En faisant un changement de variables $s_k = \sum\limits_{i=1}^{k} \xi_i$, on a :

$$\begin{aligned}
\mathbb{E}[f(S_1, \ldots, S_k) | N_t = k] &= \frac{1}{\mathbb{P}(N_t = k)} \int f(s_1, \ldots, s_k) \mathbb{1}_{s_k < t < s_{k+1}} \prod_{i=1}^{k+1} \lambda e^{-\lambda \xi_i} \,\mathrm{d}\xi_1 \ldots \mathrm{d}\xi_{k+1} \\
&= \frac{k!}{(\lambda t)^k} e^{\lambda t} \int f(s) \mathbb{1}_{\Delta_k(t)}(s) \mathbb{1}_{]t,\infty[}(s_{k+1}) \prod_{i=0}^{k} \lambda e^{-\lambda(s_{i+1} - s_i)} \,\mathrm{d}s \,\mathrm{d}s_{k+1} \\
&= \frac{k! e^{\lambda t}}{t^k} \lambda \int f(s) \mathbb{1}_{\Delta_k(t)}(s) \mathbb{1}_{]t,\infty[}(s_{k+1}) e^{-\lambda s_{k+1}} \,\mathrm{d}s \,\mathrm{d}s_{k+1} \\
&= \frac{k!}{t^k} \int f(s) \mathbb{1}_{\Delta_k(t)}(s) \left(\lambda \times \int_t^{\infty} e^{-\lambda(t - s_{k+1})} \,\mathrm{d}s_{k+1} \right) \mathrm{d}s \\
&= \frac{k!}{t^k} \int f(s) \mathbb{1}_{\Delta_k(t)}(s) \,\mathrm{d}s \,,
\end{aligned}$$

d'où le résultat souhaité. □

Remarque 74 :
En conséquence, $\{S_1, \ldots, S_{N_t}\}$ peut être vu comme un ensemble de cardinal $X \sim \mathcal{P}(\lambda t)$ aléatoire, et dont les éléments suivent tous des lois $U_j \sim \mathcal{U}([0, t])$ indépendantes.

Si on se fixe une subdivision $0 < t_1 < \cdots < t_d$, la variable $N_{t_i} - N_{t_{i-1}}$ est le nombre de sauts

S_k dans l'intervalle $[t_{i-1}, t_i]$, égal à $\sum_{j=1}^{X} \mathbb{1}_{t_{i-1} \leq U_j \leq t_i}$. Par le principe de raréfaction, on en conclut que $\left(N_{t_i} - N_{t_{i-1}}\right)_{i \leq d}$ est formé de lois de Poisson indépendantes, de paramètres $\lambda(t_i - t_{i-1})$.

4.3 Mesures aléatoires, nuages de points poissonniens

On veut désormais généraliser les ensembles de saut d'un processus de Poisson. On travaille pour ce faire sur un espace polonais $\mathbb{X}$.

Définition 75 (Espaces de mesures) **:**

On pose $\mathcal{M}^{s,+}(\mathbb{X})$ l'ensemble des mesures s-finies sur $\mathbb{X}$, c'est-à-dire de sommes dénombrables de mesures boréliennes finies. On munit cet espace de la plus petite tribu qui rend mesurables les évaluations $\mu \mapsto \mu(A)$ pour tout borélien $A \in \mathcal{B}(\mathbb{X})$.

On pose $\mathcal{M}^{\text{atom}}(\mathbb{X})$ l'ensemble des mesures atomiques, des sommes dénombrables de mesures de Dirac δ_x. C'est un ensemble mesurable dans $\mathcal{M}^{s,+}$, égal à $\left\{\mu \in \mathcal{M}^{s,+},\ \forall A \in \mathcal{B}(\mathbb{X}),\ \mu(A) \in \overline{\mathbb{N}}\right\}$.

Définition 76 :

Une mesure aléatoire de Poisson d'intensité $\mu \in \mathcal{M}^{s,+}(\mathbb{X})$ est une variable aléatoire M à valeurs dans $\mathcal{M}^{\text{atom}}$, telle que pour toute famille finie $(A_j)_{j \in J}$ de boréliens disjoints, les variables $(M(A_j))_{j \in J}$ suivent des lois de Poisson indépendantes, de paramètres $(\mu(A_j))_{j \in J}$.

Lemme 77 (Principe de superposition) **:**

Pour toute famille de mesures s-finie $(\mu_i)_{i \in I}$ dénombrable, si les M_i sont des mesures aléatoires indépendantes d'intensités μ_i, alors $M = \sum_{i \in I} M_i$ est une mesure aléatoire d'intensité $\mu = \sum_{i \in I} \mu_i$.

Démonstration. Soient $A_1, \ldots, A_N$ mesurables disjoints et $k_1, \ldots, k_N \in \mathbb{N}$. Quitte à décomposer la valeur de M en fonction de celles prises par les M_i, par produit de Cauchy on a :

$$\begin{aligned}
\mathbb{P}(\forall j \in [N],\ M(A_j) = k_j) &= \sum_{\sum_{i \in I} k_{i,j} = k_j} \mathbb{P}(\forall i \in I,\ \forall j \in [N],\ M_i(A_j) = k_{i,j}) \\
&= \sum_{\sum_{i \in I} k_{i,j} = k_j} \left(\prod_{1 \leq j \leq N} \mathbb{P}(\forall i \in I,\ M_i(A_j) = k_{i,j}) \right) \\
&= \prod_{1 \leq j \leq N} \left(\sum_{\sum_{i \in I} k_{i,j} = k_j} \mathbb{P}(\forall i \in I,\ M_i(A_j) = k_{i,j}) \right) \\
&= \prod_{1 \leq j \leq N} \mathbb{P}(M(A_j) = k_j)\,.
\end{aligned}$$

En outre, $M(A) = \sum_{i \in I} M_i(A) \overset{\mathrm{d}}{=} \mathcal{P}\left(\sum_{i \in I} \mu_i(A)\right)$ par le premier principe de superposition. □

Théorème 78 :

Pour tout $\mu \in \mathcal{M}^{s,+}(\mathbb{X})$, il existe M une mesure aléatoire de Poisson d'intensité μ. Ce processus est unique en loi, qu'on notera $MAP(\mu)$.

Démonstration. Par le principe de superposition ci-dessus, il suffit de montrer l'existence lorsque $\mu(\mathbb{X}) < \infty$, puis de conclure pour μ s-finie.

On considère alors $\overline{\mu} = \frac{\mu}{\mu(\mathbb{X})}$ la mesure normalisée. On pose $X_n \overset{\text{iid}}{\sim} \overline{\mu}$ indépendantes, et $P \sim \mathcal{P}(\mu(\mathbb{X}))$. Dans ce cas, $M = \sum_{n=1}^{P} \delta_{X_n}$ est bien une mesure aléatoire d'intensité μ. En effet, pour des boréliens disjoints $(A_j)_{j\in J}$, par le principe de raréfaction, les $M(A_j)$ sont bien des variables de Poisson de paramètres $\mu(\mathbb{X}) \times \overline{\mu}(A_j) = \mu(A_j)$.

Pour l'unicité en loi, remarquons d'abord qu'on peut étendre le résultat précédent à des familles de boréliens non-disjoints. Par exemple, pour deux boréliens A et B quelconques, on peut considérer trois boréliens $A\backslash B$, $B\backslash A$ et $A \cap B$ disjoints à la place. En conséquence, on connait la mesure par M des ensembles mesurables :

$$\left\{N \in \mathcal{M}^{\text{atom}},\, \forall j \leq r,\, N(A_j) = k_j\right\}$$

pour toute taille $r \in \mathbb{N}^*$, toute famille $A_1, \ldots, A_r \in \mathcal{B}(\mathbb{X})$ et $k_1, \ldots, k_r \in \overline{\mathbb{N}}$.

Ces mesurables forment un π-système de $\mathcal{M}^{\text{atom}}$, donc par un lemme de classes monotones, la loi de M est uniquement déterminée sur $\mathcal{M}^{\text{atom}}$, et on peut la noter $MAP(\mu)$. $\square$

Proposition 79 (Formules de Campbell, cas positif) **:**

Soient $M \sim MAP(\mu)$ et $f : \mathbb{X} \to \overline{\mathbb{R}^+}$ une fonction mesurable positive. On note $M(f)$ l'intégrale aléatoire de f contre la mesure M.

Alors $M(f)$ est une variable aléatoire réelle positive, d'espérance $\mathbb{E}[M(f)] = \mu(f)$ et sa transformée de Laplace est $\mathbb{E}\left[e^{-M(f)}\right] = \exp\left(-\mu\left(1 - e^{-f}\right)\right)$.

Démonstration. Par convergence monotone, $M(f)$ est limite croissante de $M_n(f)$, en considérant $M_n \sim MAP(\mu_n)$ pour une mesure μ_n *finie*. Si on montre le résultat dans ce cas, alors on conclura que $M(f)$ est bien une variable aléatoire réelle positive d'espérance $\mu(f)$ par convergence monotone, et on obtiendra la transformée de Laplace souhaitée par convergence dominée.

On peut donc supposer $\mu(\mathbb{X}) < \infty$. On peut exprimer f comme une limite monotone de fonctions étagées f_n. On peut à nouveau passer de f_n à f par convergence monotone ou dominée selon le cas.

Si $f = \sum_{j=1}^{k} c_j \mathbb{1}_{A_j}$ avec des A_j disjoints, alors $M(f)$ est clairement une variable aléatoire réelle positive. On a $\mathbb{E}[M(f)] = \sum_{j=1}^{k} c_j \mathbb{E}[M(A_j)] = \sum_{j=1}^{k} c_j \mu(A_j) = \mu(f)$. De même, on a :

$$\mathbb{E}\left[e^{-M(f)}\right] = \mathbb{E}\left[\prod_{j=1}^{k} e^{-c_j M(A_j)}\right] = \prod_{j=1}^{k} \mathbb{E}\left[e^{-c_j M(A_j)}\right] = \prod_{j=1}^{k} e^{\mu(A_j)\left(e^{-c_j}-1\right)} = e^{-\mu(1-e^{-f})} .$$

On en déduit les résultats souhaités. □

Proposition 80 (Formules de Campbell, cas intégrable) **:**
Soit désormais $f \in L^1(\mathbb{X}, \mu)$. Alors $\mathbb{E}[M(f)] = \mu(f)$ et $\mathbb{E}\left[e^{iM(f)}\right] = \exp\left(\mu\left(e^{if} - 1\right)\right)$.

Démonstration. Comme f est μ-intégrable, on remarque avant tout que $\mathbb{E}[M(|f|)] = \mu(|f|) < \infty$ donc f est presque-sûrement M-intégrable.

Par les mêmes arguments que précédemment, $M(f) = M(f^+) - M(f^-)$ est bien une variable aléatoire intégrable, d'espérance $\mathbb{E}[M(f)] = \mathbb{E}[M(f^+)] - \mathbb{E}[M(f^-)] = \mu(f^+) - \mu(f^-) = \mu(f)$.

Pour des fonctions étagées, on a la transformée de Fourier souhaitée par le même calcul que pour la transformée de Laplace précédente. En général, on approche f par des fonctions étagées f_n dans $L^1(\mu)$. On en déduit que $\mathbb{E}[|M(f) - M(f_n)|] \leq \mu(|f - f - n|) \to 0$. Quitte à extraire, $M(f_n)$ converge dans $L^1(\mathbb{P})$ et presque-sûrement vers $M(f)$. Rappelons que $|e^{ix} - e^{iy}| \leq |x - y|$, donc lorsque $f \in L^1(\mu)$, c'est aussi le cas de $e^{if} - 1$. Par convergence dominée on a :

$$\mathbb{E}\left[e^{iM(f)}\right] = \lim \mathbb{E}\left[e^{iM(f_n)}\right] = \exp\left(\lim \mu\left(e^{if_n} - 1\right)\right) ,$$

et $\left|\mu\left(e^{if_n} - 1\right) - \mu\left(e^{if} - 1\right)\right| = \left|\mu\left(e^{if_n} - e^{if}\right)\right| \leq \mu(|f_n - f|) \to 0$. On en déduit le résultat souhaité. □

Remarque 81 :
La loi de M est alors entièrement caractérisée par les variables $M(f)$ pour les fonctions étagées $f = \sum_{j=1}^{N} c_j \mathbb{1}_{A_j}$.

4.3.1 Propriétés de stabilité

Proposition 82 (Restriction d'une mesure) **:**
Si $M \sim MAP(\mu)$ et $Y \subset \mathbb{X}$ mesurable, alors $M(\cdot \cap Y) \sim MAP(\mu(\cdot \cap Y))$.

Proposition 83 (Passage à l'image) **:**
Si $f : \mathbb{X} \to \mathbb{Y}$ est mesurable, la transformation $f_* : \mathcal{M}^{s,+}(\mathbb{X}) \to \mathcal{M}^{s,+}(\mathbb{Y})$ est mesurable, et préserve les mesures atomiques. Dans ce cas, si $M \sim MAP(\mu)$ alors $f_* M \sim MAP(f_* \mu)$.

Proposition 84 (Stabilité par marquage) **:**

Soit I une variable aléatoire à valeurs dans $\overline{\mathbb{N}}$, et X_i des variables aléatoires (pas forcément indépendantes) telles que $M = \sum_{i=1}^{I} \delta_{X_i} \sim MAP(\mu)$.

Soient $(Y_i) \overset{\text{iid}}{\sim} \mu'$ des variables aléatoires sur un espace polonais $\mathbb{Y}$, indépendantes de I et de X. Posons $N = \sum_{i=1}^{I} \delta_{(X_i,Y_i)}$. Alors $N \sim MAP(\mu \otimes \mu')$.

Démonstration. Soit f une fonction mesurable sur $\mathbb{X} \otimes \mathbb{Y}$. On a alors :

$$\begin{aligned}
\mathbb{E}\big[e^{-N(f)}\big] &= \mathbb{E}\big[\mathbb{E}\big[e^{-N(f)}|I, X\big]\big] \\
&= \mathbb{E}\Bigg[\prod_{i\in I} \int e^{-f(X_i,y)}\,\mathrm{d}\mu'(y)\Bigg] \\
&= \mathbb{E}\Bigg[\exp\Bigg(-\sum_{i\in I}\Big(-\ln\Big(\int e^{-f(X_i,y)}\,\mathrm{d}\mu'(y)\Big)\Big)\Bigg)\Bigg].
\end{aligned}$$

Posons $F(x) = -\ln\big(\int e^{-f(x,y)}\,\mathrm{d}\mu'(y)\big)$. C'est bien une application mesurable et comme μ' est une mesure de probabilités, par l'inégalité de Jensen, on a $F(x) \geq \int f(x,y)\,\mathrm{d}\mu'(y) \geq 0$.

Par la formule de Campbell, on a :

$$\mathbb{E}\big[e^{-N(f)}\big] = \mathbb{E}\big[e^{-M(F)}\big] = \exp\big(-\mu\big(1-e^{-F}\big)\big) = \exp\bigg(-\int 1 - e^{-f}\,\mathrm{d}\mu\otimes\mu'\bigg),$$

et ce résultat caractérise la loi $MAP(\mu \otimes \mu')$. $\square$

5 Convergence de processus continus

Notre but est ici d'étudier la convergence en loi dans le cas où l'espace polonais est $\mathbb{X} = \mathcal{C}^0([0,1])$ ou bien $\mathbb{X} = \mathcal{C}^0(\mathbb{R}^+)$.

On montrera le théorème de Donsker, qui donne la convergence en loi d'une famille de processus continus vers le mouvement brownien. Pour cela, on doit pouvoir montrer la tension d'une famille de lois sur un espace $\mathbb{X}$ de fonctions continues, et donc mieux comprendre sa topologie.

5.1 Topologie des fonctions continues sur un métrique compact

Soient T un espace métrique compact et $\mathbb{X} = (\mathcal{C}^0(T, \mathbb{R}), \|.\|_\infty)$.

Proposition 85 :

L'espace $\mathbb{X}$ est polonais.

Démonstration. L'espace $\mathbb{X}$ est un espace de Banach. Si (f_n) est une suite de Cauchy, alors pour tout $t \in T$, $(f_n(t))$ en est une, converge vers $f(t)$. On peut alors établir la convergence uniforme de $f_n \overset{\|\cdot\|_\infty}{} f$. L'espace $\mathbb{X}$ est donc métrique complet.

En ce qui concerne la séparabilité de $\mathbb{X}$, on va utiliser le théorème de Stone-Weierstrass : si on a une algèbre $A \subset \mathbb{X}$ qui sépare les points, alors elle est dense dans $\mathbb{X}$. Reste à exhiber une telle algèbre dénombrable. Comme T est métrique compact, il existe une suite dénombrable dense (t_n). On peut considérer $A = \mathbb{Q}[f_i, i \in \mathbb{N}]$ où $f_i(x) = d(x, t_i)$, d'où le résultat. $\square$

5.1.1 Parties relativement compactes

Définition 86 (Module d'uniforme continuité) **:**

Pour $f \in \mathbb{X}$ et $\delta > 0$, on pose $\omega(f, \delta) = \sup\limits_{d(x,y) \leq \delta} |f(x) - f(y)|$.

