

---

–Texte–

## Agrégation limitée par diffusion interne

---

### 1 Le phénomène observé

Un fût de déchets radioactifs est enterré secrètement dans le Cantal. Au bout de quelques années, il devient poreux et laisse échapper son contenu. Pour éviter une contamination excessive, on a disposé des pièges à particules qui capturent la première particule qui passe dessus mais deviennent ensuite inertes (une nouvelle particule passe dessus sans être arrêtée). On suppose que le milieu est isotrope : une particule radioactive se déplace de la même manière dans toutes les directions. Cette particule continuera à se déplacer tant qu'elle passe sur des pièges qui ont déjà été activés et est capturée par le premier piège libre qu'elle rencontre. On souhaite connaître la forme typique des zones qui seront contaminées par cette fuite.

Pour modéliser l'évolution de la région polluée, on introduit le modèle de croissance suivant. On assimile l'espace à  $\mathbb{Z}^d$  où  $d$  vaut 1, 2 ou 3 selon le nombre de dimensions que l'on souhaite prendre en compte. Deux éléments  $x$  et  $y$  de  $\mathbb{Z}^d$  sont dits voisins, et nous noterons  $x \sim y$ , si

$$|x - y|_1 = |x_1 - y_1| + \cdots + |x_d - y_d| = 1.$$

On suppose le fût placé à l'origine, que l'on note 0. Sur chaque point du réseau est disposé un piège. Notons  $A_0$  le singleton  $\{0\}$ . Quel sera le premier piège activé ? Il s'agit de l'un des  $2d$  emplacements voisins de l'origine.

On modélise donc la trajectoire d'une particule radioactive sortant du fût par une marche aléatoire simple symétrique sur  $\mathbb{Z}^d$  issue de 0 et arrêtée lorsqu'elle sort de  $A_0$ . On note  $A_1$  l'ensemble aléatoire composé de 0 et du point où la particule est sortie. Il est évident que, pour tout  $x$  voisin de 0,  $A_1$  est égal à  $\{0, x\}$  avec probabilité  $1/2d$ . On itère ensuite le procédé. Étant donné un ensemble  $A_n \subset \mathbb{Z}^d$ , on considère une marche aléatoire symétrique  $(S_k)_{k \geq 0}$  issue de 0 arrêtée lorsqu'elle sort de  $A_n$ . On définit alors  $A_{n+1}$  comme l'ensemble des éléments de  $A_n$  et du point où est sortie la marche  $S$ .

On obtient ainsi une suite  $(A_n)_{n \geq 0}$  d'ensembles aléatoires. La proposition suivante regroupe quelques propriétés évidentes de cette suite.

**Proposition 1.1.** *La suite  $(A_n)_{n \geq 0}$  est croissante au sens de l'inclusion. Le cardinal de  $A_n$  vaut exactement  $n+1$ . Pour tout  $n \in \mathbb{N}$ , soit  $x$  et  $y$  deux éléments de  $A_n$ , alors il existe  $x_0, \dots, x_m$  éléments de  $A_n$  tels que  $x_0 = x$ ,  $x_m = y$  et  $x_i \sim x_{i+1}$  pour  $i = 0, \dots, m-1$ . On pourrait dire que l'ensemble  $A_n$  est connexe par arc dans  $\mathbb{Z}^d$ .*

On souhaite à présent étudier le comportement asymptotique de la suite  $(A_n)_n$ . On étudie en détail le cas de la dimension 1. Un résultat en dimension 2 est donné à la fin du texte.

