

# Quelques exemples de théorèmes limites fonctionnels

GRÉGORY MIERMONT

ENS de Lyon et Institut universitaire de France

Notes de cours pour ALEA'2014, CIRM, Luminy, 17–21 mars 2014.

Quand on cherche à étudier des variables aléatoires choisies dans des classes de structures combinatoires aléatoires (arbres, permutations, graphes, cartes, etc.), il est fréquent d'avoir recours à certains codages par des processus aléatoires. L'existence de théorèmes limites fonctionnels pour ces processus permet alors souvent, par contre-coup, de déterminer certaines des propriétés de grands objets aléatoires, et souvent, à donner d'un seul coup plusieurs de ces propriétés asymptotiques, là où d'autres méthodes peuvent nécessiter une étude séparée pour tel ou tel aspect de l'objet considéré (hauteur, largeur, indice de Wiener d'un arbre). Parfois, il est également possible de déterminer un analogue continu de ces objets (forme limite pour des diagrammes de Young, arbre continu aléatoire, etc.)

Le but de ces notes est de donner quelques exemples élémentaires de tels théorèmes limites fonctionnels, et de méthodes de base permettant d'obtenir de tels résultats. Nous nous concentrerons essentiellement sur le cas de la marche simple sur  $\mathbb{Z}$  et de certains de ses avatars conditionnés (pont et excursion).

## 1 Convergence de la marche aléatoire simple vers le mouvement brownien

### 1.1 La marche aléatoire simple

Soit  $X_1, X_2, \dots$  une suite de variables aléatoires réelles indépendantes et identiquement distribuées, uniformes sur  $\{-1, 1\}$ , définies sur un espace de probabilités  $(\Omega, \mathcal{F}, \mathbb{P})$ . On s'intéresse au comportement de la marche aléatoire  $(S_n, n \geq 0)$  associée définie par  $S_0 = 0$  et

$$S_n = X_1 + \dots + X_n, \quad n \geq 1.$$

Par exemple, le théorème de la limite centrale montre que pour tout  $a < b$ , on a

$$P\left(a < \frac{S_n}{\sqrt{n}} < b\right) \xrightarrow{n \rightarrow \infty} \int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx,$$

ou, de façon équivalente

$$\frac{S_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\text{loi}} N,$$

où  $N$  est une variable aléatoire gaussienne centrée réduite (on notera  $\mathcal{N}(m, \sigma^2)$  la loi gaussienne de moyenne  $m$  et de variance  $\sigma^2$ ). Ceci détermine la position asymptotique de  $S_n$  lorsque  $n \rightarrow \infty$ : on interprète le résultat précédent en disant que la loi de  $S_n$  est proche de  $\mathcal{N}(0, \sigma^2)$ .

Si l'on veut une information plus précise sur la *trajectoire* de  $(S_n, n \geq 0)$ , on peut s'intéresser à la position de la marche à des instants  $n_1, n_2, \dots, n_k$  rangés par ordre croissant. Par convention on notera aussi  $n_0 = 0$ . Les accroissements

$$S_{n_i} - S_{n_{i-1}} = X_{n_{i-1}+1} + \dots + X_{n_i}, \quad 1 \leq i \leq k$$

sont alors indépendants, et ont même loi, respectivement, que  $S_{n_i - n_{i-1}}$ . Supposons que  $n_1, \dots, n_k$  sont des suites (indexées disons par  $n$ ) et qu'il existe  $t_0 = 0 \leq t_1 \leq \dots \leq t_k$  tels que

$$t_i = \lim_{n \rightarrow \infty} \frac{n_i}{n}, \quad 1 \leq i \leq k.$$

Alors, une nouvelle utilisation du théorème de la limite centrale (multidimensionnel) montre que

$$\left( \frac{S_{n_i} - S_{n_{i-1}}}{\sqrt{n}}, 1 \leq i \leq k \right) \xrightarrow[n \rightarrow \infty]{\text{loi}} (N_1, \dots, N_k),$$

où les variables aléatoires  $N_1, \dots, N_k$  sont indépendantes, gaussiennes, de variances respectives  $t_1, t_2 - t_1, \dots, t_k - t_{k-1}$ . De façon équivalente, on a que

$$\left( \frac{S_{n_i}}{\sqrt{n}}, 1 \leq i \leq k \right) \xrightarrow[n \rightarrow \infty]{\text{loi}} (N_1, N_1 + N_2, \dots, N_1 + \dots + N_k), \quad (1)$$

ce qu'on peut interpréter comme la détermination des positions asymptotiques de la marche aléatoire considérée en un nombre fini d'instants  $n_1, \dots, n_k$ .

Peut-on aller plus loin et déterminer le comportement asymptotique de la marche aléatoire *dans son intégralité*? Au vu des résultats précédents, il est naturel de considérer une fonction aléatoire d'un paramètre réel positif  $t$ , correspondant à la marche aléatoire dont le temps est renormalisé par  $n$ , et l'espace par  $\sqrt{n}$ . Notons donc

$$S_t^{[n]} = \frac{S_{[nt]}}{\sqrt{n}}, \quad t \geq 0.$$

Si l'on en effectue une simulation, on s'aperçoit que, lorsque  $n$  augmente, la fonction  $S^{[n]}$  semble prendre la forme d'une fonction, de nature aléatoire. Cette fonction limite est appelée mouvement brownien et est un des objets les plus fondamentaux en probabilités. Précisément, un mouvement brownien standard de dimension 1 est une collection  $(B_t, t \geq 0)$  de variables aléatoires indexées par  $\mathbb{R}_+$ , vérifiant les trois propriétés suivantes

- $B_0 = 0$  p.s.,
- la fonction  $t \mapsto B_t$  est p.s. continue de  $\mathbb{R}_+$  dans  $\mathbb{R}$ ,
- pour tout  $0 = t_0 \leq t_1 \leq \dots \leq t_k$ , les variables aléatoires  $(B_{t_i} - B_{t_{i-1}}, 1 \leq i \leq k)$  sont indépendantes, respectivement de loi  $\mathcal{N}(0, t_i - t_{i-1})$ .

Il n'est pas évident *a priori* qu'une telle famille de variables aléatoires puisse être définie sur un certain espace de probabilités, mais c'est bien le cas — nous en verrons une preuve plus tard.

Avec cette définition, il est clair que la loi de  $(B_{t_1}, B_{t_2}, \dots, B_{t_k})$  est la même que celle de  $(N_1, N_1 + N_2, \dots, N_1 + \dots + N_k)$  avec les mêmes notations que ci-dessus. On reformule donc (1) en écrivant que pour tout choix de réels positifs  $t_1 \leq t_2 \leq \dots \leq t_k$ , on a

$$\left( S_{t_i}^{[n]}, 1 \leq i \leq k \right) \xrightarrow[n \rightarrow \infty]{\text{loi}} (B_{t_i}, 1 \leq i \leq k), \quad (2)$$

Si l'on prend des instants  $t_i$  de plus en plus nombreux et resserrés, cette convergence semble bien expliquer que  $S^{[n]}$  approche la trajectoire du mouvement brownien  $B$  lorsque  $n \rightarrow \infty$ , comme nous l'avons vu sur des simulations. Pourtant, les convergences (2) ne sont pas entièrement satisfaisantes de ce point de vue. Par exemple, il est naturel de conjecturer que l'on a

$$\sup_{0 \leq t \leq 1} S_t^{[n]} \xrightarrow[n \rightarrow \infty]{\text{loi}} \sup_{0 \leq t \leq 1} B_t,$$

mais cette convergence *n'est pas* une conséquence des convergences en loi (2). Disons que, même si l'on choisit des instants  $t_i$  très resserrés, la convergence (2) n'exclut pas *a priori* que le processus  $S^{[n]}$  puisse prendre des valeurs anormalement élevées entre deux de ces instants. C'est le même genre de problème qui advient lorsqu'on a convergence ponctuelle d'une suite de fonctions, et non convergence uniforme. Il nous faut donc développer un concept de convergence en loi pour des variables aléatoires qui sont des fonctions, munies de la topologie uniforme, et c'est l'objet de la première partie du cours.

## 1.2 Théorème de Prokhorov

Soit  $(E, d)$  un espace métrique, qu'on supposera séparable et complet. On note  $\mathcal{C}_b(E, \mathbb{R})$  l'ensemble des fonctions continues bornées de  $E$  dans  $\mathbb{R}$ .

**Définition 1.** Soit  $(X_n)_{n \geq 1}, X$  une suite de variables aléatoires à valeurs dans  $E$ . On dit que  $X_n$  converge en loi vers  $X$ , et on note  $X_n \Rightarrow X$ , si pour toute fonction  $f \in \mathcal{C}_b(E, \mathbb{R})$  on a  $E[f(X_n)] \rightarrow E[f(X)]$ .