Remarque 87 :

Dans le cas général, par compacité et le théorème de Heine, on a $f \in \mathbb{X}$ ssi f est uniformément continue ssi $\omega(f, \delta) \underset{\delta \to 0}{\longrightarrow} 0$.

Théorème 88 (Ascoli) **:**

Soit $\mathcal{F} \subset \mathcal{C}^0(T)$ avec T compact. Alors $\mathcal{F}$ est relativement compacte ssi :

1. $\sup\limits_{f \in \mathcal{F}} \|f\|_\infty = M < \infty$,

2. $\sup_{f \in \mathcal{F}} \omega(f, \delta) \xrightarrow[\delta \to 0]{} 0$.

Démonstration. Pour le sens direct, $\mathcal{F}$ est relativement compacte donc bornée pour la norme. En outre, si la seconde propriété n'est pas satisfaite, on peut trouver une suite (f_n) et $\varepsilon > 0$ tels que $\omega\big(f_n, \frac{1}{n}\big) \geq \varepsilon$. Quitte à extraire, par compacité, on a $f \in \overline{\mathcal{F}}$ limite de la suite. Pour tout $n \in \mathbb{N}$, on a x_n et y_n tels que $d(x_n, y_n) \leq \frac{1}{n}$ et $|f_n(x_n) - f_n(y_n)| \geq \varepsilon$. Quitte à extraire à nouveau, on a $x_n \to x$ et $y_n \to y$. On a $d(x, y) = \lim d(x_n, y_n) = 0$ donc $x = y$, mais $|f(x) - f(y)| = \lim |f_n(x_n) - f_n(y_n)| \geq \varepsilon$, une contradiction.

Réciproquement, supposons les deux points vérifiés. On fixe (t_k) dense dans T. Par extraction diagonale, de toute suite $(f_n) \in \mathcal{F}^{\mathbb{N}}$, on peut extraire une sous-suite qui converge en tout point t_k. On note $f(t_k)$ la limite. C'est une fonction *partielle*, bornée par M aussi.

Montrons alors que (f_n) est de Cauchy. Soit donc $\varepsilon > 0$. Il existe δ tel que $\sup_{n \in \mathbb{N}} \omega(f, \delta) \leq \varepsilon$. Par compacité de T, on a un rang K tel que $T \subset \bigcup_{i=1}^{K} B(t_i, \delta)$. À partir d'un rang N, pour tout $1 \leq i \leq K$, on a $|f_n(t_i) - f_m(t_i)| \leq \varepsilon$. Pour tout $t \in T$, on a $t \in B(t_i, \delta)$ pour un certain i, et donc $|f_n(t) - f_m(t)| \leq |f_n(t_i) - f_m(t_i)| + |f_n(t) - f_n(t_i)| + |f_m(t) - f_m(t_i)| \leq 3\varepsilon$. En conséquence, $\|f_n - f_m\|_\infty \leq 3\varepsilon$, la suite (f_n) est de Cauchy, donc convergente. □

5.2 Espace $\mathcal{C}^0(T)$ dans le cas σ-localement compact

Définition 89 (Espace localement compact) **:**
T est localement compact si tout point $x \in T$ admet un voisinage compact.

Définition 90 (Espace σ-compact) **:**
T est σ-compact si $T = \bigcup_{n \in \mathbb{N}} K_n$ est union dénombrable de compacts K_n.

Lemme 91 :
Soit T un espace σ-localement compact. On a les propriétés suivantes :

1. T est séparable,
2. T a une base dénombrable d'ouverts relativement compacts,
3. T a une famille *exhaustive* de compacts (K_n) : la famille est croissante, et pour tout compact $K \subset T$, on a $K \subset K_n$ à partir d'un rang.

Démonstration. Admis. □

Définition 92 (Topologie sur $\mathcal{C}^0(T)$) **:**
On munit $\mathcal{C}^0(T)$ de la convergence uniforme sur tout compact. Une base d'ouverts de cette

topologie est donnée par les :

$$V_{f,\varepsilon,K} = \left\{g \in \mathcal{C}^0(T),\ \|f-g\|_{\infty,K} < \varepsilon\right\},$$

avec K compact, $\varepsilon > 0$ et $f \in \mathcal{C}^0(T)$.

De façon équivalente, on peut se restreindre aux $\varepsilon = \frac{1}{n}$ et aux compacts exhaustifs (K_n).

Théorème 93 :
L'espace $\mathcal{C}^0(T)$ est polonais pour cette topologie, avec $d(f,g) = \sum\limits_{n\in\mathbb{N}} \frac{1}{2^n} \min\Big(1, \|f-g\|_{\infty,K_n}\Big)$.

Démonstration. L'application d définit naturellement une métrique sur l'espace. On peut vérifier que d induit la topologie précédente. En outre, toute suite de Cauchy pour cette distance induit des suites de Cauchy dans les espaces $\mathcal{C}^0(K_n)$, et on vérifie que la famille de limites obtenues se relève de façon unique en $f \in \mathcal{C}^0(T)$.

Pour tout rang fini, $\mathcal{C}^0(K_n)$ admet une famille dénombrable dense $(g_{n,m})$. Par le théorème de Tietze-Urysohn, on peut prolonger chaque $g_{n,m}$ en $f_{n,m} \in \mathcal{C}^0(T)$, avec les mêmes bornes sur la fonction, et on vérifie alors que la famille $(f_{n,m})_{n,m\in\mathbb{N}}$ est dense. □

On veut désormais savoir quelles sont les parties relativement compactes de cet espace.

Théorème 94 :
La famille $\mathcal{F} \subset \mathcal{C}^0(T)$ est relativement compacte ssi pour tout $K \subset T$ compact, la famille $\mathcal{F}_K = \{f|_K,\ f \in \mathcal{F}\}$ est relativement compacte dans $\mathcal{C}^0(K)$.

Démonstration. L'implication directe se fait par continuité des projections $\pi_K : \mathcal{C}^0(T) \to \mathcal{C}^0(K)$. Pour le sens réciproque, il suffit de constater que, si on a une suite $(f_n) \in \mathcal{F}^{\mathbb{N}}$, par extraction diagonale, on peut obtenir une suite convergente dans chaque espace $\mathcal{C}^0(K_n)$, faire un recollement des limites pour obtenir $f \in \mathcal{C}^0(T)$, et enfin vérifier que $d\big(f_{\sigma(n)}, f\big) \to 0$. □

Remarque 95 :
L'espace $\mathcal{C}^0(T)$ est en fait la *limite projective* des espaces $\mathcal{C}^0(K_n)$.

5.3 Lois sur $\mathcal{C}^0(T)$

On considère par la suite T un espace σ-localement compact.

Définition 96 (Lois fini-dimensionnelles) **:**
Soit μ une mesure de probabilité sur $\mathcal{C}^0(T)$. Pour tous $t_1, \dots, t_d \in T$, μ induit une mesure $\mu_{t_1,\dots,t_d}$ sur $\mathbb{R}^d$, la mesure image de μ par la projection continue $\pi : f \mapsto (f(t_1), \dots, f(t_d))$.

Ces mesures images sont les lois fini-dimensionnelles de μ.

Proposition 97 :
Si μ et ν ont les mêmes lois fini-dimensionnelles, alors $\mu = \nu$.

Démonstration. On veut montrer que $\mu = \nu$ sur la tribu borélienne de $\mathcal{C}^0(T)$.

On introduit $\mathcal{P}$ le tiré en arrière de la tribu cylindrique sur $\mathbb{R}^T$ via l'injection canonique $i : f \in \mathcal{C}^0(T) \mapsto (f(t))_{t\in T} \in \mathbb{R}^T$. Comme $\mathcal{P}$ est engendré par les cylindres finis, qu'on peut exprimer par des boréliens, on constate que $\mathcal{P} \subset \mathcal{B}(\mathcal{C}(T))$.

Comme μ et ν coincident sur les cylindres finis par hypothèse, $\mu|_{\mathcal{P}} = \nu|_{\mathcal{P}}$.

Reste à vérifier l'égalité entre ces deux tribus pour conclure. Il nous suffit de justifier que $V_{f,\varepsilon,K} \in \mathcal{P}$. Cela découle de l'écriture :

$$V_{f,\varepsilon,K} = \bigcup_{n\geq 1} \bigcap_{k\in\mathbb{N}} \left\{ g,\ |f(t_k) - g(t_k)| \leq \frac{n}{n+1}\varepsilon \right\}.$$

□

Corollaire 98 :
Sous réserve d'existence d'un pré-MB à trajectoires continues, il est unique en loi sur $\mathcal{C}^0(\mathbb{R}^+)$.

5.4 Critères de tension

Montrons d'abord qu'on peut se ramener au cas compact.

Proposition 99 :
Soient T σ-localement compact, et $\mathcal{P} \subset \mathcal{M}^1(\mathcal{C}^0(T))$ une famille de mesures de probabilité. Alors $\mathcal{P}$ est tendue (relativement compacte) ssi pour tout K compact, les restrictions $\mathcal{P}_K$ sont tendues.

Démonstration. Le sens direct découle de la continuité de la projection $\pi_K : \mathcal{C}^0(T) \to \mathcal{C}^0(K)$, qui induit naturellement $\pi_K^* : \mathcal{M}^1(\mathcal{C}^0(T)) \to \mathcal{M}^1(\mathcal{C}^0(K))$ continue.

Pour le sens réciproque, on fixe une suite de compacts (K_n) de T exhaustive. Pour tout $\varepsilon > 0$, pour tout rang $n \in \mathbb{N}$, par tension de $\mathcal{P}_{K_n}$, on a une famille $A_{K_n} \subset \mathcal{C}^0(K_n)$ relativement compacte, telle que $\sup\limits_{\mu\in\mathcal{P}_{K_n}} \mu\big(A_{K_n}^c\big) \leq \frac{\varepsilon}{2^n}$.

On pose alors $A = \bigcap\limits_{n\in\mathbb{N}} \pi_{K_n}^{-1}(A_{K_n})$. C'est un ensemble relativement compact car chacune des projections sur un compact K est *moralement* incluse dans une famille A_{K_n} par exhaustion, et que cette famille est relativement compacte. En outre, $\sup_{\mu \in \mathcal{P}} \mu(A^c) \leq \sum_{n \in \mathbb{N}} \sup_{\mu \in \mathcal{P}_{K_n}} \mu\left(A^c_{K_n}\right) \leq 2\varepsilon$. En conclusion, $\mathcal{P}$ est tendue. $\square$

En particulier, pour vérifier qu'une famille de lois sur $\mathcal{C}^0(\mathbb{R}^+)$ est relativement compacte, il suffira de montrer que les familles de restrictions à $\mathcal{C}^0([0,t])$ sont tendues pour $t \in \mathbb{R}^+$.

Théorème 100 :

Soit T compact. La famille $\mathcal{P} \subset \mathcal{M}^1(\mathcal{C}^0(T))$ est tendue ssi :

1. $\sup_{\mu \in \mathcal{P}} \mu(\{f,\ \|f\|_\infty \geq M\}) \underset{M \to \infty}{\longrightarrow} 0$,
2. $\forall \varepsilon > 0,\ \sup_{\mu \in \mathcal{P}}(\{f,\ \omega(f,\delta) \geq \varepsilon\}) \underset{\delta \to 0}{\longrightarrow} 0$.

Démonstration. Pour le sens direct, on fixe ε. Par tension de la famille, on a un relativement compact $A_\varepsilon \subset \mathcal{C}^0(T)$ tel que $\mu(A^c_\varepsilon) \leq \varepsilon$. En outre, en appliquant Ascoli à cette famille A_ε, les propriétés considérées sont toujours vraies à partir d'un rang. Ce résultat étant vrai pour tout ε, on en déduit la convergence souhaitée.

Pour le sens réciproque, considérons encore $\varepsilon > 0$. Il existe M tel que le premier critère est majoré par $\frac{\varepsilon}{2}$. Pour tout $n \in \mathbb{N}$, il existe δ_n tel que le second critère est majoré par $\frac{\varepsilon}{2^{n+1}}$ (avec $\varepsilon = \frac{1}{n}$ dans la formule en question, sans rapport avec le ε de la preuve).

On pose alors :

$$A = \left\{ f,\ \|f\|_\infty \leq M,\ \forall n \geq 1,\ \omega(f,\delta_n) \leq \frac{1}{n} \right\}.$$

C'est un ensemble relativement compact par Ascoli, et pour tout $\mu \in \mathcal{P}$, on a $\mu(A^c) \leq \varepsilon$. $\square$

Corollaire 101 :

En exploitant le module d'uniforme continuité sur le compact, on peut remplacer de façon équivalente le premier critère ci-dessus par :

$$\sup_{\mu \in \mathcal{P}} \mu(\{f,\ |f(x)| \geq M\}) \underset{M \to \infty}{\longrightarrow} 0\,,$$

pour un point $x \in T$ fixé quelconque.

Cette caractérisation équivalente est peu pratique à utiliser car le critère est difficile à montrer. En pratique, on va utiliser des conditions suffisantes, plus simples à mettre en évidence, qui impliquent la caractérisation précédente.

Théorème 102 (Critère de Kolmogorov-Chentsov) **:**

On se place ici sur l'espace métrique compact $[0,T] \subset \mathbb{R}^+$.

Soit (X_n) une suite de variables dans $\mathcal{C}^0([0,T])$. Si la suite de variables réelles $(X_n(0))$ est tendue, et que $\mathbb{E}[|X_n(t) - X_n(s)|^a] \leq C|t-s|^b$ (avec $a > 0$, $b > 1$ et $C > 0$) pour tout $n \in \mathbb{N}$, alors (X_n) est tendue.

Démonstration. Faisons ici la démonstration dans le cas $a \geq 1$.

La première hypothèse permet directement de satisfaire la première condition. Pour satisfaire la seconde condition, quitte à faire un changement d'échelle, on peut supposer que $T = 1$.

On introduit les variables $\xi_{n,k} = \sup_{i=0}^{2^k-1} \left|X_n\left(\frac{i+1}{2^k}\right) - X_n\left(\frac{i}{2^k}\right)\right|$. Par hypothèse, on obtient :

$$\mathbb{E}\left[\xi_{n,k}^a\right] \leq \sum_{i=0}^{2^k-1} \mathbb{E}\left[\left|X_n\left(\frac{i+1}{2^n}\right) - X_n\left(\frac{i}{2^n}\right)\right|^a\right] \leq 2^k \times C\left(\frac{1}{2^k}\right)^b = C2^{k(1-b)}.$$

Si $|s-t| \leq \frac{1}{2^k}$, alors $|X_n(s) - X_n(t)| \leq 2\sum_{l=k}^{\infty} \xi_{n,l}$. En conséquence :

$$\begin{aligned}
\mathbb{E}\left[\sup_{|s-t|\leq\frac{1}{2^k}} |X_n(s) - X_n(t)|^a\right]^{\frac{1}{a}} &\leq 2\left\|\sum_{l\geq k} E[\xi_{n,l}]\right\|_a \\
&\leq 2\sum_{l\geq k} \|\xi_{n,l}\|_a \\
&\leq C^{\frac{1}{a}} \sum_{l\geq k} \left(\frac{1}{2^{\frac{b-1}{a}}}\right)^l \\
&= O\left(\frac{1}{2^{k\frac{b-1}{a}}}\right).
\end{aligned}$$

En d'autres termes,$\mathbb{E}\left[\omega\left(X_n, \frac{1}{2^k}\right)^a\right] = O\left(\frac{1}{2^{k(b-1)}}\right)$.

Par inégalité de Markov, $\mathbb{P}\left(\omega\left(X_n, \frac{1}{2^k}\right) \geq \varepsilon\right) = O\left(\frac{1}{\varepsilon^a 2^{k(b-1)}}\right)$, uniformément en n, donc le second critère est vérifié. □

Remarque 103 :

Dans le cas $0 < a < 1$, on doit exploiter des relations du type $\mathbb{E}[|X+Y|^a] \leq \mathbb{E}[|X|^a] + \mathbb{E}[|Y|^a]$ plutôt que des propriétés de $\|.\|_a$.

Théorème 104 (Donsker) **:**

Soient $(A_n)_{n\in\mathbb{N}}$ des variables iid centrées réduites. On définit le processus continu et affine par morceaux :

$$C^{(n)}(t) = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{\lfloor nt\rfloor} A_i + \{nt\}A_{\lfloor nt+1\rfloor}\right),$$

où $\{x\} = x - \lfloor x\rfloor$ est la partie fractionnaire. Alors $C^{(n)} \rightharpoonup B$ converge en loi vers le mouvement brownien B, l'unique processus continu qui est un pré-MB.

Démonstration. On va ici faire une preuve dans le cas de variables L^4, avec $\mathbb{E}[A^4] = M_4 < \infty$.

