Convergence of Probability Measure and Optimal Transport
M2 RI
Lecture notes

R. Chhaibi and C. Pellegrini
Institut de Mathématiques.
Université Paul Sabatier.
31062 Toulouse cedex 09.

2 octobre 2024

Table des matières

Chapitre 1

Introduction/Notations/Preliminary

1.1 Notations

On the whole we shall work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The random variable will take value on a set E. The set E will denote a Polish space, that is a metrizable complete separable space. A separable space is a topological space that contains a dense and at most countable subset, i.e., it contains a finite or countable set of points whose closure is equal to the entire topological space.

— We denote by $\mathcal{B}(E)$ the set of Borel sets on E.

— $\mathcal{M}_1(E)$ will denote the set of probability measures on E

— $C(E)$ will denote the set of continuous functions from E to $\mathbb{R}$

— $C_c(E) \subset C(E)$ will denote the continuous functions on E with compact support

— $C_b(E) \subset C(E)$ will denote the bounded continuous functions on E

— If $X : \Omega \to E$ is a random variable (r.v), we denote $\mathcal{L}(X)$ the law of X and we denote

$$\mathbb{E}(f(X)) = \int_E f(x) d\mu(x),$$

where $\mu = \mathcal{L}(X)$. We shall also use the notation

$$\mu(f) = \int_E f(x) d\mu(x)$$

— For a r.v X valued in $\mathbb{R}^d$, we denote

$$\Phi_X(\xi) = \mathbb{E}[e^{i\langle \xi, X \rangle}],$$

for all $\xi \in \mathbb{R}^d$. This is the so-called characteristic function.

Recall the usual definition of convergence in probability measure

Definition 1. A sequence of random variable (X_n) defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and valued in E converge in law (or in distribution) to a r.v X if

$$\lim_n \mathbb{E}(f(X_n)) = \mathbb{E}(f(X)),$$

for all $f \in C_b(E)$. We denote

$$X_n \xrightarrow[n \to \infty]{\mathcal{L}} X$$

Let us recall the link between convergence in law and vague or tight convergence

Definition 2. A sequence of measure (μ_n) converges tightly to a measure μ if

$$\int f d\mu_n \underset{n\to\infty}{\to} \int f d\mu$$

for all $f \in C_b(E)$.

A sequence of measure (μ_n) converges vaguely to a measure μ if

$$\int f d\mu_n \underset{n\to\infty}{\to} \int f d\mu$$

for all $f \in C_c(E)$

It is then clear that

$$\text{Tight CV} \Rightarrow \text{Vague CV}$$

but the converse is not true. Indeed consider (δ_n).

We leave the following result as an exercize

$$\text{Vague CV} + \mu_n(\mathbb{R}^d) \to \mu(\mathbb{R}^d) \Rightarrow \text{Tight CV}$$

Recal that if V is a topological vector space, the space V^* denote the set of continuous linear form.

Definition 3. We say that a sequence (x_n) converges weakly to x in V if for all $\ell \in V^*$

$$\ell(x_n) \underset{n\to\infty}{\longrightarrow} \ell(x)$$

We say that a sequence ℓ_n converges $*$ weakly to ℓ on V^* if for all $x \in V$

$$\ell(x_n) := \langle l_n, x\rangle \underset{n\to\infty}{\longrightarrow} \langle \ell, x\rangle.$$

Recall that when $V = (C_0(E), \|.\|_\infty)$, we have V^* is the set of Borelian measure which are signed and finite.

Then essentially we have

$$\begin{aligned} \text{CV in law} &\equiv \text{Tight CV on } \mathcal{M}_1(E) &(1.1)\\ &\equiv \text{Vague CV on } \mathcal{M}_1(E) \text{ towards limits on } \mathcal{M}_1(E) &(1.2)\\ &\equiv *\text{weak CV with } V = (C_0(E), \|.\|_\infty) \text{ restricted to } \mathcal{M}_1(E) &(1.3)\end{aligned}$$

The following proposition can be useful

Proposition 1. *Let H be a dense subset of $\big(C_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. Let μ and $(\mu_n)_{n\geq 0}$, probability measures on $\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$.*

Then

$$\mu_n \overset{tightly}{\underset{n\to\infty}{\longrightarrow}} \mu \quad \Leftrightarrow \quad \forall \varphi \in H,\ \lim_{n\to\infty} \int \varphi\, d\mu_n = \int \varphi\, d\mu.$$

1.2 Warm-up

In this section, the aim is to motivate the whole lecture by making a first link between the well-known Central Limit Theorem and the Brownian motion. The complete link figured out by the so called Donsker's Theorem will be a red thread during this lecture. We start by recalling the result of central limit theorem by giving different proofs. A important result used in the proof is the Levy Theorem.

Theorem 2. (Lévy).

1. *If* $X_n \xrightarrow{\mathcal{L}} X$ *then* $\forall \xi \in \mathbb{R}^d$, $\lim_{n\to\infty} \Phi_{X_n}(\xi) = \Phi_X(\xi)$.
2. *If there exists a function* $\Phi : \mathbb{R}^d \to \mathbb{C}$ *continuous in 0 such that for all* $\xi \in \mathbb{R}^d$,

$$\lim_{n\to\infty} \Phi_{X_n}(\xi) = \Phi(\xi),$$

then there exists a unique probability measure μ *on* $\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$ *such that* $\Phi = \hat{\mu}$. *Moreover* $\mu_n \xrightarrow{tight} \mu$ *and if* X *is a r.v of law* μ *we have* $X_n \xrightarrow{\mathcal{L}} X$.

Theorem 3. *Let* $(X_n)_{n\geq 0}$ *be a sequence of i.i.d r.v valued in* $\mathbb{R}$ *with*

$$\mathbb{E}[X_1] = 0, \quad \mathbb{E}[X_1^2] = \sigma^2.$$

Denote for all $n \in \mathbb{N}^*$

$$S_n = \sum_{i=1}^n X_i,$$

then

$$\frac{S_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

We shall need a Lemma which we shall see later is related to the relative compactness.

Lemma 4.

$$\lim_{K\to\infty} \sup_{n\geq 1} \mathbb{P}\left[\left|\frac{S_n}{\sqrt{n}}\right| \geq K\right] = 0$$

Démonstration. It is a consequence of Byenaimé Tchebychev. □

Démonstration. **First proof : with characteristic function** By independance, we have for all $t \in \mathbb{R}$

$$\begin{aligned} \mathbb{E}\left[e^{it\frac{S_n}{\sqrt{n}}}\right] &= \left(\mathbb{E}[e^{it\frac{X_1}{\sqrt{n}}}]\right)^n \\ &= \left(1 - \frac{t^2}{2n}(1 + \circ(1))\right)^n \end{aligned}$$

Then

$$\lim_{n\to\infty} \mathbb{E}[e^{it\frac{S_n}{\sqrt{n}}}] = e^{-t^2/2} = \mathbb{E}[e^{it\mathcal{N}}].$$

The Levy Theorem yields the result. Note that here we have identified the limit characteristic function then it is automatically continuous in 0. Hidden by this continuity there is an important fact of relative compactness.

Second proof : with Lindberg swapping trick. For this proof, we shall suppose that $\mathbb{E}[|X_1|^3] < \infty$. An obvious observation is that if $X_1 \sim \mathcal{N}(0, \sigma^2)$ then

$$\frac{S_n}{\sqrt{n}} \sim \mathcal{N}(0, \sigma^2)$$

and the convergence is straightforward. Now let us consider $\phi \in C_c^3(\mathbb{R})$, and consider a sequence of i.i.d r.v $(\mathcal{N}_i)_{i \in \mathbb{N}^*}$ where $\mathcal{N}_1 \sim \mathcal{N}(0, \sigma^2)$. We suppose them independent of the sequence $(X_i)_{i \in \mathbb{N}^*}$. Let $\mathcal{N} \sim \mathcal{N}(0, \sigma^2)$, as already said we have

$$\mathcal{N} \stackrel{\mathcal{L}}{=} \frac{\sum_{i=1}^{n} \mathcal{N}_i}{\sqrt{n}} := S_n^{\mathcal{N}}$$

Let us define for $i = 1, \ldots, n$

$$S_n^i = \sum_{j=1}^{i} X_j + \sum_{j=i+1}^{n} \mathcal{N}_j$$

Let us check that we have

$$S_n = S_n^n, \quad S_n^{\mathcal{N}} = S_n^0$$

and

$$S_n^{i+1} = \hat{S}_n^i + X_{i+1}, \quad S_n^i = \hat{S}_n^i + \mathcal{N}_{i+1}$$

where we have defined

$$\hat{S}_n^i = \sum_{j=1}^{i} X_j + \sum_{j=i+2}^{n} \mathcal{N}_j \quad \perp\!\!\!\perp \quad (X_{i+1}, \mathcal{N}_{i+1}),$$

where the symbol $\perp\!\!\!\perp$ means independent.

Using a telescopic sum and a Taylor formula, we have

$$\begin{aligned}
\mathbb{E}\left[\phi\left(\frac{S_n}{\sqrt{n}}\right) - \phi(\mathcal{N})\right] &= \mathbb{E}\left[\phi\left(\frac{S_n}{\sqrt{n}}\right) - \phi\left(\frac{S_n^0}{\sqrt{n}}\right)\right] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left[\phi\left(\frac{S_n^{i+1}}{\sqrt{n}}\right) - \phi\left(\frac{S_n^i}{\sqrt{n}}\right)\right] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left[\phi\left(\frac{\hat{S}_n^i + X_{i+1}}{\sqrt{n}}\right) - \phi\left(\frac{\hat{S}_n^i + \mathcal{N}_{i+1}}{\sqrt{n}}\right)\right] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left[\phi'\left(\frac{\hat{S}_n^i}{\sqrt{n}}\right)\left(\frac{X_{i+1} - \mathcal{N}_{i+1}}{\sqrt{n}}\right) + \frac{1}{2}\phi''\left(\frac{\hat{S}_n^i}{\sqrt{n}}\right)\frac{X_{i+1}^2 - \mathcal{N}_{i+1}^2}{n}\right. \\
&\quad \left. + \mathcal{O}(\|\phi'''\|_\infty)\frac{|X_i|^3 + |\mathcal{N}_{i+1}|^3}{n^{3/2}}\right] \\
&= \mathcal{O}(\|\phi'''\|_\infty) \sum_{i=0}^{n-1} \mathbb{E}\left[\frac{|X_i|^3 + |\mathcal{N}_{i+1}|^3}{n^{3/2}}\right] \\
&= \mathcal{O}(\|\phi'''\|_\infty)\mathbb{E}\left[\frac{|X_1|^3 + |\mathcal{N}_1|^3}{\sqrt{n}}\right]
\end{aligned}$$

Then

$$\lim_{n\to\infty}\mathbb{E}\left[\phi\left(\frac{S_n}{\sqrt{n}}\right)-\phi(\mathcal{N})\right]=0$$

By density of C_c into C_c^3 for the norm $\|.\|_\infty$ this limit holds for all $\phi\in C_c$. The passage from C_c to C_b is more tricky. First note that $\bar{C}_c=C_0$ and then does not give C_b. But we shall use the Lemma. Indeed let $\phi\in C_b$. For all $K\geq 0$ there exists a function $f_K\in C^c$ such that $f_K=1$ on $[-K,K]$ and $f_K=0$ on $[-(K+1),K+1]^c$. then we can consider

$$\phi=\phi f_K+\phi(1-f_K)$$

such that $\phi f_K\in C_c$

$$\begin{aligned}
&\left|\mathbb{E}\left[\phi\left(\frac{S_n}{\sqrt{n}}\right)\right]-\mathbb{E}[\phi(\mathcal{N})]\right|\\
\leq\ &\left|\mathbb{E}\left[\phi f_K\left(\frac{S_n}{\sqrt{n}}\right)\right]-\mathbb{E}[\phi f_K(\mathcal{N})]\right|+\left|\mathbb{E}\left[\phi\left(\frac{S_n}{\sqrt{n}}\right)(1-f_K)\right]\right|+\left|\mathbb{E}\left[\phi\left(\mathcal{N}\right)(1-f_K)\right]\right|\\
\leq\ &\left|\mathbb{E}\left[\phi f_K\left(\frac{S_n}{\sqrt{n}}\right)\right]-\mathbb{E}[\phi f_K(\mathcal{N})]\right|+\|\Phi\|_\infty\left(\sup_{n\geq 1}\mathbb{P}\left[\left|\frac{S_n}{\sqrt{n}}\right|\geq K\right]+\mathbb{P}[|\mathcal{N}|\geq K]\right)
\end{aligned}$$

Then take the limsup in $n\to\infty$ next in $K\to\infty$ yields the result. □

Theorem 5. *Let $0=t_0<t_1<\ldots<t_k$, then*

$$\left(\frac{S_{\lfloor nt_j\rfloor}}{\sqrt{n}},j=1,\ldots,k\right)\xrightarrow[n\to\infty]{\mathcal{L}}\left(\sum_{i=1}^{j}\sqrt{t_i-t_{i-1}}\mathcal{N}_i,j=1,\ldots,k\right)$$

or in an equivalent way

$$\left(\frac{S_{\lfloor nt_j\rfloor}-S_{\lfloor nt_{j-1}\rfloor}}{\sqrt{n}},j=1,\ldots,k\right)\xrightarrow[n\to\infty]{\mathcal{L}}\left(\sqrt{t_j-t_{j-1}}\mathcal{N}_j,j=1,\ldots,k\right)$$

But we could look at $(\frac{S_{\lfloor nt\rfloor}}{\sqrt{n}},t\geq 0)$ or the affine interpolation $(S_t^n,t\geq 0)$. It is then natural to address the question wether a limit theorem can be established for the whole trajectory and not only for a finite number of time.