Le théorème de Prokhorov, qui est sans doute l'outil le plus fondamental pour les théorèmes limite fonctionnels, donne un critère de « relative compacité » pour la convergence en loi. Pour bien comprendre ce théorème, nous commençons par énoncer le résultat fondamental suivant.

**Théorème 2. [Riesz, Banach-Alaoglu]** Si  $E$  est compact, alors de toute suite de variable aléatoires à valeurs dans  $E$ , on peut extraire une sous-suite qui converge en loi.

**Définition 3.** La suite  $(X_n, n \geq 1)$  de variables aléatoires à valeurs dans  $E$  est dite tendue si pour tout  $\varepsilon > 0$ , il existe un compact  $K_\varepsilon \subseteq E$  tel que

$$\inf_{n \geq 1} P(X_n \in K_\varepsilon) \geq 1 - \varepsilon.$$

Autrement dit, une suite de variables aléatoires est tendue si l'on peut trouver un seul et même compact qui concentre la plupart de la masse des lois de tous les éléments de cette suite.

**Théorème 4. [Prokhorov]** Soit  $(X_n, n \geq 1)$  une suite tendue de variables aléatoires à valeurs dans  $E$ . Alors il existe une sous-suite de  $(X_n, n \geq 1)$  qui converge en loi.

Ce théorème est en général utilisé conjointement avec la proposition suivante, qui est une conséquence du fait que la convergence en loi correspond à une topologie métrisable (dite topologie étroite) sur l'ensemble des mesures de probabilités sur  $E$ .

**Proposition 5.** Soit  $(X_n, n \geq 1)$  une suite de variables aléatoires à valeurs dans  $E$ . On a que  $X_n \Rightarrow X$  si et seulement si de toute suite extraite, on peut réextraire une sous-sous-suite qui converge en loi vers  $X$ .

Nous appliquerons ce théorème dans le cas particulier où  $E = \mathcal{C}([0, 1], \mathbb{R})$ , muni de la norme uniforme, et que l'on notera désormais  $\mathcal{C}$ . Nous donnons un critère utile de tension pour des variables aléatoires à valeurs dans  $\mathcal{C}$ . Dans l'énoncé suivant, pour  $x \in \mathcal{C}$ , la quantité

$$N_\alpha(x) = \sup_{s, t \in [0, 1], s \neq t} \frac{|x(s) - x(t)|}{|t - s|^\alpha},$$

est appelée la *norme de Hölder d'exposant  $\alpha$*  de  $x$ .

**Théorème 6. [Critère de Kolmogorov]** On considère une suite  $(X_n)_{n \geq 1}$  de variables aléatoires à valeurs dans  $\mathcal{C}$ . On suppose

- que la suite  $(X_n(0))_{n \geq 1}$  est tendue, et
- qu'il existe  $\beta, p, C > 0$  tel que

$$\sup_{n \geq 1} \mathbb{E}[|X_n(s) - X_n(t)|^p] \leq C |t - s|^{1+\beta}, \quad \text{pour tout } s, t \in [0, 1].$$

Alors  $(X_n)_{n \geq 1}$  est tendue. Plus précisément, pour tout  $\alpha \in ]0, \beta/p[$  fixé et pour tout  $\varepsilon > 0$ , il existe  $M > 0$  tel que

$$\sup_{n \geq 1} \mathbb{P}(N_\alpha(X_n) > M) \leq \varepsilon. \quad (3)$$

Pour montrer ce résultat, nous avons recours à un lemme intermédiaire. Notons  $\mathcal{D}_k = \{i/2^k : 0 \leq i \leq 2^k\}$  les rationnels dyadiques de niveau  $k$  dans  $[0, 1]$ .

**Lemme 7.** Soit  $f: [0, 1] \rightarrow \mathbb{R}$  une fonction. Supposons qu'il existe  $K, \alpha > 0$  pour lesquels, pour tout  $k \geq 0$  et tout  $s, t \in \mathcal{D}_k$  consécutifs, on ait  $|f(s) - f(t)| \leq K |t - s|^\alpha$ . Alors pour tout  $s, t \in \bigcup_{n \geq 0} \mathcal{D}_n$  on a

$$|f(s) - f(t)| \leq \frac{2K}{1 - 2^{-\alpha}} |t - s|^\alpha.$$

En particulier, si  $f \in \mathcal{C}$ , on a

$$N_\alpha(f) \leq \frac{2K}{1 - 2^{-\alpha}}.$$

**Preuve.** Soit  $s, t$  deux nombres dans  $\bigcup_{k \geq 0} \mathcal{D}_k$ . Alors il existe un unique entier  $r \geq 0$  tel que  $2^{-r-1} \leq |t - s| < 2^{-r}$ . Sans perte de généralité, on peut supposer que  $s \leq t$ . On distingue deux cas.

Dans un premier cas, il existe  $i \in \{0, 1, \dots, 2^r - 1\}$  tel que  $i 2^{-r} \leq s, t \leq (i+1) 2^{-r}$ . Dans ce cas, les écritures dyadiques de  $s$  et  $t$  sont de la forme

$$s = \frac{i}{2^r} + \sum_{k \geq r+1} \frac{i_k(s)}{2^k}, \quad t = \frac{i}{2^r} + \sum_{k \geq r+1} \frac{i_k(t)}{2^k}$$

où les suites  $(i_k(s))_{k \geq r+1}$  et  $(i_k(t))_{k \geq r+1}$  sont à valeurs dans  $\{0, 1\}$ , et nulles à partir d'un certain rang. On a alors, en posant

$$s_l = \frac{i}{2^r} + \sum_{k=r+1}^{r+l} \frac{i_k(s)}{2^k},$$

que  $s_l$  et  $s_{l+1}$  sont ou bien des nombres dyadiques consécutifs dans  $\mathcal{D}_{r+l+1}$ , ou bien égaux (ce qu'ils sont pour tout  $l$  assez grand, puisqu'alors  $s_l = s$ ). On a donc par hypothèse

$$\begin{aligned} |f(s) - f(i 2^{-r})| &\leq \sum_{l \geq 0} |f(s_l) - f(s_{l+1})| \leq \sum_{l \geq 0} K 2^{-(l+r+1)\alpha} \\ &= \frac{K 2^{-(r+1)\alpha}}{1 - 2^{-\alpha}} \leq \frac{K}{1 - 2^{-\alpha}} |t - s|^\alpha \end{aligned}$$

La même majoration est vraie pour  $t$  à la place de  $s$ , et on déduit que

$$|f(s) - f(t)| \leq \frac{2K}{1 - 2^{-\alpha}} |t - s|^\alpha.$$

Dans un second cas, il existe  $i \in \{1, 2, \dots, 2^r - 1\}$  tel que  $(i - 1) 2^{-r} \leq s \leq i 2^{-r} \leq t \leq (i + 1) 2^{-r}$ . On écrit alors plutôt

$$s = \frac{i}{2^r} - \sum_{k \geq r+1} \frac{i'_k(s)}{2^k}, \quad t = \frac{i}{2^r} + \sum_{k \geq r+1} \frac{i_k(t)}{2^k},$$

où de la même manière que précédemment,  $(i'_k(s))_{k \geq r+1}$  est une suite de  $\{0, 1\}$  nulle à partir d'un certain rang. Le même raisonnement que ci-dessus permet alors de conclure à la même majoration pour  $|f(s) - f(t)|$ .