On va d'une part montrer la convergence des lois fini-dimensionnelles, et d'autre part la tension de la famille via le critère de Kolmogorov.

On peut décomposer $C^{(n)} = D^{(n)} + R^{(n)}$, avec $D^{(n)}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} A_i$. Pour le terme de reste, on a naturellement $R^{(n)}(t) \xrightarrow{\mathbb{P}} 0$. On a alors :

$$\begin{array}{rcl}\left(C^{(n)}(t_1), C^{(n)}(t_2) - C^{(n)}(t_1), \ldots\right) & = & \left(D^{(n)}(t_1), D^{(n)}(t_2) - D^{(n)}(t_1), \ldots\right) \\ & + & \left(R^{(n)}(t_1), R^{(n)}(t_2) - R^{(n)}(t_1), \ldots\right).\end{array}$$

Le vecteur en R converge en probabilité vers 0. Les coordonnées de D sont alors indépendantes deux à deux par construction, et chacune converge en loi vers la gaussienne souhaitée par le TCL. On en déduit bien la convergence des lois fini-dimensionnelles de $\left(C^{(n)}\right)$ vers un pré-MB.

On a toujours $C^{(n)}(0) = 0$, donc la première partie du critère de Kolmogorov est satisfaite.

Reste à justifier l'autre partie du critère pour obtenir la tension. On va le montrer grâce aux moments d'ordre 4, avec $a = 4$ et $b = 2$. Soient donc $s < t \in \mathbb{R}^+$.

Si $t - s > \frac{1}{n}$, on a $C^{(n)}(t) - C^{(n)}(s) = \sum_{k=\lfloor ns \rfloor + 2}^{\lfloor nt \rfloor} \frac{A_k}{\sqrt{n}} + \frac{\{nt\}}{\sqrt{n}} A_{\lfloor nt \rfloor + 1} + \frac{1 - \{ns\}}{\sqrt{n}} A_{\lfloor ns \rfloor + 1}$. Les variables (A_k) sont indépendantes, donc on peut majorer le moment d'ordre 4 de la somme par :

$$\sum_{k=\lfloor ns \rfloor + 2}^{\lfloor nt \rfloor} \frac{M_4}{n^2} + 2\frac{M_4}{n^2} + \binom{\lfloor nt \rfloor - \lfloor ns \rfloor}{2} \frac{\mathbb{E}[A^2]^2}{n^2} = O\left(\frac{\left(\lfloor nt \rfloor - \lfloor ns \rfloor\right)^2}{n^2}\right) = O\left((t-s)^2\right),$$

uniformément en n.

Si $t - s \leq \frac{1}{n}$ on distingue deux cas, selon si t et s tombent dans la même subdivision ou dans deux subdivisions consécutives. Si on tombe dans le même pas de la subdivision, on a directement $C^{(n)}(t) - C^{(n)}(s) = \frac{n(t-s)}{\sqrt{n}} A_k \leq \sqrt{t-s} A_k$, donc on majore le moment d'ordre 4 par $M_4 (t-s)^2$. Pour deux subdivisions consécutives, on conclut par un calcul similaire.

On peut donc appliquer le critère de Kolmogorov. La suite est tendue, admet une unique valeur d'adhérence par convergence des moments fini-dimensionnels, donc converge en loi vers B. Ceci garantit au passage *l'existence* de B, dont on avait jusqu'ici uniquement justifié l'unicité sous réserve d'existence. $\square$

Remarque 105 :

Pour obtenir une convergence en loi, on a réellement besoin de la tension de la famille, la convergence des lois fini-dimensionnelles ne suffit pas.

Considérons par exemple $X_n(t) = \min(1, n \times |U - t|)$ où $U \sim \mathcal{U}([0,1])$. La limite des lois fini-dimensionnelles est toujours le vecteur constant égal à 1. Si X_n convergeait en loi, sa limite serait $X = 1$. Cependant, $f \mapsto \min(f)$ est une application continue, et $\min(X) = 1$ n'est pas la limite en loi de $\min(X_n) = 0$.

5.5 Applications du théorème de Donsker

Lemme 106 (Critère raffiné de convergence en loi) **:**

Soit π_n une suite de mesures de probabilité sur un espace polonais $\mathbb{X}$, qui converge en loi vers π, et $h : \mathbb{X} \to \mathbb{Y}$ une transformation mesurable.

Si il existe $A \subset \mathbb{X}$ tel que $\pi(A) = 1$ et h continue sur A, alors $h_*\pi_n$ converge en loi vers $h_*\pi$.

Démonstration. Soit D l'ensemble des points de discontinuité de h.

Si F est fermé, alors $D^c \cap \overline{h^{-1}(F)} \subset h^{-1}(F)$. En effet, si $x \in \overline{h^{-1}(F)}$ n'est pas dans D, on peut en particulier approcher x par $(x_n) \in h^{-1}(F)^{\mathbb{N}}$. Comme x est un point de continuité et que F est fermé, en particulier $f(x) = \lim f(x_n) \in F$, d'où l'inclusion.

On peut alors appliquer le théorème de Portemanteau. Pour tout fermé $F \subset \mathbb{Y}$, on a $\overline{\lim}\, h_*\pi_n(F) = \overline{\lim}\, \pi_n(h^{-1}(F)) \leq \overline{\lim}\, \pi_n\left(\overline{h^{-1}(F)}\right) \leq \pi\left(\overline{h^{-1}(F)}\right) = \pi(h^{-1}(F))$, d'où enfin la convergence en loi souhaitée. □

Proposition 107 (Loi de l'arcsinus) **:**

Soit B un MB usuel, et $T = \sup\{t \in [0,1], X_t = 0\}$. Alors T a la densité $\frac{1}{\pi\sqrt{x(1-x)}}$, ou de façon équivalente :

$$\mathbb{P}(T \leq t) = \frac{2}{\pi} \arcsin\left(\sqrt{t}\right).$$

Démonstration. Pour ce faire, on peut étudier le dernier passage en 0 d'une somme de variables $U_n \overset{\text{iid}}{\sim} \mathcal{U}(\{\pm 1\})$, de façon combinatoire. On pose $D_{2n} = \max\left\{k \in [\![0, 2n]\!], X_k := \sum_{i=1}^{k} U_i = 0\right\}$, et on s'intéresse à la loi de $\frac{D_{2n}}{2n}$.

Lorsque $D_{2n} = 2j$, on a $X_{2j} = 0$, puis on part d'une nouvelle marche aléatoire initialisée en 0, pour laquelle on veut que le *premier* temps de retour en 0 soit *strictement* supérieur à $2(n-j)$. Par propriété de Markov faible :

$$\mathbb{P}(D_{2n} = 2j) = \mathbb{P}(X_{2j} = 0) \times \mathbb{P}\left(X_1 \neq 0, \ldots, X_{2(n-j)} \neq 0\right).$$

Par un principe de réflexion, appliqué à l'étude du premier temps de retour, on peut montrer que le terme de droite est égal à $\mathbb{P}\left(X_{2(n-j)} = 0\right)$.

Admettons pour l'instant que $\mathbb{P}(T \leq t) = \lim_{n\to\infty} \mathbb{P}\left(\frac{D_{2n}}{2n} \geq t\right)$. Alors en utilisant la bonne version de Stirling, on obtient des sommes de Riemann, qui approchent l'intégrale $\int_0^t \frac{1}{\pi\sqrt{x(1-x)}}\,\mathrm{d}x$, d'où la loi souhaitée sur T.

A priori, dans le cas général, l'application h qui à f associe son dernier passage en 0 n'est pas continue. Par exemple, $f_n(x) = \frac{x}{n}$ converge vers 0 uniformément, mais $h(0) = 1$ n'est pas la limite des $h(f_n) = 0$. On peut cependant montrer la convergence en loi de $h\big(C^{(n)}\big)$ ver $h(B)$, où $C^{(n)}$ sont les marches aléatoires renormalisées, pour appliquer le lemme précédent.

En effet, on peut montrer que le mouvement brownien B est presque-sûrement sans zéros isolés sur $[0,1]$. En particulier, son dernier zéro T est un point d'accumulation à gauche, dans l'intervalle $]0,1[$, et prend des valeurs du signe opposé au signe (constant) sur $]h(x),1[$, et ce arbitrairement près de T. Dans une boule $B_{\|\cdot\|_\infty}(f,\delta)$ assez petite, on peut en particulier garantir que $T(g)$ est ε-proche de $T(f)$, ce qui garantit enfin la continuité presque-sûre de h en $B(\omega)$. □

On peut également montrer le résultat directement grâce aux symétries du mouvement brownien, sans passer par les symétries de la marche aléatoire, car alors $T \leq u$ ssi le premier passage en 0 de B après u est au moins 1.

Proposition 108 :
Soit B le MB usuel. On pose $S_t = \sup\limits_{u \in [0,t]} B_u$. Alors $(2S - B) \overset{\mathrm{d}}{=} \mathrm{Bessel}(3)$ a la loi de la norme d'un mouvement brownien sur $\mathbb{R}^3$.

6 Processus discontinus

Définition 109 :
On pose $\mathcal{D}(I)$ l'espace des fonctions càdlàg sur I.

On aimerait trouver une topologie sur $\mathcal{D}(I)$ qui rend l'espace polonais. On commence par étudier le cas $I = [0,T]$. On a précédemment caractérisé la continuité sur un compact par son module d'uniforme continuité. On va par la suite introduire une notion similaire dans le cas de processus discontinus.

6.1 Caractérisation des fonctions càdlàg

Définition 110 :
Posons $w(f,J) = \sup\limits_{s,t \in J} |f(s) - f(t)|$ sur l'intervalle J, et :

$$w'(f,\delta) = \inf_{\substack{0=t_0<t_1<\cdots<t_d=T \\ t_{i+1} \geq t_i+\delta}} \max_{i=1}^{d} w(f,[t_{i-1},t_i[) \,.$$

Autrement dit, avec w on regarde l'amplitude des variations sur un intervalle J, et avec w' on regarde l'amplitude minimale parmi les intervalles induits par des subdivisions de pas *au moins* égal à δ. Par définition, $\delta \mapsto w'(f,\delta)$ est décroissante.

Proposition 111 :

On a f càdlàg ssi $w(f,\delta) \underset{\delta\to 0}{\longrightarrow} 0$.

Démonstration. Pour le sens direct, il nous suffit de montrer que pour tout ε, il existe une subdivision ε-adaptée, pour laquelle $w(f,[t_{i-1},t_i[) \leq \varepsilon$ sur chacun des intervalles. Soit U le supremum des $t \in [0,T]$ pour lesquels il existe une subdivision de $[0,t]$ qui est ε-adaptée. En particulier, on sait que $U > 0$ par continuité à droite en 0. Si $U < T$, on peut trouver un $\delta > 0$ pour lequel $U + \delta < T$ et $w(f,[U,U+\delta[) < \varepsilon$ par continuité à droite, et $w(f,[U-\delta,U[) < \varepsilon$ par limite à gauche. Par définition de U en tant que supremum, il existe une subdivision ε-adaptée qui dépasse $U-\delta$. Quitte à tronquer cette subdivision en $U-\delta$, et à ajouter les intervalles $[U-\delta,U[$ et $[U,U+\delta[$, on obtient une nouvelle subdivision ε-adaptée, qui atteint $U+\delta > U$. C'est une contradiction, donc nécessairement $U = T$.

Réciproquement, si $w(f,\delta) \to 0$, on a en particulier des subdivisions ε-adaptées pour tout $\varepsilon > 0$. En partant d'une subdivision ε-adaptée, pour tout $s \in [0,T]$, il existe donc $\delta(s,\varepsilon)$ tel que $\sup_{t\in[s,s+\delta]} |f(t) - f(s)| \leq \varepsilon$, et $\sup_{t,t'\in]s-\delta,s[} |f(t) - f(t')| \leq \varepsilon$. Si s est un des points de la subdivision, prendre le pas de la subdivision convient, sinon on considère la distance de s à ladite subdivision.

Ceci étant vrai pour tout $\varepsilon > 0$, f est bien continue à droite en s, et vérifie une propriété de Cauchy sur un voisinage à gauche privé de s, donc est limitée à gauche en s : f est càdlàg. □

Corollaire 112 :

Si f est càdlàg sur $[0,T]$, alors $\|f\|_\infty < \infty$ et f admet un nombre au plus dénombrable de discontinuités.

Démonstration. Pour le premier point, remarquons que en prenant une subdivision ε-adaptée, alors $\|f\|_\infty \leq \max_{i=0}^d |f(t_i)| + \varepsilon < \infty$.

Pour le second point, on montre plus fortement que pour tout seuil ε, il existe un nombre fini de discontinuités plus grandes que ε. En effet, si on note $D_{f,\varepsilon}$ ces points de discontinuité, alors $D_{f,\varepsilon}$ est nécessairement inclus dans l'ensemble fini des pas d'une subdivision $\frac{\varepsilon}{2}$-adaptée par exemple. On a alors l'ensemble des discontinuités $D_f = \bigcup_{n\in\mathbb{N}} D_{f,\frac{1}{n}}$ fini ou dénombrable. □

Définition 113 (Distance sur $\mathcal{D}(I)$) **:**

On note $Homeo^+$ l'ensemble des homéomorphismes strictement croissants sur $[0,T]$. Cet ensemble est en particulier stable par passage à l'inverse.

On pose alors $d(f,g) = \min_{\psi\in Homeo^+} \max(\|f - g\circ\psi\|_\infty, \|\psi - \mathrm{Id}\|_\infty)$.

Lemme 114 :
L'application d est une distance sur $\mathcal{D}([0,T])$.

Démonstration. Pour la symétrie, il suffit de remarquer que composer par un homéomorphisme ne change pas la norme $\|.\|_\infty$, donc $\|\psi - \mathrm{Id}\|_\infty = \|\psi^{-1} - \mathrm{Id}\|_\infty$ et :

$$\|f - g \circ \psi\|_\infty = \left\|f \circ \psi^{-1} - g\right\|_\infty .$$

Quitte à utiliser la bijection $\psi \mapsto \psi^{-1}$ sur $Homeo^+$, on passe alors de $d(f,g)$ à $d(g,f)$, d'où la symétrie.

Pour l'inégalité triangulaire, on remarque par exemple que :

$$\|\psi_2 \circ \psi_1 - \mathrm{Id}\|_\infty \leq \|\psi_2 \circ \psi_1 - \psi_1\|_\infty + \|\psi_1 - \mathrm{Id}\|_\infty = \|\psi_2 - \mathrm{Id}\|_\infty + \|\psi_1 - \mathrm{Id}\|_\infty ,$$

et de même pour l'autre terme. Comme le maximum de la somme est plus petit que la somme des maxima, en prenant ψ_1 et ψ_2 qui approchent $d(f,g)$ et $d(g,h)$ à ε près, on peut alors majorer par :

$$d(f,h) \leq \max(\|f - h \circ \psi_2 \circ \psi_1\|_\infty, \|\psi_2 \circ \psi_1 - \mathrm{Id}\|_\infty) \leq d(f,g) + \varepsilon + d(g,h) + \varepsilon ,$$

d'où l'inégalité voulue pour $\varepsilon \to 0$.

Si $d(f,g) = 0$, alors naturellement, on a $\psi_n \in Homeo^+$ qui approche $d(f,g)$ à $\frac{1}{n}$ près. Par continuité à droite de f en t, pour tout $\varepsilon > 0$, on a $\delta > 0$ tel que $\|f - f(t)\|_{\infty,[t,t+\delta[} \leq \varepsilon$, et de même pour g.

On peut supposer qu'on a $\psi_n(t) \geq t$ une infinité de fois. Si ce n'est pas le cas, quitte à remplacer la suite (ψ_n) par (ψ_n^{-1}) et à intervertir f et g, on obtient la propriété. Quitte à extraire, on se ramène donc à $\psi_n(t) \geq t$ pour tout rang n.

Dans ce cas, à partir d'un certain rang, $\psi_n(t) \in [t, t+\delta[$ par convergence uniforme vers l'identité, et donc $|g \circ \psi_n(t) - g(t)| \leq \varepsilon$. Par convergence uniforme de $g \circ \psi_n$ vers f, on a finalement $|f(t) - g(t)| \leq \|f - g \circ \psi_n\|_\infty + |g \circ \psi_n(t) - g(t)| \leq 2\varepsilon$ à partir d'un rang, d'où $f(t) = g(t)$. □

Remarque 115 :
Sur cet espace, l'addition n'est *pas* continue.

Démonstration. Avec $T = 2$, $f_n = \mathbb{1}_{[0,\frac{n-1}{n}[}$ et $g_n = \mathbb{1}_{[\frac{n+1}{n},2]}$, on vérifie que $f_n \to f = \mathbb{1}_{[0,1[}$ et $g_n \to g = \mathbb{1}_{[1,2]}$. Cependant, $f + g = 1$ et $d(f_n + g_n, f + g) \geq 1$ car $f_n + g_n$ visite forcément 0 à un moment où à un autre. □

Remarque 116 :
La distance d ne rend pas l'espace complet.