## 2 Le modèle unidimensionnel

On suppose que la propagation se fait selon un axe horizontal. On se place donc dans le cas où  $d = 1$ . L'ensemble initial  $A_0$  est le singleton  $\{0\}$  puis  $A_1$  est égal à  $\{0, 1\}$  avec probabilité  $1/2$  et  $\{-1, 0\}$  avec probabilité  $1/2$  etc. Notons  $G_n = \min A_n$  et  $D_n = \max A_n$ . L'ensemble  $A_n$  est de la forme  $A_n = \{G_n, G_n + 1, \dots, D_n - 1, D_n\}$ . Puisque le cardinal de  $A_n$  est  $n + 1$ , on a  $D_n - G_n = n$ . Ainsi,  $A_n$  est caractérisé par  $X_n = D_n + G_n$  et, en particulier,

$$D_n = \frac{X_n + n}{2} \quad \text{et} \quad G_n = \frac{X_n - n}{2}.$$

Les accroissements  $(X_{n+1} - X_n)_{n \geq 0}$  ne peuvent prendre que les valeurs  $-1$  ou  $1$ . Plus précisément,  $X_{n+1} = X_n - 1$  si la marche issue de  $0$  atteint  $G_n - 1$  avant  $D_n + 1$ , et  $X_{n+1} = X_n + 1$  si la marche issue de  $0$  atteint  $D_n + 1$  avant  $G_n - 1$ . On peut obtenir la loi de cette suite de variables aléatoires en s'appuyant sur un résultat très classique rappelé ici.

**Proposition 2.1.** *Soit  $S$  une marche aléatoire simple sur  $\mathbb{Z}$  issue de  $0$  et soit  $a$  et  $b$  deux entiers distincts, le premier négatif, le second positif. On note  $T_i = \inf \{n \geq 0, S_n = i\}$  pour  $i = a, b$  et  $T = T_a \wedge T_b$ . Alors,  $T_a$  et  $T_b$  sont finis presque sûrement,  $T$  est intégrable et*

$$\mathbb{P}(T = T_a) = 1 - \mathbb{P}(T = T_b) = \frac{b}{b - a} \quad \text{et} \quad \mathbb{E}(T) = -ab.$$

*Démonstration.* C'est le problème classique de la ruine du joueur. □

La proposition ci-dessous montre que la suite  $(X_n)_{n \geq 0}$  est une chaîne de Markov inhomogène sur  $\mathbb{N}$ , c'est-à-dire une chaîne de Markov dont la matrice de transition dépend du temps.

**Proposition 2.2.** *La suite  $(X_n)_{n \geq 0}$  est une chaîne de Markov inhomogène sur  $\mathbb{N}$  issue de  $0$  dont les transitions sont décrites par les relations suivantes : pour  $i \leq n$ ,*

$$\mathbb{P}(X_{n+1} = i - 1 | X_n = i) = \frac{n + 2 + i}{2(n + 2)} \quad \text{et} \quad \mathbb{P}(X_{n+1} = i + 1 | X_n = i) = \frac{n + 2 - i}{2(n + 2)}.$$

*Démonstration.* Si  $X_n = i$  alors  $D_n = (i + n)/2$  et  $G_n = (i - n)/2$ . La probabilité  $\mathbb{P}(X_{n+1} = i - 1 | X_n = i)$  est donc égale à la probabilité que la marche aléatoire simple issue de  $0$  atteigne  $(i - n)/2 - 1$  avant  $(i + n)/2 + 1$ . □

*Remarque 2.3.* Lorsque  $X_n$  est strictement positif,  $X_{n+1} - X_n$  a une probabilité plus forte de valoir  $-1$  que de valoir  $+1$ . La tendance s'inverse lorsque  $X_n$  est strictement négatif. On peut donc penser que le processus aura davantage tendance à revenir en  $0$  qu'une marche aléatoire simple symétrique (dont l'accroissement vaut plus ou moins  $1$  avec probabilité  $1/2$ ).

### 3 La convergence presque sûre

On établit ici le premier résultat (relativement intuitif) de convergence en temps long pour le processus  $X$ .

**Proposition 3.1.**

$$\frac{X_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

L'idée de la preuve est de rendre rigoureuse l'intuition de la remarque 2.3 en comparant (de manière déterministe) une trajectoire de  $X$  et une trajectoire de la marche aléatoire simple. La proposition 3.1 est un corollaire immédiat du théorème suivant.