Comme l'ensemble  $\bigcup_n \mathcal{D}_n$  est dense dans  $[0, 1]$ , si  $f \in \mathcal{C}$ , nous en déduisons que la majoration reste valable pour tout  $s, t \in [0, 1]$ . Ceci termine la preuve du lemme.  $\square$

Nous prouvons à présent la proposition 6. Posons

$$Z_{n,k} = \max_{0 \leq i \leq 2^k - 1} |X_n((i + 1) 2^{-k}) - X_n(i 2^{-k})|.$$

On a alors, par inégalité de Markov,

$$\begin{aligned} \mathbb{P}(Z_{n,k} \geq K 2^{-k\alpha}) &\leq 2^k \max_{0 \leq i \leq 2^k - 1} \mathbb{P}(|X_n((i + 1) 2^{-k}) - X_n(i 2^{-k})| \geq K 2^{-k\alpha}) \\ &\leq 2^k \max_{0 \leq i \leq 2^k - 1} \frac{\mathbb{E}[|X_n((i + 1) 2^{-k}) - X_n(i 2^{-k})|^p]}{K^p 2^{-pk\alpha}} \end{aligned}$$

Par hypothèse, cette dernière quantité est majorée par  $C K^{-p} 2^{-(\beta - p\alpha)k}$ . Si l'on choisit  $\alpha \in ]0, \beta/p[$ , on en déduit que ces majorants sont sommables sur  $k$ . Le lemme 7 implique que

$$\begin{aligned} \mathbb{P}\left(N_\alpha(X_n) > \frac{2K}{1 - 2^{-\alpha}}\right) &\leq \mathbb{P}(\exists k \geq 0: Z_{n,k} > K 2^{-k\alpha}) \\ &\leq \frac{C}{K^p} \sum_{k \geq 0} 2^{-(\beta - p\alpha)k} \leq \frac{C'}{K^p}, \end{aligned}$$

pour une constante  $C' \in ]0, \infty[$ , qui ne dépend que de  $\alpha, \beta, p, C$  mais non de  $n$ . Le dernier majorant converge vers 0 lorsque  $K \rightarrow \infty$ , d'où l'on déduit que pour tout  $\varepsilon > 0$ , il existe  $M > 0$  tel que (3) est vraie. Ceci, conjointement à la première hypothèse de la proposition 6, implique la tension de  $(X_n)_{n \geq 1}$ . En effet, pour tout  $M > 0$ , l'ensemble des  $f \in \mathcal{C}$  tels que  $|f(0)| \leq M$  et  $N_\alpha(f) \leq M$  est une famille équicontinue, qui est donc relativement compacte par le théorème d'Ascoli-Arzelà.

### 1.3 Le théorème de Donsker

Soit  $X_1, X_2, \dots$  une suite i.i.d. de variables aléatoires uniformes sur  $\{-1, 1\}$ . On pose  $S_0 = 0$ , et pour  $n \geq 1$ ,  $S_n = X_1 + \dots + X_n$ . On étend alors  $S$  en une fonction continue de  $\mathbb{R}_+$  dans  $\mathbb{R}$  par la formule

$$S_t = (1 - \{t\}) S_{[t]} + \{t\} S_{[t]+1}, \quad t \geq 0,$$

où  $\{t\} = t - [t]$  est la partie fractionnaire de  $t$ . Enfin, pour tout  $n \geq 1$ , on pose

$$S_t^{(n)} = \frac{S_{nt}}{\sqrt{n}}, \quad 0 \leq t \leq 1.$$

On voit ainsi  $S^{(n)}$  comme une variable aléatoire à valeurs dans  $\mathcal{C}$ .

**Théorème 8. [Donsker]** *La suite  $(S^{(n)})_{n \geq 1}$  converge en loi dans  $\mathcal{C}$  vers le mouvement brownien standard unidimensionnel  $(B_t)_{0 \leq t \leq 1}$ .*

*La loi de ce dernier est caractérisée par le fait que pour tout  $0 = t_0 \leq t_1 \leq \dots \leq t_k$ , les variables  $B_{t_i} - B_{t_{i-1}}$ ,  $1 \leq i \leq k$  sont indépendantes, respectivement de loi gaussienne  $\mathcal{N}(0, (t_i - t_{i-1}) I_d)$ .*

En fait, le théorème de Donsker est vrai dans le cadre beaucoup plus général où les variables  $X_i, i \geq 1$  sont indépendantes et de même loi, centrées, et de variance finie. Il se généralise également à des marches aléatoires à valeurs dans  $\mathbb{R}^d$ . La preuve dans le cas particulier de la marche au plus proche voisin étant un peu plus simple, c'est cette version qu'on démontrera ici. Nous laissons en exercice au lecteur de montrer ce théorème par les mêmes méthodes, en supposant que les variables aléatoires  $X_1, X_2, \dots$  sont i.i.d., centrées, et vérifient  $\mathbb{E}[X_1^4] < \infty$ . La dernière hypothèse n'est pas nécessaire, mais s'en débarrasser nécessite des critères un peu plus fins que celui de Kolmogorov.

**Étape 1: convergence des lois marginales de dimension finie.** Nous fixons des instants  $0 = t_0 < t_1 < \dots < t_k \leq 1$ . Nous avons déjà vu que

$$\left( \frac{S_{\lfloor nt_i \rfloor} - S_{\lfloor nt_{i-1} \rfloor}}{\sqrt{n}}, 1 \leq i \leq k \right) \xrightarrow[n \rightarrow \infty]{\text{loi}} (N_1, N_2, \dots, N_k),$$

où les variables aléatoires  $N_1, \dots, N_k$  sont indépendantes, respectivement de loi  $\mathcal{N}(0, (t_i - t_{i-1}) I_d)$ . Ceci provient d'une application du théorème centrale limite vectoriel. Par image continue, on en déduit que

$$\left( \frac{S_{\lfloor nt_i \rfloor}}{\sqrt{n}}, 1 \leq i \leq k \right) \xrightarrow[n \rightarrow \infty]{\text{loi}} (N_1, N_1 + N_2, \dots, N_1 + \dots + N_k).$$

Enfin, remarquons que

$$\left| S_t^{(n)} - \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}} \right| \leq \frac{|X_{\lfloor nt \rfloor}|}{\sqrt{n}},$$

qui converge vers 0 en probabilité pour tout  $t$  fixé. Pour cette raison, on a la convergence en probabilité:

$$\left( S_{t_i}^{(n)} - \frac{S_{\lfloor nt_i \rfloor}}{\sqrt{n}}, 1 \leq i \leq k \right) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

Par le lemme de Slutsky, on en déduit que

$$\left( S_{t_i}^{(n)}, 1 \leq i \leq k \right) \xrightarrow[n \rightarrow \infty]{\text{loi}} (B_{t_i}, 1 \leq i \leq k).$$

**Étape 2: tension.** Nous montrons enfin que la suite  $(S^{(n)})_{n \geq 1}$  est tendue dans  $\mathcal{C}$ . Pour cela, on applique le critère de Kolmogorov. Soit  $p$  un entier. Si  $s$  et  $t$  sont dans un même intervalle de la forme  $[k/n, (k+1)/n]$ , on a alors

$$\mathbb{E} \left[ \left| S_t^{(n)} - S_s^{(n)} \right|^{2p} \right] = \left( \frac{|nt - ns|}{\sqrt{n}} \right)^{2p} \leq |t - s|^p.$$

Sinon, en supposant en plus,  $s < t$ , il existe deux entiers  $k$  et  $k'$  tels que  $k - 1 \leq ns \leq k \leq k' \leq nt \leq k' + 1$ , et on a alors

$$\begin{aligned} \mathbb{E} \left[ \left| S_t^{(n)} - S_s^{(n)} \right|^{2p} \right] &\leq 2^{2p} \left( |s - k/n|^p + \mathbb{E} \left[ \left| \frac{S_{k'} - S_k}{\sqrt{n}} \right|^{2p} \right] + |t - k'/n|^p \right) \\ &\leq 2^{2p+1} |t - s|^p + 2^{2p} \mathbb{E} \left[ \left| \frac{S_{k'} - S_k}{\sqrt{n}} \right|^{2p} \right] \end{aligned}$$

Or un rapide calcul montre que  $\mathbb{E}[S_k^{2p}] \leq C_p k^p$  pour une constante  $C_p > 0$  et pour tout  $k \geq 0$ . En effet, on a  $\mathbb{E}[S_k^{2p}] = \sum \mathbb{E}[X_{i_1} \dots X_{i_{2p}}]$  où la somme porte sur tous les indices  $i_1, \dots, i_p$  entre 1 et  $n$ , et les espérances concernées sont non-nulles si et seulement si on peut appairer ces indices deux par deux. D'où l'on déduit finalement qu'il existe une constante  $C'_p$  telle que pour tout  $s, t \in [0, 1]$  on ait :

$$\mathbb{E} \left[ \left| S_t^{(n)} - S_s^{(n)} \right|^{2p} \right] \leq C'_p |t - s|^p.$$

En constatant en plus que  $S_0^{(n)} = 0$  pour tout  $n$ , on peut donc appliquer le critère de Kolmogorov.

**Conclusion.** Nous avons montré que la suite  $(S^{(n)})_{n \geq 1}$  est tendue, et converge au sens des marginales de dimension finie vers le mouvement brownien. Toute limite en loi de  $S^{(n)}$  le long d'une sous-suite a donc même loi que le mouvement brownien. Donc  $S^{(n)} \Rightarrow B$ .

**Remarque 9.** De façon tacite, on a utilisé le fait que les lois marginales de dimension finie d'une variable aléatoire  $X$  à valeurs dans  $\mathcal{C}$ , c'est-à-dire les lois des variables aléatoires  $(X(t_1), \dots, X(t_k))$  pour tous les choix de  $k \geq 1$  et  $t_1, \dots, t_k \in [0, 1]$ , caractérisent la loi de  $X$ . Ceci est donné par l'exercice suivant.