Then let us speak about the limit object which is the so-called Brownian motion.

Definition 4. A one dimensionnal standard Brownian motion is a stochastic process $(B_t,t\geq 0)$ such that

1. $B_0=0$
2. $t\to B_t$ are almost surely continuous from $\mathbb{R}_+$ to $\mathbb{R}$
3. For all $0=t_0<t_1<\ldots<t_k$, $(B_{t_j}-B_{t_{j-1}},j=1,\ldots,k)$ are independant increments of law $\mathcal{N}(0,t_j-t_{j-1})$

Within this definition, it is not clear that such a process exists. Assuming the existence we want to promote the previous theorem to

$$(S_t^n,t\geq 0)\xrightarrow[n\to\infty]{\mathcal{L}}(B_t,t\geq 0)$$

which means that for all continuous and bounded functions $F : C(\mathbb{R}_+) \mapsto \mathbb{R}$ we have

$$\mathbb{E}\left[F\left(\frac{S_{\lfloor n. \rfloor}}{\sqrt{n}}\right)\right] \to_{n\to\infty} \mathbb{E}[F(B_.)].$$

By example, for $k > 0$

$$F(f) = \min(\sup_{[0,T]} |f(t)|, k)$$

Chapitre 2

Tight convergence in $\mathcal{M}_1(E)$

2.1 Polish space, examples and counter-examples

The typical set that we shall consider during the whole lecture are Polish spaces (as already mentioned). Recall the exact definition.

Definition 5. A topological polish space E is said to be polish when

1. **Metrisable :** there exists $d : E \times E \to \mathbb{R}_+$ a distance which induce the same topology than the one of E.
2. **Complete :** i.e all Cauchy sequences are convergent.
3. **Separable :** there exists a dense and countable family $\{x_n, n \in \mathbb{N}\}$ in E, that is $\overline{\{x_n, n \in \mathbb{N}\}} = E$

Remark 1. Note that sometimes the distance is explicit and tractable. In this sense (E, d) is a natural metric space. Nevertheless, when the distance is neither explicit nor tractable, knowing that E is metrizable allows to exploit several properties (namely convergent subsequences).

Remark 2. The separability can be addressed from a topological point of view or from a metric point of view :

— Metric : for all $\varepsilon > 0$, for all $x \in E$ there exists $n \in \mathbb{N}$, such that $d(x_n, x) \leq \varepsilon$.

— Topologic : for all $x \in E$ and for all neighborhood V of E, there exists $n \in \mathbb{N}$ such that $x_n \in V$.

Let us present some examples

— $E = \mathbb{R}^d$ with its natural topology is Polish. The distance is $d(x, y) = \|x - y\|_2$. The separability comes from $\mathbb{Q}^d$

— All compact metric spaces (E, d) is Polish.

— For all $T > 0$, the space $E = C([0, T], \mathbb{R})$ is Polish for $d = \|.\|_\infty$. Consider $\mathbb{Q}[X] = \bigcup_n \mathbb{Q}_n[X]$ is countable. For the dense character, one can invoque the Stone Weierstrass Theorem.

— A Hilbert space $(E, \|.\|_2)$ with a countable basis is Polish.

— A product of Polish space is Polish

— $C(\mathbb{R}_+, \mathbb{R}^d)$ equipped with the uniform convergence on compact is Polish.

For counterexample

Lemma 6. *Let (E, d) a metric space. Suppose that there exist $\{x_i\}_{i\in I}$ such that there exist $\delta > 0$ such that for all $(i, j) \in I^2$ $i \neq j \Rightarrow d(x_i, x_j) \geq \delta$. Then if I is not countable then E is not separable.*

— $E = \ell^\infty(\mathbb{N})$ with the uniform topology. Consider $\{0, 1\}^{\mathbb{N}}$.
— $E = (L^\infty(\mathbb{R}), \|.\|_\infty)$. Indeed one can inject isometrically $(\ell^\infty(\mathbb{N}), \|.\|_\infty)$
— $E = (C_b(\mathbb{R}_+), \|.\|_\infty)$

2.2 Tightness and Prokhorov Theorem

The notion of tightness is crucial in the rest of the lecture. This notion is enlighted by the Prokhorov Theorem that we shall prove later since it needs a fine understanding of the topology of $\mathcal{M}_1(E)$.

Definition 6. A family $\mathcal{F} \subset \mathcal{M}_1(E)$ is said to be tight if for all $\varepsilon > 0$ there exists a compact set $K \subset E$ such that

$$\sup_{\mu\in\mathcal{F}} \mu(E \setminus K) \leq \varepsilon$$

In other words up to a small error the mass of all measures $\mu \in \mathcal{F}$ is concentrated on a compact set which is common for all measures $\mu \in \mathcal{F}$.

Proposition 7. *Let E be a Polish space. All finite family of $\mathcal{M}_1(E)$ is tight.*

Démonstration. Assume that we have shown that singletons are tight. Then consider $\mathcal{F} = \{\mu_1, \dots, \mu_k\}$, and denote K_i the compact associated to μ_i. The compact set $K = \bigcup_{i=1}^n K_i$ will satisfy the tightness property for $\mathcal{F}$.

Let us show now that a singleton is tight. Let $\varepsilon > 0$, since E is separable there exist a dense sequence (x_n) and for all $k > 0$ we have

$$E = \bigcup_n B_F\left(x_n, \frac{1}{k}\right)$$

By increasing limit, for all $k \geq 1$ there exist N_k such that we have

$$\mu\left(\bigcup_{n=1}^{N_k} B_F\left(x_n, \frac{1}{k}\right)\right) \geq 1 - \frac{\varepsilon}{2^k}$$

Then let us define

$$K = \bigcap_k \bigcup_{n=1}^{N_k} B_F\left(x_n, \frac{1}{k}\right)$$

Let us show that it is a compact set. This is a closed set then it is complete. Note that if a set is pre-compact and complete then it is compact. It is then sufficient to show that for all $\delta > 0$ there exist I with $|I| < \infty$ and

$$K \subset \bigcup_{i\in I} B(y_i, \delta)$$

which is obvious by construction of K. The tightness property is obvious. □

In general showing that an infinite family of measures is tight is a difficult task, sometimes an art. The following Theorem shows the importance of tightness

Theorem 8 (Prokhorov). *Let E be a Polish space and $\mathcal{F} \subset \mathcal{M}_1(E)$.*
The family $\mathcal{F}$ is tight if and only if $\mathcal{F}$ is relatively compact for the tight topology.

The proof is postponed at the end of the chapter.

Remark 3. The implication $\Rightarrow$ needs the Polish property whereas the implication $\Leftarrow$ does not need this property.

Remark 4. As we shall see later the tight convergence is metrisable and we shall present adequate metric. Then the compactness in $\mathcal{M}_1(E)$ can be addressed by using the sequential aspect of compactness.

Proposition 9. *Let $\{\mu_n, n \in \mathbb{N}^*\} \subset \mathcal{M}_1(E)$. If $\{\mu_n, n \in \mathbb{N}^*\}$ is a tight family and if there exists a measure μ such that for all subsequence $(\mu_{\phi(n)})$ one can extract a subsequence which converges weakly to μ then (μ_n) converges weakly to μ.*

2.3 Properties and topology of $\mathcal{M}_1(E)$

Proposition 10. *Let E be a metrisable space and $\mu \in \mathcal{M}_1(E)$ then μ is entirely determined by its values on open sets or in equivalent way on closed sets or equivalently on Lipschitz function.*
More precisely for all Borel set A

$$\begin{aligned} \mu(A) &= \inf\{\mu(O), A \subset O \text{ open set}\} \\ &= \sup\{\mu(O), A \supset F \text{ closed set}\} \end{aligned}$$

Moreover if E is Polish one can replace closed by compact set.

Démonstration. Let us denote $\mathcal{A} \subset \mathcal{B}(E)$ the set of Borelian set such that

$$\begin{aligned} \mu(A) &=^{(1)} \inf\{\mu(O), A \subset O \text{ open set}\} \\ &=^{(2)} \sup\{\mu(F), A \supset F \text{ closed set}\} \end{aligned}$$

First step : Let us show that $\mathcal{A}$ contains the open set. This is obvious for (1). Let A be an open set let us show that $\mu(A)$ satisfies (2). For all $k \in \mathbb{N}^*$, we denote

$$F_k = \left\{ x \in E | d(x, A^c) \geq \frac{1}{k} \right\},$$

where $x \mapsto d(x, Z) = \inf_{z \in Z} d(x, z)$ is continuous. Then F_k is closed by continuity and (F_k) is non decreasing sequence of sets. Since A is open, we have

$$\bigcup_{k \geq 1} F_k = A$$

and

$$\mu(A) = \lim \mu(F_k) \leq \sup\{\mu(F) | F \subset A, \ F \text{ closed}\}$$

The other inequality is straightforward.

Second step : $\mathcal{A}$ is stable by complementary. This is straightforward by remarking that $1-\mu(Z)=\mu(Z^c)$ for all subsets Z and $-\inf\{...\}=sup\{-...\}$.

Third step : Stability by countable union. Let $(A_i)_{i\in\mathbb{N}^*}\in\mathcal{A}^{\mathbb{N}}$. For all $\varepsilon>0$ there exist O_i open and F_i closed such that $F_i\subset A_i\subset O_i$

$$\mu(O_i)-\frac{\varepsilon}{2^i}\leq\mu(A_i)\leq\mu(F_i)-\frac{\varepsilon}{2^i}$$

Then $\bigcup A_i\subset\bigcup O_i$ which is an open set and we have

$$\begin{aligned}\mu(\bigcup_{i\in\mathbb{N}^*}O_i)-\mu(\bigcup_{i\in\mathbb{N}^*}A_i) &= \mu(\bigcup_{i\in\mathbb{N}^*}O_i\setminus\bigcup_{i\in\mathbb{N}^*}A_i)\\ &\leq \mu(\bigcup_{i\in\mathbb{N}^*}O_i\setminus A_i)\\ &\leq \sum_{i\in\mathbb{N}^*}\mu(O_i\setminus A_i)\\ &\leq \sum_{i\in\mathbb{N}^*}\frac{\varepsilon}{2^i}=\varepsilon\end{aligned}$$

Then, at this stage, we have

$$\mu(\bigcup_{i\in\mathbb{N}^*}O_i)-\varepsilon\leq\mu(\bigcup_{i\in\mathbb{N}^*}A_i)$$

For the other inequality, we cannot consider $\bigcup_i F_i$ since this set is not necessarily closed. Nevertheless for all $N\in\mathbb{N}^*$

$$\bigcup_{i=1}^N F_i\subset\bigcup_i A_i$$

We shall use that

$$\mu(\bigcup_i F_i)=\lim_N\mu(\bigcup_{i=1}^N F_i)$$

and if $B\subset D\subset A$ then $A\setminus B=(A\setminus D)\cup(D\setminus B)$

$$\begin{aligned}\mu(\bigcup_{i\in\mathbb{N}^*}A_i\setminus\bigcup_{i=1}^N F_i) &\leq \mu(\bigcup_{i\in\mathbb{N}^*}F_i\setminus\bigcup_{i=1}^N F_i)+\mu(\bigcup_{i\in\mathbb{N}^*}A_i\setminus\bigcup_{i\in\mathbb{N}^*}F_i)\\ &\leq \varepsilon+\sum_i\mu(A_i\setminus F_i)\\ &\leq 2\varepsilon,\end{aligned}$$

for N sufficiently large. Then $\bigcup_i A_i\in\mathcal{A}$.

As a conclusion, we have proved that $\mathcal{A}$ is a σ-algebra that contains the open sets, then $\mathcal{A}=\mathcal{B}(E)$

Concerning the Lipschitz functions. Let F be a closed set and consider

$$f_{K,F}(x)=(1-Kd(x,F))^+$$

We have

$$\mathbf{1}_F(x) \leq f_{K,F}(x) \leq 1$$

and

$$\lim_K f_{K,F}(x) = \mathbf{1}_F(x),$$

for all $x \in E$. The function $f_{K,F}$ is Lipschitz. By dominated convergence Theorem

$$\mu(F) = \mu(\mathbf{1}_F) = \lim_K \mu(f_{K,F})$$

and the result holds.