Démonstration. Soit $f_n = \mathbb{1}_{[0,\frac{1}{2n}[}$. C'est une suite de Cauchy pour d, mais elle ne converge pas. En effet, on peut montrer que la convergence pour cette topologie implique la convergence L^1, et $\|f_n\|_1 \to 0$, mais $d(f_n, 0) \geq 1$. $\square$

Proposition 117 :

L'espace $(\mathcal{D}([0,T]), d)$ est polonais.

Démonstration. Commençons par montrer la séparabilité de la topologie. Pour cela, on peut toujours approcher f par des fonctions en escalier, à valeurs dans $\mathbb{Q}$, et avec un nombre fini de sauts dans $\mathbb{Q} \cap [0,T]$.

Pour la complétude, considérons $d'(f,g) = \inf\limits_{\varphi \in Homeo^+} \max(\|f - g\circ\psi\|_\infty, \|\psi\|_H)$ avec la norme opérateur $\|\psi\|_H = \sup\limits_{s<t\in[0,T]} \left|\ln\left(\frac{\psi(t)-\psi(s)}{t-s}\right)\right| \in \overline{\mathbb{R}^+}$, qui maximise le logarithme de la pente des cordes de la courbe. On peut vérifier que $\|.\|_H$ induit une distance $d_H(\psi_1,\psi_2) = \left\|\psi_1 \circ \psi_2^{-1}\right\|$ sur $Homeo^+$, dont on déduit que d' est bien une distance sur $\mathcal{D}([0,T])$.

On va d'une part montrer que d et d' correspondent à la même topologie, et que d' est complète. Notons qu'on a d'abord introduit d car cette distance est plus *naturelle*, plus aisée à manipuler en pratique.

Notons que $|\psi(t) - t| = |t| \times \left|\frac{\psi(t)-\psi(0)}{t-0} - 1\right|$ donc $\|\psi - \mathrm{Id}\|_\infty \leq T\left(e^{\|\psi\|_H} - 1\right)$. En conséquence, si $d'(f_n, f) \underset{n\to\infty}{\longrightarrow} 0$, alors $d(f_n, f) \underset{n\to\infty}{\longrightarrow} 0$.

Le sens réciproque, assez naturellement, est moins aisé, et on n'a pas de majoration explicite, ce qui nous donnerait une équivalence entre les suites de Cauchy pour d et d', ce qu'on cherche précisément à éviter.

Supposons que $d(f_n, f) \underset{n\to\infty}{\longrightarrow} 0$. On a des homéomorphismes croissants ψ_n, qui convergent uniformément vers l'identité, tels que $\|f_n - f\circ\psi_n\|_\infty \to 0$.

Soit $(t_i)_{1\leq i\leq d}$ une partition ε-adaptée pour f. On pose $\Theta_n : \psi_n^{-1}(t_i) \mapsto t_i$ qu'on prolonge de façon affine par morceaux. On a naturellement $\Theta_n \in Homeo^+$, et comme elle est affine par morceaux, maximiser le logarithme des pentes des cordes revient juste à prendre le maximum des logarithmes des pentes de chaque morceau affine. En clair :

$$\|\Theta_n\|_H = \max_{i=1}^{d} \left|\ln\left(\frac{t_{i+1} - t_i}{\psi_n^{-1}(t_{i+1}) - \psi_n^{-1}(t_i)}\right)\right| \underset{n\to\infty}{\longrightarrow} 0\,,$$

car on a la convergence de $\psi_n(t)$ vers t ponctuelle sur chacun des pas de la subdivision, donc a fortiori convergence des pentes vers 1 sur chacun des d morceaux affines.

On a alors $\|f_n - f\circ\Theta_n\|_\infty \leq \|f_n - f\circ\psi_n\|_\infty + \|f\circ\psi_n - f\circ\Theta_n\|_\infty$. Le terme de gauche tend vers 0 par hypothèse. Pour tout $s \in [\psi_n^{-1}(t_i), \psi_n^{-1}(t_{i+1})[$, on a $\psi_n(s), \Theta_n(s) \in [t_i, t_{i+1}[$ et donc $|f\circ\psi_n(s) - f\circ\Theta_n(s)| \leq \varepsilon$ car (t_i) est ε-adaptée. Le terme de droite est donc majoré par ε.

En conséquence, si $d(f_n, f) \to 0$ alors $d'(f_n, f) \to 0$. Les deux distances définissent bien la même notion de convergence, donc la même topologie.

Reste à justifier la complétude de d'. Soit $(f_n)_{n\in\mathbb{N}}$ de Cauchy pour d'. Quitte à extraire g de f, on suppose que $d(g_n, g_{n+1}) < \frac{1}{2^n}$. Pour tout n, on a $\Theta_n \in Homeo^+$ telle que $\|\Theta_n\|_H < \frac{1}{2^n}$.

On pose $\psi_{k,n} = \Theta_k \circ \cdots \circ \Theta_{k+n}$. À k fixé, on vérifie que $(\psi_{k,n})_{n\in\mathbb{N}}$ est une suite de Cauchy pour $\|.\|_\infty$, donc converge uniformément vers ψ_k continue, et $\|\psi_k\|_H \leq \frac{1}{2^k} < \infty$, donc $\psi_k \in Homeo^+$. En outre, la suite converge vers l'identité, et $\Theta_k \circ \psi_{k+1} = \psi_k$.

On donc $\|g_k \circ \psi_k - g_{k+1} \circ \psi_{k+1}\|_\infty \leq \frac{1}{2^k}$. La suite $(g_k \circ \psi_k)_{k\in\mathbb{N}}$ converge donc uniformément vers une limite g. En particulier, g est càdlàg en tant que limite uniforme de fonctions càdlàg. Dans ce cas, on a $d'(g_k, g) \to 0$, d'où finalement la convergence de la suite de Cauchy initiale. □

6.2 Théorème d'Ascoli sur $\mathcal{D}([0,T])$

Théorème 118 (Ascoli) **:**

Soit $\mathcal{F} \subset \mathcal{D}([0,T])$. Alors $\mathcal{F}$ est relativement compacte ssi :

1. $M = \sup\limits_{f\in\mathcal{F}} \|f\|_\infty < \infty$,
2. $\sup\limits_{f\in\mathcal{F}} w'(f,\delta) \underset{\delta\to 0}{\longrightarrow} 0$.

Démonstration. On va seulement montrer l'implication réciproque. Il nous suffit de montrer que $\mathcal{F}$ est d'-précompacte, pour conclure par complétude.

Soient $f \in \mathcal{F}$ et $(t_i)_{1\leq i\leq d}$ une subdivision ε-adaptée, δ-espacée, où le $\delta(\varepsilon)$ ne dépend pas de la fonction f choisie.

On peut modifier la subdivision en s'alignant sur des points $s_i \in \left\{\frac{n}{2^k}T,\, n \leq 2^k\right\}$. Si Θ_f envoie s sur t, prolongée de façon affine par morceaux, alors les pentes de Θ_f sont toutes comprises dans l'intervalle $\left[\frac{\delta}{\delta+\frac{T}{2^k}}, \frac{\delta}{\delta-\frac{T}{2^k}}\right]$. En particulier, pour k assez grand, on peut garantir $\|\Theta_f\|_H \leq \varepsilon$.

On approche alors $f \circ \Theta_f$ par une fonction g constante sur chaque intervalle $\left[\frac{n}{2^k}T, \frac{n+1}{2^k}T\right[$, à valeurs dans $\frac{M}{2^k} \times [\![-2^k, 2^k]\!]$, de sorte que $\|f \circ \Theta_f - g\|_\infty \leq \frac{M}{2^k} + \varepsilon$, d'où $d'(f,g) \leq 3\varepsilon$ pour k assez grand.

À k fixé, il n'existe qu'un nombre fini de telles fonctions g, donc pour tout $\varepsilon > 0$, on a une famille *finie* de boules de rayon 3ε, centrées en les fonctions g, qui recouvre $\mathcal{F}$. La famille est donc bien précompacte. □

6.3 Critères de tension dans $\mathcal{D}$

6.3.1 Caractérisation des lois sur $\mathcal{D}([0,T])$

Dans le cas continu, la loi π était entièrement caractérisée par les projections fini-dimensionnelles $\pi_{\vec{t}} := \pi_{t_1<\cdots<t_d}$. Ici, $\pi_{\vec{t}} : f \mapsto (f(t_i))_{1\leq i\leq d}$ n'est pas continue, ce qui complique la tâche.

Lemme 119 :

L'application $\pi_{\vec{t}}$ est mesurable.

Démonstration. Pour montrer la mesurabilité de cette application, il suffit de montrer qu'elle est mesurable coordonnée par coordonnée, donc on se ramène au cas π_t avec $t \in [0,T]$.

Si $t = 0$ ou $t = T$, par définition de la topologie, nos homéomorphismes dans $Homeo^+$ envoient 0 sur 0 et T sur T, donc π_0 et π_T sont continues.

Si $t < T$, par continuité à droite, on a $\pi_t(f) = f(t) = \lim \int_t^{t+\varepsilon} f(s)\frac{\mathrm{d}s}{\varepsilon} =: \pi_{t,\varepsilon}(f)$. On a donc $\pi_t = \lim\limits_{\varepsilon\to 0} \pi_{t,\varepsilon}$. Il suffit de montrer que les applications $\pi_{t,\varepsilon}$ sont continues pour qu'elles soient mesurables, et donc leur limite π_t aussi.

Soit (f_n) qui converge vers f. On a donc des homéomorphismes $\psi_n \in Homeo^+$ tels que $\|f_n - f\circ\psi_n\|_\infty \to 0$ et $\|\psi_n\|_H \to 0$. On a $\pi_{t,\varepsilon}(f_n) = \frac{1}{\varepsilon}\int_t^{t+\varepsilon} f_n(s)\,\mathrm{d}s = \frac{1}{\varepsilon}\int_{\psi_n^{-1}(t)}^{\psi_n^{-1}(t+\varepsilon)} f_n \circ \psi_n^{-1} g_n(u)\,\mathrm{d}u$ avec le changement de variable $s = \psi_n^{-1}(u)$, donc $\mathrm{d}_s = g_n(u)\,\mathrm{d}u$ pour une fonction g_n dont les pentes sont dans $\left[e^{-\|\psi_n\|_H}, e^{\|\psi_n\|_H}\right]$. En conséquence, on a bien $\pi_{t,\varepsilon}(f_n) \to \pi_{t,\varepsilon}(f)$. □

Lemme 120 :

Si f est continue en t, alors π_t est continue en f.

Démonstration. Si (f_n) converge vers f, alors $|f_n(t) - f(t)| \leq \|f_n - f\circ\psi_n\|_\infty + |f\circ\psi_n(t) - f(t)|$. Le terme de gauche converge vers 0 par convergence de la suite (f_n), celui de droite par continuité de f en t. □

Lemme 121 :

Soit μ une loi sur $\mathcal{D}([0,T])$. On pose $D_\mu = \{t \in [0,T],\ \mu(\{f,\ f \text{ discontinue en } t\}) > 0\}$. Alors cet ensemble est au plus dénombrable.

Démonstration. On pose $D_{\mu,\delta,\varepsilon} = \{t \in [0,T],\ \mu(\{f,\ |f(t) - f(t^-)| \geq \delta\}) \geq \varepsilon\}$. L'ensemble D_μ est une union dénombrable de tels ensembles, donc il nous suffit de montrer que chaque tel ensemble est *fini*.

Supposons $D_{\mu,\delta,\varepsilon}$ infini, qui contient une suite (t_n) injective. On a alors :

$$\mu\big(\{f,\ \big|f(t_n)-f(t_n^-)\big|\geq\delta \text{ pour une infinité de } n\}\big)=\mu\left(\bigcap_{n\geq 1}\bigcup_{m\geq n}\{f,\ \big|f(t_m)-f(t_m^-)\big|\geq\delta\}\right),$$

et cette mesure est limite décroissante des mesures aux rangs n finis fixés. Pour chaque valeur de n, la mesure du terme à l'intérieure est au moins ε, donc la mesure limite aussi. Ceci contredit le fait que, pour *toute* fonction càdlàg, le nombre de discontinuités au seuil δ est *fini*. □

Lemme 122 :

Une loi μ sur les fonctions càdlàg est déterminée par ses lois fini-dimensionnelles, à l'exclusion éventuelle d'un ensemble dénombrable d'instants D.

Démonstration. La preuve est la même que dans le cas continu, via l'égalité entre la tribu borélienne sur $\mathcal{D}([0,T])$, la tribu produit sur $\mathbb{R}^{[0,T]}$ et la tribu produit sur $\mathbb{R}^{[0,T]\setminus D}$. □

Théorème 123 :

Soit (μ_n) une suite de mesures sur $\mathcal{D}([0,T])$. On a équivalence entre les propriétés :

1. (μ_n) converge en loi vers une probabilité μ sur les fonctions càdlàg,
2. (μ_n) est tendue, et il existe un ensemble $D\subset[0,T]$ dénombrable, tel que pour toute famille $t_1,\dots,t_d\in[0,T]\setminus D$, on a convergence de la loi fini-dimensionnelle $\pi_{\vec{t},*}\mu_n$.

Démonstration. Pour le sens direct, (μ_n) converge en loi vers μ, donc la suite est tendue. Par les lemmes précédents, on a un ensemble D_μ dénombrable d'instants en dehors duquel $\mu(\{f \text{ continue en } t\})=1$. Pour un tel t, on a π_t continue sur un ensemble de mesure 1, d'où $\pi_{t,*}\mu_n\rightharpoonup\pi_{t,*}\mu$. On procède de même pour $\pi_{\vec{t}}$.

Pour le sens réciproque, il suffit de remarquer que pour toute sous-suite convergente, on a les mêmes lois fini-dimensionnelles à la limite, à l'exclusion d'une famille dénombrable de paramètres, d'où l'unicité de la valeur d'adhérence en loi, et donc la convergence de la suite par tension. □

6.3.2 Critères de tension sur $\mathcal{D}([0,T])$

Théorème 124 :

Une famille $\mathcal{P}$ de lois est tendue ssi :

1. $\sup_{\mu\in\mathcal{P}}\mu(\{\|f\|_\infty\geq M\})\underset{M\to\infty}{\longrightarrow}0$,
2. $\sup_{\mu\in\mathcal{P}}\mu(\{w'(f,\delta)\geq\varepsilon\})\underset{\delta\to 0^+}{\longrightarrow}0$ pour tout $\varepsilon>0$.

Idée. On adapte les arguments du cas continu, en combinant Ascoli et Prohorov. □

Remarque 125 (Cas des processus de Poisson) **:**

Un processus de Poisson d'intensité λ est un processus N càdlàg, de sorte que pour tous instants $t_1 < \cdots < t_d$, la famille $\left(N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_d} - N_{t_{d-1}}\right)$ est constituées de variables de Poisson indépendantes, de paramètres $\lambda(t_{i+1} - t_i)$.

D'après ce qui précède, sous réserve d'existence, ce processus est unique en loi dans $\mathcal{D}([0,T])$ puisqu'on connait ses lois fini-dimensionnelles.

On se ramène au cas $T = 1$. Soit $X_t^{(n)} = \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{1}_{U_k \leq \frac{\lambda}{n}}$. Par construction, les $\left(X_{t_{i+1}}^{(n)} - X_{t_i}^{(n)}\right)$ sont indépendants à n fixé, et $X_{t_{i+1}}^{(n)} - X_{t_i}^{(n)} = \mathcal{B}\left(\lfloor nt_{i+1} \rfloor - \lfloor nt_i \rfloor, \frac{\lambda}{n}\right) \rightharpoonup \mathcal{P}(\lambda(t_{i+1} - t_i))$ par la loi des petits nombres vue précédemment.

Si on montre la tension de $\left(X^{(n)}\right)$, alors on aura bien sa convergence en loi vers le processus de Poisson d'intensité λ. Remarquons que $\left\|X^{(n)}\right\|_\infty = X_1^{(n)} \sim \mathcal{B}\left(n, \frac{\lambda}{n}\right)$. Cette famille converge en loi vers $\mathcal{P}(\lambda)$, elle est tendue, donc a fortiori $\sup_{n \in \mathbb{N}} \mathbb{P}\left(\left\|X^{(n)}\right\|_\infty \geq M\right) \underset{M \to 0}{\longrightarrow} 0$.