**Exercice 1.** Sur l'ensemble  $\mathcal{C}$ , on peut définir naturellement deux tribus :

- la tribu borélienne associée à la topologie de  $(\mathcal{C}, \|\cdot\|_\infty)$ ,
- la tribu produit, qui est la plus petite tribu rendant mesurables les applications  $f \rightarrow f(t)$  de  $\mathcal{C}$  dans  $\mathbb{R}$  pour tout  $t \in [0, 1]$ .

Montrer que ces deux tribus coïncident. En déduire que la loi d'une variable aléatoire  $X$  à valeurs dans  $\mathcal{C}$  est caractérisée par ses lois marginales de dimension finie.

## 2 Variantes conditionnées

### 2.1 Ponts

Le pont brownien est intuitivement le processus  $(B_t)_{0 \leq t \leq 1}$  conditionné par l'événement  $\{B_1 = 0\}$ . Il existe plusieurs façons équivalentes de définir un tel conditionnement singulier (l'événement en question est de probabilité nulle), mais celle que nous retiendront utilise l'absolue continuité par rapport au mouvement brownien. Soit  $p_t(x) = \exp(-x^2/2t)/\sqrt{2\pi t}$  la densité de la loi de  $B_t$ .

**Théorème 10.** La suite de variables aléatoires  $S^{(2n)}$ , respectivement conditionnées par l'événement  $S_1^{(2n)} = 0$ , converge en loi dans  $\mathcal{C}$  vers un processus  $(\mathbf{b}_t)_{0 \leq t \leq 1}$  dont les marginales de dimension finie sont données par le fait que  $\mathbf{b}_0 = \mathbf{b}_1 = 0$  p.s., et pour  $0 < t_1 \leq t_2 \leq \dots \leq t_k < 1$ ,

$$\frac{\mathbb{P}(\mathbf{b}_{t_1} \in dx_1, \dots, \mathbf{b}_{t_k} \in dx_k)}{dx_1 \cdots dx_k} = \sqrt{2\pi} p_{t_1}(x_1) p_{t_2-t_1}(x_2-x_1) \cdots p_{t_k-t_{k-1}}(x_k-x_{k-1}) p_{1-t_k}(-x_k).$$

**Exercice 2.** Montrer que si  $B$  est un mouvement brownien, le processus  $(B_t - tB_1)_{0 \leq t \leq 1}$  a même loi que  $\mathbf{b}$ . On utilisera le fait que  $(B_t - tB_1)_{t \geq 0}$  est un processus gaussien dont on calculera la fonction de covariance.

Notons dans un premier temps que si  $F$  est une fonction mesurable par rapport à la tribu engendrée par rapport à  $S_0, S_1, \dots, S_k$  avec  $k < n$ , on a, si  $P_k(r) = \mathbb{P}(S_k = r)$ ,

$$\mathbb{E}[F | S_{2n} = 0] = \frac{\mathbb{E}[F \mathbf{1}_{\{S_{2n}=0\}}]}{\mathbb{P}(S_{2n}=0)} = \frac{\mathbb{E}[F P_{2n-k}(-S_k)]}{P_{2n}(0)}.$$

où l'on a appliqué la propriété de Markov de  $(S_m)_{m \geq 0}$ . On a alors recours à un théorème de la limite locale stipulant que

$$\delta_k = \sup_{r \in \mathbb{Z}} \left| \sqrt{k} (P_k(r) + P_k(r+1)) - 2p_1\left(\frac{r}{\sqrt{k}}\right) \right| \xrightarrow{k \rightarrow \infty} 0. \quad (4)$$

Noter que dans le cas particulier qui nous intéresse, on peut obtenir ce résultat par une utilisation d'un raffinement de la formule de Stirling, puisque si  $k$  et  $r$  ont même parité, on a

$$P_k(r) = \frac{1}{2^r} \binom{k}{\frac{k+r}{2}}.$$

On note aussi que, des deux quantités  $P_k(r)$  et  $P_k(r+1)$  apparaissant dans la formule ci-dessus, une seule est non nulle.

On en déduit que, si  $F$  est une fonction mesurable bornée

$$\left| \mathbb{E} \left[ F \frac{P_{2n-k}(-S_k)}{P_{2n}(0)} \right] - \mathbb{E} \left[ F \frac{p_1(-S_k/\sqrt{2n-k})}{\sqrt{1-k/2n} \cdot p_1(0)} \right] \right| \leq C \|F\|_\infty \delta_k$$

Par conséquent, si  $G: \mathbb{R}^k \rightarrow \mathbb{R}$  est une fonction continue bornée, et si  $k = k_n$  vérifie  $k_n/n \rightarrow t_k$ , on voit aisément, en utilisant la convergence au sens des marginales de dimension finie de  $S^{(2n)}$  vers le mouvement brownien, que

$$\mathbb{E} \left[ G \left( S_{t_1}^{(2n)}, \dots, S_{t_k}^{(2n)} \right) \mid S_1^{(2n)} = 0 \right] \xrightarrow{n \rightarrow \infty} \mathbb{E} \left[ G(B_{t_1}, \dots, B_{t_k}) \frac{p_1(-B_{t_k}/\sqrt{1-t_k})}{p_1(0)\sqrt{1-t_k}} \right] = \mathbb{E}[G(\mathbf{b}_{t_1}, \dots, \mathbf{b}_{t_k})].$$

Nous montrons à présent la tension de la suite  $S^{(2n)}$  conditionnée par  $S_1^{(2n)} = 0$ . Pour cela, on applique à nouveau le critère de Kolmogorov. Pour cela, soit  $k, k' \in \{0, 1, \dots, n\}$ . Alors pour tout entier  $p$ , en utilisant à nouveau (4), on a l'existence d'une constante  $C$  universelle telle que

$$\mathbb{E}[|S_k - S_{k'}|^{2p} \mid S_{2n} = 0] = \mathbb{E} \left[ |S_k - S_{k'}|^{2p} \frac{P_n(-S_n)}{P_{2n}(0)} \right] \leq C \mathbb{E}[|S_k - S_{k'}|^{2p}].$$

De cela, on déduit immédiatement que  $\mathbb{E} \left[ \left| S_t^{(2n)} - S_s^{(2n)} \right|^{2p} \mid S_1^{(2n)} = 0 \right] \leq C_p |t - s|^p$  pour une constante  $C_p > 0$ , dès que  $s, t \in [0, 1/2]$ . Ceci reste donc vrai pour  $s, t \in [1/2, 1]$  par un simple argument de symétrie. Enfin, si  $s \in [0, 1/2]$  et  $t \in [1/2, 1]$ , on a

$$\mathbb{E} \left[ \left| S_t^{(2n)} - S_s^{(2n)} \right|^{2p} \mid S_1^{(2n)} = 0 \right] \leq 2^{2p} C_p (|t - 1/2|^p + |1/2 - s|^p) \leq 2^{2p+1} C_p |t - s|^p.$$

On peut donc appliquer le critère de Kolmogorov.

Notons que le même résultat s'applique trivialement si l'on considère le processus  $S^{(2n+1)}$  conditionné par  $S_1^{(2n+1)} = -1/\sqrt{2n+1}$ , ce qui correspond à conditionner par  $S_{2n+1} = -1$  au lieu de  $S_{2n} = 0$ .

## 2.2 Excursions

**Théorème 11.** *La suite de variables aléatoires  $S^{(2n)}$ , respectivement conditionnées par l'événement  $\{S^{(2n)} \geq 0, S_1^{(2n)} = 0\}$ , converge en loi dans  $\mathcal{C}$  vers un processus continu  $(\mathbf{e}_t)_{0 \leq t \leq 1}$ . Ce dernier est obtenu à partir du pont brownien  $\mathbf{b}$  par la transformation de Vervaat : si  $t_*$  est l'instant p.s. unique de  $[0, 1]$  où  $\mathbf{b}$  atteint son minimum, on a*

$$\mathbf{e}_t = \mathbf{b}_{t+t_*} - \mathbf{b}_{t_*}, \quad 0 \leq t \leq 1,$$

où l'on note  $\mathbf{b}_t = \mathbf{b}_{t-1}$  si  $t \in [1, 2]$ .



