

# PROBABILITÉS NUMÉRIQUES

Éditeurs

Nicolas BOULEAU

Denis TALAY

Édité par



INSTITUT NATIONAL DE RECHERCHE  
EN INFORMATIQUE ET EN AUTOMATIQUE  
DOMAINE DE VOLUCEAU - ROCQUENCOURT - FRANCE

faut savoir si la valeur précédente de  $Y(t_{k+1})$  est acceptable ou non. Il faut donc savoir comment les erreurs locales se propagent le long de la trajectoire. Supposons qu'elles soient assez petites pour que cette propagation soit linéaire. En remplaçant  $X$  par  $X + \epsilon Y$  dans (9) et en faisant tendre  $\epsilon$  vers zéro, on a

$$dY = A(t)Y dt + B(t)Y dW_t \quad \text{avec} \quad A(t) = \dot{a}(X(t)) \quad B(t) = \dot{\sigma}(X(t)) \quad (15)$$

Par une première solution approchée de l'E.D.S. avec un pas fixe assez petit, on peut alors résoudre de façon rétrograde (15). Dans notre exemple on a facilement

$$Y(t) = \frac{Y(T)X(t)}{X(T)}$$

Il faut maintenant choisir  $Y(t)$  afin de minimiser l'erreur finale  $Y(T)$ . Le choix optimal, qui minimise la somme des erreurs absolues, est de laisser chaque erreur locale fournir une part égale de l'erreur finale. Ainsi dans notre exemple, si l'on veut une précision finale de  $\epsilon$  (i.e.  $|Y(T)| \leq \epsilon$ ), il faut prendre  $|Y(t)| < \epsilon |X(t)X(T)|^{-1}$ , c'est à dire ici  $|\Delta W_{k+1}|^2 < C$ . Des résultats numériques montrent la pertinence d'une telle méthode ([26]).

### II.3.3.2 Contrôle spatial de la diffusion

L'idée est d'utiliser des pas de discrétisation aléatoires afin par exemple de permettre un contrôle spatial de la diffusion. Pour cela nous construisons en même temps la subdivision de  $[0, T]$  et le schéma de discrétisation. Introduisons à cet effet un schéma de type Euler introduit par Bichteler [10] dans un cadre plus général : le schéma Euler est défini par  $X_0^h = 0$ ,  $\tau_0 = 0$  et

$$\begin{aligned} \tau_{k+1} &= \tau_k + \inf\{t > 0, |b(X_{\tau_k}^h)t + \sigma(X_{\tau_k}^h)(W_{t+\tau_k} - W_{\tau_k})| = \sqrt{h}\}, \\ X_{\tau_{k+1}}^h &= X_{\tau_k}^h + b(X_{\tau_k}^h)(\tau_{k+1} - \tau_k) + \sigma(X_{\tau_k}^h)(W_{\tau_{k+1}} - W_{\tau_k}), \\ \forall l \in [\tau_k, \tau_{k+1}], X_t^h &= X_{\tau_k}^h + \frac{t - \tau_k}{\tau_{k+1} - \tau_k} (X_{\tau_{k+1}}^h - X_{\tau_k}^h) \end{aligned}$$

Ainsi défini ce schéma vérifie  $|\bar{X}_{\tau_{k+1}}^h - \bar{X}_{\tau_k}^h| = \sqrt{h}$  et  $h$  apparaît comme un pas de discrétisation spatial homogène à un temps. On peut montrer pour ce schéma un résultat de convergence tout à fait similaire à celui sur le schéma d'Euler classique ([24]). De plus il est possible de simuler numériquement un tel schéma, au moins dans le cas d'un brownien monodimensionnel, car il suffit pour cela de savoir simuler le temps de sortie d'une intervalle pour un brownien avec drift, et le brownien à la sortie de cet intervalle. On utilise pour cela des méthodes classiques de simulation de densités données par des séries ([21]). Un tel schéma permet par exemple d'approcher une diffusion très près de son temps d'explosion, ou d'un point de dégradation. D'autres schémas analogues sont utilisables, par exemple quand les  $\tau_k$  contrôlent spatialement le brownien ([18], [24]).

## II.4 SUR LA METHODE DU SHIFT EN SIMULATION

(M. Ben Alaya)

Du point de vue des simulations numériques en grande dimension ou pour les processus aléatoires, l'usage du théorème ergodique ponctuel de Birkhoff pour l'opérateur de décalage, apparaît comme une méthode performante à bien des égards (voir Bouleau [11]).