When E is Polish, then μ is tight, then for all $\varepsilon > 0$, there exist K_ε such that

$$\mu(E \setminus K_\varepsilon) \leq \varepsilon$$

Then if F is closed and $F \subset A$ with $\mu(A) \leq \mu(F) + \varepsilon$, then $F' = F \cap K_\varepsilon$ is a compact set such that $F' \subset F \subset A$ and

$$\mu(A) \leq \mu(F') + 2\varepsilon,$$

which yields the result. □

The following Theorem gives a criterion for tight convergence.

Theorem 11 (Portmanteau)**.** *Let $(\mu_n, n \in \mathbb{N}^*)$ a sequence of probability measures in $\mathcal{M}_1(E)$ and $\mu \in \mathcal{M}_1(E)$ with (E, d) a metric space. The following assertions are equivalent*

1. *In the tight convergence topology, i.e for all bounded and continuous functions we have*
$$\mu_n(f) \to \mu(f)$$
2. *For all bounded and uniformly continuous functions f*
$$\mu_n(f) \to \mu(f)$$
3. *For all bounded and Lipschitz functions*
$$\mu_n(f) \to \mu(f)$$
4. *For all open set $O \subset E$*
$$\liminf_n \mu_n(O) \geq \mu(O)$$
5. *For all closed set $F \subset E$*
$$\limsup_n \mu_n(F) \leq \mu(F)$$
6. *For all Borel set with $\mu(\delta A) = O$*
$$\lim_n \mu_n(A) = \mu(A)$$

Démonstration. 4) $\Leftrightarrow$ 5) It comes from the fact that the complement of a closed set is an open set, and vice versa.

Montrons que 3) $\Rightarrow$ 4). Let O be an open set. Consider the Lipschitz bounded function

$$\phi_k(x) = \min(1, kd(x, O^c))$$

We have $\phi_k \leq \mathbf{1}_O$ and $\phi_k \to_k \mathbf{1}_O$. Then for all n

$$\mu_n(O) \geq \int \phi_k d\mu_n$$

and then

$$\liminf_n \mu_n(O) \geq \liminf_n \int \phi_k d\mu_n = \int \phi_k d\mu$$

Then using dominated convergence Theorem, taking k to infinity yields

$$\liminf_n \mu_n(O) \geq \liminf_k \int \phi_k d\mu_n = \int \mathbf{1}_O d\mu = \mu(O)$$

Show that 4) and 5) implies 6). If $\mu(\partial B) = 0$, since $\partial B = \bar{B} \setminus \mathring{B}$ then

$$\mu(B) = \mu(\mathring{B}) = \mu(\bar{B})$$

then

$$\liminf \mu_n(B) \geq \liminf \mu_n(\mathring{B}) \geq \mu(\mathring{B}) = \mu(B) \quad (2.1)$$

$$\limsup \mu_n(B) \leq \limsup \mu_n(\bar{B}) \leq \mu(\bar{B}) = \mu(B) \quad (2.2)$$

Then

$$\liminf \mu_n(B) = \limsup \mu_n(B) = \lim \mu_n(B) = \mu(B)$$

Show finally that 6) $\Rightarrow$ 1). Let $\phi \in C_b(\mathbb{R}^d)$. Up to considering the negative and the positive part we can suppose that $\phi \geq 0$. We then have

$$\begin{aligned} \int_{\mathbb{R}^d} \phi d\mu_n &= \int_{\mathbb{R}^d} \left(\int_0^{\|\phi\|_\infty} \mathbf{1}_{y \leq \phi(x)} dy \right) d\mu_n(x) \\ &= \int_0^{\|\phi\|_\infty} \mu_n(\{x : \phi(x) \geq y\}) dy \quad (Fubini) \end{aligned}$$

At this stage let us consider $F_y = \{x : \phi(x) \geq y\}$ which contains the open set $\{x : \phi(x) > y\}$. We then have

$$\partial F_y \subset F_y \setminus \{x : \phi(x) > y\} = \phi^{-1}(\{y\}).$$

Then $\mu(\partial F_y) = 0$ for almost all y. Indeed

$$\{y : \mu(\partial F_y) > 0\} \subset \{y : \mu(\phi^{-1}(\{y\})) > 0\}$$

which are the atom of the image measure of μ by ϕ. Then $\{y : \mu(\partial F_y) > 0\}$ is at most countable. For almost all y

$$\mu_n(F_y) \to \mu(F_y)$$

and by dominated convergence Theorem

$$\int_0^{\|\phi\|_\infty} \mu_n(\{x : \phi(x) \geq y\}) dy \to \int_0^{\|\phi\|_\infty} \mu(\{x : \phi(x) \geq y\}) dy = \int_{\mathbb{R}^d} \phi d\mu$$

which yields the result. □

2.4 Some considerations on $\mathbb{R}$ and $\mathbb{R}^d$

The case of $\mathbb{R}$ and $\mathbb{R}^d$ are rich enough and usual results are easy to show

Proposition 12. *Let X and $(X_n)_{n\geq 0}$ be real random variables. Then, $X_n \xrightarrow{\mathcal{L}} X$ if and only if for every $t \in \mathbb{R}$ where F_X is continuous, we have*

$$\lim_{n\to\infty} F_{X_n}(t) = F_X(t).$$

Démonstration. The implication $\Rightarrow$ is a consequence of the Portmanteau Theorem. For the converse, we will prove point 2) of the Portmanteau Theorem. Let's consider an open interval $O =]a, b[$ (the general case follows from a countable union). Since the number of discontinuity points of F_X is countable, we can find a decreasing sequence (a_k) converging to a such that F_X is continuous at a_k for all k, and an increasing sequence (b_k) converging to b such that F_X is continuous at b_k. By right and left continuity, we have that :

$$\begin{aligned} \limsup F_{X_n}(a) &\leq \limsup F_{X_n}(a_k) = F_X(a_k) \to F_X(a) \\ \liminf F_{X_n}(b^-) &\geq \liminf F_{X_n}(b_k) = F_X(b_k) \to F_X(b-) \end{aligned}$$

Then

$$\begin{aligned} \liminf P_{X_n}(]a, b[) &= \liminf F_{X_n}(b^-) - F_{X_n}(a) \\ &\geq F_X(b^-) - F_X(a) = P_X(]a, b[) \end{aligned}$$

and we get the result. □

The Prokhorov Theorem is also easy.

Theorem 13 (Prokhorov)**.** *Let $(\mu_n)_{n\in\mathbb{N}}$ be a tight sequence of probability measures on $\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$. Then, it is possible to extract a subsequence that converges tightly to a probability measure.*

Démonstration. According to the Banach-Alaoglu theorem, we can extract a subsequence (μ_{n_k}) that converges weak-$*$ to a positive measure μ. For any function $f \in C_0(\mathbb{R}^d)$, we have :

$$\int f d\mu_{n_k} \to \int f d\mu.$$

On a que $\mu(\mathbb{R}^d) \leq 1$. To conclude, it suffices to show that $\mu(\mathbb{R}^d) = 1$. Let $\varepsilon > 0$; we can then find a compact set K such that :

$$\sup_{n_k} \mu_{n_k}(K) \geq 1 - \varepsilon.$$

We can find $f \in C_0(\mathbb{R}^d)$ such that $\mathbf{1} \geq f \geq \mathbf{1}_K$. We then have

$$\mu(\mathbb{R}^d) \geq \int f d\mu = \lim \int f d\mu_{n_k} \geq \lim \int \mathbf{1}_K d\mu_{n_k} \geq 1 - \varepsilon$$

Now, as we let ε tend towards 0, we have $\mu(\mathbb{R}^d) \geq 1$, and the result is proven. □

Then we can now prove the Levy Theorem.

Theorem 14. (Lévy).

1. *If* $X_n \xrightarrow{\mathcal{L}} X$*, then for all* $\xi \in \mathbb{R}^d$*,*
2. *If there exists a function* $\Phi : \mathbb{R}^d \to \mathbb{C}$ *continuous at 0, such that for all* $\xi \in \mathbb{R}^d$*,*

$$\lim_{n\to\infty} \Phi_{X_n}(\xi) = \Phi(\xi),$$

then there exists a probability measure μ *on* $\left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right)$ *such that* $\Phi = \hat{\mu}$*. Moreover,* $\mu_n \xrightarrow{tight} \mu$ *and if* X *is a random variable with distribution* μ*, then* $X_n \xrightarrow{\mathcal{L}} X$*.*

Démonstration. Point 1) is evident by considering the real and imaginary parts of $x \mapsto e^{ix}$.

For Point 2), it suffices to show that μ_n is tight. Indeed, the Prokhorov theorem will then provide the existence of a convergent subsequence to a measure μ. Thus, Φ is necessarily the characteristic function of μ. Thus, Φ must be the characteristic function of μ. In this way, all the limit points will be equal to μ (as they will have the same characteristic function).

To demonstrate tightness, we will show that $P_{X_n}([-M, M])$ uniformly converges to 1 as M tends to infinity. On the set $[-M, M]^c$, the quantity

$$1 - \frac{\sin(X/M)}{X/M}$$

is bounded by a certain constant α independent of M. Therefore, we have

$$\begin{aligned}
P_{X_n}([-M, M]^c) &= E\left(\mathbf{1}_{[-M,M]^c}(X_n)\right) \\
&\leq \frac{1}{\alpha} E\left(1 - \frac{\sin(X_n/M)}{X_n/M}\right) \\
&\leq \frac{1}{\alpha} E\left(\frac{M}{2}\int_{-1/M}^{1/M} (1 - e^{iX_n t})dt\right) \\
&\leq \frac{1}{\alpha}\frac{M}{2}\int_{-1/M}^{1/M} (1 - \Phi_{X_n}(t))dt
\end{aligned}$$

When n goes to infinity, by dominated convergence, we have

$$\frac{M}{2}\int_{-1/M}^{1/M} (1 - \Phi_{X_n}(t)) \to \frac{M}{2}\int_{-1/M}^{1/M} (1 - \Phi(t))$$

Then

$$\limsup_n P_{X_n}([-M, M]^c) \leq \frac{1}{\alpha}\frac{M}{2}\int_{-1/M}^{1/M} (1 - \Phi(t))$$

Thus there exist N such that

$$\sup_{n\geq N} P_{X_n}([-M, M]^c) \leq \frac{1}{\alpha}\frac{M}{2}\int_{-1/M}^{1/M} (1 - \Phi(t))$$

Or

$$\lim_M \frac{M}{2}\int_{-1/M}^{1/M} (1 - \Phi(t)) = 1 - \Phi(0) = 0$$

then $P_{X_n}([-M, M]^c)$ converges uniformly to 0 and we get the result. □

Theorem 15. (Lévy)(Version faible). *$X_n \xrightarrow{\mathcal{L}} X$ ssi $\forall \xi \in \mathbb{R}^d$, $\lim_{n\to\infty} \Phi_{X_n}(\xi) = \Phi_X(\xi)$.*

2.5 Distance on $\mathcal{M}_1(E)$

We shall define two distances on $\mathcal{M}_1(E)$ and then show that they give the tight topology. We then consider a Polish space E equipped with a metric d.

Before defining the Levy Prokhorov distance let us define the ε neighborhood. For a set A and ε we define

$$A^\varepsilon = \{x \in E \ s.t \ \exists y \in A \ s.t \ d(x,y) \le \varepsilon\} = \bigcup_{y \in A} B(y,\varepsilon)$$

Proposition 16 (Levy Prokhorov Distance)**.** *Let $(\mu,\nu) \in \mathcal{M}_1(E)^2$ we define*

$$\begin{aligned} \rho(\mu,\nu) &:= \inf\{\varepsilon > 0 | \forall B \in \mathcal{B}(E), \mu(B) \le \nu(B^\varepsilon) + \varepsilon\} \quad (2.3)\\ &= \inf\{\varepsilon > 0 | \forall F \ closed \ \mu(F) \le \nu(F^\varepsilon) + \varepsilon\}. \quad (2.4)\end{aligned}$$

which is a distance on $\mathcal{M}_1(E)$ called Levy Prokhorov distance.

Démonstration. Let us start by proving that we can concentrate on closed set. It is sufficient to show that

$$\begin{aligned} & \forall B \in \mathcal{B}(E), \mu(B) \le \nu(B^\varepsilon) + \varepsilon \quad (2.5)\\ \Leftrightarrow \ & \forall F \text{ closed}, \ \mu(F) \le \nu(F^\varepsilon) + \varepsilon \quad (2.6)\end{aligned}$$

The first implication is straightforward. For the reverse, let us remark that $B^\varepsilon = \bar{B}^\varepsilon$. Then we have

$$\mu(B) \le \mu(\bar{B}^\varepsilon) \le \nu(\bar{B}^\varepsilon) + \varepsilon = \nu(B^\varepsilon) + \varepsilon$$

Now let us check that it is a distance.