On regarde maintenant la loi des sauts de $X^{(n)}$, conditionnellement à $X_1^{(n)} = k$. Ces sauts ont lieu en des instants $\left\{\frac{i}{n},\, 1 \leq i \leq n\right\}$. L'ensemble des sauts est uniforme parmi les $\binom{n}{k}$ possibilités. Si on a deux temps de saut espacés au plus de δ, cela revient à choisir $\binom{n}{k-1}$ temps de saut quelconques, puis de choisir le dernier parmi les δn points à droite d'un des $k-1$ points déjà choisis. Autrement dit, la probabilité d'avoir deux sauts espacés de moins de δ conditionnellement à $X_1^{(n)} = k$ est majorable par :

$$\frac{n\delta(k-1)\binom{n}{k-1}}{\binom{n}{k}} = \frac{k(k-1)n\delta}{n-k+1} \leq k^2\delta\left(1 + \frac{k-1}{n-k-1}\right) \leq \delta k^3\,,$$

et donc la probabilité qu'il y ait deux sauts espacés de moins de δ, en oubliant le conditionnement, est majorée par $\delta \sum_{k \in \mathbb{N}} k^3 \mathbb{P}\left(\mathcal{B}\left(n, \frac{\lambda}{n}\right) = k\right)$. Comme la loi de Poisson admet un moment d'ordre 3 fini, on peut majorer la somme indépendamment de n, et donc la limite pour $\delta \to 0$ est bien 0.

6.3.3 Critère de Kolmogorov discontinu, applications

Théorème 126 (Critère de Kolmogorov discontinu) **:**

Soit $\left(X^{(n)}\right)$ une suite de processus càdlàg. Si :

1. la suite $\left(\left\|X^{(n)}\right\|_{\infty,[0,T]}\right)_{n \in \mathbb{N}}$ est tendue,
2. il existe des constantes $C > 0$, $a > 0$ et $b > 1$ telles que, pour tous $r < s < t \in [0,T]$, on a l'inégalité $\mathbb{E}\left[\left|X_t^{(n)} - X_s^{(n)}\right|^a \left|X_s^{(n)} - X_r^{(n)}\right|^a\right] \leq C|t-r|^b$,

alors la suite $\left(X^{(n)}\right)_{n \in \mathbb{N}}$ est tendue.

Démonstration. Admis. □

Proposition 127 (Donsker discontinu) **:**

Soit $X_t^{(n)} = \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} A_k$, avec (A_k) est variables iid centrées réduites. Alors on a $X^{(n)} \overset{\mathcal{D}([0,T])}{\rightharpoonup} B$, le processus converge vers un pré-mouvement brownien à trajectoires càdlàg.

Démonstration. Pour la tension, il suffit ici de remarquer que :

$$\mathbb{E}\left[\left|X_t^{(n)} - X_s^{(n)}\right|^2 \left|X_s^{(n)} - X_r^{(n)}\right|^2\right] = \frac{\lfloor nt \rfloor - \lfloor ns \rfloor}{n} \times \frac{\lfloor ns \rfloor - \lfloor nr \rfloor}{n} \leq C(t-r)^2 .$$

On peut donc appliquer le critère de Kolmogorov discontinu. $\square$

Remarque 128 :

Avec le même argument que pour Donsker, si U_k sont des variables iid de loi $\mathcal{U}([0,1])$, alors $X_t^{(n)} = \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{1}_{U_k \leq \frac{\lambda}{n}}$ induit une suite tendue.

6.4 Cas de $\mathcal{D}(\mathbb{R}^+)$

Remarque 129 :

On aimerait obtenir une topologie projective sur $\mathcal{D}(\mathbb{R}^+)$.

Autrement dit, on veut une topologie sur $\mathcal{D}(\mathbb{R}^+)$, telle que $f_k \to f$ ssi $\pi_n^\infty(f_k) \to \pi_n^\infty(f)$, avec un système de projections $\pi_n^\infty \mathcal{D}(\mathbb{R}^+) \to \mathcal{D}([0,n])$ et $\pi_n^{n+1} : \mathcal{D}([0,n+1]) \to \mathcal{D}([0,n])$ tel que $\pi_n^\infty = \pi_n^{n+1} \circ \pi_{n+1}^\infty$, de sorte que le diagramme associé soit commutatif.

Remarque 130 :

Les projections naïves $\pi_n(f) = f|_{[0,n]}$ ne sont pas adaptées, car la convergence de $(\pi_{n+1}(f_k))$ n'implique pas la convergence de $(\pi_n(f_k))$, à cause de l'éventuelle discontinuité de la limite en n.

Définition 131 :

Soit $\delta_n(x) = \mathbb{1}_{x<n-1} + \mathbb{1}_{x \geq n-1}(n-x)^+$, la fonction continue, affine par morceaux, qui vaut 1 jusqu'à $n-1$ et 0 après n.

Si $f \in \mathcal{D}(\mathbb{R}^+)$, alors $\pi_n(f) = f \times \delta_n \in \mathcal{D}([0,n])$ dans le sens où elle est nulle sur $[n, \infty[$.

La famille de projections $(\pi_n)_{n\in\mathbb{N}}$ est commutative.

Théorème 132 :

On munit $\mathcal{D}(\mathbb{R}^+)$ de la topologie induite par le système projectif (π_n) défini précédemment. Alors on a les propriétés :

1. $K \subset \mathcal{D}(\mathbb{R}^+)$ relativement compact ssi les projections $\pi_n(K)$ le sont,
2. la famille de lois $\mathcal{P}$ est tendue ssi les projections $\pi_{n,*}(\mathcal{P})$ le sont.

Démonstration. Admis. □

7 Processus de Poisson composés, processus de Lévy

7.1 Compléments sur les mesures aléatoires de Poisson

On va voir quelques applications de la stabilité des mesures aléatoires.

Proposition 133 (Sélection de points dans un nuage) **:**

Soit $M = \sum_{i=1}^{I} \delta_{X_i} \sim MAP(\mu)$. Si $Y_i \overset{\text{iid}}{\sim} \mathcal{B}(p)$, alors $\widetilde{M} = \sum_{i=1}^{I} Y_i \delta_{X_i} \sim MAP(p\mu)$.

Démonstration. On a vu que $\sum_{i=1}^{I} \delta_{X_i, Y_i}$ est une mesure aléatoire d'intensité $\mu \otimes \mathcal{B}(p)$, par stabilité par marquage. En conséquence, par stabilité par restriction, $\sum_{i=1}^{I} \delta_{X_i, Y_i} Y_i$ est une mesure aléatoire d'intensité $\mu \otimes p\delta_1$. Finalement, par stabilité par image, en projetant sur la première coordonnée, on obtient $\sum_{i=1}^{I} \delta_{X_i} Y_i$ une mesure aléatoire d'intensité $p\mu$. □

Remarque 134 (Percolation continue) **:**

On peut construire des variables (X_n) telles que $\sum_{n=1}^{\infty} \delta_{X_n}$ est une mesure aléatoire sur $\mathbb{R}^d$, d'intensité λ (la mesure de Lebesgue).

On considère une suite de rayons aléatoires (R_n) iid de loi μ. On étudie alors l'ensemble aléatoire $A = \bigcup_{n=1}^{\infty} B(X_n, R_n)$.

Soit $x \in \mathbb{R}^d$ fixé. On peut montrer que N_x, le nombre de boules de A qui rencontrent x, suit la loi de Poisson de paramètre $\mathbb{E}[\lambda(B(0, R_n))]$. En effet, quitte à marquer chaque point X_n par son rayon R_n, on a $\widetilde{M} = \sum_{n=1}^{\infty} \delta_{X_n, R_n}$ d'intensité $\lambda \otimes \mu$. Soit $f : (y, r) \mapsto \mathbb{1}_{d(x,y) \leq r}$. Par passage à l'image, $N_x = f_* \widetilde{M}(1)$. C'est donc bien une variable de Poisson, de paramètre :

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^+} f(y, r) \,\mathrm{d}y \,\mathrm{d}\mu(r) = \int_{\mathbb{R}^+} \lambda(B(0, r)) \,\mathrm{d}\mu(r) = \lambda(B_d(0, 1)) \times \mathbb{E}\left[R^d\right].$$

Si μ a un moment d'ordre d, alors avec probabilité non nulle, $N_x = 0$ et donc A ne couvre pas tout l'espace. En outre, comme $\left\{A = \mathbb{R}^d\right\}$ est invariant par translation, par un argument d'ergodicité, on a $\mathbb{P}\left(A = \mathbb{R}^d\right) \in \{0, 1\}$, d'où $\mathbb{P}\left(A = \mathbb{R}^d\right) = 0$.

Réciproquement, si $\mathbb{E}\left[R^d\right] = \infty$, alors presque-sûrement, pour tout $x \in \mathbb{Q}^d$, on a $N_x = \infty$. Le même argument marche en remplaçant R par $R' = (R-1)^+$. Presque-sûrement, $\bigcup_{n \in \mathbb{N}} B(X_n, R_n - 1)$ atteint $\mathbb{Q}^d$ tout entier, et donc quitte à épaissir ces boules, par densité, on a bien $A = \mathbb{R}^d$.

7.2 Processus ponctuel de Poisson

Définition 135 :

Soit μ une mesure s-finie sur un espace polonais $\mathbb{X}$. On définit le processus ponctuel de Poisson d'intensité μ comme la mesure aléatoire de Poisson sur $\mathbb{X} \times \mathbb{R}^+$ d'intensité $\mu \otimes \lambda$.

Lemme 136 :

Une mesure aléatoire d'intesité μ est presque-sûrement sans points multiples ssi μ est sans atomes.

Démonstration. Pour le sens direct, par contraposée, si $\mu(\{x\}) = \lambda > 0$, alors $M(\{x\}) \sim \mathcal{P}(\lambda)$ est non nul avec probabilité > 0.

Réciproquement, il suffit de montrer l'implication lorsque μ est finie sans atomes pour obtenir le cas général. Dans ce cas, on peut écrire $M = \sum_{i=1}^{N} \delta_{X_i}$, avec N aléatoire et les $X_i \sim \frac{\mu}{\mu(\mathbb{X})}$ iid. Pour tous $i \neq j$, comme la loi des X_i est sans atomes, on a $\mathbb{P}(X_i = X_j) = 0$, d'où le résultat. $\square$

Corollaire 137 :

Soit M un processus ponctuel de Poisson. Presque-sûrement, pour tout $t \in \mathbb{R}^+$, on a $M(\mathbb{X} \times \{t\}) \in \{0, 1\}$.

Démonstration. Il suffit de montrer le résultat pour $t < T \in \mathbb{N}$, puis de prendre la limite croissante pour $T \to \infty$. Dans ce cas on considère $A_n = \bigcup_{k=1}^{n} \left\{ M\left(\mathbb{X} \times \left[T\frac{k-1}{n}, T\frac{k}{n}\right[\right) \geq 2 \right\}$. Si $M(\mathbb{X} \times \{t\}) \geq 2$ pour un certain t, alors a fortiori A_n est réalisé pour tout n.

Si on montre que $\mathbb{P}(A_n) \to 0$, on en déduit le résultat souhaité. Or, les tranches de temps étant disjointes, les n évènements considérés dans A_n sont indépendants, donc on a l'égalité $\mathbb{P}(A_n) = n \times \mathbb{P}\left(\mathcal{P}\left(\frac{T\mu(\mathbb{X})}{n}\right) \geq 2\right)$. Lorsque $\mu(\mathbb{X}) < \infty$, les termes en $\frac{1}{n}$ dans la probabilité se compensent, donc on obtient $\mathbb{P}(A_n) = O\left(\frac{1}{n}\right)$. Ceci conclut la preuve dans ce cas.

En général, si l'intensité μ est s-finie, on l'approche par une suite croissante de mesures finies. En parallèle, on peut approcher M par une suite croissante de processus ponctuels, associés aux mesures finies sur $\mathbb{X}$. Par convergence monotone, on obtient alors le résultat. $\square$

Définition 138 :
On ajoute à $\mathbb{X}$ un point cimetière isolé $\dagger$.

Sous l'évènement presque-sûr précédent, pour $N \sim MAP(\mu \otimes \mathrm{d}t)$, on peut définir un processus X par $X_t = x$ lorsque $N(\{(x,t)\}) = N(\mathbb{X} \times \{t\}) = 1$ et $X_t = \dagger$ sinon.

La famille (X_t) décrit entièrement N, et on note $X \sim PPP(\mu)$.

Lemme 139 (Loi des temps de saut) **:**
Soient $(T_k)_{k \geq 1}$ les temps de saut du processus de Poisson P d'intensité λ. Alors $N = \sum\limits_{k=1}^{\infty} \delta_{T_k}$ est une mesure aléatoire sur $\mathbb{R}^+$, de loi $MAP(\lambda \, \mathrm{d}t)$. En outre, les accroissements $(T_{k+1} - T_k)_{k \geq 0}$ sont iid, de loi $\mathcal{E}(\lambda)$.

Démonstration. Posons $M = P_t \sim \mathcal{P}(\lambda t)$. On a déjà vu que, conditionnellement à $M = k$, l'ensemble *désordonné* $\{T_1, \ldots, T_k\}$ a la loi d'un ensemble $\{U_1, \ldots, U_k\}$ de k variables uniformes indépendantes.

Pour une famille $(B_i)_{1 \leq i \leq r}$ de r boréliens disjoints de $[0,t]$, $N(B_i)$ est le nombre de sauts de P qui arrivent à des temps *dans* B_i. On a $(N(B_1), \ldots, N(B_r)) \overset{\mathrm{d}}{=} \left(\sum\limits_{j=1}^{M} \mathbb{1}_{U_j \in B_1}, \ldots, \sum\limits_{j=1}^{M} \mathbb{1}_{U_j \in B_r} \right)$. En conséquence, N a bien les bonnes lois fini-dimensionnelles.

D'autre part, on a naturellement $\mathbb{P}(T_1 > t) = \mathbb{P}(P_t = 0) = e^{-\lambda t}$ donc T_1 a la loi exponentielle souhaitée. On peut alors montrer que les accroissements sont iid, ce qui conclut la preuve. $\square$

Théorème 140 :
Si $\mu(\mathbb{X}) = \infty$, alors $\{t, \, X_t \neq \dagger\}$ est dense dans $\mathbb{R}^+$.

Si $\mu(\mathbb{X}) < \infty$, alors on peut construire la loi de (X_t) comme suit. On considère $(T_k)_{k \geq 1}$ les temps de saut d'un processus de Poisson d'intensité $\mu(\mathbb{X})$, et $(Y_k) \overset{\mathrm{iid}}{\sim} \mu_{norm} := \frac{\mu}{\mu(\mathbb{X})}$ indépendants de T. On a alors $X_{T_k} = Y_k$ et $X_t = \dagger$ si $t \notin \{T_k, \, k \geq 1\}$. Dans ce cas, on a $X \sim PPP(\mu)$.

Démonstration. Pour le premier point, pour tout intervalle rationnel $[a,b]$, on a alors p.s. l'égalité $N(\mathbb{X} \times [a,b]) = \infty$. Comme la mesure d'une tranche vaut au plus 1, on en déduit que l'ensemble $[a,b]$ contient une infinité de temps t où $X_t \neq \dagger$, d'où la densité souhaitée.

Pour le second point, comme $\mu(\mathbb{X}) < \infty$, on peut voir N comme un marquage du processus de Poisson P d'intensité $\mu(\mathbb{X}) \, \mathrm{d}t$ sur $\mathbb{R}^+$.

Considérons $(T_k)_{k \in \mathbb{N}}$ les instants de saut de P. On a vu que $P \overset{\mathrm{d}}{=} \sum\limits_{k=1}^{\infty} \delta_{T_k}$, et la suite T est presque-sûrement *strictement* croissante. Par marquage, en considérant des variables $Y_k \sim \mu_{norm}$

iid et indépendantes de T, on a bien :

$$\sum_{k=1}^{\infty} \delta_{T_k, Y_k} \sim MAP(\mu_{norm} \otimes \mu(\mathbb{X})\,\mathrm{d}t) = MAP(\mu \otimes \mathrm{d}t) = PPP(\mu)\,.$$

Avec cette décomposition de N, on a directement $\{t,\, X_t \neq \dagger\} = \{T_k,\, k \in \mathbb{N}^*\}$, et $X_{T_k} = Y_k$, ce qui conclut la preuve. □

7.3 Processus de Poisson composés

Définition 141 (Accroissements d'une fonction) **:**

Soit $X \sim PPP(\mu)$ pour une mesure s-finie μ, à valeurs dans $\mathbb{X} \sqcup \{\dagger\}$.

Pour $f \in L^1(\mathbb{X}, \mu)$, on considère $\sum\limits_{s \in [0,t],\, X_s \neq \dagger} f(X_s)$.

En termes de $N \sim MAP(\mu \otimes \mathrm{d}t)$, cette quantité est $\int f(x) \mathbb{1}_{[0,t]}(s)\,\mathrm{d}N(x,s) = N_t(f)$.

Théorème 142 :

$(N_t(f))_{t \geq 0}$ est défini p.s., c'est un processus càdlàg, dans $\mathcal{D}(\mathbb{R}^+)$.

Pour $s \in \mathbb{R}^+$, $(N_{t+s}(f) - N_s(f))_{t \geq 0}$ est indépendant de $\sigma(N_u(f), u \leq s)$ et a la loi de $(N_t(f))$.