Pour montrer ceci, on a recours à un analogue discret de la transformée de Vervaat, qui repose sur le « lemme cyclique ». Notons  $\mathcal{P}_n$  (pour « pont ») l'ensemble des suites  $(y_1, y_2, \dots, y_{2n+1})$  d'éléments de  $\{-1, 1\}$  de somme  $-1$ , et  $\mathcal{E}_n$  (pour « excursion ») le sous-ensemble des suites de  $\mathcal{P}_n$  dont les sommes partielles  $y_1 + \dots + y_i$  sont toutes positives pour  $0 \leq i \leq 2n$ . Pour  $\mathbf{y} \in \mathcal{P}_n$  et  $i \in \{0, 1, \dots, 2n\}$ , on note  $\mathbf{y}^{(i)}$  la suite  $(y_{i+1}, y_{i+2}, \dots, y_{i+2n+1})$ , où par convention on pose  $y_{j+2n+1} = y_j$  pour  $j \in \{0, 1, \dots, 2n\}$ .

**Lemme 12.** *Soit  $\mathbf{y} = (y_1, \dots, y_{2n+1})$  un élément de  $\mathcal{P}_n$ . Alors il existe un unique  $n_* \in \{0, 1, \dots, 2n\}$  tel que  $\mathbf{y}^{(n_*)} \in \mathcal{E}_n$ . De plus,  $n_*$  est le premier indice  $i \in \{0, 1, \dots, 2n\}$  tel que*

$$y_1 + \dots + y_i = \min \{y_1 + \dots + y_j : j \in \{0, 1, \dots, 2n\}\}.$$

**Preuve.** Supposons  $n > 0$ . Si l'on voit les termes de  $\mathbf{y}$  comme ordonnés de façon cyclique, il est clair qu'il existe un indice  $i$  tel que  $y_i = 1$  et tel que  $y_{i+1} = y_{i+2} = -1$  pour que la somme totale soit négative. Mais il est alors clair que ni  $\mathbf{y}^{(i)}$  ni  $\mathbf{y}^{(i+1)}$  n'est dans  $\mathcal{E}_n$ , car le premier terme de ces deux suites est  $-1$ . On considère la suite  $\tilde{\mathbf{y}} \in \mathcal{P}_{n-1}$  obtenue de  $\mathbf{y}$  en enlevant les termes  $y_i, y_{i+1}$ . Par récurrence, on voit que l'on peut se ramener à la suite  $(-1) \in \mathcal{E}_0$ , pour laquelle le résultat est trivial.

On en déduit le résultat suivant.

**Lemme 13.** *Soit  $Y = (Y_1, \dots, Y_{2n+1})$  une suite de variables aléatoires à valeurs dans  $\mathcal{P}_n$ , dont la loi est invariante par permutation cyclique, c'est-à-dire que la loi de  $Y^{(i)}$  est la même que celle de  $Y$  pour tout  $i \in \{0, 1, \dots, 2n\}$ . Alors*

$$\mathbb{P}(Y \in \mathcal{E}_n) = \frac{1}{2n+1}.$$

*De plus, la suite  $Y$  conditionnée à être dans  $\mathcal{E}_n$  a même loi que  $Y^{(n_*)}$ , où  $n_*$  est défini dans le lemme précédent.*

**Preuve.** On a

$$\mathbb{P}(Y \in \mathcal{E}_n) = \frac{1}{2n+1} \sum_{i=0}^{2n} \mathbb{P}(Y^{(i)} \in \mathcal{E}_n) = \frac{1}{2n+1} \mathbb{E} \left[ \sum_{i=0}^{2n} \mathbf{1}_{\{Y^{(i)} \in \mathcal{E}_n\}} \right].$$

Par le lemme précédent, exactement une des indicatrices de la somme est égale à 1. Par ailleurs, on a pour tout  $\mathbf{y} \in \mathcal{E}_n$ ,

$$\mathbb{P}(Y = \mathbf{y} | Y \in \mathcal{E}_n) = \frac{1}{\mathbb{P}(Y \in \mathcal{E}_n)} \cdot \frac{1}{2n+1} \mathbb{E} \left[ \sum_{i=0}^{2n} \mathbf{1}_{\{Y^{(i)} = \mathbf{y}\}} \right] = \mathbb{P}(Y^{(n_*)} = \mathbf{y}).$$

Le Théorème 11 suit facilement du théorème 10 et du lemme 13. En effet, on définit, sur l'espace  $\mathcal{C}$ , l'application  $f \mapsto Vf$  dite transformation de Vervaat, définie comme suit. Soit  $t_* = \inf \{t \in [0, 1] : f(t_*) = \inf_{[0,1]} f\}$ , et

$$Vf(t) = \begin{cases} f(t+t_*) - f(t_*) & \text{si } t \in [0, 1-t_*] \\ f(t+t_*-1) - f(0) + f(1) - f(t_*) & \text{si } t \in [1-t_*, 1]. \end{cases}$$

Il est alors aisé de voir que  $V: \mathcal{C} \rightarrow \mathcal{C}$  est continue en tous les points  $f$  de  $\mathcal{C}$  pour lesquels  $\inf f$  est atteint uniquement en  $t_*$ . C'est le cas du pont brownien presque sûrement.

**Exercice 3.** Le montrer ! On pourra commencer par montrer que le mouvement brownien atteint son minimum sur  $[0, 1]$  en un unique point, puis obtenir le même résultat pour le pont par un argument d'absolue continuité. On rappelle que la loi de  $\inf_{0 \leq s \leq t} B_s$  admet une densité par rapport à la mesure de Lebesgue, pour tout  $t > 0$ .

On a aisément que  $(S_0, S_1, \dots, S_{2n})$  conditionnée à être positive et par  $S_{2n} = 0$  a même loi que la restriction aux  $2n + 1$  premiers indices de  $(S_0, S_1, \dots, S_{2n+1})$  conditionnée à être positive jusqu'au temps  $2n$ , et à valoir  $-1$  au temps  $2n + 1$ . On a alors que  $S^{(2n+1)}$  sachant que  $(X_1, \dots, X_{2n+1}) \in \mathcal{E}_n$  a même loi que  $VS^{(2n+1)}$  sachant que  $(X_1, \dots, X_{2n+1}) \in \mathcal{P}_n$ . Le théorème 10 implique que cette dernière suite converge en loi vers  $V\mathbf{b}$  où  $\mathbf{b}$  est un pont brownien, en vertu du résultat bien connu suivant :

**Proposition 14.** *La suite de variables aléatoires  $Z_n$  à valeurs dans l'espace topologique  $E$  converge en loi vers  $Z$  si et seulement si l'on a  $\mathbb{E}[f(Z_n)] \rightarrow \mathbb{E}[f(Z)]$  pour toute fonction  $f$  qui est continue en  $\mathbb{P}_Z$ -presque tout point, où  $\mathbb{P}_Z$  est la loi de  $Z$ .*

En dépit de son côté économe, la méthode précédente présente un certain nombre d'inconvénients. Le premier est qu'elle n'est pas la plus robuste : le fait que la marche aléatoire prenne des valeurs entières, et plus précisément que les pas de cette marche soient dans  $\{-1, 0, 1, 2, 3, \dots\}$ , est important. Le second inconvénient est que l'on n'obtient pas ainsi de description de l'excursion brownienne par ses lois marginales de dimension finie. Nous les donnons donc maintenant.

**Proposition 15.** *Les marginales de dimension finie de  $e$  sont données par le fait que  $e_0 = e_1 = 0$  p.s., et pour tout  $0 < t_1 < t_2 < \dots < t_k < 1$ , on a*

$$\frac{\mathbb{P}(e_{t_1} \in dx_1, \dots, e_{t_k} \in dx_k)}{dx_1 \cdots dx_k} = 2\sqrt{2\pi} q_{x_1}(t_1) p_{t_2-t_1}^+(x_1, x_2) \cdots p_{t_k-t_{k-1}}^+(x_{k-1}, x_k) q_{x_k}(1-t_k),$$

où l'on a noté, pour  $t > 0$  et  $x \geq 0$ ,

$$q_x(t) = \frac{x}{t} p_t(x) = \frac{x}{\sqrt{2\pi t^3}} \exp(-x^2/2t),$$

et pour  $t > 0$  et  $x, y \geq 0$ ,

$$p_t^+(x, y) = p_t(x-y) - p_t(x+y).$$

La preuve sera vue en exercice.

### 3 Quelques mots sur les processus discontinus

L'étude précédente de la convergence en loi dans l'espace  $\mathcal{C}$  ne permet pas de traiter les exemples naturels de processus discontinus. Par exemple le processus de Poisson, qui est la limite rééchelonnée naturelle du processus de Bernoulli (processus défini sur  $\mathbb{N}$ , et qui augmente de 1 avec probabilité  $1/n$  à chaque pas) lorsque  $n \rightarrow \infty$ . Ou encore, les processus de Lévy, dont les accroissements sont indépendants et de lois stationnaires, qui sont les analogues discrets des marches aléatoires.