En effet les particularités de son implémentation informatique évitent de tirer des trajectoires complètement différentes. Le but de ce papier est d'expliquer au lecteur, par des exemples inspirés de problèmes pratiques, l'intérêt de cette méthode et de donner les résultats, élaborés au cours de ma thèse, sur la vitesse de convergence pour des fonctions que l'on rencontre couramment en pratique.

Dans la suite on notera par  $[0, 1]^N$  l'espace produit de l'intervalle  $[0, 1]$ ,  $\lambda = dx^{\otimes N}$  la mesure de Lebesgue produit définie sur  $[0, 1]^N$  et  $\theta$  la fonction de décalage, appelée aussi opérateur de shift, définie sur  $[0, 1]^N$  par

$$\theta(U_1, U_2, \dots, U_k, \dots) = (U_2, U_3, \dots, U_{k+1}, \dots).$$

Il est clair que  $\theta(\lambda) = \lambda$ , où  $\theta(\lambda)$  est la mesure image. On vérifie (cf. Krengel [36]) que le système dynamique  $([0, 1]^N, \mathcal{B}([0, 1]^{\otimes N}), \lambda, \theta)$  est ergodique. Par conséquent, d'après le théorème ergodique ponctuel de Birkhoff on a :

$$\frac{1}{N} \sum_{n=0}^{N-1} F \circ \theta^n \longrightarrow \mathbf{E}F \quad \lambda - p.s., \quad (16)$$

pour toute fonctionnelle  $F$   $\lambda$ -intégrable sur  $[0, 1]^N$ .

### II.4.1 Présentation de la méthode

La modélisation des phénomènes aléatoires, fait souvent intervenir des processus à temps discret, ainsi : les problèmes de transport de particules, la discrétisation des équations différentielles stochastiques, etc. Très souvent ces processus sont des chaînes de Markov. Nous allons décrire comment simplement la simulation d'une telle chaîne par la méthode du shift.

Soit  $X_n$  une chaîne de Markov définie par :

$$X_0 = x, \quad X_{n+1} = H(X_n, \eta_n, h((l_{nd+1}, \dots, l_{(n+1)d}))),$$

où les  $(l_n)$  sont des variables aléatoires i.i.d. de loi uniforme sur  $[0, 1]$ . Ainsi le processus  $(X_n)_{n \in \mathbf{N}}$  admet une représentation sur  $([0, 1]^{\mathbf{N}}, dx^{\otimes \mathbf{N}})$ .

On considère le temps d'arrêt  $T = \inf\{n \geq 1; X_n \in A\}$  et les fonctionnelles associées  $F = G(X_T, T)$ . L'espérance  $EG(X_T, T)$  peut être calculée par le théorème de Birkhoff, puisque

$$EG(X_T, T) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} EG(X_T, T) \theta^{nd} \quad \lambda - p.s., \quad (17)$$

dès que  $G(X_T, T) \in L^1$ . L'implémentation de cet algorithme exploite de façon essentielle la notion de pointeur (ou tout autre procédé équivalent de gestion de piles). Pour appliquer la formule (2) nous avons à calculer  $G(X_T, T)$  aux points successifs  $U, \theta^d(U), \theta^{2d}(U), \dots, d \in [0, 1]^N$ , ces points sont des suites de points de  $[0, 1]$

$$\begin{aligned} U &= (U_1, U_2, \dots, U_d, \dots, U_k, \dots) \\ \theta^d(U) &= (U_{d+1}, U_{d+2}, \dots, U_{2d}, \dots, U_{d+k}, \dots) \\ &\vdots \end{aligned}$$

Alors si le temps d'arrêt  $T$  est fini, ce que nous supposons ici, le calcul  $G(X_T, T)$  ne dépend que d'un nombre fini de coordonnées, le long de chaque trajectoire  $U = (U_1, U_2, \dots, U_k, \dots)$  : la suite  $U$  étant calculée jusqu'à l'indice  $k = dT(U)$  où  $T(U)$  est déterminé par le test

$$\begin{aligned} < \text{test} > \{ X_n \notin A &\Rightarrow n < T(U) \\ X_n \in A &\Rightarrow n = T(U) \end{aligned}$$