La loi des accroissements de $(N_t(f))_{t \geq 0}$ est entièrement déterminée par les :

$$\mathbb{E}\left[e^{i\xi N_t(f)}\right] = \exp\left(t \int_{\mathbb{X}} \left(e^{i\xi f(x)} - 1\right) \mathrm{d}\mu(x) \right).$$

Démonstration. À t fixé, on a $\mathbb{E}[N_t(|f|)] = t\|f\|_{L^1(\mu)} < \infty$. Ceci implique *a fortiori* que pour tout $s \leq t$, on a $|N_s(f)| \leq N_t(|f|)$ fini p.s. En passant à l'intersection pour des temps arbitrairement grands, on en déduit que, presque-sûrement, $(N_t(f))$ est bien défini, à partir d'une intégrale. Dans ce cas, à ω fixé, la nature càdlàg de $(N_t(f))$ se vérifie par convergence dominée.

Par construction de l'objet, par définition de la mesure aléatoire N, on a naturellement $N_{s+t}(f) - N_s(f)$ mesurable dans $\sigma\big(N|_{\mathbb{X} \times]s,\infty[}\big)$, donc indépendant de $\sigma\big(N|_{\mathbb{X} \times [0,s]}\big)$, et *a fortiori* de la famille $(N_u(f))_{u \leq s}$.

Quitte à considérer la mesure image, on peut toujours se ramener au cas réel, via :

$$(N_t(f, \mathbb{X}, \mu)) \overset{\mathrm{d}}{=} (N_t(\mathrm{Id}, \mathbb{R}, f_*\mu))\,.$$

Dans ce cas, il est facile de vérifier les propriétés restantes, puis de conclure via :

$$\exp\left(t\int_{\mathbb{X}}\left(e^{i\xi f(x)}-1\right)\mathrm{d}\mu(x)\right)=\exp\left(t\int_{\mathbb{R}}\left(e^{i\xi y}-1\right)\mathrm{d}f_*\mu(y)\right).$$

$\square$

Définition 143 (Processus de Poisson composé) **:**
Le processus de Poisson composé d'intensité ν sur $\mathbb{R}$, telle que $\nu(|x|)<\infty$, est le processus càdlàg défini par $N\sim MAP(\nu\otimes\mathrm{d}t)$, via $N_t(\mathrm{Id})=\int\limits_{\mathbb{R}\times[0,t]}x\,\mathrm{d}N(x,s)=\sum\limits_{s\in[0,t],\,X_s\neq\dagger}X_s$, avec le processus $X\sim PPP(\nu)$. On note alors simplement $N=(N_t)_{t\geq 0}\sim PPC(\nu)$.

Remarque 144 (Hiérarchie des objets de Poisson) **:**
À partir de variables de Poisson, d'intensité $\lambda\in\mathbb{R}^+$, on obtient les temps de saut d'un processus de Poisson d'intensité λ.

Un processus ponctuel de Poisson, d'intensité μ, est un cas particulier de mesure aléatoire de Poisson, d'intensité $\mu\otimes\mathrm{d}t$. Les temps de saut du processus ponctuel, en particulier, correspondent à un processus de Poisson d'intensité $\mu(\mathbb{X})$.

On peut enfin obtenir un processus de Poisson composé d'intensité ν à partir du processus ponctuel de Poisson correspondant.

Remarque 145 (Affaiblissement des hypothèses) **:**
Lorsque $\nu(|x|)<\infty$, on a $\mathbb{E}[N_t]=t\nu(x)$ bien défini.

Plus généralement, supposons $\nu(|x|\wedge 1)<\infty$. On peut décomposer ν en sa restriction à $[-1,1]$ et sa restriction au complémentaire. Par hypothèse, on a encore $PPC\left(\nu|_{[-1,1]}\right)$ bien défini. Ce terme correspond à $N_t^-=\sum\limits_{s\in[0,t],\,|X_s|<1}X_s$, qui est intégrable. L'autre somme, autrement dit $N_t^+=\sum\limits_{s\in[0,t],\,|X_s|\geq 1}X_s$, est la somme d'un nombre aléatoire de termes de loi $\mathcal{P}(t\nu(\mathbb{R}\backslash[-1,1]))$, finis p.s. Cette somme est finie p.s., ce qui garantit que N_t^+ est bien défini, mais plus forcément intégrable. Le processus N ainsi obtenu est encore càdlàg, à accroissements indépendants et stationnaires, mais plus nécessairement intégrables.

Comme $e^{i\xi x}-1=O(1\wedge|x|)$ sur $\mathbb{R}$ tout entier, on a $\mathbb{E}\left[e^{i\xi N_t}\right]=\exp\left(t\int_{\mathbb{R}}\left(e^{i\xi x}-1\right)\mathrm{d}\nu(x)\right)$ bien défini.

8 Processus de Lévy

Définition 146 :

Un processus de Lévy est un processus réel X càdlàg, qui vérifie les propriétés :

1. $X_0 = 0$.
2. Accroissements indépendants : $(X_{t+s} - X_s)_{t\geq 0}$ est indépendante de $(X_u)_{u\leq s}$.
3. Stationnarité : $(X_{t+s} - X_s)_{t\geq 0} \overset{\mathrm{d}}{=} X$.

Remarque 147 :

Les points 2 et 3 de la définition précédente sont équivalentes à dire que pour tous temps $t_1 < \ldots t_d$, la famille $\left(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_d} - X_{t_{d-1}}\right)$ est constitué de variables indépendantes, et $X_t - X_s \overset{\mathrm{d}}{=} X_{t-s}$.

En conséquence, la loi du processus de $(X_t)_{t\geq 1}$ est entièrement déterminée par la loi de X_1.

Remarque 148 :

Les processus affines, les mouvements browniens, les processus de Poisson, et les processus de Poisson composés sont des cas particuliers de processus de Lévy.

On verra que les processus de Lévy continus correspondent à des mouvements browniens avec un drift affine. On verra que tout processus de Lévy est limite de sommes de processus continus et de processus de Poisson composés. On affaiblira plus tard ce résultat, pour voir tout processus de Lévy comme limite de processus de Poisson composés.

Définition 149 (Loi infiniment divisible) **:**

Une loi sur $\mathbb{R}$ est dite infiniment divisible lorsque, pour tout $n \geq 1$, il existe ρ_n sur $\mathbb{R}$ telle que $\mu = \rho_n^{*n}$. Autrement dit, il existe $(X_i) \overset{\text{iid}}{\sim} \rho_n$ telles que $\sum\limits_{i=1}^{n} X_i \overset{\mathrm{d}}{=} \mu$.

Lemme 150 :

Les lois marginales d'un processus de Lévy sont infiniment divisibles.

Démonstration. Pour tout $t > 0$, on a $X_t = \sum\limits_{k=0}^{n-1} X_{t\frac{k+1}{n}} - X_{t\frac{k}{n}}$, égal en loi à la somme de n variables $X_{\frac{t}{n}}$ indépendantes. □

Remarque 151 :

Une gaussienne $\mathcal{N}(m, \sigma^2)$ est infiniment divisible, convolution de n gaussiennes $\mathcal{N}\left(\frac{m}{n}, \frac{\sigma^2}{n}\right)$.

Une loi de Poisson $\mu = \mathcal{P}(\lambda)$ est infiniment divisible, convolution de n lois $\mathcal{P}\left(\frac{\lambda}{n}\right)$.

Une loi géométrique $\mathcal{G}(p)$ est infiniment divisible, à partir de lois binomiales négatives définies par $\rho_n(k) = \frac{\prod_{j=0}^{k-1}\left(-\frac{1}{n} - j\right)}{k!}(-1)^k p^{\frac{1}{n}}(1-p)^k$. On vérifie cette propriété via $\widehat{\mu * \nu} = \widehat{\mu} * \widehat{\nu}$, et en calculant explicitement $\widehat{\rho_n}(\xi) = \left(\frac{p}{1-(1-p)e^{i\xi}}\right)^{\frac{1}{n}}$.

Les lois $\Gamma(\lambda, c)$, à densité $\frac{\lambda^c x^{c-1} e^{-\lambda x}}{\Gamma(c)}$ pour $x \in \mathbb{R}^+$, est infiniment divisible. En effet on vérifie que sa transformée de Fourier est $\widehat{\mu}(\xi) = \left(\frac{\lambda}{\lambda - i\xi}\right)^c$, donc convolution de n lois $\Gamma\left(\lambda, \frac{c}{n}\right)$.

Les lois de Cauchy $\mathcal{C}(m, \gamma)$, à densité $\frac{1}{\pi\gamma\left(1+\left(\frac{x-m}{\gamma}\right)^2\right)}$ sur $\mathbb{R}^+$, a pour transformée de Fourier $\widehat{\mu}(\xi) = e^{im\xi - \gamma|\xi|}$, donc convolution de n lois $\mathcal{C}\left(\frac{m}{n}, \frac{\gamma}{n}\right)$.

Lemme 152 :

Si μ est infiniment divisible, à support borné, alors $\mu = \delta_a$ est un Dirac en un certain point.

Démonstration. Si $\mu([-M, M]) = 1$, alors ρ_n est à support dans $\left[-\frac{M}{n}, \frac{M}{n}\right]$. Comme la loi est à support bornée, elle a une variance finie. En outre, $\mathrm{Var}(\mu) = n\mathrm{Var}(\rho_n) \leq n\left(\frac{M}{n}\right)^2$, donc pour $n \to \infty$ on a $\mathrm{Var}(\mu) = 0$, μ est bien une constante. □

Corollaire 153 :

Les lois uniformes $\mathcal{U}([a, b])$ pour $a < b$ ne sont pas infiniment divisibles.

Théorème 154 :

Si μ est infiniment divisible, alors $\widehat{\mu}$ ne s'annule pas sur $\mathbb{R}$.

Démonstration. Soient $X, X' \sim \mu$, indépendantes, et on note $\Phi_X = \Phi_{X'} = \widehat{\mu}$ leur transformée de Fourier. Soit alors $Z = X - X'$. On a $\Phi_Z(\xi) = \Phi_X(\xi)\Phi_{X'}(-\xi) = |\widehat{\mu}(\xi)|^2 \geq 0$.

De même, pour X_n et X'_n de loi ρ_n, indépendantes, et $Z_n = X_n - X'_n$, on a $\Phi_{Z_n} = |\widehat{\rho_n}(\xi)|^2 \geq 0$. En outre, on a $\Phi_{Z_n}^n = \Phi_Z$. On a déjà établi que :

$$\mathbb{P}\left(|Z_n| \geq \frac{2}{\delta}\right) \leq \frac{1}{\delta}\int_{-\delta}^{\delta} (1 - \Phi_{Z_n}(\xi))d\xi = \frac{1}{\delta}\int_{-\delta}^{\delta}\left(1 - \Phi_Z^{\frac{1}{n}}(\xi)\right)d\xi\,,$$

or $0 \leq 1 - \Phi_Z^{\frac{1}{n}} \leq 1 - \Phi_Z$, et ce terme converge vers 0 pour $n \to \infty$, donc $\mathbb{P}\left(|Z_n| \geq \frac{2}{\delta}\right) \xrightarrow[n\to\infty]{} 0$ par convergence dominée.

En conséquence, la suite des lois $(Z_n)_{n\geq 0}$ est tendue. On a donc une sous-suite qui converge en loi, vers une variable Y. En conséquence, la limite $\Phi_Z(\xi)^{\frac{1}{\sigma(n)}} \to \mathbb{1}_{\Phi_Z(\xi)\neq 0} = \Phi_Y(\xi)$ est une fonction continue, qui vaut 1 en 0, donc constante égale à 1. Finalement, $\widehat{\mu}$ ne s'annule pas. □

Lemme 155 (Lemme de relèvement) **:**

Soit $f : \mathbb{R} \to \mathbb{C}$ continue, qui ne s'annule pas. Alors il existe une application $g : \mathbb{R} \to \mathbb{C}$

continue, unique à $2i\pi\mathbb{Z}$ près, telle que $f = e^g$.

Définition 156 (Exposant de Lévy-Khintchine) **:**
Si μ est infiniment divisible, on appelle exposant de Lévy-Khintchine l'unique application $\psi \in \mathcal{C}^0(\mathbb{R}, \mathbb{C})$ telle que $\psi(0) = 0$ et $\widehat{\mu} = e^{\psi}$.

Lemme 157 :
Si μ est infiniment divisible, alors la loi ρ_n est unique, et sa transformée de Fourier est égale à $e^{\frac{\psi}{n}}$. A fortiori, ρ_n est infiniment divisible, et son exposant est $\frac{\psi}{n}$.

Remarque 158 :
On a déjà établi les exposants de certaines lois infiniment divisibles classiques :

Loi	Exposant
$\mathcal{N}(m, \sigma^2)$	$im\xi - \frac{\sigma^2\xi^2}{2}$
$\mathcal{P}(\lambda)$	$\lambda\big(e^{i\xi} - 1\big)$
$\Gamma(\lambda, c)$	$c\ln\big(1 - \frac{i\xi}{\lambda}\big)$
$\mathcal{C}(m, \gamma)$	$im\xi - \gamma\lvert\xi\rvert$

Proposition 159 :
Toute loi infiniment divisible est une limite de lois de Poisson composées évaluées en $t = 1$.

Démonstration. Il suffit de considérer $X^{(n)} \sim PPC(\mu_n)$, où $\mu_n = n\rho_n$. Pour un tel processus :

$$\begin{array}{rcl} \widehat{\mu_n}(\xi) & = & \exp\big(\int\big(e^{ix\xi} - 1\big)\,\mathrm{d}\mu_n(x)\big) \\ & = & \exp\big(n\int\big(e^{ix\xi} - 1\big)\,\mathrm{d}\rho_n(x)\big) \\ & = & \exp(n\widehat{\rho}_n - 1) \\ & = & \exp\Big(ne^{\frac{\psi}{n}} - 1\Big) \\ & \to & e^{\psi}\,. \end{array}$$

On a convergence simple de $\widehat{\mu_n}$ vers $\widehat{\mu}$, d'où la convergence en loi souhaitée. $\square$

9 Formule de Lévy-Khintchine, décomposition de Lévy-Itô

9.1 Formule de Lévy

Théorème 160 (Formule de Lévy-Khintchine) **:**
L'exposant d'une loi infiniment divisible μ s'écrit de manière unique :

$$\psi(\xi) = ia\xi - \frac{\sigma^2\xi^2}{2} + \int\limits_{\mathbb{R}\setminus\{0\}} \left(e^{ix\xi} - 1 - ix\xi\mathbb{1}_{|x|<1}\right) \mathrm{d}\nu(x)\,,$$

avec $a \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}^+$, et ν une mesure positive sur $\mathbb{R}^*$ (autrement dit sans atome en 0) pour laquelle $\min(1, x^2) \in L^1(\nu)$, ou autrement dit $(1 \wedge |x|) \in L^2(\nu)$.

On dit que ν est la mesure de Lévy de μ.

Démonstration. La démonstration va se faire en quatre temps. Commençons par montrer que l'ensemble des lois infiniment divisibles (ID) est fermé, pour la convergence en loi. Soit (μ_k) une suite de fonctions ID, qui converge en loi vers μ. Pour $n \geq 1$, on a $\mu_k = \rho_k^{*n}$. La suite $(\rho_k)_{k\in\mathbb{N}}$ est tendue. En effet, $\rho_k([-M, M]^c) \leq \mu_k([-nM, nM])^{\frac{1}{n}}$, et le majorant tend uniformément vers 0 pour $M \to \infty$. En outre, pour toute valeur d'adhérence ρ de la suite, on a $\rho^{*n} = \mu$ par compatibilité de la convolution avec la convergence en loi. Ceci montre que $(\rho_k)_{k\geq 0}$ converge en loi vers ρ, et que μ est ID.

Montrons désormais que tout triplet (a, σ^2, ν) correspond à une loi ID. Le terme $ia\xi - \frac{\sigma^2\xi^2}{2}$ est l'exposant associé à $\mathcal{N}(a, \sigma^2)$, qui est une loi ID. On s'est ainsi ramené au cas $(0, 0, \nu)$. On découpe $\mathbb{R}^*$ en tranches, via $T_n = \left]-\frac{1}{n}, -\frac{1}{n+1}\right] \sqcup \left[\frac{1}{n+1}, \frac{1}{n}\right[$, avec la convention $\frac{1}{0} = +\infty$. Soit alors $\nu_n = \nu|_{T_n}$. Par hypothèse, on a ν_n de masse finie (d'où ν s-finie), donc la variable $X_n = P(\nu_n) - \int\limits_{\mathbb{R}} \mathbb{1}_{|x|<1} x \,\mathrm{d}\nu_n(x)$ est bien définie. On vérifie que X_n correspond à une loi ID, d'exposant $\psi_n(\xi) = \int\limits_{T_n} \left(e^{ix\xi} - 1 - \mathbb{1}_{|x|<1} ix\xi\right) \mathrm{d}\nu_n(x)$. Comme ν intègre x^2 en 0, on a la convergence simple de la suite ψ_n vers $\psi(\xi) = \sum\limits_{n\in\mathbb{N}} \psi_n(x)$. En conséquence, $\sum\limits_{i=0}^{N} X_n$ converge en loi vers une variable dont l'exposant est donné par le triplet $(0, 0, \nu)$. Comme chaque terme de la suite est ID, par le premier point de la démonstration, la limite l'est également.