**Théorème 3.2.** *Le processus  $(|X_n|)_{n \geq 0}$  est une chaîne de Markov inhomogène et  $(|X_n|/n)_n$  converge vers  $0$  presque sûrement.*

*Démonstration.* En discutant les valeurs de  $X_n$ , on a facilement que  $|X_{n+1}| = 1$  si  $|X_n| = 0$  et, que, si  $|X_n| > 0$ ,

$$|X_{n+1}| = \begin{cases} |X_n| - 1 & \text{avec probabilité } \frac{1}{2} + \frac{|X_n|}{2(n+2)}, \\ |X_n| + 1 & \text{avec probabilité } \frac{1}{2} - \frac{|X_n|}{2(n+2)}. \end{cases}$$

On va construire à présent deux processus  $(Y_n)_{n \geq 0}$  et  $(Z_n)_{n \geq 0}$  dont les lois respectives sont celles de  $(|X_n|)_n$  et celle de la valeur absolue d'une marche aléatoire simple issue de  $0$  tels que, presque sûrement, pour tout  $n \in \mathbb{N}$ ,  $0 \leq Y_n \leq Z_n$ . C'est ce que l'on appelle le couplage. Pour cela, on se donne une suite  $(U_n)_{n \geq 0}$  de variables aléatoires indépendantes et identiquement distribuées de loi uniforme sur  $[0, 1]$  et on pose :  $Y_0 = Z_0 = 0$  et pour  $n \geq 0$ ,

$$Y_{n+1} = \begin{cases} Y_n + 2\mathbf{1}_{\{U_n < (n+s+Y_n)/2(n+s)\}} - 1 & \text{si } Y_n > 0, \\ 1 & \text{si } Y_n = 0, \end{cases}$$

et

$$Z_{n+1} = \begin{cases} Z_n + 2\mathbf{1}_{\{U_n < 1/2\}} - 1 & \text{si } Z_n > 0, \\ 1 & \text{si } Z_n = 0. \end{cases}$$

Il est clair que les deux processus ci-dessus ont bien les lois annoncées et que  $Y_0 \leq Z_0$ . De plus, si  $0 < Y_n \leq Z_n$  alors, par construction,  $Y_{n+1} \leq Z_{n+1}$ . Enfin, si  $0 = Y_n \leq Z_n$  alors,

$Z_n$  peut être nul et dans ce cas,  $Y_{n+1} \leq Z_{n+1}$  ou  $Z_n \geq 2$  car  $Z_n$  et  $Y_n$  ont même parité. Ceci assure le résultat. Enfin,  $(Z_n/n)_n$  converge presque sûrement vers 0 en vertu de la loi forte des grands nombres.  $\square$

## 4 La convergence en loi

On s'intéresse à présent aux fluctuations de  $X_n/n$  autour de sa limite. L'objet de cette section est de donner des éléments de preuve du résultat suivant.

**Proposition 4.1.**

$$\frac{X_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1/3).$$

*Remarque 4.2.* La vitesse de convergence en loi est la même que pour la marche aléatoire simple et la loi limite est encore gaussienne. Pourtant sa variance est plus petite que dans le cas identiquement distribué. Cela provient de la tendance auto-stabilisatrice de chaîne  $X$ .

Avant d'attaquer toute preuve de ce résultat, on peut se poser la question du comportement des premiers moments de  $X_n/n$  quand  $n$  tend vers l'infini. On remarque tout d'abord que  $X_n$  est centrée. En fait, tous les moments impairs de  $X_n$  sont centrés. Que se passe-t-il pour le moment d'ordre 2? Pour  $n \geq 0$ , on a

$$\begin{aligned} \mathbb{E}[X_{n+1}^2 | \mathcal{F}_n] &= \frac{n+s-X_n}{2(n+s)}(X_n+1)^2 + \frac{n+s+X_n}{2(n+s)}(X_n-1)^2 \\ &= \frac{1}{2}[(X_n+1)^2 + (X_n-1)^2] - \frac{X_n}{2(n+s)}[(X_n+1)^2 - (X_n-1)^2] \\ &= 1 + \frac{n+s-2}{n+s}X_n^2. \end{aligned}$$