Supposons qu'on ait calculé  $G(X_T, T)$  au point  $U = (U_1, U_2, \dots)$ , et que les calculs des termes intermédiaires,  $Y_n = h(U_{(n-1)d+1}, \dots, U_{nd})$ ,  $1 \leq n \leq T(U)$ , aient été mis dans des pointeurs comme indiqué dans la figure 1. La simulation de  $G(X_T, T)$  sur la trajectoire  $\theta^d(U) = (U_{d+1}, U_{d+2}, \dots)$  sera plus simple puisque le calculs intermédiaire des  $Y_i$  associés à  $\theta^d(U)$  est fait au moins partiellement. En effet soit on a  $T(\theta^d(U)) < T(U)$  auquel cas aucun calcul supplémentaire n'est requis ou alors on a simplement à rallonger la chaîne entre  $T(U)$  et  $T(\theta^d(U))$  auquel cas nous avons la figure 2 :

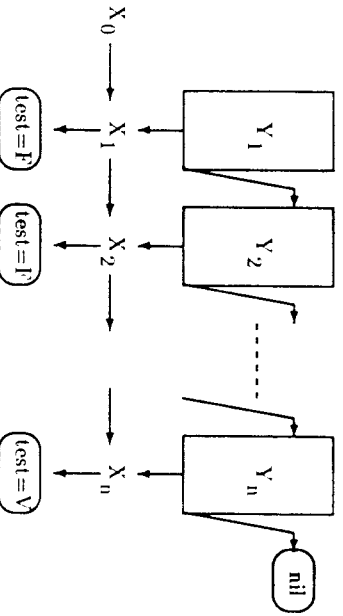


Figure 1.

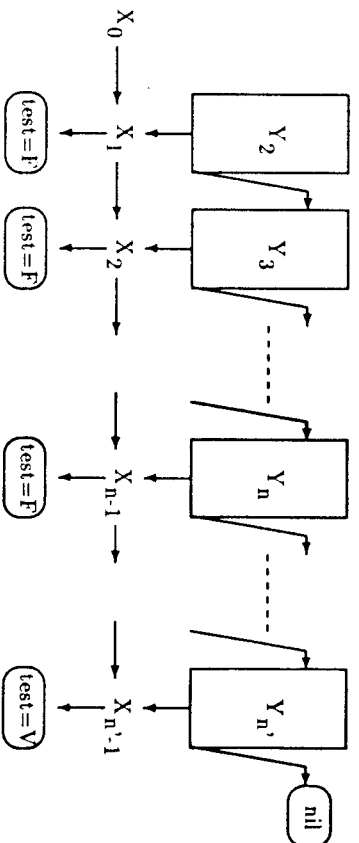


Figure 2.

Dans ces diagrammes les flèches représentent essentiellement le calcul de :

$$X_{n+1} = H(X_n, n, h(U_{nd+1}, \dots, U_{(n+1)d})) = H(X_n, n, Y_{n+1}).$$

Il est clair que les performances de cette méthode proviennent de la boîte de stockage qui évite de recommencer inutilement les calculs partiels des variables  $Y_i$ . Ce stockage doit être fait évidemment lors du rallongement de la suite lorsque ce dernier intervient comme expliqué ci-dessus.

L'utilisation de cette méthode nécessite une estimation de la vitesse de convergence, telle que la loi du logarithme itéré et le théorème de la limite centrale, dont le théorème ergodique ponctuel ne dispose en général pas. Rappelons deux résultats pour l'opérateur du shift sur  $([0, 1]^N, \mathcal{B}([0, 1]^N), \lambda, \theta)$  (cf. Krengel [36]), dont on déduit que la vitesse de convergence peut être arbitrairement lente ou arbitrairement proche de  $0(\frac{1}{N})$ .

D'une part : pour toute suite  $(\alpha_n)_{n \in \mathbb{N}}$ ,  $\alpha_n > 0$ ,  $\alpha_n \rightarrow 0$ , il existe une fonction continue  $F$  de  $[0, 1]^N$  à valeurs réelles telle que

$$\frac{1}{\alpha N} \left| \frac{1}{N} \sum_{n=0}^{N-1} F \circ \theta^n - \mathbb{E}F \right| \rightarrow \infty, \quad \lambda - p.s..$$

D'autre part : pour toute suite  $(c_n)_{n \in \mathbb{N}}$ ,  $c_n > 0$ , croissant vers l'infini, avec  $c_1 \geq 2$ , il existe un ensemble mesurable  $A$  avec  $P(A) = \frac{1}{2}$ , pour lequel

$$\forall N \quad \left| \frac{1}{N} \sum_{n=0}^{N-1} 1_A \circ \theta^n - \frac{1}{2} \right| \leq \frac{c_N}{N}.$$

Dans la deuxième partie, on montre des résultats sur la vitesse de convergence pour des fonctions contrairement employées en pratique.