First let us show that $\rho(\mu,\mu) = 0$. We have that for all $\varepsilon > 0$

$$\forall B \in \mathcal{B}(E), \mu(B) \le \mu(B^\varepsilon) + \varepsilon$$

Then for all $\varepsilon > 0$, $\rho(\mu,\mu) \le \varepsilon$ which yields $\rho(\mu,\mu) = 0$.

In order to show that ρ is symmetric. We just have to show that for all (μ,ν)

$$\rho(\mu,\nu) \le \rho(\nu,\mu)$$

To this end let us remark that

$$\forall B \in \mathcal{B}(E), ((B^\varepsilon)^c)^\varepsilon \subset B^c$$

Then

$$\begin{aligned} & \varepsilon < \rho(\mu,\nu)\\ \Rightarrow \ & \exists B, \mu(B) > \nu(B^\varepsilon) + \varepsilon\\ \Rightarrow \ & \exists B, \mu(B^c) + \varepsilon < \nu((B^\varepsilon)^c)\\ \Rightarrow \ & \exists B, \mu(((B^\varepsilon)^c)^\varepsilon) + \varepsilon \le \mu(B^c) + \varepsilon < \nu((B^\varepsilon)^c)\\ \Rightarrow \ & \exists C = (B^\varepsilon)^c, \nu(C) > \nu(C^\varepsilon) + \varepsilon\\ \Rightarrow \ & \varepsilon < \rho(\nu,\mu)\end{aligned}$$

Then it is clear that $\rho(\mu,\nu) \leq \rho(\nu,\mu)$.

Let us show now that if $\rho(\mu,\nu) = 0 \Rightarrow \mu = \nu$. To this end, for all $k \geq 1$ and all closed set F, we have

$$\mu(F) \leq \nu(F^{1/k}) + \frac{1}{k}$$

Then

$$\begin{aligned}
\nu(F) &= \nu(\bigcap_k F^{1/k}) \\
&= \lim_k \nu(F^{1/k}) \\
&\geq \mu(F)
\end{aligned}$$

and then $\mu(F) \leq \nu(F)$ and by symmetry $\nu(F) \leq \mu(F)$ which yields $\mu = \nu$.

Now it remains to show the triangular inequality. Let $(\lambda,\mu,\nu) \in \mathcal{M}_1(E)^3$. Let $\varepsilon, \delta > 0$ such that

$$\rho(\lambda,\mu) < \varepsilon, \quad \rho(\mu,\nu) < \delta$$

We have for all $B \in \mathcal{B}(E)$

$$\lambda(B) \leq \mu(B^\varepsilon) + \varepsilon \leq \nu((B^\varepsilon)^\delta) + \varepsilon + \delta$$

and therefore $\rho(\lambda,\nu) \leq \varepsilon + \delta$. Then we get the triangular inequality. □

Proposition 17 (Kantorovich-Rubinstein). *Let (E,d) be a metric space, define for a function f*

$$\|f\|_{BL} = \|f\|_\infty + \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x,y)}$$

The set $BL = \{f : E \mapsto \mathbb{R} | \|f\|_{BL} < \infty\}$ corresponds to the bounded Lipschitz functions.

The quantity

$$\beta(\mu,\nu) = \sup_{\|f\|_{BL} \leq 1} \left| \int_E f d\mu - \int_E f d\nu \right|$$

for $(\mu,\nu) \in \mathcal{M}_1(E)^2$ defines a distance on $\mathcal{M}_1(E)$ called Kantorovich-Rubinstein distance.

Démonstration. The only non-trivial fact is that $\beta(\mu,\nu) = 0$ implies that for all bounded Lipschitz function, we have $\mu(f) = \nu(f)$. Using the regularity of measure on (E,d), this implies that $\mu = \nu$ □

Before expressing the main result, let us recall the Theorem of Ascoli-Arzela. To this end let us recall the definition of the continuity modulus. For a function $f : (E,d) \to \mathbb{R}$

$$\omega_\delta(f) := \sup_{d(t,s) \leq \delta} |f(t) - f(s)|$$

Recall that f is uniformly continuous if and only if $\lim_{\delta \to 0} \omega_\delta(f) = 0$.

Theorem 18 (Arzela-Ascoli). *Let (K,d) be a metric space and $C(K) := \{f : K \mapsto \mathbb{R} \text{ continuous}\}$.*

We have $A \subset C(K)$ is relatively compact if and only if

1. For all $t \in K$

$$\sup_{f \in A} |f(t)| < \infty$$

2. Equicontinuity

$$\sup_{f \in A} \omega_\delta(f) \underset{\delta \to 0}{\to} 0$$

Remark 5. (1') : If $K \subset \mathbb{R}^d$ we just need that $\exists t_0 \in K$ such that $\sup_{f \in A} |f(t_0)| < \infty$

Démonstration. Let us show that (1) comes from (1') and (2). Define $t_k = t_0 + \frac{(t-t_0)k}{n}$ define $\Delta t = \frac{d(t,t_0)}{n}$, we have for $t \in K$

$$\begin{aligned} |f(t)| &\leq |f(t_0)| + \sum_{k=1}^{n} |f(t_k) - f(t_{k-1})| \\ &\leq |f(t_0)| + n\omega_{\Delta t}(f) \\ &\leq |f(t_0)| + n\omega_{\frac{diam(K)}{n}}(f) \end{aligned}$$

Then

$$\sup_{f \in A} |f(t)| \leq |f(t_0)| + n \sup_{f \in A} \omega_{\frac{diam(K)}{n}}(f).$$

This quantity is bounded for n sufficiently large.

Let us show that the relative compactness implies (1) and (2). If A is relatively compact then it is pre-compact. Therefore for all $\varepsilon > 0$ there exists n and $f_1, \ldots, f_n \in A$ such that

$$A \subset \bigcup_{i=1}^{n} B(f_i, \varepsilon).$$

In particular for all $f \in A$ tehre exist j such that

$$\|f\|_\infty \leq \|f - f_i\|_\infty + \|f_i\|_\infty$$

Then

$$\sup_{f \in A} \|f\|_\infty \leq \varepsilon + \sup_{i=1,\ldots,n} \|f_i\|_\infty$$

Then A is uniformly bounded which yields (1).

Now for all $f \in A$ and all $(t, s) \in K^2$

$$|f(t) - f(s)| \leq |f_i(t) - f_i(s)| + |f(s) - f_i(s)| + |f(t) - f_i(t)|$$

Then

$$\omega_\delta(f) \leq 2\varepsilon + \omega_\delta(f_i)$$

Then

$$\sup_{f \in A} \omega_\delta(f) \leq 2\varepsilon + \sup_{i=1,\ldots,n} \omega_\delta(f_i)$$

Now the Heine Theorem implies that

$$\lim_{\delta} \sup_{i=1,\ldots,n} \omega_\delta(f_i) = 0$$

And this implies that

$$\lim_{\delta}\sup \sup_{f\in A}\omega_\delta(f) \le 2\varepsilon$$

Now let us show that (1) + (2) implies that A is relatively compact. Let us use the sequential characterization. Consider (f_n) a sequence in A and let us show that there exists a convergent subsequence. The equicontinuity implies that for all $n \ge 1$ there exists $\delta_n \le \frac{1}{n}$ such that

$$\sup_{f\in A}\omega_n(f) \le \frac{1}{n}$$

Now since K is compact there exist a finite family $t_{n,j}$ such that

$$K \subset \bigcup_j B(t_{n,j}, 1/n).$$

Since A is uniformly bounded the set $(f_m(t_{n,j}); m \ge 1)$ is bounded Using Bolzano Weierstrass there exist ϕ_1 such that

$$(f_{\phi_1(m)}(t_{1,j}), m \ge 1)$$

is convergent. Then one can extract ϕ_2 such that

$$(f_{\phi_2\circ\phi_1(m)}(t_{l,j}), m \ge 1)$$

is convergent for $l = 1, 2$ and with a diagonal argument, there exist ϕ such that

$$(f_{\phi(m)}(t_{l,j}), m \ge 1)$$

is convergent for $l \ge 1$

Now let us define the dense subset

$$K' = \{t_{n,j}, n \ge 1, j \in J_n\} \subset K$$

For all $t' \in K'$ define

$$f(t') = \lim_n f_{\phi(n)}(t')$$

Note that f is uniformly continuous on K' since for all $|t - s| \le \delta$

$$|f(t) - f(s)| = \lim |f_{\phi(n)}(t) - f_{\phi(n)}(s)| \le \sup_{g\in A}\omega_\delta(g)$$

Then there exists a unique extension f wich is continuous on K, then uniformly continuous. Let $t \in K$. We have for $t' = t_{n,j}$ such that

$$|t - t'| \le \delta_n \le \frac{1}{n}$$

$$|f(t) - f_{\phi(n)}(t)| \le |f(t) - f(t')| + |f(t') - f_{\phi(n)}(t')| + |f_{\phi(n)}(t') - f_{\phi(n)}(t)| \le \omega_{\delta_n}(f) + \frac{2}{n}$$

and the convergence in norm holds. □

Now, we are in the position to express the main Theorem relating tight convergence and the measure β and ρ

Theorem 19 (The distance β and ρ metrizes $\mathcal{M}_1(E)$)**.** *The following assertion are equivalent*

1. $\mu_n \longrightarrow \mu$ *tightly*
2. $\beta(\mu_n, \mu) \longrightarrow 0$
3. $\rho(\mu_n, \mu) \longrightarrow 0$

Démonstration. Let us show that 1) $\Rightarrow$ 2). Since μ is tight, for all $\varepsilon > 0$, there exists K_ε such that $\mu(K) \geq 1 - \varepsilon$. Be aware that at this stage, we cannot use Prokhorov. Then let K^ε the open ε-neighborhood et by Portmanteau

$$\liminf_n \mu_n(K^\varepsilon) \geq \mu(K^\varepsilon) \geq 1 - \varepsilon,$$

then for n sufficiently large $\mu_n(K^\varepsilon) \geq 1 - 2\varepsilon$.

Now let $A = \{f, \|f\|_{BL} \leq 1\}$ and let $A_K = \{f_{|K}, f \in A\}$. By Ascoli A_K is relatively compact. Then there exist $k = k_\varepsilon$ and $f_1, \ldots, f_n$ such that

$$A_K \subset \bigcup_{i=1}^{k} B(f_i, \varepsilon)$$

which implies that for all f there exists f_j such that

$$\sup_{t \in K} |f(t) - f_j(t)| \leq \varepsilon$$

Then for all $f \in A$ there exist j such that

$$\begin{aligned} \sup_{t \in K^\varepsilon} |f(t) - f_j(t)| &\leq \sup_{s \in K} |f(s) - f_j(s)| + \sup_{t \in K^\varepsilon} \inf_{s \in K} |f(t) - f(s)| \\ &\quad + \sup_{t \in K^\varepsilon} \inf_{s \in K} |f_j(t) - f_j(s)| \\ &\leq 3\varepsilon \end{aligned}$$

We have used the Lipschitziannity in the last two terms. Then for all $f \in A$, there exists j such that

$$\begin{aligned} |\mu_n(f) - \mu(f)| &\leq |\mu_n(f - f_j)| + |\mu(f - f_j)| + |\mu_n(f_j) - \mu(f_j)| \\ &\leq |\mu_n((f - f_j)\mathbf{1}_{K^\varepsilon})| + \|f - f_j\|_\infty \mu_n((K^\varepsilon)^c) \\ &\quad + |\mu((f - f_j)\mathbf{1}_K)| + \|f - f_j\|_\infty \mu((K)^c) \\ &\quad + |\mu_n(f_j) - \mu(f_j)| \\ &\leq |\mu_n(f_j) - \mu(f_j)| + 10\varepsilon \end{aligned}$$

Then

$$\beta(\mu_n, \mu) \leq 10\varepsilon + \sup_{1 \leq j \leq k} |\mu_n(f_j) - \mu(f_j)|$$

This way, since $|\mu_n(f_j) - \mu(f_j)| \to 0$ for all $j = 1, \ldots, k$

$$\limsup_n \beta(\mu_n, \mu) \leq 10\varepsilon$$

and then

$$\lim_n \beta(\mu_n, \mu) = 0.$$

Let us now show that (3) $\Rightarrow$ (4). To do this, we will demonstrate that it is possible to compare the distances ρ and β. Indeed we shall show that

$$\rho(\mu, \nu) \leq 2\sqrt{\beta(\mu, \nu)}$$

Let $B \in \mathcal{B}(E)$ and $\mu, \nu \in \mathcal{M}_\infty(E)$, without loss of generality, we can assume that $\mu(B) \leqslant \nu(B)$. Let $\varepsilon > 0$, and consider the function g defined as

$$g(x) = \max\left(0, 1 - \frac{1}{\varepsilon} d(x, B)\right) \text{ for all } x \in E.$$

Recall that $g \in \mathrm{BL}(E)$ with $\|g\|_{\mathrm{BL}} \leqslant 1 + \frac{1}{\varepsilon}$ and $\mathbf{1}_B \leqslant g \leqslant \mathbf{1}_{B\varepsilon}$. The first term 1 corresponds to the $\|.\|_\infty$ and the term $\frac{1}{\varepsilon}$ corresponds to the Lipschtiz constant. Therefore, we have the following inequalities :

$$\begin{aligned}\mu(B) \leqslant \nu(B) \leqslant \int_E g d\nu &\leqslant \int_E g d\mu + \left(1 + \frac{1}{\varepsilon}\right) \beta(\mu, \nu) \\ &\leqslant \mu(B_\varepsilon) + \left(1 + \frac{1}{\varepsilon}\right) \beta(\mu, \nu) \\ &\leqslant \mu(B_\delta) + \delta,\end{aligned}$$

where $\delta = \max\left(\varepsilon, \left(1 + \frac{1}{\varepsilon}\right) \beta(\mu, \nu)\right)$. Therefore, by the definition of ρ, we have :

$$\rho(\mu, \nu) \leqslant \delta.$$

Two cases arise :
If $\beta(\mu, \nu) < 1$, and we choose ε such that $\beta(\mu, \nu) < \varepsilon^2 < 1$, then we deduce that :

$$\left(1 + \frac{1}{\varepsilon}\right) \beta(\mu, \nu) \leqslant \varepsilon + \varepsilon^2 < 2\varepsilon.$$

Hence, $\rho(\mu, \nu) \leqslant 2\varepsilon$. By taking the limit as ε tends to $\beta(\mu, \nu)$, we obtain the following inequality :

$$\rho(\mu, \nu) \leqslant 2\sqrt{\beta(\mu, \nu)}.$$

Now, if $\beta(\mu, \nu) \geqslant 1$, we deduce that $\rho(\mu, \nu) \leqslant 2\sqrt{\beta(\mu, \nu)}$ since $\rho(\mu, \nu) \leqslant 1$.
In all cases, we can conclude that (3) $\Rightarrow$ (4).