Notons  $x_n(2) = \mathbb{E}(X_n^2)$ . En prenant l'espérance dans la relation ci-dessus, on obtient la relation suivante :

$$x_{n+1}(2) = 1 + \frac{n+s-2}{n+s}x_n(2).$$

Une simple récurrence permet d'écrire

$$x_n(2) = 1 + \frac{1}{(n+s)(n+s-1)} \sum_{k=1}^{n-1} (k+s)(k+s-1) + \frac{s(s-1)}{(n+s)(n+s-1)}x_0(2)$$

et de conclure que  $(x_n(2)/n)_{n \geq 0}$  converge vers  $1/3$ .

On pourrait de même montrer que, pour tout  $k \geq 0$ , la suite  $(x_n(2k)/n^k)_{n \geq 1}$  admet une limite  $\mu(2k)$  quand  $n$  tend vers l'infini et que la suite  $(\mu(2k))_{k \in \mathbb{N}}$  est solution de

$$\mu(0) = 1 \quad \text{et} \quad \mu(2k+2) = \frac{2k+1}{3}\mu(2k). \quad (1)$$

**Proposition 4.3.** *La seule mesure dont les moments impairs sont nuls et les moments pairs vérifient (1) est la loi  $\mathcal{N}(0, 1/3)$ .*

Un peu plus de travail (et un contrôle plus fin) permettrait de déduire de cette convergence des moments de  $X_n/\sqrt{n}$  la convergence en loi une mesure gaussienne centrée de variance  $1/3$ .

Un autre moyen d'établir la convergence annoncé est d'utiliser un raisonnement basé sur le théorème limite central pour des martingales de carré intégrable.

**Proposition 4.4.** *Le processus  $(M_n)_{n \geq 0}$  défini par  $M_n = (n+1)X_n$  est une martingale de carré intégrable qui vérifie*

$$\mathbb{E}[(M_{n+1} - M_n)^2 | \mathcal{F}_n] = (n+2)^2 - X_n^2.$$

On associe à  $M$  le processus  $(\langle M \rangle_n)_n$  défini par

$$\langle M \rangle_0 = 0 \quad \text{et} \quad \langle M \rangle_{n+1} = \langle M \rangle_n + \mathbb{E}[(M_{n+1} - M_n)^2 | \mathcal{F}_n].$$

**Proposition 4.5.** *Le processus  $\langle M \rangle$  possède les propriétés suivantes :*

- il est croissant,
- il est prévisible :  $\langle M \rangle_n$  est  $\mathcal{F}_{n-1}$  mesurable,
- $(M_n^2 - \langle M \rangle_n)_n$  est une martingale,
- il vérifie

$$\frac{\langle M \rangle_n}{n^3} \xrightarrow[n \rightarrow \infty]{p.s.} \frac{1}{3}.$$

*Démonstration.* Les trois premiers points sont évidents. Pour le troisième, on remarque que

$$\langle M \rangle_n = \frac{n^3}{3} - \sum_{k=0}^{n-1} X_k^2 + o(n^3).$$

De plus

$$\frac{1}{n^3} \sum_{k=1}^{n-1} X_k^2 \leq \frac{1}{n} \sum_{k=1}^{n-1} \frac{X_k^2}{k^2}.$$

Puisque  $X_n/n$  converge vers 0 presque sûrement, le lemme de Cesaro fournit le resultat annoncé.  $\square$

Rappelons le théorème limite central pour les martingales de carré intégrable (que l'on admettra).