Montrons maintenant l'unicité d'un tel triplet. On a $\frac{\psi(\xi)}{\xi^2} = \frac{ia}{\xi} - \frac{\sigma^2}{2} + \int\limits_{\mathbb{R}} \frac{e^{i\xi x} - 1 - \mathbb{1}_{|x|<1} ix\xi}{\xi^2} \mathrm{d}\nu(x)$. Lorsque $\xi \to \infty$, le terme de gauche tend vers 0. Le terme dans l'intégrale converge *simplement* vers 0, et on peut le dominer comme un $O(1 \wedge x^2)$. Autrement dit, $\sigma^2 = -2 \times \lim\limits_{\xi\to\infty} \frac{\psi(\xi)}{\xi^2}$ est uniquement déterminé par ψ. Soit alors $\pi = \psi + \frac{\sigma^2\xi^2}{2}$, caractérisée par ψ. En vérifiant que les grandeurs considérées sont bien intégrables, par Fubini :

$$\int_{-1}^{1} \pi(\xi) - \pi(\xi + t)\,\mathrm{d}t = \int_{\mathbb{R}} \int_{-1}^{1} e^{ix\xi} - e^{ix(\xi+t)}\,\mathrm{d}t\,\mathrm{d}\nu(x) = 2\int_{\mathbb{R}} e^{i\xi x}(1 - \operatorname{sinc}(x))\,\mathrm{d}\nu(x)\,.$$

En conséquence, μ caractérise entièrement la mesure $(1 - \operatorname{sinc}(x))\,\mathrm{d}\nu(x)$, via sa transformée de Fourier. Le seul point d'annulation de la densité $(1 - \operatorname{sinc}(x))$ est en 0, où ν n'a pas d'atome par hypothèse, donc μ caractérise ν de façon unique. En retranchant le terme intégral à π, on obtient finalement l'application $\xi \mapsto ia\xi$, ce qui détermine entièrement le coefficient a.

Reste à vérifier que toute loi admet effectivement un tel triplet. Admettons pour l'instant le lemme technique qui suit : une limite en loi de mesures ID à triplets est encore une mesure à triplet. Dans ce cas, on peut conclure par la convergence des lois de Poisson associées aux $PPC(n\rho_n)$, correspondant aux triplets $\big(a_n = n\rho_n(x\mathbb{1}_{|x|<1}), 0, n\rho_n\big)$, montrée dans la proposition précédente. □

Lemme 161 (Lemme technique) **:**

Soient μ_k associées aux triplets (a_k, σ_k^2, ν_k). On introduit une fonction de coupure c continue, qui vaut 1 au voisinage de 0 et s'annule en dehors de $[-1,1]$. Posons $\widetilde{\sigma_k}^2 = \sigma_k^2 + \int_{\mathbb{R}} x^2 c(x)\,\mathrm{d}\nu_k(x)$. Alors $(\mu_k)_{k\in\mathbb{N}}$ converge en loi vers μ ssi il existe un triplet (a, σ^2, ν) tel que :

1. $a_k \underset{k\to\infty}{\longrightarrow} a$,
2. $\widetilde{\sigma_k}^2 \underset{k\to\infty}{\longrightarrow} \widetilde{\sigma}^2 := \sigma^2 + \int_{\mathbb{R}} x^2 c(x)\,\mathrm{d}\nu(x)$,
3. pour toute fonction $f(x) = g(x) \times \min(1, x^2)$, avec g continue, bornée, nulle en 0, on a $\nu_k(f) \underset{k\to\infty}{\longrightarrow} \nu(f)$.

Dans ce cas, μ est une loi ID, issue du triplet (a, σ^2, ν) (modulo la présence éventuelle d'un atome en 0 pour ν). En outre, on a $\sigma^2 \geq \overline{\lim_{k\to\infty}}\, \sigma_k^2$. On peut cependant avoir un saut à la limite, d'où la nécessité de passer par une fonction de coupure pour réguler cette convergence.

Démonstration. Commençons par montrer le sens réciproque. On a :

$$\begin{aligned}
\psi_k(\xi) &= ia_k\xi - \sigma_k^2\tfrac{\xi^2}{2} + \int_{\mathbb{R}} e^{ix\xi} - 1 - \mathbb{1}_{|x|<1} ix\xi\,\mathrm{d}\nu_k(x) \\
&= ia_k\xi - \widetilde{\sigma_k}^2\tfrac{\xi^2}{2} + \int_{\mathbb{R}} e^{ix\xi} - 1 - \mathbb{1}_{|x|<1} ix\xi + \tfrac{\xi^2 x^2}{2}\,\mathrm{d}\nu_k(x)\,.
\end{aligned}$$

On peut en outre mettre la fonction dans l'intégrale sous la forme $g(x)\min(1, x^2)$. En utilisant nos hypothèses, on a donc la convergence simple de ψ_k vers :

$$\psi(\xi) = ia\xi - \frac{\xi^2}{2}\left(\widetilde{\sigma}^2 - \int_{\mathbb{R}} x^2 c(x)\,\mathrm{d}\nu(x)\right) + \int_{\mathbb{R}} e^{ix\xi} - 1 - \mathbb{1}_{|x|<1} ix\xi\,\mathrm{d}\nu(x)\,.$$

Si on montre que $\sigma^2 = \widetilde{\sigma}^2 - \int_{\mathbb{R}} x^2 c(x)\,\mathrm{d}\nu(x) \geq 0$, alors on aura bien ψ associée au triplet de Lévy (a, σ^2, ν), donc l'exposant d'une loi μ ID vers laquelle (μ_k) converge en loi.

On introduit pour ce faire une autre fonction seuil c_ε qui s'annule en 0 et vaut 1 en dehors de $[-\varepsilon, \varepsilon]$, prolongée de façon affine. Soit ρ_k la mesure finie, à densité $\min(1, x^2)$ par rapport à ν_k. Soient alors $\rho_{k,\varepsilon}$ à densité c_ε par rapport à ρ_k. On définit de même ρ et les ρ_ε par rapport à ν. Sous nos hypothèses, on a en particulier convergence en loi de $\rho_{k,\varepsilon}$ vers ρ_ε. En conséquence :

$$\begin{aligned}
\widetilde{\sigma}^2 &= \lim \widetilde{\sigma_k}^2 \\
&= \lim\left(\sigma_k^2 + \int_{\mathbb{R}} c(x) x^2 \,\mathrm{d}\nu_k(x)\right) \\
&\geq \overline{\lim}\left(\sigma_k^2 + \int_{\mathbb{R}} c(x)\,\mathrm{d}\rho_{\varepsilon,k}(x)\right) \\
&\geq \overline{\lim}\,\sigma_k^2 + \int_{\mathbb{R}} c(x)\,\mathrm{d}\rho_\varepsilon(x)\,.
\end{aligned}$$

Ceci étant vrai pour tout $\varepsilon > 0$, à la limite, par convergence monotone :

$$\widetilde{\sigma}^2 \geq \overline{\lim}\,\sigma_k^2 + \int_{\mathbb{R}} c(x) \min\bigl(1, x^2\bigr)\,\mathrm{d}\nu(x) = \overline{\lim}\,\sigma_k^2 + \int_{\mathbb{R}} c(x) x^2 \,\mathrm{d}\nu(x)\,,$$

donc $\sigma^2 \geq \overline{\lim}\,\sigma_k^2 \geq 0$, ce qui conclut l'implication.

Pour le sens direct, en conservant les mêmes notations, montrons que $(\rho_k)_{k\in\mathbb{N}}$ est une suite tendue de mesures finies. On veut montrer que $\sup\limits_{k\in\mathbb{N}} \rho_k(\mathbb{R}) < \infty$, et que $\sup\limits_{k\in\mathbb{N}} \rho([-M,M]^c) \underset{M\to\infty}{\longrightarrow} 0$. Par un calcul direct, pour tout $t \geq 0$, on a :

$$-\frac{1}{2t}\int_{-t}^{t} \psi_k(\xi)\,\mathrm{d}\xi = \sigma_k^2 \times \frac{t^2}{6} + \int_{\mathbb{R}} (1 - \operatorname{sinc}(tx))\,\mathrm{d}\nu_k(x)\,.$$

Comme $1 - \operatorname{sinc}(y) \geq \frac{\min\left(1,y^2\right)}{6}$, on a $-\frac{1}{2}\int_{-1}^{1} \psi_k(\xi)\,\mathrm{d}\xi \geq \frac{1}{6}\int_{\mathbb{R}} \min(1, x^2)\,\mathrm{d}\nu_k(x) \geq \frac{1}{6}\rho_k(\mathbb{R})$ en $t = 1$. Par convergence en loi de (μ_k), on a convergence localement uniforme des exposants ψ_k vers ψ. On a donc convergence de la suite $\left(-\frac{1}{2t}\int_{-t}^{t} \psi_k(\xi)\,\mathrm{d}\xi\right)$, donc une majoration uniforme sur les $\rho_k(\mathbb{R})$.

Par le même argument, la suite (σ_k^2) est bornée. En outre, $\int_{\mathbb{R}} x^2 c(x)\,\mathrm{d}\nu_k(x) \leq \rho_k([-1,1])$ est naturellement bornée, donc la suite $\left(\widetilde{\sigma_k}^2\right)$ l'est également.

Comme $\psi(0) = 0$, à nouveau par convergence localement uniforme, pour tout $\varepsilon > 0$, il existe un seuil $t > 0$ en deçà duquel, pour tout $k \in \mathbb{N}$, $\left|-\frac{1}{2t}\int_{-t}^{t} \psi_k(\xi)\,\mathrm{d}\xi\right| \leq \varepsilon$. En conséquence, on a l'inégalité :

$$\varepsilon \geq \frac{1}{6}\int_{\mathbb{R}} \min\big(1,(tx)^2\big)\,\mathrm{d}\nu_k(x) \geq \frac{1}{6}\int_{|x|>\frac{1}{t}} \mathrm{d}\nu_k(x) \geq \frac{1}{6}\times\rho_k\left(\left[-\frac{1}{t},\frac{1}{t}\right]^c\right),$$

d'où le résultat souhaité lorsque $M=\frac{1}{t}\to\infty$. La suite (ρ_k) est donc tendue.

Soit alors ρ une valeur d'adhérence de la suite. Quitte à extraire une sous-suite, on suppose non seulement que $\rho_{\theta(k)} \rightharpoonup \rho$ mais aussi que $\widetilde{\sigma_{\theta(k)}}^2 \longrightarrow \widetilde{\sigma}^2$. On a alors la convergence faible de $\big(\nu_{\theta(k)}\big)$ vers $\nu = \frac{\mathbb{1}_{x\neq 0}}{1\wedge x^2}\rho$, contre les fonctions $f(x)=g(x)\min(1,x^2)$ telles que $g(0)=0$ et g continue bornée. Quitte à joindre ces deux convergences avec la convergence simple $\psi_k(1)\underset{k\to\infty}{\longrightarrow}\psi(1)$, on a nécessairement convergence de $\big(a_{\theta(k)}\big)$ vers une limite a. Autrement dit, pour la sous-suite $\big(\psi_{\theta(k)}\big)$, le sens réciproque s'applique, donc (a,σ^2,ρ) est l'unique triplet de Lévy de la loi μ, qui est donc ID.

En d'autres termes, on a montré l'unicité de la valeur d'adhérence de la suite tendue (ρ_k), qui est donc convergente en loi. On peut alors reproduire ce raisonnement pour montrer l'unicité de la valeur d'adhérence de $\big(\widetilde{\sigma_k}^2\big)$, et en déduire finalement la convergence de (a_k). $\square$

On a donc établi l'équivalence entre les lois ID et les triplets de Lévy. On a vu, en outre, que tout processus de Lévy correspond à une famille de lois ID (μ_t). Montrons maintenant qu'un tel processus est en fait caractérisé par une unique loi ID.

9.2 Décomposition de Lévy

Proposition 162 :

Soit (X_t) un processus de Lévy, de lois marginales $(\mu_t)_{t\geq 0}$, toutes ID. Il existe un exposant ψ tel que, pour tout $t\geq 0$, on a $\widehat{\mu_t}=e^{t\psi}$, autrement dit $t\psi$ est l'exposant de μ_t.

Démonstration. Soit ψ l'exposant de μ_1. Comme $\mu_{\frac{1}{q}}^{*q}=\mu_1$, on en déduit que $\widehat{\mu_{\frac{1}{q}}}=e^{\frac{\psi}{q}}$, la fonction $\frac{1}{q}\psi$ est l'exposant de $\mu_{\frac{1}{q}}$. En passant à la puissance p, on en déduit pour tout rationnel $\frac{p}{q}$, $\frac{p}{q}\psi$ est l'exposant de $\mu_{\frac{p}{q}}$.

Reste donc à montrer la continuité à droite de $t\mapsto\mu_t$, pour la convergence en loi, pour conclure. Si $t_n\to t^+$ par valeurs supérieures, on a $\mu_{t_n}=\mu_t*\mu_{t_n-t}$. Il suffit donc de justifier que si $\varepsilon_n\to 0$, alors $\mu_{\varepsilon_n}\rightharpoonup\delta_0$. Ceci vient du fait que $\mu_0=\delta_0$, et que les trajectoires de X sont continues à droite en 0. Autrement dit, $X_t \underset{t\to 0^+}{\xrightarrow{\text{p.s.}}} 0$, donc *a fortiori* X_t converge en loi vers X_0. $\square$

En conséquence, tout processus de Lévy est entièrement déterminé en loi par un triplet. Il n'est pour l'instant pas évident que cette injection est une bijection, qu'à tout triplet correspond un processus de Lévy.

Théorème 163 (Décomposition de Lévy-Itô) **:**

Soit (a, σ^2, ν) un triplet admissible. Alors il existe un processus de Lévy X associé à ce triplet.

Démonstration. Soit B un MB usuel. Alors $(at + \sigma B_t)$ est un processus de Lévy, associé au triplet $(a, \sigma^2, 0)$. Construisons désormais un processus de Lévy indépendant de B, associé au triplet $(0, 0, \nu)$, ce qui terminera la preuve (le triplet d'une somme finie est la somme des triplets).

Considérons à nouveau $(T_n)_{n\in\mathbb{N}}$ le découpage de $\mathbb{R}^*$ en tranches. Pour chaque tranche, on introduit un processus de Poisson composé, indépendant des autres, d'intensité ν_n. On pose alors $P_t^0 = PPC(\nu_0)$, et $P_t^n = PPC(\nu_n) - t \int\limits_{\mathbb{R}} x \,\mathrm{d}\nu_n(x)$ si $n \geq 1$.

Notre candidat est le processus $X = \sum\limits_{n=0}^{\infty} P^n$, a priori mal défini. Soit la somme partielle $S^N = \sum\limits_{n=1}^{N} P^n$, encore un processus de Lévy. Montrons que $\left(S^N\right)_{N\in\mathbb{N}}$ converge en loi dans $\mathcal{D}(\mathbb{R}^+)$.

Pour une mesure finie ν, on a $\ln\left(\mathbb{E}\left[e^{i\xi PPC(\nu)_t}\right]\right) = t \int\limits_{\mathbb{R}} \left(e^{ix\xi} - 1\right) \mathrm{d}\nu(x)$. La dérivée selon ξ, en 0, donne $i\mathbb{E}[PPC(\nu)_t] = it \int x \,\mathrm{d}\nu(x)$. La dérivée seconde donne $-\mathrm{Var}(PPC(\nu)_t) = -t \int x^2 \,\mathrm{d}\nu(x)$ en $\xi = 0$.

Montrons maintenant que (S_N) vérifie le critère de Kolmogorov discontinu :

$$\begin{aligned}
\mathbb{E}\left[\left(S_t^N - S_s^N\right)^2\left(S_s^N - S_r^N\right)^2\right] &= \mathbb{E}\left[\left(S_t^N - S_s^N\right)^2\right]\mathbb{E}\left[\left(S_s^N - S_r^N\right)^2\right] \\
&= \left(\sum_{n=1}^{N} \mathbb{E}\left[(P_t^n - P_s^n)^2\right]\right)\left(\sum_{n=1}^{N} \mathbb{E}\left[(P_s^n - P_r^n)^2\right]\right) \\
&= \left(\sum_{n=1}^{N} \mathbb{E}\left[\left(P_{t-s}^n\right)^2\right]\right)\left(\sum_{n=1}^{N} \mathbb{E}\left[\left(P_{s-r}^n\right)^2\right]\right) \\
&\leq (t-s)(s-r)\left(\int_{-1}^{1} \mathrm{d}\nu(x)\right)^2 \\
&\leq C(t-r)^2\,,
\end{aligned}$$

car on regarde une somme de variables centrées, indépendantes. La famille $\left(S^N\right)$ est donc tendue, et ses lois fini-dimensionnelles sont convergentes, d'où la convergence en loi souhaitée vers un certain processus S^∞. Par convergence simple des exposants, on en déduit que ce processus qui correspond au triplet $\left(0, 0, \nu|_{[-1,1]}\right)$. Quitte à considérer $P^0 + S^\infty$, on obtient alors le processus souhaité. □

Remarque 164 (Convergence localement uniforme) **:**

On peut en réalité obtenir une propriété plus forte, à savoir la convergence presque-sûre localement uniforme de S^N vers S^∞.