It remains to prove that 3) $\Rightarrow$ 1). To this end by the Portmanteau Theorem it is sufficient to prove that for all closed set F

$$\limsup \mu_n(F) \leq \mu(F)$$

By hypothesis for all $\varepsilon > 0$ there exist n_0 such that for all $n \geq n_0$ $\rho(\mu_n, \mu) \leq \varepsilon$. Then for all closed set F

$$\mu_n(F) \leq \mu(F^\varepsilon) + \varepsilon$$

which yields

$$\limsup \mu_n(F) \leq \mu(F^\varepsilon) + \varepsilon$$

Taking $\varepsilon \to 0$ and (F^ε) a decreasing sequence we get

$$\limsup \mu_n(F) \leq \mu(F)$$

□

Then we can prove the Prokhorov Theorem

Theorem 20 (Prokhorov)**.** *Let E be a Polish space and $\mathcal{F} \subset \mathcal{M}_1(E)$.*
The family $\mathcal{F}$ is tight if and only if $\mathcal{F}$ is relatively compact for the tight topology.

Démonstration. Let us start by the sense $\Leftarrow$ Note that up to replace $\mathcal{F}$ by $\bar{\mathcal{F}}$, we can suppose that $\mathcal{F}$ is compact. Note also that $\bar{\mathcal{F}}$ tight implies that $\mathcal{F}$ is tight. Let us start with a Lemma

Lemma 21. *Is $\mathcal{F}$ is compact and*

$$E = \bigcup_n \uparrow O_n,$$

is a non decreasing union of open sets. Then for all ε there exists N such that for all $\mu \in \mathcal{F}$

$$\mu(O_N) \geq 1 - \varepsilon$$

Démonstration. By contradiction there exist $\varepsilon > 0$ such that for all $N \in \mathbb{N}$ there exist $\mu_N \in \mathcal{F}$ with

$$\mu_N(O_N) < 1 - \varepsilon$$

This gives us a sequence (μ_N). Since $\mathcal{F}$ is compact and since $\mathcal{M}_1(E)$ is a metric space. By the sequential characterization there exist an extraction ϕ such that, in the tight topology

$$\mu_{\phi(N)} \to \mu.$$

By the Portmanteau Theorem for all n

$$\begin{aligned} \mu(O_n) &\leq \liminf_N \mu_{\phi(N)}(O_n) \\ &\leq \liminf_N \mu_{\phi(N)}(O_{\phi(N)}) \quad O_n \subset O_{\phi(N)} \quad N >> 1 \\ &\leq 1 - \varepsilon \end{aligned}$$

Since (O_n) is increasing, we get that

$$\mu(\bigcup_n O_n) \leq 1 - \varepsilon$$

which contradicts that $E = \bigcup_n \uparrow O_n$ □

Let us come back to the proof that $\mathcal{F}$ compact $\Rightarrow \mathcal{F}$ tight. Let $\varepsilon > 0$ and $\{x_n, n \geq 1\}$ a dense sequence. For $k \in \mathbb{N}^*$ being fixed

$$E = \bigcup_n B(x_n, \frac{1}{k})$$

From the Lemma there exist $N_k(\varepsilon)$ such that fro all $\mu \in \mathcal{F}$

$$\mu(\bigcup_{n=1}^{N_k(\varepsilon)} B(x_n, 1/k) > 1 - \frac{\varepsilon}{2^k}$$

The set

$$K \bigcap_{k\geq 1} \bigcup_{n=1}^{N_k(\varepsilon)} B_F(x_n, 1/k)$$

satisfies for all $\mu \in \mathcal{F}$

$$\begin{aligned} \mu(K^c) &= \mu(\bigcup_k \left(\bigcup_{n=1}^{N_k(\varepsilon)} B_F(x_n, 1/k)\right)^c) \\ &\leq \mu(\bigcup_k \left(\bigcup_{n=1}^{N_k(\varepsilon)} B(x_n, 1/k)\right)^c) \\ &\leq \sum_k \frac{\varepsilon}{2^k} = \varepsilon \end{aligned}$$

Furthermore note that K is pre-compact (can be covered by a finite number of ball of radius $1/k$ for all k) and it is clear that $\bar{K} = K$ then K is complete since E is complete. It follows that K is compact and then $\mathcal{F}$ is tight.

For the sense $\Rightarrow$ let us atrt with the case where E is compact. We shall need this proposition.

Proposition 22. *If (K, d) is a metric compact set. Then from any sequence (μ_n) in $\mathcal{M}_1(K)$ one can extract a convergent subsequence.*

Démonstration. Somehow it is an abstract version of the real case $K = [a, b]$. Here we consider K equipped with d, $C(K)$ equipped with $\|.\|_\infty$ and $\{x_n, n \geq 1\}$ a dense sequence in K. We consider as well $\mathcal{A} = \mathbb{Q}[f_k, k \geq 1]$ where $f_k(.) = d(., x_k)$ as the algebra genrated by (f_k).

Star by showing that $\mathcal{A}$ is dense. To thsi end we shall show the following abstract version of Stone-Weierstrass. $\mathcal{A}$ is a sub-algebra that separates the points i.e $\forall (x, y) \in K \times K x \neq y \Rightarrow \exists f \in \mathcal{A} f(x) \neq f(y) \Leftrightarrow \mathcal{A}$ is dense in $C(K)$. Let $(x, y) \in K \times K$, we have

$$\begin{aligned} & \forall f \in \mathcal{A} f(x) = f(y) \\ \Rightarrow\ & \forall k \geq 1, f_k(x) = f_k(y) \\ \Leftrightarrow\ & \forall k \geq 1 d(x, x_k) = d(y, x_k) \\ \Rightarrow\ & \text{for} x_k \to y 0 = d(y, y) = \lim d(y, x_k) = \lim d(x, x_k) = d(x, y) \\ \Rightarrow\ & x = y \end{aligned}$$

Note that $\mathcal{A}$ is countable. Let us proceed by making a diagonal extraction. For all $f \in \mathcal{A}$ we have

$$|\mu_n(f)| \leq \|f\|_\infty < \infty$$

For all f one can extract, then by countability we can make a diagonal extarction ϕ such that

$$\mu(\phi(n))(f) \to M(f)$$

Clearly $M : \mathcal{A} \mapsto \mathbb{R}$ is linear. For all $f \in \mathcal{A}$

$$|M(f)| \leq \|f\|_\infty$$

Then M is a linear form on $\mathcal{A}$ which is uniformly continuous then there exists a unique linear continuous extension on $C(K)$.

let us show that for all $g \in C(K)$

$$M(g) = \lim \mu_{\phi(n)}(g)$$

Indeed let $(g_k) \in \mathcal{A}$ such that $g_k \to g$ in $\|.\|_\infty$. We have

$$|\mu_{\phi(n)}(g_k) - \mu_{\phi(n)}(g)| \leq \|g_k - g\| \to 0$$

Let u be any adherence value of $(\mu_{\phi(n)}(g))$, we then have

$$|M(g_k) - u| \leq \|g_k - g\| \to 0$$

and then there exist a unique adherence value

$$u = \lim_k M(g_k) = M(g)$$

Note that if we knew that

$$M(g) = \mu(g)$$

for $\mu \in \mathcal{M}_1(E)$, we would have finished. To conclude we can check that

$$M(1) = 1, \quad f \geq 0 \Rightarrow M(f) \geq 0$$

□

Come back to the general case. Since (μ_n) is tight, there exist an increasing sequence of compact sets $(K_j)_{j \geqslant 1}$ (i.e. $K_j \subset K_{j+1}$ for all $j \geqslant 1$) such that

$$\forall n \in \mathbb{N}, \quad \mu_n(K_j) \geqslant 1 - \frac{1}{j} \text{ pour tout } n \in \mathbb{N}$$

Indeed, for that it is sufficient to choose a family of compact sets $(L_j)_{j \in \mathbb{N}}$ such that $\mu_n(L_j) \geqslant 1 - \frac{1}{j}$ by tightness and then $K_j = \cup_{i \leqslant j} L_i$; the family of compact set $(K_j)_{j \in \mathbb{N}}$ satisfied what we wanted.

Next, for all $j \geqslant 1$ and $n \in \mathbb{N}$, denote by $\nu_n^j = \mu_{n|K_j}$ the restriction of μ_n to K_j.

$$\nu_n^j = \frac{\mathbf{1}_k \mu_n}{\mu_n(K_j}$$

By the last proposition for all j (ν_n^j) converges up to an extraction. By diagonally extracting

$$\nu_{\phi(n)}^j \underset{n\to+\infty}{\to} \nu^j \quad \text{for all} \quad j \geqslant 1$$

Up to extracting again we have

$$\mu_{\phi(n)}(K_j) \to \mu_j \geq 1 - \frac{1}{2j}$$

$$\begin{aligned}
\mu_{j+1}\nu^{j+1}(B^{2\varepsilon} \cap K_{j+1}) &\geq \mu_{j+1}\nu^{j+1}(\bar{B}^{\varepsilon} \cap K_{j+1}) \\
&\geq \limsup_n \mu_{\phi(n)}(K_{j+1})\nu_{\phi(n)}^{j+1}(\bar{B}^{\varepsilon} \cap K_{j+1}) \\
&= \limsup_n \mu_{\phi(n)}(\bar{B}^{\varepsilon} \cap K_{j+1}) \\
&\geq \limsup_n \mu_{\phi(n)}(\bar{B}^{\varepsilon} \cap K_j) \\
&= \mu_j \limsup_n \nu^j(\bar{B}^{\varepsilon} \cap K_j) \\
&\geq \mu_j \liminf_n \nu^j(B^{\varepsilon} \cap K_j) \\
&\geq \mu_j\nu^j(B^{\varepsilon} \cap K_j)
\end{aligned}$$

Taking $\varepsilon \to 0$ yields

$$\mu_{j+1}\nu^{j+1}(\bar{B} \cap K_{j+1}) \geq \mu_j\nu^j(\bar{B} \cap K_j)$$

This yields that fro all closed set F

$$\mu_{j+1}\nu^{j+1}(F \cap K_{j+1}) \geq \mu_j\nu^j(F \cap K_j)$$

By the regularity of the measures for all the Borel sets B

$$\mu_{j+1}\nu^{j+1}(B \cap K_{j+1}) \geq \mu_j\nu^j(B \cap K_j)$$

Therefore by using this increasing fact one can define μ such that

$$\mu(B) = \lim_j \mu_j\nu^j(B \cap K_j) = \lim_j \mu_j\nu^j(B \cap K_j),$$

since $1 - \frac{1}{2j} \leq \mu_j \leq 1$

The application $\mu \in \mathcal{M}_1(E)$. Indeed

$$\mu(E) = \lim_j \mu_j\nu^j(K_j) = 1$$

Let (B_i) disjoint

$$\begin{aligned}
\mu(\bigsqcup_i B_i) &= \lim_j \mu_j\nu^j((\bigsqcup_i B_i) \cap K_j) \\
&= \lim_j \sum_i \mu_j(\nu^j((B_i) \cap K_j) \\
&= \sum_i \lim_j \mu_j(\nu^j((B_i) \cap K_j) \quad \text{monotone convergence}
\end{aligned}$$