### II.4.2 Vitesse de convergence

On considère toujours le système dynamique  $([0, 1]^N, \mathcal{B}([0, 1]^{\otimes N}), \lambda, \theta)$ . On suppose que  $\int_{[0, 1]^N} F d\lambda = 0$  et on pose

$$\sigma^2 = \sigma^2(F) = \|F\|_2^2 + 2\sum_{k=1}^{\infty} \langle F, \theta^k F \rangle, \quad (18)$$

où  $\langle \cdot, \cdot \rangle$  est le produit scalaire sur  $L^2([0, 1]^N)$  et  $\|\cdot\|_2$  est la norme sur  $L^2([0, 1]^N)$  :

$$\langle F, G \rangle = \int_{[0, 1]^N} FG d\lambda \quad \|F\|_2 = \left( \int_{[0, 1]^N} F^2 d\lambda \right)^{\frac{1}{2}}.$$

Pour plus de généralité, on considérera, un temps d'arrêt  $T$  de  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  et une fonctionnelle  $F, \mathcal{F}_T$ -mesurable, avec  $\mathcal{F}_n = \sigma(U_1, \dots, U_n)$ . On a alors les résultats suivants :

**Lemma II.4.1** Si  $T$  admet un moment d'ordre  $p > 2$  alors pour toute  $F \in L^2([0, 1]^N, \mathcal{F}_T)$ , telle que  $\int_{[0, 1]^N} F = 0$ , on a

$$| \langle F, \theta^k F \rangle | \leq \frac{\left( \int_{[0, 1]^N} |F|^2 d\lambda \right) (\mathbb{E}T^p)^{\frac{1}{2}}}{k^{\frac{p}{2}}},$$

et donc  $\sigma^2(F)$  est une série absolument convergente.

**Preuve:** Voir référence [5], (preuve de la proposition 3). ■

A partir de ce lemme, on obtient un théorème de la limite centrale.

**Théorème II.4.1** Si  $T$  admet un moment d'ordre  $p > 2$  alors pour toute  $F \in L^2([0, 1]^N, \mathcal{F}_T)$ , telle que  $\int_{[0, 1]^N} F = 0$  et  $\sigma^2(F) > 0$  on a

$$\frac{1}{\sigma\sqrt{N}} \sum_{n=0}^{N-1} F \circ \theta^n \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1), \quad (19)$$

où  $\mathcal{N}(0, 1)$  est la loi normale centrée réduite.

**Preuve:** Soit  $F \in L^2([0, 1]^N, \mathcal{F}_T)$  et  $l \in \mathbb{N}$ , si on écrit que

$$F = (F_{1T \leq l} - \mathbb{E}F_{1T \leq l}) + (F_{1T > l} - \mathbb{E}F_{1T > l}),$$

on a

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} F \circ \theta^n = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (F_{1T \leq l} - \mathbb{E}F_{1T \leq l}) \circ \theta^n + \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (F_{1T > l} - \mathbb{E}F_{1T > l}) \circ \theta^n.$$

Le premier terme à gauche de l'égalité converge en loi, lorsque  $N$  tend vers l'infini, à  $\mathcal{N}(0, \sigma^2(F_{1T \leq l} - \mathbb{E}F_{1T \leq l}))$ , pour tout  $l \in \mathbb{N}$ , le second terme, lui, vérifie :

$$\lim_{l \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbf{P} \left\{ \left| \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (F_{1T > l} - \mathbb{E}F_{1T > l}) \circ \theta^n \right| \geq \epsilon \right\} = 0$$

la preuve est achevée, en montrant que  $\sigma^2(F_{1T \leq l} - \mathbb{E}F_{1T \leq l})$  converge vers  $\sigma^2$ . Pour une démonstration complète voir [6].