Remarquons que pour $n \geq 1$, les processus P_n et donc S_N sont des martingales L^2, et à trajectoires càdlàg. En utilisant une généralisation de l'inégalité de Doob, on a :

$$\mathbb{E}\left[\sup_{0\leq s\leq t}|M_s|^2\right] \leq 4\mathbb{E}\left[M_t^2\right].$$

Dès lors, on en déduit ici que :

$$\mathbb{E}\left[\sup_{0\leq s\leq t}\left(S_s^{N+P} - S_s^N\right)^2\right] \leq 4t\int_{-\frac{1}{N}}^{\frac{1}{N}} x^2\,\mathrm{d}\nu(x)\,,$$

donc pour tout $t \geq 0$, on a $\mathbb{E}\left[\left\|S^N - S^M\right\|_{\infty,[0,t]}^2\right] \underset{M,N\to\infty}{\longrightarrow} 0$, ce qui permet à terme d'obtenir la convergence annoncée.

9.3 Triplet de Lévy et régularité du processus associé

Soit X un processus de Lévy associé au triplet (a, σ^2, ν). Lorsque $\min(1, |x|) \in L^1(\nu)$, on peut réécrire $X_t = \left(a - \int_{-1}^{1} x\,\mathrm{d}\nu(x)\right)t + \sigma B_t + PPC(\nu)_t$.

On va par la suite justifier la hiérarchisation suivante :

	ν	Régularité
Niveau 0	$\nu = 0$	MB avec drift, continu.
Niveau 1	$\nu(\mathbb{R}) < \infty$	MB+$PPC(\nu)$, ensemble de sauts discret, trajectoires càdlàg.
Niveau 2	$\nu(\mathbb{R}) = \infty$ mais $1 \wedge \|x\| \in L^1(\nu)$	MB+$PPC(\nu)$, ensemble de sauts dense, variations finies ssi $\sigma^2 = 0$.
Niveau 3	$1 \wedge \|x\| \in L^2(\nu)$ mais $1 \wedge \|x\| \notin L^1(\nu)$	Pas de décomposition MB+PPC, ensemble de sauts dense, variations localement infinies.

Théorème 165 :

Soit (X_t) un processus de Lévy, de triplet (a, σ^2, ν). On pose $\Delta X_s = X_s - X_{s-}$ la valeur du saut de X en s. On pose maintenant $M = \sum_{s,\, \Delta X_s \neq 0} \delta_{s, \Delta X_s}$.

Alors M est une mesure aléatoire de Poisson sur $\mathbb{R}^+ \times \mathbb{R}$, de loi $MAP(\lambda \otimes \nu)$.

Démonstration. Utilisons la décomposition de Lévy-Itô de X : $X_t = at + \sigma B_t + \sum_{n=0}^{\infty} \widetilde{P}_t^{\nu_n}$, où les $\widetilde{P}$ sont des processus de Poisson composés, éventuellement compensés.

La composante brownienne et le drift sont continus, sans sauts. Chaque processus $\widetilde{P}^{\nu_n}$ a des sauts de loi $MAP(\lambda \otimes \nu_n)$, bien définie. Par indépendance des processus, on n'a p.s. jamais deux sauts simultanés. En conséquence, la mesure des sauts de X est la somme de celles des processus $\widetilde{P}^{\nu_n}$. Par superposition des mesures aléatoires, on a le résultat voulu. $\square$

Corollaire 166 :

Un processus de Lévy est continu p.s. ssi $\nu = 0$. On a ainsi caractérisé le niveau 0.

Remarque 167 :

On peut utiliser cette caractérisation pour définir les mouvements browniens (avec drift) comme des processus de Lévy continus.

Ceci permet également d'étendre la notion de mouvements browniens sur des groupes, comme $SO_n(\mathbb{R})$, et par extension à des objets sur lesquels les groupes agissent, comme la sphère.

Corollaire 168 :

Lorsque $\nu \neq 0$, l'ensemble des sauts est infini. Si $\nu(\mathbb{R}) < \infty$, alors l'ensemble des sauts est tout de même discret. On a ainsi caractérisé le niveau 1. Sinon, l'ensemble des sauts est presque-sûrement dense dans $\mathbb{R}^+$.

Il reste donc à étudier les variations du processus pour distinguer les niveaux 2 et 3.

Lemme 169 :

On a $\mathbb{E}\left[\sum_{a \leq s \leq b} \mathbb{1}_{|\Delta X_s| \leq 1} |\Delta X_s|\right] = (b-a) \int_{-1}^{1} |x| \, \mathrm{d}\nu(x)$.

Lorsque $1 \wedge |x| \in L^1(\nu)$, en particulier, $\sum_{a \leq s \leq b} |\Delta X_s| < \infty$.

Démonstration. Pour la tranche T_n, on pose $A_n(X, a, b) = \sum_{a \leq s \leq b} \mathbb{1}_{|\Delta X_s| \in T_n} |\Delta X_s|$. Comme A_n est la quantité de variations d'une loi de Poisson composée, on a vu que $\mathbb{E}[A_n] = (b-a) \int |x| \, \mathrm{d}\nu_n(x)$, et $\mathrm{Var}(A_n) = (b-a) \int x^2 \, \mathrm{d}\nu_n(x)$. En passant à la somme sur toutes les tranches (sauf T_0), on a bien l'égalité annoncée.

En particulier, lorsque $1 \wedge |x| \in L^1(\nu)$, on a $\mathbb{E}\left[\sum_{a \leq s \leq b} \mathbb{1}_{|\Delta X_s| \leq 1} |\Delta X_s|\right] = (b-a)\nu\big(|x|\mathbb{1}_{|x|\leq 1}\big) < \infty$. Cette grandeur est intégrable, donc finie p.s. En outre, $\nu_0(1) < \infty$, donc le nombre de sauts plus grands que 1 est fini p.s. En conséquence, quitte à ajouter ce nombre fini de sauts, on a bien $\sum_{a \leq s \leq b} |\Delta X_s| < \infty$ p.s. □

Définition 170 (Variations de X) **:**

On pose $var(X, [a, b]) = \sup_{a=t_0<\cdots<t_p=b} \sum_{j=1}^{p} \left|X_{t_j} - X_{t_{j-1}}\right|$. Attention à ne pas confondre var les variations du processus avec Var la variance d'une variable.

On dit que X est à variations localement bornées lorsque $var(X, [a, b]) < \infty$ pour tous $a < b$.

Par croissance, il suffit de vérifier la propriété p.s. sur un intervalle $[a, b]$ quelconque pour l'obtenir simultanément sur tout intervalle.

Proposition 171 :

Supposons $1 \wedge |x| \in L^2(\nu)$, mais $\nu(\mathbb{R}) = \infty$. Autrement dit, on considère le cas d'un ensemble de sauts dense.

Alors X est à variations localement finies ssi $1 \wedge |x| \in L^1(\nu)$ et $\sigma^2 = 0$.

Démonstration. La variation issue du drift est toujours finie.

Si $1 \wedge |x| \in L^1(\nu)$, alors $\sum_{a \leq s \leq b} |\Delta X_s| < \infty$ traduit la variation du processus X issue des sauts induits par ν. Dans ce cas, comme le mouvement brownien est à variations infinies, on a X lui-même à variations finies ssi $\sigma^2 = 0$.

En revanche, lorsque $\sigma^2 \neq 0$, on ajoute un brownien à variations infinies à un processus VF, donc on obtient un processus à variations localement finies.

Lorsque $1 \wedge |x| \in L^2(\nu)$ n'est pas intégrable, alors $\sum_{a \leq s \leq b} |\Delta X_s| = \infty$ p.s. dès que $a < b$. En effet, quitte à faire des paquets de tranches, on se ramène à $B_n = \sum_{k=k_n}^{k_{n+1}-1} A_n$, une somme de variables indépendantes, telle que $\mathbb{E}[B_n;] \geq 1$. Par l'inégalité de Bienaymé-Chebychev, on a alors :

$$\mathbb{P}\left(|B_n - \mathbb{E}[B_n]| \geq \frac{1}{2}\right) \leq 4\mathrm{Var}(B_n) = 4 \sum_{k=k_n}^{k_{n+1}-1} \mathrm{Var}(A_k)\,.$$

Comme ν intègre x^2 en 0, cette famille est bien sommable. Par le lemme de Borel-Cantelli, les évènements ne sont pas réalisés à partir d'un rang. À partir de ce rang, on a a fortiori $B_n \geq \mathbb{E}[B_n] - \frac{1}{2} \geq \frac{1}{2}$, et alors $\sum_{a \leq s \leq b} |\Delta X_s|$, en tant que somme des B_n, est infini p.s.

Ainsi, lorsque $\sigma^2 = 0$, on ajoute un processus à variations finies et un processus à variations infinies, donc les variations sont encore localement infinies.

On admettra le cas restant, lorsque $\sigma^2 \neq 0$, qui est un peu plus subtil à justifier, car il faut vérifier que les variations infinies du mouvement brownien, continu, et des sauts, ne viennent pas se compenser. □

9.4 Cas des subordinateurs

Définition 172 :
Un subordinateur est un processus de Lévy qui est croissant p.s. sur $\mathbb{R}^+$.

Remarque 173 :
Un tel processus est à variations bornées, car on a $var(X, [a, b]) = X_b - X_a < \infty$ p.s. D'après la classification précédente, on a donc $\sigma^2 = 0$ et $1 \wedge |x| \in L^1(\nu)$.

On pourra alors écrire $X_t = mt + P_t^{\nu}$.

Proposition 174 :
X est un subordinateur ssi $m \geq 0$ et $\nu(\mathbb{R}^-) = 0$.

Démonstration. Le sens réciproque est évident.

Pour le sens direct, supposons d'abord que $m \leq 0$. Pour ε assez petit :

$$\mathbb{E}\left[\sum_{s\leq 1} \mathbb{1}_{|\Delta X_s|\leq\varepsilon}|\Delta X_s|\right] = \int_{-\varepsilon}^{\varepsilon} |x| \, \mathrm{d}\nu(x) \leq \frac{|m|}{2} \,,$$

et donc, avec probabilité p positive, on a $\sum\limits_{s\leq 1} \mathbb{1}_{|\Delta X_s|\leq\varepsilon}|\Delta X_s| \leq \frac{|m|}{2}$. En outre, avec probabilité $q > 0$, $\sum\limits_{s\leq 1} \mathbb{1}_{|\Delta X_s|>\varepsilon}|\Delta X_s| = 0$. En conséquence, par indépendance entre les deux sommes, avec probabilité au moins $pq > 0$, on a $X_1 \leq m + \frac{|m|}{2} \leq \frac{m}{2} < 0$, ce qui contredit la croissance de X, donc $m \geq 0$.

Supposons désormais que $\nu(\mathbb{R}^-) > 0$. On a donc $\varepsilon > 0$ tel que $\nu(]-\infty, \varepsilon]) = \lambda > 0$. Avec probabilité $p > 0$, on a $\sum\limits_{s\leq 1} \mathbb{1}_{\Delta X_s\geq\varepsilon}\Delta X_s = 0$. Avec probabilité $q > 0$ on a $\sum\limits_{s\leq 1} \mathbb{1}_{|\Delta X_s|\leq\varepsilon}|\Delta X_s| \leq K$. Avec probabilité $r > 0$ on a plus de $\frac{2(K+m)}{\varepsilon}$ sauts de valeur au plus $-\varepsilon$ pour $s \in [0, 1]$. En conséquence, par indépendance de ces tranches, avec probabilité au moins $pqr > 0$, on a :

$$X_1 \leq m + K - \varepsilon\frac{2(K + m)}{\varepsilon} = -(K + m) < 0 \,,$$

ce qui contredit à nouveau les propriétés de X, d'où $\nu(\mathbb{R}^-) = 0$. □

Corollaire 175 :

En termes d'exposant, on a X un subordinateur ssi $\psi(\xi) = im\xi + \int_0^\infty \left(e^{ix\xi} - 1\right) \mathrm{d}\nu(x)$, ssi pour tout $u \geq 0$ on a :

$$\mathbb{E}\left[e^{-uX_t}\right] = \exp\left(t\left(-mu + \int_0^\infty \left(e^{-xu} - 1\right) \mathrm{d}\nu(x)\right)\right).$$

9.4.1 Application aux temps d'atteinte d'un MB

Soit B un MB issu de 0. On pose $T_x = \inf\{t \geq 0,\, B_t = x\}$ le temps d'atteinte de x, fini p.s.

Lemme 176 :

Le processus $(T_x)_{x\geq 0}$ est un subordinateur.

Démonstration. Naturellement, $x \mapsto T_x$ est croissante, et admet des limites à gauche. Avec les propriétés du mouvement brownien, pour tout $\eta > 0$, on obtient un $\varepsilon > 0$ pour lequel $B_{T_x+\varepsilon} = B_{T_x} + \eta = x + \eta$. En conséquence, $T_{x+\eta} \leq T_x + \varepsilon$. Ceci est vrai pour tout ε, donc par croissance, on a la continuité à droite en x.

On a montré que, pour tout x fixé, T est p.s. càdlàg en x. À modification près, on peut obtenir T un processus à trajectoires globalement càdlàg.

Par propriété de Markov forte, $B_t^x := B_{T_x+t} - B_{T_x}$ est un MB, indépendant de $\mathcal{F}_{T_x}$, et donc de T_x. En conséquence, $T_{x+y} = T_y(B^x) + T_x(B)$ est la somme de deux temps d'atteinte indépendants, avec les lois voulues. On a bien T un processus à accroissements indépendants, stationnaire. □

Remarque 177 (Exposant de T) **:**

On veut désormais calculer l'exposant du processus.

Le processus $\left(e^{uB_t - u^2 t/2}\right)_{t\geq 0}$ est une martingale. Si on l'arrête en T_x, elle est bornée, donc par le théorème d'arrêt :

$$\mathbb{E}\left[e^{ux - u^2 T_x/2}\right] = 1\,,$$

d'où $\mathbb{E}\left[e^{-\lambda T_x}\right] = e^{-x\sqrt{2\lambda}}$. Cette méthode nous donne l'exposant, mais pas son triplet.

Par principe de réflexion, on a :

$$\mathbb{P}(T_x \leq t) = \mathbb{P}(S_t \geq x) = 2\mathbb{P}(B_t \geq x) = 2\int_x^\infty \frac{e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}}\,\mathrm{d}y = 2\int_{\frac{x}{\sqrt{t}}}^\infty e^{-\frac{y^2}{2}} \frac{\mathrm{d}y}{\sqrt{2\pi}}\,,$$

donc T_x a pour densité $f_{T_x}(t) = \frac{xe^{-x^2/2t}}{t\sqrt{2\pi t}}$. La loi de T_1 est aussi la limite en loi de $PC\left(nT_{\frac{1}{n}}\right)$.

L'exposant de T_1 s'exprime donc comme la limite :

$$\int \left(e^{it\xi}-1\right)\frac{e^{-\frac{1}{2n^2t}}}{t\sqrt{2\pi t}}\,\mathrm{d}t \longrightarrow \int_0^\infty \frac{e^{it\xi}-1}{t\sqrt{2\pi t}}\,\mathrm{d}t\,,$$

par convergence dominée. On peut vérifier que $\mathrm{d}\nu(t)=\frac{\mathrm{d}t}{t\sqrt{2\pi t}}$ intègre $1\wedge|x|$. On a donc explicité un subordinateur sans drift, de mesure ν.

9.4.2 Application au MB en dimension 2

Soient B^1 et B^2 deux MB issus de 0 indépendants. On pose T_x le temps d'atteinte de x par B^1. Quelle est la loi de $Y_x := B^2_{T_x}$? On peut vérifier que (Y_x) est un processus de Lévy.

Remarque 178 (Exposant de Y) **:**

On peut écrire Y_x comme une gaussiennes dont la variance est elle-même aléatoire. Autrement dit, $Y_x \overset{\mathrm{d}}{=} \sqrt{T_x}G$ où $G\sim\mathcal{N}(0,1)$ indépendante de T_x. En passant à la transformée de Fourier :

$$\mathbb{E}\left[e^{i\xi Y_1}\right]=\mathbb{E}\left[e^{i\xi\sqrt{T_1}G}\right]=\mathbb{E}\left[e^{-\frac{\xi^2T_1}{2}}\right]=e^{-|\xi|}\,.$$

On a donc Y_1 une variable de Cauchy, de densité $\frac{1}{\pi(1+y^2)}$. On peut vérifier que sa mesure de Lévy est sous la forme $\mathrm{d}\nu(t)=C\times\frac{\mathrm{d}t}{t^2}$.