In ordee to conclude, so we must demonstrate that it is possible to extract a subsequence of $(\mu_n)_{n\in\mathbb{N}}$ such that this subsequence converges tightly towardsc μ. In other words, for all $f \in C_b(E)$ we shall show

$$\lim_{n\to+\infty}\int_E f d\mu_n = \int_E f d\mu.$$

Without loss of generality we can suppose that $0 \leqslant f \leqslant 1$. For all $n \in \mathbb{N}$ and for all $j \geqslant 1$, we get

$$\begin{aligned}\left|\int_E f d\mu_n - \int_E f d\mu\right| &\leqslant \left|\int_E f d\mu_n - \int_{K_j} f d\nu^j\right| + \left|\int_{K_j} f d\nu^j - \int_E f d\mu\right| \\ &\leqslant \left|\int_{K_j} f d\mu_n - \int_{K_j} f d\nu^j\right| + \mu_n\left(K_j^c\right) + \left|\int_{K_j} f d\nu^j - \int_E f d\mu\right| \\ &\leqslant \left|\int_{K_j} f d\nu_n^j - \int_{K_j} f d\nu^j\right| + \frac{1}{j} + \left|\int_{K_j} f d\nu^j - \int_E f d\mu\right|.\end{aligned}$$

Since $\mu_{n|K_j} = \nu_n^j$ and $\mu_n\left(K_j^c\right) \leqslant \frac{1}{j}$ by tightness. Next, since $(\nu_n^j)_{n\in\mathbb{N}}$ converges tightly to ν^j, we deduce that

$$\begin{aligned}\limsup_{n\to+\infty}\left|\int_E f d\mu_n - \int_E f d\mu\right| &\leqslant \frac{1}{j} + \left|\int_{K_j} f d\nu^j - \int_E f d\mu\right| \\ &\leqslant \frac{1}{j} + \left|\nu^j\left(K_j\right) - \mu(E)\right|.\end{aligned}$$

Now j to infinity allows to conclude. Next by definition og μ (with $B = E$) we get

$$\int_{K_j} f d\nu^j \underset{j\to+\infty}{\uparrow} \int_E f d\mu.$$

Then $\lim_{n\to+\infty}\int_E f d\mu_n = \int_E f d\mu$ which finishes the proof. □

Let us express a corollary

Corollary 23. *The metric space $(\mathcal{M}_1(E), \beta)$ is complete.*

This is also true for β replace by ρ.

Démonstration. .Let $(\mu_n)_{n\in\mathbb{N}}$ be a Cauchy sequence of (E, β) then it is pre-compact. By Prokhorov there exist a convergent subsequence. Since a Cauchy sequence has at most one limit point, we can deduce that it converges, and therefore, the space is complete (with respect to the metric induced by ρ or β) □

Chapitre 3

Functional Limit theorems

3.1 Goal

Proposition 24. *If E is a Polish space then*

$$\mathcal{C} = C([0,T], E)$$

is a Polish space.

Démonstration. The metrisability and the completness are usual. □

Then it is reasonable to consider constructing probability measure in $\mathcal{M}_1(\mathcal{C})$

Theorem 25 (Existence and uniqueness of Brownian motion)**.** *Let $E = \mathbb{R}$ (or $E = \mathbb{R}^d$). There exist a unique measure $\mathbb{W} \in \mathcal{M}_1(\mathcal{C})$) such that if X is r.v valued in E with $\mathcal{L}(X) = \mathbb{W}$, we have*

- $X_0 = 0$
- $t \mapsto X_t$ *is almost surely continuous, i.e* $X \in \mathcal{C}$
- $\forall 0 = t_0 < t_1 < \ldots, t_k, k \in \mathbb{N}$ $(X_{t_i} - X_{t_{i-1}}, i = 1, \ldots, k)$ *are independent increments such that fro all* $i = 1, \ldots, k$

$$X_{t_i} - X_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1}), (\mathcal{N}(0, (t_i - t_{i-1})I_d, E = \mathbb{R}^d)$$

We call $\mathbb{W}$ the Wiener measure. All r.v X such that $\mathcal{L}(X) = \mathbb{W}$ is called a standard Brownian motion.

The Brownian motion is a essential ingredient in modern probability that allows to construct other interesting processes. This is also a universal process that appears naturally as the limit process of the central limit Theorem.

Come back to the warmup. Let (ξ_i) be a sequence of i.i.d r.v such that

$$\mathbb{E}\xi_1 = 0, \quad \mathbb{E}\xi_1^2 = 1$$

Denote

$$S_n = \sum_{i=1}^{n} \xi_i$$

and

$$S_t^{[n]} = \frac{S_{\lfloor tn \rfloor}}{\sqrt{n}} + (nt - \lfloor nt \rfloor) \frac{S_{\lfloor tn \rfloor + 1} - S_{\lfloor tn \rfloor}}{\sqrt{n}}$$

We have the following theorem

Theorem 26 (Donsker Invariant Principle)**.** *We have the following convergence in distribution*

$$(S_t^{[n]}; 0 \leq t \leq T) \underset{n \to \infty}{\overset{\mathcal{L}}{\to}} (X_t, 0 \leq t \leq T),$$

where X is a Brownian motion.

In an equivalent way, in the tight topology

$$\mathcal{L}(S^{[n]}) \underset{n \to \infty}{\to} \mathbb{W}.$$

This convergence can be also established in the space $\mathcal{D}$ where $\mathcal{D}$ holds for discontinuous function ; more precisely càdlàg (right continuous and left limit). This won't be addressed in these notes.

We shall show the result by assuming $\mathbb{E}\xi_1^{2+\varepsilon} < \infty$.

Before establishing this theorem we shall present the usual procedure to show convergence in distribution. Essentially there are two steps : tightness and identification of the limit process. For the identification, usually we study the marginals.

3.2 Marginals and Product σ-algebra

On $\mathcal{C}$ there is a natural application called projections. For all $0 \leq t \leq T$ define

$$\begin{array}{rccc} \pi_t : & \mathcal{C} & \to & E \\ & f & \mapsto & f(t) \end{array}$$

Definition 7. Let X be a continuous process. Let $0 =\leq t_1 < \ldots, t_k \leq T, k \in \mathbb{N}$, then

$$\mathcal{L}(X_{t_1}, \ldots, X_{t_k})$$

is called the k-dimensional marginal or the finite dimensional marginals.

In other words if $\mathcal{L}(X) = \mu \in \mathcal{M}_1(\mathcal{C})$ then

$$\mathcal{L}(X_{t_1}, \ldots, X_{t_k}) = (\pi_{t_1}, \ldots, \pi_{t_k}) \# \mu,$$

where the last symbol means mesure image.

Recall that $\mathcal{C}$ is endowed with its Borel σ-algebra generated bu the open sets for the uniform topology. Another σ-algebra is also natural

Definition 8 (Cylindric σ-algebra)**.** The cylindric σ-algebra $\mathcal{C}yl(\mathcal{C})$ is defined by

$$\mathcal{C}yl(\mathcal{C}) = \sigma(\pi_t, 0 \leq t \leq T),$$

this is the σ-algebra generated by the projection applications.

We have the following theorem

Theorem 27. *The Borel and the cylindric σ algebras are the same*

$$\mathcal{C}yl(\mathcal{C}) = \mathcal{B}(\mathcal{C})$$

Démonstration. The inclusion $\subset$ is obvious since the application π_t are continuous, then for all $0 \leq t \leq T$ $\pi_t^{-1}(O) \in \mathcal{B}(\mathcal{C})$ for all open sets and therefore

$$\mathcal{C}yl(\mathcal{C}) = \sigma(\pi_t^{-1}(O), O \text{ open}) \subset \mathcal{B}(\mathcal{C})$$

For the inclusion $\supset$, let us show that all open ball $B(f;r) \in \mathcal{B}(\mathcal{C})$ satisfies $B(f;r) \in \mathcal{C}yl(\mathcal{C})$.

Let $f \in \mathcal{C}, r > 0$

$$\begin{aligned}
\overline{B(f;r)} &= \{g \in \mathcal{C} | \quad |g - f|_\infty \leqslant r\} \\
&= \{g \in \mathcal{C} \mid \forall t \in [0,T], d(f(t); g(t)) \leqslant r\} \\
&= \{g \in \mathcal{C} \mid \forall t \in \mathbb{Q} \cap [0,T], d(f(t), g(t)) \leq r\} \quad \text{continuity} \\
&= \bigcap_{t \in \mathbb{Q}\cap[0,T]} \{g \in \mathcal{C} \mid g(t) \in \overline{B(f(t);r)}\} \\
& \bigcap_{t \in \mathbb{Q}\cap[0,T]} \{g \in \mathcal{C} \mid \pi_t(g) \in \overline{B(f(t);r)}\} \\
&= \bigcap_{t \in \mathbb{Q}\cap[0,T]} \pi_t^{-1}(\underbrace{\overline{B(f(t);r)}})
\end{aligned}$$

Since $\overline{B(f(t);r)} \in \mathcal{B}(E)$ then $\pi_t^{-1}(\overline{B(f(t);r)}) \in \mathcal{C}yl(\mathcal{C})$ and then $\overline{B(f;r)} \in \mathcal{C}yl(\mathcal{C})$. As a consequence

$$B(f;r) = \bigcup_{\substack{r' < r \\ r' \in \mathbb{R}}} \overline{B(f;r)} \in \mathcal{C}yl(\mathcal{C})$$

Finally all open set in a Polish space E is a countable union of open ball

$$U = \bigcup_{k=1}^{\infty} B\left(x_k; \varepsilon_k\right),$$

where (x_k) is a dense sequence and $B\left(x_k; \varepsilon_k\right) \subset U$ for all k.

Then $\mathcal{C}yl(\mathcal{C})$ contains the open set. As a consequence

$$\mathcal{C}yl(\mathcal{C}) \supset \mathcal{B}(E)$$

□

Corollary 28. — *All law $\mu \in \mathcal{M}_1(\mathcal{C})$ is uniquely determined by its marginals.*

— $\mu_n \underset{n\to+\infty}{\to} \mu \in \mathcal{M}_1(\mathcal{C})$ *tightly if and only if*

$$\begin{cases} a) \quad (\mu_n; n \geq 1) \text{ is tight} \\ b) \quad CV \text{ of finite dimensional marginals towards the one of } \mu \qquad \text{(for all subsequences)} \\ i.e\, (\pi_{t_1}, \ldots, \pi_{t_k})\, \#\mu_{\varphi(n)} \overset{n\to+\infty}{\longrightarrow} (\pi_{t_1}, \ldots, \pi_{t_k})\, \#\mu \quad \forall 0 \leqslant t_1 < \cdots < t_k \end{cases}$$

- In more probabilistic words. Let X, Y two processes on $\mathcal{C}$ then $X \stackrel{\mathcal{L}}{=} Y$ if and only if $\forall k \geqslant 1\ \forall 0 < t_1 < t_2 < \ldots < t_k$

$$(X_{t_1}, X_{t_2}, \ldots, X_{t_k}) \stackrel{\mathcal{L}}{=} (Y_{t_1}, \ldots, Y_{t_k})$$

- Let (X_n) be a sequence of processes in $\mathcal{C}$ then

$$X_n \xrightarrow{\mathcal{L}} X$$

if and only if

$$\begin{cases} (X_n; n \geqslant 1) \text{ is a tight family} \\ CV \text{ of marginales for all subsequences} \end{cases}$$

The convergence of marginales means that for all extraction ϕ, if $(X_{\phi(n)})$ converges in distribution then $\forall k \geqslant 1, \forall 0 \leqslant t_1 < t_2 < \cdots < t_k$,

$$\left(X_{\phi(n)}(t_1), \ldots, X_{\phi(n)}(t_k)\right) \underset{n\to+\infty}{\xrightarrow{\mathcal{L}}} (X_{t_1}, \ldots, X_{t_n})$$

Démonstration. For the first point let us show that if μ and μ' have the same marginals then $\mu = \mu'$. To this end let us consider $B_1, \ldots, B_k$ k Borel sets and let $0 \le t_1 < \ldots < t_k$ and define

$$A = \{f \in \mathcal{C} | f(t_i) \in B_i, i = 1, \ldots, k\} = \bigcap_i \pi_{t_i}^{-1}(B_i)$$

Then we have

$$\begin{aligned} \mu(A) &= (\pi_{t_1}, \ldots, \pi_{t_k})\#\mu(B_1 \times \ldots \times B_k) \\ &= (\pi_{t_1}, \ldots, \pi_{t_k})\#\mu'(B_1 \times \ldots \times B_k) \\ &= \mu'(A) \end{aligned}$$

Then the set

$$\mathcal{T} = \{A \in \mathcal{B}(\mathcal{C}) | \mu(A) = \mu'(A)\}$$

contains the sets that generates $\mathcal{C}yl(\mathcal{C})$ which is stable by intersection, then it is a class monotone. Then $\mathcal{T} \supset \mathcal{C}yl(\mathcal{C})$ and the first point follows.