Pour traiter ces exemples, on peut avoir recours à l'espace de Skorokhod, souvent noté  $\mathcal{D}$ , des fonctions sur  $[0, 1]$  qui sont continues à droite et admettent des limites à gauche en tout point (on demande aussi la continuité à gauche en 1). On parle de fonctions « càdlàg ». On munit  $\mathcal{D}$  d'une topologie dite de Skorokhod, pour laquelle une suite  $f_n$  converge vers  $f$  si et seulement si pour tout  $n$ , il existe  $\lambda_n$  une fonction continue strictement croissante de  $[0, 1]$  sur  $[0, 1]$ , telle que

$$\|f_n \circ \lambda_n - f\|_\infty \xrightarrow{n \rightarrow \infty} 0 \quad \text{et} \quad \|\lambda_n - \text{Id}_{[0,1]}\|_\infty \xrightarrow{n \rightarrow \infty} 0.$$

Intuitivement, deux fonctions sont proches pour cette topologie si, en décalant légèrement les sauts de la première, on obtient une fonction uniformément proche de la seconde.

Les sous-ensembles compacts de  $\mathcal{D}$  sont plus compliqués à déterminer que ceux de  $\mathcal{C}$ . Il existe néanmoins des critères de tension très efficaces lorsque l'on considère des processus de Markov, dont le critère d'Aldous est un exemple emblématique. Soit donc  $(X_n)_{n \geq 1}$  une suite de processus de  $\mathcal{D}$ . On note  $\mathcal{T}_n$  l'ensemble des temps d'arrêt associés à  $X_n$ , c'est-à-dire les variables aléatoires  $\tau$  à valeurs dans  $[0, 1]$  telles que pour tout  $t \in [0, 1]$ , l'événement  $\{\tau \leq t\}$  est mesurable par rapport à la tribu  $\sigma(X_n(s), 0 \leq s \leq t)$ .

**Théorème 16.** *Avec les notations précédentes, on suppose que pour tout  $\varepsilon > 0$ ,*

$$\limsup_{n \rightarrow \infty} \sup_{\tau \in \mathcal{T}_n} \sup_{0 \leq \theta \leq \delta} \mathbb{P}(|X_n((\tau + \theta) \wedge 1) - X_n(\tau)| > \varepsilon) \xrightarrow{\delta \downarrow 0} 0.$$

*On suppose également que  $(X_n(t))_{n \geq 0}$  est tendue pour tout  $t \in [0, 1]$ . Alors il existe une sous-suite de  $(X_n)_{n \geq 1}$  qui converge en loi dans l'espace de Skorokhod  $\mathcal{D}$ .*

En utilisant la propriété de Markov des marches aléatoires, le lecteur familier avec les rudiments de chaînes de Markov pourra par exemple déduire de ce critère le théorème de Donsker général sous la forme suivante.

**Théorème 17.** *Soit  $X_1, X_2, \dots$  une suite de variables aléatoires réelles indépendantes et de même loi, telles que  $\mathbb{E}[X_1] = 0$ ,  $\text{Var}(X_1) = 1$ . On pose  $S_0 = 0$ ,  $S_n = X_1 + \dots + X_n$  pour  $n \geq 1$ , et*

$$S_t^{[n]} = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}, \quad t \in [0, 1[, S_1^{[n]} = S_{1-}^{[n]}.$$

*Alors  $S^{[n]}$  converge en loi dans l'espace de Skorokhod vers le mouvement brownien.*

On remarque que, dans cet exemple, la limite elle-même est continue (on parle d'une suite  $\mathcal{C}$ -tendue). Néanmoins, le théorème se généralise avec une preuve très similaire au cas où  $X_1$  n'est plus de variance finie (ni nécessairement intégrable), mais est telle que les moyennes recentrées et renormalisées

$$\frac{X_1 + \dots + X_n - b_n}{a_n}$$

convergent en loi lorsque  $n \rightarrow \infty$  pour des suites  $(a_n)$  et  $(b_n)$  bien choisies. Les lois limites qui apparaissent sont appelées *lois stables*.

## 4 Topologie de Gromov-Hausdorff

### 4.1 Définitions

Soit  $(Z, \delta)$  un espace métrique. Rappelons que la distance de Hausdorff  $\delta_H$  entre deux parties fermées  $A$  et  $B$  de  $Z$  est donnée par

$$\delta_H(A, B) = \inf \{ \varepsilon > 0 : A \subset V_\varepsilon(B) \text{ et } B \subset V_\varepsilon(A) \}$$

où  $V_\varepsilon(A) = \{z \in Z : \delta(z, A) \leq \varepsilon\}$ . Si  $(X, d)$  et  $(X', d')$  sont deux espaces métriques compacts, on définit leur distance de Gromov-Hausdorff comme

$$d_{\text{GH}}((X, d), (X', d')) = \inf \{ \delta_H(\varphi(X), \varphi'(X')) \}$$

où l'infimum est pris sur tous les espaces métriques  $(Z, \delta)$  et tous les plongements isométriques  $\varphi: X \rightarrow Z$  et  $\varphi': X' \rightarrow Z$ . En particulier, si  $(X, d)$  et  $(X', d')$  sont isométriques, on voit immédiatement que leur distance Gromov-Hausdorff est nulle (prendre  $(Z, \delta) = (X, d)$ ,  $\varphi$  l'identité de  $X$  et  $\varphi'$  une isométrie de  $X'$  sur  $X$ ). Plus généralement,  $d_{\text{GH}}((X, d), (X', d'))$  est une fonction des classes d'isométries des espaces métriques  $(X, d)$  et  $(X', d')$ , c'est-à-dire de ces espaces vus à isométrie près. Souvent, on note  $d_{\text{GH}}(X, X')$  au lieu de  $d_{\text{GH}}((X, d), (X', d'))$ , les distances étant considérées implicites.

**Théorème 18.** *Soit  $\mathbb{M}$  l'ensemble des classes d'isométries d'espaces métriques compacts. Alors  $(\mathbb{M}, d_{\text{GH}})$  est un espace métrique séparable complet.*

L'infimum apparaissant dans la définition est pris sur un ensemble gigantesque. En fait, pour être rigoureux, il ne s'agit même pas à proprement parler d'un ensemble, puisque la classe de tous les espaces métriques n'est pas un ensemble ! Ce problème peut être néanmoins contourné en se restreignant aux classes d'isométries d'espaces métriques compacts, qui elles forment bien un ensemble. Dans tous les cas, la définition ne donne pas une méthode très concrète pour évaluer la distance de Gromov-Hausdorff. On peut avoir recours à la définition alternative suivante.

Si  $X$  et  $X'$  sont deux ensembles, on dit que le sous-ensemble  $C \subset X \times X'$  est une correspondance si pour tout  $x \in X$  il existe  $x' \in X'$  tel que  $(x, x') \in C$ , et pour tout  $x' \in X'$  il existe un  $x \in X$  tel que  $(x, x') \in C$ . Si on se donne une telle correspondance, et que  $X, X'$  sont munis de deux métriques  $d$  et  $d'$ , la distortion de  $C$  par rapport à  $d, d'$  est la quantité

$$\text{dis}(C) = \sup \{|d(x, y) - d'(x', y')| : (x, x'), (y, y') \in C\}.$$

En général, on ne mentionne pas les distances  $d, d'$  dans les définitions.

**Proposition 19.** *Soit  $(X, d), (X', d')$  deux espaces métriques compacts. Alors*

$$d_{\text{GH}}(X, X') = \frac{1}{2} \inf_{C \in \mathcal{C}(X, X')} \text{dis}(C), \quad (5)$$

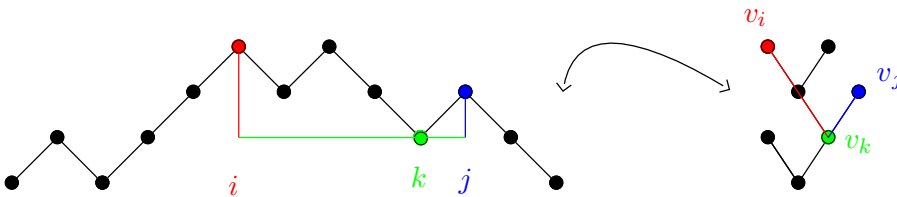
où  $\mathcal{C}(X, X')$  est l'ensemble des correspondances entre  $X$  et  $X'$ .