**Théorème 4.6.** *Soit  $(M_n)_n$  une martingale de carré intégrable et soit  $(a_n)_n$  une suite réelle, positive, déterministe croissante vers l'infini. On pose  $\Delta M_n := M_n - M_{n-1}$  et on suppose que*

1. Il existe  $\lambda \geq 0$  déterministe tel que  $a_n^{-1} \langle M \rangle_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \lambda$ .
2. Pour tout  $\varepsilon > 0$ ,  $\frac{1}{a_n} \sum_{k=1}^n \mathbb{E}(\Delta M_k^2 \mathbb{I}_{\{|\Delta M_k| \geq \varepsilon \sqrt{a_n}\}} | \mathcal{F}_{k-1}) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0$  (condition de Lindeberg).

Alors, on a

$$\frac{1}{\sqrt{a_n}} M_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \lambda) \quad \text{et si } \lambda > 0 \quad \sqrt{a_n} \frac{M_n}{\langle M \rangle_n} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \lambda^{-1}).$$

Ce résultat permet d'obtenir la convergence en loi de la proposition 4.1. On ne cherchera pas à démontrer que la condition de Lindeberg est satisfaite dans le modèle présent.

## 5 Le modèle d'agrégation limitée en dimension deux

On se place à présent sur  $\mathbb{Z}^2$ . Comme en dimension 1, on imagine que la forme limite de l'ensemble aléatoire doit être relativement régulière puisque la marche aléatoire issue de l'origine aura « tendance » à atteindre l'ensemble aléatoire  $A_n^c$  en des points qui ont beaucoup de voisins dans  $A_n$ . Le théorème suivant montre que c'est effectivement le cas : l'ensemble  $A_n$  ressemble à la boule euclidienne de volume comparable.

**Théorème 5.1.** *Pour tout  $\varepsilon > 0$ , avec probabilité 1, il existe  $n_0$  tel que, pour tout  $n \geq n_0$ ,*

$$B\left(0, (1 - \varepsilon)\sqrt{\frac{n}{\pi}}\right) \subset A_n \subset B\left(0, (1 + \varepsilon)\sqrt{\frac{n}{\pi}}\right),$$

où  $B(0, r)$  désigne le disque dans  $\mathbb{R}^2$  centré en l'origine et de rayon  $r$ .

## 6 Suggestions

1. On pourra présenter le modèle et démontrer en particulier la proposition 1.1.
2. On pourra démontrer la proposition 2.2.
3. On pourra s'intéresser à des éléments de la preuve de la proposition 4.1. On pourra par exemple montrer la convergence du moment d'ordre 2 ou démontrer la proposition 4.3.
4. On pourra démontrer les propositions 4.4 et 4.5.
5. On pourra illustrer par la simulation les propositions 3.1 et 4.1.
6. On pourra illustrer par la simulation le théorème 5.1 et faire le lien avec les résultats en dimension 1.

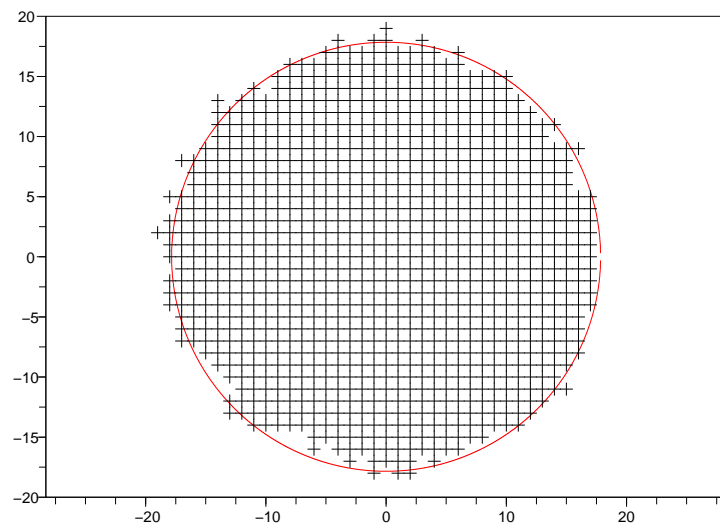


FIG. 1 – Une réalisation de  $A_{1000}$ .