Dans la suite on montre un résultat de type loi du logarithme itéré. ■

**Théorème II.4.2** Si  $T$  admet un moment d'ordre  $p > 2$  alors pour toute  $F \in L^2([0, 1]^N, \mathcal{F}_T)$ , telle que  $\int_{[0, 1]^N} F = 0$ , on a

$$\forall \epsilon > 0 \quad \frac{1}{N} \sum_{n=0}^{N-1} F \circ \theta^n = o(N^{-\frac{1}{2}} (\log(N))^{\frac{1}{2} + \epsilon}) \quad \lambda - p.s. \quad (20)$$

**Preuve :** Ce résultat est vrai pour toute fonction dans  $L^2$ , vérifiant  $\sigma^2$  fini. En effet sous cette hypothèse on montre que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \int_{[0, 1]^N} \left| \sum_{n=0}^{N-1} F \circ \theta^n \right|^2 d\lambda = \sigma^2,$$

soit encore

$$\int_{[0, 1]^N} \left| \sum_{n=0}^{N-1} F \circ \theta^n \right|^2 d\lambda = O(N),$$

et d'après un théorème de Gál et Koksma [27] on a le résultat énoncé. Pour une démonstration complète voir [5]. ■

#### Remarques :

1/On sait que si  $(X_n)_{n \in \mathbb{N}}$  est une suite de variables aléatoires réelles indépendantes équilibrées telle que  $\mathbb{E}X_1^2 < \infty$ , alors

$$\frac{1}{N} (X_1 + \dots + X_N) = O(N^{-\frac{1}{2}} (\log \log(N))^{\frac{1}{2}}),$$

L'hypothèse d'indépendance a été affaiblie dans plusieurs travaux (cf. Berger [8]).

2/Il est évident que les estimations ci-dessus, connues sous le nom de propriété du logarithme itéré, sont plus fortes que le résultat du théorème 2.

L'intérêt de notre résultat est de donner une estimation assez voisine du logarithme itéré mais sous des hypothèses plus faibles et assez naturelles dans le cadre de la simulation.

### II.4.3 Conclusion

Les résultats établis ci-dessus montrent que la méthode du shift pour le calcul de l'espérance de fonctionnelles de processus aléatoires a une vitesse asymptotique du même ordre de grandeur en pratique que la loi des grands nombres pour une large classe de fonctionnelles parmi les plus couramment rencontrées. Ceci renforce l'intérêt de cette méthode dont l'avantage essentiel, il convient de le souligner, réside dans l'économie de calculs que permet son implémentation informatique.

## Bibliographie

- [1] L. ARNOLD. *Stochastic Differential Equations*. Wiley, New-York, 1974.
- [2] R. AZENCOTT. Formule de Taylor stochastique et développement asymptotique d'intégrales de Feynmann. In *Sem. Prob. XVI supp. Géométrie Différentielle Stochastique*, volume 921 of *Lect. Notes in Math.*, pages 237–285. Springer, 1982.
- [3] F.J. BELFANTE. *A Survey of Hidden Variable Theory*. Pergamon Press, 1973.
- [4] M. BEN ALAYA. La méthode du shift en simulation. *Ce volume*, 1991.
- [5] M. BEN ALAYA. Sur le calcul approché d'espérances de variables aléatoires dépendant d'un temps d'arrêt. *Rapport LAM-ENPC*, 1991.
- [6] M. BEN ALAYA. Théorème de limite centrale et vitesse expérimentales pour le calcul d'espérances par l'opérateur de décalage. *Rapport CERMA*, 1991.
- [7] G. BEN AROUS. Flots et séries de Taylor stochastiques. *Probab. Th. Rel. Fields*, 81:29–77, 1989.
- [8] E. BERGER. An almost sure invariance principle for stationary sequences of Banach space valued random variables. *Probab. Th. Rel. Fields*, 84:161–201, 1990.
- [9] R.N. BHATTACHARYA. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 60:185–201, 1982.
- [10] K. BICHTLER. Stochastic integration and  $L^p$  theory of semimartingales. *Annals of Probability*, 9(1):49–89, 1981.
- [11] N. BOULEAU. On effective computation of expectations in large or infinite dimension. *Journal of Computational and Applied Mathematics*, 31:23–34, 1990.
- [12] N. BOULEAU. Irregular and simulatable functionals on Wiener space. *Ce volume*, 1991.
- [13] N. BOULEAU, F. HIRSCH. *Dirichlet Forms and Analysis on Wiener Space*. De Gruyter, 1991.