Il est en fait bien plus facile de montrer que  $d_{\text{GH}}$  est une distance en utilisant cette dernière définition. On a par ailleurs :

**Proposition 20.** *L'espace  $(\mathbb{M}, d_{\text{GH}})$  est séparable et complet.*

## 4.2 Application à la convergence d'arbres aléatoires

Nous donnons maintenant un exemple important d'utilisation de la topologie de Gromov-Hausdorff et de ses liens avec les théorèmes limites fonctionnels discutés précédemment. Soit  $T_n$  un arbre plan uniforme aléatoire enraciné à  $n$  arêtes. Ce dernier est naturellement codé par un chemin de Dyck de longueur  $2n$ , selon un procédé bien connu consistant à faire le tour de l'arbre de gauche à droite, en commençant par la racine : au temps  $i \in \{0, 1, 2, \dots, 2n\}$ , on note  $v_i$  le sommet de l'arbre qui est visité (chaque sommet est visité autant de fois que son degré, sauf la racine qui est visitée une fois de plus).



Il est clair que ce chemin de Dyck aléatoire, que l'on note  $C_n$ , a la même loi que l'interpolation linéaire de la marche aléatoire simple  $(S_t, 0 \leq t \leq 2n)$  conditionnée à être positive et à valoir 0 en  $2n$ . Par ailleurs, la distance de graphe entre  $v_i$  et  $v_j$  dans  $T_n$  est donnée explicitement par la formule

$$d_{T_n}(v_i, v_j) = d_{C_n}(i, j),$$

où l'on a noté, pour une fonction  $f: [0, \sigma] \rightarrow \mathbb{R}_+$  continue définie sur un intervalle compact, et telle que  $f(0) = f(\sigma) = 0$ ,

$$d_f(s, t) = f(s) + f(t) - 2\check{f}(s, t),$$

où enfin  $\check{f}(s, t) = \min \{f(u) : s \wedge t \leq u \leq s \vee t\}$ . Sur le dessin, la quantité  $d_{C_n}(i, j)$  correspond à la somme des longueurs des segments rouge et bleus, et un instant  $k$  où le minimum définissant  $\check{C}_n(i, j)$  est réalisé est tel que  $v_k$  est le plus récent ancêtre commun à  $v_i$  et  $v_j$ .

La fonction  $d_f$  définie sur  $[0, \sigma]^2$  est une pseudo-distance : elle vérifie les mêmes conditions qu'une distance à l'exception du fait qu'elle ne sépare pas les points. Par exemple, il est toujours vrai que  $d_f(0, \sigma) = 0$ , même si  $0 \neq \sigma$  sauf cas triviaux.

On définit alors un vrai espace métrique en prenant le quotient  $\mathcal{T}_f$  de  $[0, \sigma]$  par la relation d'équivalence  $\{(s, t) \in [0, \sigma]^2 : d_f(s, t) = 0\} = \{d_f = 0\}$ . Si on la munit de la distance induite par  $d_f$ , que l'on note encore  $d_f$ , on a que  $(\mathcal{T}_f, d_f)$  est un espace métrique compact. C'est même un  $\mathbb{R}$ -arbre, c'est-à-dire

- qu'il est géodésique, c'est-à-dire que toute paire de points  $x$  et  $y$  peut être reliée par un plongement isométrique  $[0, d_f(x, y)] \rightarrow \mathcal{T}_f$ , et
- qu'il n'admet pas de cycle, c'est-à-dire de fonction continue injective du cercle  $\mathbb{S}^1 \rightarrow \mathcal{T}_f$ .

Les  $\mathbb{R}$ -arbres constituent la généralisation naturelle des arbres combinatoires, du moins si l'on veut voir les arbres comme des espaces métriques (ce qui peut avoir pour conséquence de perdre un peu de leur structure naturelle).

**Proposition 21.** *Soit  $f$  et  $f'$  deux fonctions continues de  $[0, 1]$  dans  $\mathbb{R}_+$ , telles que  $f(0) = f(1) = f'(0) = f'(1) = 0$ . Alors*

$$d_{\text{GH}}(\mathcal{T}_f, \mathcal{T}_{f'}) \leq 2 \|f - f'\|_{\infty}.$$

**Preuve.** On construit une correspondance entre  $\mathcal{T}_f$  et  $\mathcal{T}_{f'}$  en notant  $p, p'$  les projections canoniques de  $[0, 1]$  sur  $\mathcal{T}_f, \mathcal{T}_{f'}$ , et en posant  $C = \{(p(t), p'(t)) : t \in [0, 1]\}$ . Il nous faut évaluer sa distortion, qui vaut clairement

$$\text{dis}(C) = \sup_{s, t \in [0, 1]} |f(s) - f'(s) + f(t) - f'(t) - 2(\check{f}(s, t) - \check{f}'(s, t))| \leq 4 \|f - f'\|_{\infty}.$$

On conclut par la proposition 19.

**Corollaire 22.** *Soit  $T_n$  un arbre plan uniforme enraciné à  $n$  arêtes, et  $V(T_n)$  l'ensemble de ses sommets. Alors on a la convergence en loi pour la topologie de Gromov-Hausdorff :*

$$\left( V(T_n), \frac{d_{T_n}}{\sqrt{2n}} \right) \xrightarrow[n \rightarrow \infty]{\text{loi}} (\mathcal{T}_e, d_e)$$

où  $e$  est une excursion brownienne normalisée.

**Preuve.** Tout d'abord, on constate que  $d_{\text{GH}}((V(T_n), d_{T_n}), (\mathcal{T}_{C_n}, d_{C_n})) \leq 1/2$ , comme on peut s'en apercevoir en prenant la correspondance  $R = \{(v_{\lfloor 2ns \rfloor}, p(s)) : s \in [0, 1]\}$ , où l'on rappelle que  $v_i$  est le  $i$ -ème sommet de  $T_n$  visité par ordre de contour, et  $p: [0, 1] \rightarrow \mathcal{T}_{C_n}$  est la projection canonique. Il est facile de voir que  $\text{dis}(R) = 1$ . Ensuite, on constate que, si  $C_{(n)}(s) = C_n(2ns)/\sqrt{2n}$ , alors

$$\left( \mathcal{T}_{C_n}, \frac{d_{C_n}}{\sqrt{2n}} \right) \stackrel{\text{isom}}{=} (\mathcal{T}_{C_{(n)}}, d_{C_{(n)}}).$$

On conclut alors en utilisant la convergence en loi  $C_{(n)} \Rightarrow e$  et la proposition 21, qui implique en particulier que  $f \mapsto (\mathcal{T}_f, d_f)$  est continue de  $\{f \in \mathcal{C} : f \geq 0, f(0) = f(1) = 0\}$  vers  $(\mathbb{M}, d_{\text{GH}})$ .

Le  $\mathbb{R}$ -arbre aléatoire limite  $(\mathcal{T}_e, d_e)$  est appelé l'arbre continu aléatoire d'Aldous. Il intervient dans de très nombreux problèmes asymptotiques impliquant des structures combinatoires aléatoires (arbres, fonction de hachage, processus de fragmentation, cartes aléatoires, laminations, dissections, etc.).

## 5 Quelques références

### Théorèmes limites fonctionnels

- P. Billingsley, *Convergence of probability measures*, 2ème édition, Wiley (1999).

Une des références les plus agréables à lire, et déjà très complète.

- S. Ethier & T. Kurtz, *Markov processes. Characterization and convergence*, Wiley, 1986.
- J. Jacod, A. Shiryaev, *Limit theorems for stochastic processes*, 2ème édition, Springer, 2003.

Ces deux dernières références se focalisent sur des objets relativement différents (processus de Markov pour la première, semimartingales pour la seconde), et sont assez exigeantes, mais très complètes.

Pour une introduction plus concise à ces questions, et d'autres références, on pourra aussi consulter des notes de cours de M2 disponibles sur ma page web <http://perso.ens-lyon.fr/gregory.miermont/thlim.pdf>.

Le théorème de convergence d'une marche aléatoire à pas i.i.d. de variance finie est dû à Donsker (1951). Une version conditionnée pour les excursions apparaît dans Kaigh (1976). La preuve proposée ici utilise l'observation due à Vervaat (1979) que pont et excursion sont liés par une permutation cyclique des accroissements, ce qui est vrai aussi pour des marches aléatoires à pas entiers supérieurs à  $-1$ .

### Utilisation en probabilités de la topologie de Gromov-Hausdorff

La topologie de Gromov-Hausdorff a été introduite par Gromov dans le but d'étudier des structures géométriques moins régulières que les variétés riemanniennes. On pourra consulter son foisonnant ouvrage (dont la première version française, trois fois plus courte, remonte à 1981).

- M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, réimprimée de l'édition de 2001, Birkhäuser, 2007

Une autre référence utile est l'ouvrage suivant

- D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry*, AMS, 2001.