For the second point. Since $\mathcal{M}_1(\mathcal{C})$ is metrisable we get that

$$\mu_n \longrightarrow^{tight} \mu$$

if and only if $\{\mu_n, n \ge 1\} \cup \{\mu\}$ is compact in $\mathcal{M}_1(\mathcal{C}\}$ and the unique limit point is μ. By Prokhorov this holds if and only if $\{\mu_n, n \ge 1\}$ is tight and for all extraction ϕ

$$\mu_{\phi(n)} \to \nu$$

implies $\mu = \nu$ (unique limit point). By the first point this holds if and only if $\{\mu_n, n \ge 1\}$ is tight and for all extraction ϕ

$$(\pi_{t_1}, \ldots, \pi_{t_k})\#\mu_{\varphi(n)} \stackrel{n\to+\infty}{\longrightarrow} (\pi_{t_1}, \ldots, \pi_{t_k})\#\mu$$

□

Remark 6. Note that

$$X_n \xrightarrow{\mathcal{L}} X$$

implies that for all $0 \leq t_1 < \ldots < t_k$ we have

$$(X_n(t_1), \ldots, X_n(t_k)) \xrightarrow[n\to+\infty]{\mathcal{L}} (X_{t_1}, \ldots, X_{t_n})$$

3.3 Kolmogorov criterion

Theorem 29 (Kolmogorov Criterion)**.** *Let $E = \mathbb{R}^d$. For $\alpha > 0$ and for $f : [0,T] \to E$, we define*

$$N_\alpha(f) = \sup_{s\neq t} \frac{|f(t) - f(s)|}{|t-s|^\alpha},$$

the α Hölder norm.

Let (X_n) a sequence of r.v in $\mathcal{C}$, satisfying

- *$(X_n(0), n \geq 1)$ is tight*
- *$\exists p, c, \beta > 0, \mathbb{E}[|X_n(t) - X_n(s)|^p] \leq c|t-s|^{1+\beta}$*

Then $(X_n, n \geq 1)$ is tight and more precisely, for all $0 < \alpha < \beta/p$ $(N_\alpha(X_n), n \geq 1)$ is tight.

Démonstration. Let us prove the result for $d = 1$ and let us suppose that $T = 1$. let us start by showing that

$$\begin{cases} (X_n(0), n \geq 1) & \text{tight} \\ (N_\alpha(X_n), n \geq 1) & \text{tight} \end{cases} \Rightarrow (X_n, n \geq 1) \quad \text{tight}$$

Recall that we have to show that for all $\varepsilon > 0$, there exists a compact $K \subset \mathcal{C}$ such that

$$\sup_n \mathbb{P}(X_n \in K^c) \leq \varepsilon$$

By Arzela-Ascoli a compacts et is characterized by

- The boundeness in 0, i.e $\exists M > 0, \forall f \in K, |f(0)| \leq M$
- The equi-continuity

$$\sup_{f\in K} \omega_\delta(f) \underset{\delta\to 0}{\to} 0$$

Note that for all $\alpha > 0$ if there exists $M' > 0$ and $K \subset \{f \in \mathcal{C} | N_\alpha(f) \leq M'\}$ then K is equicontinuous. Therefore let us consider

$$K = \{f \in \mathcal{C} |\ |f(0)| \leq M, N_\alpha(f) \leq M'\}.$$

Let us show that

$$\begin{cases} (X_n(0), n \geq 1) & \text{tight} \\ (N_\alpha(X_n), n \geq 1) & \text{tight} \end{cases} \Rightarrow (X_n, n \geq 1) \quad \text{tight.}$$

We have to show that for all $\varepsilon > 0, \exists K$ compact

$$\sup_{n\geq 1} \mathbb{P}(X_n \in K^c) \leq \varepsilon$$

which is implied by the fact that for all $\varepsilon > 0, \exists M, M' > 0$

$$\sup_{n\geq 1} \mathbb{P}(|X_n(0)| \geq M, N_\alpha(X_n) \geq M') \leq \varepsilon,$$

which is implied by the tightness of $(X_n(0), n \geq 1)$ and $(N_\alpha(X_n), n \geq 1)$.

Let us start by showing that for all $0 < \alpha < \beta/p$ $(N_\alpha(X_n), n \geq 1)$ is tight. This needs a preliminary analysis. Let define for all $m \in \mathbb{N}^*$

$$\begin{aligned} D_m &:= \text{the dyadic numbers of order } m \\ &= \{\frac{k}{2^m}, 0 \leq k < 2^m\} \end{aligned}$$

Note that $card(D_m) = 2^m$. Define

$$K_m(X) := \sup_{t\in D_m} |X_{t+2^{-m}} - X_t|.$$

Let us show that

$$\mathbb{E}[|X_t - X_s|^p] \leq c|t-s|^{1+\beta} \Rightarrow \mathbb{E}K_m(X)^p \leq c2^{-m\beta}$$

Indeed

$$\begin{aligned} \mathbb{E}K_m(X)^p &= \mathbb{E}\sup_{t\in D_m} |X_{t+2^{-m}} - X_t|^p \\ &\leq \sum_{t\in D_m} \mathbb{E}|X_{t+2^{-m}} - X_t|^p \\ &\leq cCard(D_m)(2^{-m})^{1+\beta} \\ &= c2^{-m\beta} \end{aligned}$$

Now let us show that for all $X \in \mathcal{C}$

$$N_\alpha(X) \leq 2\sum_{m\geq 0} K_m(X)2^{\alpha m}.$$

By continuity

$$N_\alpha(X) = \sup_{s\neq t, (s,t)\in\left(\bigcup_m D_m\right)^2} \frac{|X_t - X_s|}{|t-s|^\alpha}$$

If $s < t$ with $s, t \in \bigcup_m D_m$, there exists $m > 0$ such that

$$2^{-(m+1)} \leq t - s \leq 2^{-m}$$

This way one can write

$$[s,t[= \bigsqcup_{i=0}^{N-1} [\tau_i, \tau_{i+1}[$$

with $s = \tau_0 < \tau_1 < \ldots < \tau_N = t$ with $\tau_i \in D_k, k \geq m$. All the pairs (τ_i, τ_{i+1}) belongs to the same D_{k_0} and $\tau_{i+1} - \tau_i = 2^{-k_0}$. This decomposition can be done in such way that $|\tau_{i+1} - \tau_i| = |\tau_{j+1} - \tau_j|$ for at most 2 indices.

Then we have

$$\begin{aligned}|X_t - X_s| &\leq \sum_{i=0}^{N-1} |X_{\tau_{i+1}} - X_{\tau_i}| \\ &\leq 2 \sum_{k \geq m+1} K_k(X)\end{aligned}$$

Then

$$\begin{aligned}\frac{|X_t - X_s|}{|t-s|^\alpha} &\leq 2 \sum_{k \geq m+1} K_k(X) 2^{(m+1)\alpha} \\ &\leq 2 \sum_{k \geq m+1} K_k(X) 2^{k\alpha} \\ &\leq 2 \sum_{k \geq 0} K_k(X) 2^{k\alpha}\end{aligned}$$

The last term is independent of (t, s) therefore

$$N_\alpha(X) \leq 2 \sum_{k \geq 0} K_k(X) 2^{k\alpha}$$

Now we should show that

$$\left(2 \sum_{k \geq 0} K_k(X_n) 2^{k\alpha}, n \geq 1 \right)$$

is tight. Using Markov inequality we have

$$\mathbb{P}\left(2 \sum_{k \geq 0} K_k(X_n) 2^{k\alpha} \geq M \right) \leq \frac{\mathbb{E}\left[\left| 2 \sum_{k \geq 0} K_k(X_n) 2^{k\alpha} \right|^p \right]}{M^p}$$

Let us show that $\mathbb{E}\left[\left| \sum_{k \geq 0} K_k(X_n) 2^{k\alpha} \right|^p \right]$ is uniformly bounded in n.

To this end we have

$$\begin{aligned}\left(\mathbb{E}\left[\left| \sum_{k \geq 0} K_k(X_n) 2^{k\alpha} \right|^p \right] \right)^{1/p} &\leq \sum_{k \geq 0} 2^{\alpha k} \left(\mathbb{E}(K_m(X_n))^p \right)^{1/p} \\ &\leq \sum_{k \geq 0} 2^{\alpha k} \left(\mathbb{E} c 2^{-k\beta} \right)^{1/p} \\ &\leq c^{1/p} \sum_{k \geq 0} 2^{k\left(\alpha - \frac{\beta}{p}\right)}\end{aligned}$$

which is finite and idependent of n for all $0 < \alpha < \frac{\beta}{p}$.

This way there exists C such that

$$\sup_n \mathbb{P}\left(2 \sum_{k \geq 0} K_k(X_n) 2^{k\alpha} \geq M \right) \leq \frac{C}{M^p}$$

and the tightness follows. $\square$

3.4 Proof of the announced results

3.4.1 Existence and uniqueness of the Brownian motion

Let us quickly present construction of Brownian motion. To do so, consider the Hilbert space $L^2([0,1], d\lambda)$ where λ denotes the Lebesgue measure, and let $(e_i)_{i\in\mathbb{N}}$ be an orthonormal basis of H. For example, one can choose the basis defined using Haar wavelets. These are obtained by translation and dilation of the following function

$$H(t) = \begin{cases} 1 \text{ si } 0 \leqslant t < \frac{1}{2} \\ -1 \text{ si } \frac{1}{2} \leqslant t \leqslant 1 \\ 0 \text{ sinon} \end{cases}$$

Let us then note that for every integer $n \in \mathbb{N}$, there exists a unique pair $(j,k) \in \mathbb{N}^2$ such that

$$n = 2^j + k \quad \text{with} \quad 0 \leqslant k < 2^j.$$

For such integers, we then define, for all $t \in [0,1]$,

$$e_n(t) = 2^{j/2} H\left(2^j t - k\right) \quad \text{for} \quad n = 2^j + k \quad \text{with} \quad j \geqslant 0 \quad \text{and} \quad 0 \leqslant k < 2^j,$$

and impose $e_0(t) = 1$. The family thus defined $(e_n)_{n\geqslant 0}$ is an orthonormal basis of $L^2[0,1]$.

From such a family, it is possible to construct a Schauder basis $(v_n)_n$ of $C[0,1]$ (i.e., a dense family in a Banach space). To do this, we define

$$v_n(t) = \int_0^t e_n(u)\, du \quad \text{for all} \quad t \in [0,1].$$

Moreover, if $d : [0,1] \to \mathbb{R}$ is defined by

$$d(t) = \begin{cases} 2t & \text{if } 0 \leqslant t < \frac{1}{2} \\ 2(1-t) & \text{if } \frac{1}{2} \leqslant t \leqslant 1 \\ 0 & \text{otherwise} \end{cases}$$

and $d_n(t) = d\left(2^j t - k\right)$ for $n = 2^j + k$ with $0 \leqslant k < 2^j$. It is then possible to show that

$$v_n(t) = \lambda_n d_n(t)$$

with $\lambda_0 = 1$ and, for $n \geqslant 1$, $\lambda_n = \frac{1}{2} 2^{-j/2}$ where $n = 2^j + k$ with $0 \leqslant k < 2^j$ and $j \geqslant 0$.

From this family, it is possible to construct Brownian motion. Indeed, let $(g_i)_{i\in\mathbb{N}}$ be a sequence of standard Gaussian random variables, independent and identically distributed (i.e., $g_1 \sim \mathcal{N}(0,1)$) on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and let us define

$$X_t^n = \sum_{i=1}^n g_i v_i(t) \quad \text{with} \quad t \in [0,1].$$

This sequence of random functions satisfies the following convergence :

Theorem 30. *Almost surely, $(X^n)_{n\in\mathbb{N}}$ converges uniformly on $[0,1]$ to a limit X which is a Brownian motion.*

DÃ©monstration. First, (assuming the limit exists), let us see what happens at the level of the covariance functions and quickly verify that

$$\lim_{n\to+\infty} \mathbb{E}\left[X_s^n X_t^n\right] = \mathbb{E}\left[X_s X_t\right] = \min(s,t).$$

To do this, we will use the fact that the variables $(g_i)_{i\in\mathbb{N}}$ are independent and identically distributed standard Gaussian random variables.

$$\begin{aligned}
\mathbb{E}\left[X_s^n X_t^n\right] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} g_i v_i(s)\right)\left(\sum_{j=1}^{n} g_j v_j(t)\right)\right] \\
&= \sum_{i=1}^{n} v_i(s) v_i(t) \\
&= \sum_{i=1}^{n} \left(\int_0^s e_i(u)\,du\right)\left(\int_0^t e_i(u)\,du\right) \\
&= \sum_{i=1}^{n} \left\langle e_i, 1_{[0,s]}\right\rangle_{L^2(d\mathbb{P})} \left\langle e_i, 1_{[0,t]}\right\rangle_{L^2(d\mathbb{P})}.
\end{aligned}$$

To conclude, it suffices to use Parseval's theorem (see [204]), which assures us that

$$\lim_{n\to+\infty} \sum_{i=1}^{n} \left\langle e_i, 1_{[0,s]}\right\rangle_{L^2} \left\langle e_i, 1_{[0,t]}\right\rangle_{L^2} = \left\langle 1_{[0,s]}, 1_{[0,t]}\right\rangle_{L^2} = \min(s,t).$$

This construction will be employed again when we prove the Schilder theorem in Chapter 4.

Let us now demonstrate the uniform convergence. For this, we will need the following lemma.

Lemma 31. *In the framework described in Section 2.6, there exists a random variable C such that, almost surely, for all $n \geqslant 2$, we have*

$$|g_n| \leqslant C\sqrt{\log n} \quad \textit{et} \quad \mathbb{P}(C<\infty) = 1.$$

DÃ©monstration. For $x \geqslant 1$, we have

$$\begin{aligned}
\mathbb{P}\left(|g_n| \geqslant x\right) &= \frac{2}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2} du \\
&\leqslant \sqrt{\frac{2}{\pi}} \int_x^\infty u e^{-u^2/2} du = e^{-x^2/2}\sqrt{\frac{2}{\pi}}.
\end{aligned}$$

Thus, for all $\alpha > 1$, we obtain

$$\mathbb{P}\left(|g_n| \geqslant \sqrt{2\alpha \log n}\right) \leqslant \exp(-\alpha \log n)\sqrt{\frac{2}{\pi}} = n^{-\alpha}\sqrt{\frac{2}{\pi}}.$$

Moreover, since $\alpha > 1$, this last quantity is summable, and the Borel-Cantelli lemma 1.1.1 then implies that

$$\mathbb{P}\Big(|g_n| \geqslant \sqrt{2\alpha \log n} \quad \text{infinitely often}\Big) = 0.$$

Consequently, the random variable $C = \sup_{2\leqslant n\leqslant\infty} \frac{|g_n|}{\sqrt{\log n}}$ is finite almost surely. Armed with this lemma, we can establish the uniform convergence of $X_n(t)$ to X_t. To do this, we will show that, almost surely, the remainder of the series (in absolute value) converges uniformly to 0.

First, observe that for all $n \in [2^j, 2^{j+1}[$ we have :

- $\log n < j+1$

- for all $t \in [0,1], d_n(t) \neq 0$ for at most one integer $n \in [2^j, 2^{j+1}[$ (this localization of the support fully justifies the use of the functions d_n).

This is why, using the previous lemma, we have, for all $M \geqslant 2^J$ with $J \geqslant 1$,

$$\begin{aligned}\sum_{n=M}^{\infty} |g_n|\, v_n(t) = \sum_{n=M}^{\infty} |g_n|\, \lambda_n d_n(t) &\leqslant C \sum_{n=M}^{\infty} \lambda_n \sqrt{\log n} d_n(t) \\ &\leqslant C \sum_{j=J}^{\infty} \sum_{k=0}^{2^j-1} \frac{1}{2} 2^{-j/2} \sqrt{j+1} d_{2^j+l}(t) \\ &= C \sum_{j=J}^{\infty} \frac{1}{2} 2^{-j/2} \sqrt{j+1}\end{aligned}$$

Since $0 \leqslant d_n(t) \leqslant 1$ for all $n \in \mathbb{N}$ and for all $t \in [0,1]$, the last term above tends to zero as $J \to +\infty$. This implies that (almost surely) X_t^n converges uniformly ; furthermore, the limiting process $(X_t)_{t\in[0,1]}$ is continuous (almost surely) since the functions $t \mapsto X_t^n$ are continuous for all $n \in \mathbb{N}$.

The proof is not completely finished. It remains to show that it is a Gaussian process. Note that the form of the covariance function easily allows us to show that the increments of the process are independent. Using this observation, for any $0 \leqslant t_1 < t_2 < \ldots < t_d \leqslant 1$, it is not difficult to show (by considering the characteristic function) that the vector

$$(X_{t_1}, \ldots, X_{t_d})$$

is a Gaussian vector. These details are left to the reader's attention and can be found, for example.

3.4.2 proof of Donsker

To prove the Theorem we shall show the convergence of finite dimensional law and the tightness.

Finite dimensional Laws

Pour identifier la limite Ã©ventuelle, nous allons utiliser les lois (en dimension finie) induites par

Proposition 2.5.2. Soient $0 \leqslant t_1 < t_2 < \ldots < t_d \leqslant 1$ alors

$$\left(X_{t_1}^n, X_{t_2}^n, \ldots, X_{t_d}^n\right) \xrightarrow{\mathcal{L}} G \quad \text{dans} \quad \mathbb{R}^d \quad \text{lorsque} \quad n \to \infty$$

ou G est un vecteur gaussien centrÃ© de matrice covariance $\Gamma = (\Gamma_{ij})_{1 \leqslant i,j \leqslant d}$ avec $\Gamma_{ij} = \min(t_i, t_j)$ pour tout $1 \leqslant i, j \leqslant d$.

DÃ©monstration. Pour simplifier, nous ferons la dÃ©monstration lorsque $d = 2$. Le cas gÃ©nÃ©ral est similaire mais les notations alourdissent inutilement la preuve.

Tout d'abord, observons que pour tout $t \in [0, 1]$

$$\left| X_t^n - \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor} \right| \leqslant \frac{1}{\sqrt{n}} \left| V_{\lfloor nt \rfloor + 1} \right|.$$

Moreover, pour tout $\varepsilon > 0$, puisque les variables ont mÃªme loi,

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \left| V_{\lfloor nt \rfloor + 1} \right| > \varepsilon \right) = \mathbb{P}\left(|V_1| > \sqrt{n}\varepsilon \right)$$

D'ou, d'aprÃ¨s l'inÃ©galitÃ© de Tchebychev, $\frac{1}{\sqrt{n}} V_{\lfloor nt \rfloor + 1} \xrightarrow{\mathbb{P}} 0$ lorsque $n \to +\infty$. C'est donc Ã©galement le cas de $X_t^n - \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}$.

En consÃ©quence, la norme euclidienne $\left\| \left(X_{t_1}^n, X_{t_2}^n\right) - \left(\frac{1}{\sqrt{n}} S_{\lfloor nt_1 \rfloor}, \frac{1}{\sqrt{n}} S_{\lfloor nt_2 \rfloor} \right) \right\|_2$ tend vers 0 en probabilitÃ©.

Pour poursuivre notre Ã©tude, nous utiliserons le lemme ci-dessous.

Lemme 2.5.1. Soient Y_n, X_n et X des variables alÃ©atoires rÃ©elles. Supposons que $Y_n \xrightarrow{\mathcal{L}} X$ et $|X_n - Y_n| \xrightarrow{\mathbb{P}} 0$ lorsque $n \to +\infty$. Alors $X_n \xrightarrow{\mathcal{L}} X$ lorsque $n \to +\infty$.

Au vu de ce qui prÃ©cÃ¨de, le prÃ©cÃ©dent lemme nous assure que nous pouvons travailler avec $\frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}$ plutÃ´t que X_t^n.

Puisque nous sommes en dimension finie, nous pouvons utiliser la transformÃ©e de Fourier pour Ã©tablir un rÃ©sultat de convergence en loi. Ici, pour $(u_1, u_2) \in \mathbb{R}^2$, nous avons

$$\begin{aligned}
\mathbb{E}\left[e^{iu_1 \frac{1}{\sqrt{n}} S_{\lfloor nt_1 \rfloor} + iu_2 \frac{1}{\sqrt{n}} S_{\lfloor nt_2 \rfloor}} \right] &= \mathbb{E}\left[e^{i(u_1+u_2) \frac{1}{\sqrt{n}} S_{\lfloor nt_1 \rfloor} + iu_2 \frac{1}{\sqrt{n}} \left(S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor} \right)} \right] \\
&= \mathbb{E}\left[e^{i(u_1+u_2) \frac{\sqrt{t_1}}{\sqrt{nt_1}} S_{\lfloor nt_1 \rfloor}} \right] \mathbb{E}\left[e^{iu_2 \frac{1}{\sqrt{n}} \left(S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor} \right)} \right]
\end{aligned}$$

puisque $S_{\lfloor nt_1 \rfloor}$ et $\left(S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor} \right)$ sont des variables alÃ©atoires indÃ©pendantes (car $t_1 < t_2$). Nous pouvons alors utiliser le thÃ©orÃ¨me 1.5.2 de la limite centrale dans $\mathbb{R}^2$ qui nous assure que

$$\lim_{n \to +\infty} \mathbb{E}\left[e^{i(u_1+u_2) \frac{\sqrt{t_1}}{\sqrt{nt_1}} S_{\lfloor nt_1 \rfloor}} \right] \mathbb{E}\left[e^{iu_2 \frac{1}{\sqrt{n}} \left(S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor} \right)} \right] = e^{-(u_1+u_2)^2 \frac{t_1}{2}} \times e^{-u_2^2 \frac{t_2 - t_1}{2}}$$

En d'autres termes,

$$\left(X_{t_1}^n, X_{t_2}^n\right) \xrightarrow{\mathcal{L}} G \quad \text{lorsque} \quad n \to +\infty$$

ou G est un vecteur gaussien centrÃ© de matrice covariance $\Gamma = \begin{pmatrix} t_1 & t_1 \\ t_1 & t_2 \end{pmatrix}$.

Remarque. L'Ã©tape prÃ©cÃ©dente nous montre la forme de la structure de covariance d'un Ã©ventuel processus limite X :

$$\mathbb{E}\left[X_s X_t\right] = \min(s,t) \quad \text{pour tout} \quad s,t \in [0,1]$$

Tightness

In order to show the tightness, we shall use the Kolmogorov criterion. We apply this Theorem for $X_n = S^{[n]}$ and $p = 4$. Note that $p = 2$ will imply that $\beta = 0$ and then it is not sufficient.

Let us compute

$$\mathbb{E}\left[\left|S_t^{[n]} - S_s^{[n]}\right|^4\right]$$

We shall consider two cases

— $\exists k \in \mathbb{N}, sn \leq k < tn$

— $\exists k \in \mathbb{N}, k \leq sn < tn \leq k+1$

The second case is easy. For the first

$$\left|S_t^{[n]} - S_s^{[n]}\right| \leq \frac{S_{\lfloor tn \rfloor} - S_{\lceil sn \rceil}}{\sqrt{n}} + \frac{|tn - \lfloor tn \rfloor|}{\sqrt{n}}|\xi_{\lfloor tn \rfloor + 1}| + \frac{|\lceil sn \rceil - sn|}{\sqrt{n}}|\xi_{\lceil sn \rceil}|$$

At this stage, we have

$$\mathbb{E}\left(\sum_{\lceil sn \rceil + 1}^{\lfloor tn \rfloor} \xi_k\right)^4 = \mathbb{E}|\xi_1|^4(\lfloor tn \rfloor - \lceil sn \rceil) + (\mathbb{E}|\xi_1|^2)^2 \frac{(\lfloor tn \rfloor - \lceil sn \rceil)(\lfloor tn \rfloor - \lceil sn \rceil - 1)}{2}$$

Now let us observe that for all $p \geq 1$, for all $k \geq 1$ for all $a_i > 0$

$$\sum_{i=1}^k a_i^p \leq \left(\sum_{i=1}^k a_i\right)^p \leq p \sum_{i=1}^k a_i^p$$

We have now

$$\begin{aligned}
&\mathbb{E}\left[\left|S_t^{[n]} - S_s^{[n]}\right|^4\right] \\
\leq\ & \frac{4}{n^2}\left[\mathbb{E}|S_{\lfloor tn \rfloor} - S_{\lceil sn \rceil}|^4 + |tn - \lfloor tn \rfloor|^4 \mathbb{E}\xi_1^4 + |sn - \lceil tn \rceil|^4 \mathbb{E}\xi_1^4\right] \\
\leq\ & \frac{C_\xi}{n^2}\left[\frac{1}{2}(\lfloor tn \rfloor - \lceil sn \rceil) + \frac{(\lfloor tn \rfloor - \lceil sn \rceil)(\lfloor tn \rfloor - \lceil sn \rceil)}{2} + |tn - \lfloor tn \rfloor|^4 + |\lceil sn \rceil - sn|^4\right] \\
\leq\ & \frac{C_\xi}{n^2}\left[(\lfloor tn \rfloor - \lceil sn \rceil)^2 + |tn - \lfloor tn \rfloor|^2 + |\lceil sn \rceil - sn|^2\right] \\
\leq\ & \frac{C_\xi}{n^2}\left(\lfloor tn \rfloor - \lceil sn \rceil + tn - \lfloor tn \rfloor + \lceil sn \rceil - sn\right)^2 \\
\leq\ & \frac{C_\xi}{n^2}\left(tn - sn\right)^2 = C_\xi (t-s)^2
\end{aligned}$$

and the Kolmogorov criterion is satisfied with $p = 4$, $\beta = 1$ and $c = C_\xi$.

Chapitre 4

Optimal transport