L'introduction de la topologie de Gromov-Hausdorff en probabilité a été popularisée par Evans d'une part (on pourra par exemple trouver, dans un papier de Evans, Pitman et Winter de PTRF de 2006, une preuve de la complétude de la distance  $d_{GH}$ , dans le cas particulier non-restrictif des  $\mathbb{R}$ -arbres), et Duquesne et Le Gall d'autre part dans le contexte des processus de branchement (PTRF 2005). L'idée de coder la limite d'échelle des arbres généalogiques par des processus remonte à Aldous (1993). On pourra consulter les articles d'exposition de Le Gall disponibles sur sa page web. En particulier

- J.-F. Le Gall, Random trees and applications, *Probab. Surveys.*, 2005
- J.-F. Le Gall, G. Miermont, Scaling limits of random trees and planar maps, Clay Math. Proc., AMS, 2012

La dernière référence s'intéresse plus particulièrement à la convergence des cartes aléatoires. Enfin, pour l'approche générale consistant à coder des objets combinatoires par des processus, qui contient beaucoup des idées exposées ici on pourra consulter les notes de cours de Saint-Flour de Pitman :

- J. Pitman, *Combinatorial stochastic processes*, Springer, 2006.

## 6 Exercices

### 1. Processus croissants

Soit  $(X_n, n \geq 0)$  et  $X$  des processus croissants, continus, de  $[0, 1]$  dans  $\mathbb{R}$ . Montrer que  $X_n$  converge en loi vers  $X$  (noté  $X_n \Rightarrow X$ ) dans  $\mathcal{C}$  si et seulement si  $X_n$  converge vers  $X$  au sens des marginales de dimension finie.

On pourra fixer  $\varepsilon > 0$  et chercher à majorer la probabilité

$$\mathbb{P}(\omega(X_n, \delta) \geq \varepsilon)$$

et on utilisera que, si  $Y_n \Rightarrow Y$  alors pour tout ensemble fermé  $F$ ,

$$\limsup_{n \rightarrow \infty} \mathbb{P}(Y_n \in F) \leq \mathbb{P}(Y \in F).$$

### 2. Aire sous l'excursion

Soit  $T_n$  un arbre plan uniforme enraciné à  $n$  arêtes, et  $C_n$  son processus de contour. Comme dans le cours, on note  $C_{(n)}(t) = C_n(2nt)/\sqrt{2n}$ . Montrer que

$$\int_0^1 C_{(n)}(t) dt \Rightarrow \int_0^1 e(t) dt,$$

où  $e$  est l'excursion brownienne normalisée.

On note  $W(T_n)$  la somme des distances des sommets de  $T_n$  à la racine. Montrer que lorsque  $n \rightarrow \infty$ ,

$$\frac{W(T_n)}{\sqrt{2n^3}} \Rightarrow \int_0^1 e(t) dt.$$

### 3. Un brownien peut en cacher un autre

En revenant à la définition du mouvement brownien par ses marginales de dimension finie, ou en utilisant le théorème de Donsker, montrer que si  $(B_t, t \geq 0)$  est un mouvement brownien standard dans  $\mathbb{R}^d$ , alors pour tout  $\lambda > 0$ , le processus

$$\left( \frac{1}{\sqrt{\lambda}} B_{\lambda t}, t \geq 0 \right)$$

est aussi un mouvement brownien standard. Montrer également que pour tout  $t > 0$  fixé, le processus

$$B^{(t)} = (B_{t+s} - B_t, s \geq 0)$$

est un mouvement brownien standard, et que ce processus est indépendant du processus  $(B_u, 0 \leq u \leq t)$  (propriété de Markov simple).

Posons, pour tout  $c \in ]0, 1[$ ,

$$\beta_t^{(c)} = \frac{1}{\sqrt{c}} B_{ct}, \quad 0 \leq t \leq 1.$$

(i) Montrer que la famille  $(\beta^{(c)}, c > 0)$  de variables aléatoires à valeurs dans  $\mathcal{C}$  est uniformément tendue.

(ii) Montrer l'égalité en loi

$$((B_{c+t}, 0 \leq t \leq 1), \beta^{(c)}) \stackrel{\text{loi}}{=} ((\sqrt{c} \beta_1' + B_t, 0 \leq t \leq 1), \beta'),$$



où  $\beta'$  est un processus indépendant de  $B$ , et de même loi.

(iii) On se donne des réels positifs  $0 < t_1 < t_2 < \dots < t_k \leq 1$  et  $0 < t'_1 < \dots < t'_{k'} \leq 1$ . Montrer la convergence en loi

$$(B_{t_1}, \dots, B_{t_k}, \beta_{t'_1}^{(c)}, \dots, \beta_{t'_{k'}}^{(c)}) \xrightarrow[c \downarrow 0]{\text{loi}} (B_{t_1}, \dots, B_{t_k}, \beta_{t'_1}', \dots, \beta_{t'_{k'}}'),$$

avec les mêmes notations que dans la question précédente. Montrer que cela reste vrai si  $t_1 = 0$  ou  $t'_1 = 0$ .

(iv) En conclure, toujours avec les mêmes notations, que  $(B, \beta^{(c)})$  converge en loi dans  $\mathcal{C}$  vers  $(B, \beta')$  lorsque  $c \rightarrow 0$ . Expliquer pourquoi ce résultat peut paraître surprenant

#### 4. Plus sur le lemme cyclique

Soit  $\mathbf{y} = (y_1, \dots, y_n)$  une suite de  $\{-1, 1\}^n$  telle que  $y_1 + \dots + y_n = -k$ . Montrer qu'il existe exactement  $k$  indices  $n_1, \dots, n_k$  distincts dans  $\{0, 1, \dots, n-1\}$  tels que  $\mathbf{y}^{(n_1)}, \dots, \mathbf{y}^{(n_k)}$  sont dans l'ensemble

$$\mathcal{E}_{n,k} = \{(y_1, \dots, y_n) \in \{-1, 1\}^n : y_1 + \dots + y_n = -k, \forall i \in \{0, 1, \dots, n-1\}, y_1 + \dots + y_i > -k\}.$$

En déduire que si  $Y = (Y_1, \dots, Y_n)$  sont des variables aléatoires cycliquement échangeables (c'est-à-dire que  $Y$  a même loi que  $Y^{(i)}$  pour tout  $i$ , avec les notations du cours), alors

$$\mathbb{P}(Y \in \mathcal{E}_{n,k}) = \frac{k}{n} \mathbb{P}(Y_1 + \dots + Y_n = -k).$$

#### 5. Marginales de dimension finie de l'excursion brownienne

Soit  $X_1, X_2, \dots$ , une suite de variables aléatoires i.i.d. uniformes dans  $\{-1, 1\}$ , et  $S_n = X_1 + \dots + X_n$ . On rappelle que  $P_n(k) = \mathbb{P}(S_n = k)$ .

(i) On pose, pour  $k, k' \in \mathbb{N}$  et  $n > 0$ ,

$$P_n^+(k, k') = \mathbb{P}(k + S_n = k'; \forall m \in \{0, 1, \dots, n-1\}, k + S_m \geq 0).$$

Montrer que  $P_n^+(k, k') = P_n(k' - k) - P_n(k' + k + 2)$  en utilisant le principe de réflexion, qui donne une bijection naturelle entre les chemins sur réseau reliant  $k$  à  $k'$  en  $n$  pas et ayant au moins une valeur strictement négative, et les chemins reliant  $k$  à  $-k' - 2$  en  $n$  pas.

(ii) On pose  $Q_k(n) = \mathbb{P}(S_n = k; \forall m \in \{0, 1, \dots, n\}, S_m \geq 0)$ . Montrer à l'aide du lemme cyclique que

$$Q_k(n) = 2\mathbb{P}((X_1, \dots, X_n, -1) \in \mathcal{E}_{n+1, k+1}) = 2 \cdot \frac{k+1}{n+1} P_{n+1}(-k-1).$$

(iii) Montrer que l'on a que  $\mathbb{P}(S_{2n_1} = 2r_1, \dots, S_{2n_k} = 2r_k \mid (X_1, \dots, X_{2n}) \in \mathcal{E}_n)$  égale

$$\frac{Q_{2r_1}(2n_1) P_{2(n_2 - n_1)}^+(2r_1, 2r_2) \cdots P_{2(n_k - n_{k-1})}^+(2r_{k-1}, 2r_k) Q_{2r_k}(2(n - n_k))}{2Q_1(2n+1)}$$

(iv) En déduire les marginales de dimension finie de  $e$ .

#### 6. Un calcul simple de distance de Gromov-Hausdorff

Soit  $(X, d)$  un espace métrique compact, et  $\epsilon$  l'espace métrique constitué d'un unique point. Calculer  $d_{\text{GH}}(X, \epsilon)$ .