Algorithmes de Monte Carlo par chaînes de Markov et méthodes particulaires

Benjamin JOURDAIN

4 mars 2025

Table des matières

Chapitre 1

Algorithme de Metropolis Hastings et convergence de chaînes de Markov

1.1 noyau markovien et chaîne de Markov

On munit l'espace d'états E d'une tribu $\mathcal{E}$.

Définition 1.1.1. — *On appelle noyau markovien sur* $(E,\mathcal{E})$*, une application* $P : E \times \mathcal{E} \mapsto [0,1]$ *telle que*

- — $\forall x \in E$, $\mathcal{E} \ni A \mapsto P(x,A)$ *est une probabilité,*
- — $\forall A \in \mathcal{E}$, $E \ni x \mapsto P(x,A)$ *est mesurable.*

— *On appelle chaîne de Markov de noyau de transition* P*, un processus* $(X_n)_{n\in\mathbb{N}}$ *à valeurs dans* E *adapté à la filtration* $(\mathcal{F}_n)_{n\geq 0}$ *tel que*

$$\forall n \in \mathbb{N},\ \forall A \in \mathcal{E},\ \mathbb{P}(X_{n+1} \in A|\mathcal{F}_n) = P(X_n, A).$$

On adoptera les notations suivantes :

- — $\mathcal{P}(E)$ désigne l'ensemble des probabilités sur $(E,\mathcal{E})$,
- — pour $\varphi : E \to \mathbb{R}$ mesurable positive ou telle que $\forall x \in E$, $\int_E |\varphi(y)|P(x,dy) < +\infty$, on note $P\varphi$ la fonction de E dans $\mathbb{R} \cup \{+\infty\}$ définie par $P\varphi(x) = \int_E \varphi(y)P(x,dy)$,
- — pour μ dans l'espace $\mathcal{P}(E)$ des probabilités sur $(E,\mathcal{E})$, on note μP la probabilité sur $(E,\mathcal{E})$ définie par $\forall A \in \mathcal{E},\ \mu P(A) = \int_E P(x,A)\mu(dx)$,
- — la probabilité $\pi \in \mathcal{P}(E)$ est dite invariante par P si $\pi P = \pi$,
- — pour μ une mesure positive sur $(E,\mathcal{E})$ et $\varphi : E \to \mathbb{R}$ mesurable positive ou telle que $\int_E |\varphi(x)|\mu(dx) < \infty$, on note $\mu(\varphi) = \int_E \varphi(x)\mu(dx)$,
- — pour P et Q deux noyaux markoviens sur $(E,\mathcal{E})$, on note PQ le noyau markovien sur $(E,\mathcal{E})$ défini par $\forall x \in E,\ \forall A \in \mathcal{E},\ PQ(x,A) = \int_E Q(y,A)P(x,dy)$ et par $(P^n)_{n\in N}$ la suite de noyaux définis par $P^0(x,A) = 1_A(x)$ (i.e. $P^0(x,dy) = \delta_x(dy)$) et la relation de récurrence $P^n := PP^{n-1} = P^{n-1}P$ pour $n \geq 1$. Notons que $P^1 = P$.

Si $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de noyau de transition P avec X_0 de loi de probabilité μ, alors pour tout $n \in \mathbb{N}$, la loi de X_n est μP^n.

Exercice 1.1.2. *Montrer cette propriété (on pourra commencer par remarquer que pour* $\varphi : E \to \mathbb{R}$ *mesurable bornée et* $k \in \mathbb{N}$, $\mathbb{E}[\varphi(X_{k+1})|\mathcal{F}_k] = P\varphi(X_k)$*).*

1.2 Algorithme de Metropolis-Hastings

Cet algorithme est utilisé lorsque, sur l'espace d'états E, on veut simuler suivant la probabilité π de la forme $\pi(dx) = \frac{\eta(x)\lambda(dx)}{\int_E \eta(y)\lambda(dy)}$ où

— λ est une mesure positive de référence sur $(E, \mathcal{E})$,

— η est une fonction mesurable de E dans $\mathbb{R}_+$ telle que $\int_E \eta(x)\lambda(dx) \in]0, \infty[$.

Il va permettre de construire une chaîne de Markov de probabilité invariante π sans nécessairement connaître la constante de normalisation $\frac{1}{\int_E \eta(y)\lambda(dy)}$ de la densité η. Parmi ses très nombreuses applications, on peut notamment citer :

— dans le cas où λ est la mesure de Lebesgue sur $E = \mathbb{R}^d$, la simulation en physique statistique de la probabilité de Boltzmann-Gibbs de densité proportionnelle à $\eta(x) = e^{-\frac{1}{k_B T}V(x)}$ où k_B désigne la constante de Boltzmann, T la température et $V : \mathbb{R}^d \to \mathbb{R}$ est la fonction potentiel,

— dans le cas où E est l'espace des valeurs possibles d'un paramètre θ en statistique bayésienne, la simulation suivant la loi à posteriori de ce paramètre sachant que l'on a observé $Y = y$: si on note $p_{Y|\Theta}(y|\theta)$ la densité de l'observation Y lorsque le paramètre vaut θ et $p_\Theta(\theta)$ la densité à priori de Θ par rapport à λ, la formule de Bayes assure que sa densité à posteriori est

$$p_{\Theta|Y}(\theta|y) = \frac{p_{Y|\Theta}(y|\theta)p_\Theta(\theta)}{\int_E p_{Y|\Theta}(y|\vartheta)p_\Theta(\vartheta)\lambda(d\vartheta)}.$$

Notons que dans ces deux exemples, le calcul de la constante de normalisation est difficile.

Soit $q : E \times E \to \mathbb{R}_+$ une fonction mesurable telle que $\forall x \in E$, $\int_E q(x,y)\lambda(dy) = 1$ ($Q(x,dy) = q(x,y)\lambda(dy)$ est alors un noyau markovien) et on sait simuler suivant la probabilité $q(x,y)\lambda(dy)$. On pose

$$\alpha(x,y) = \begin{cases} \min\left(1, \frac{\eta(y)q(y,x)}{\eta(x)q(x,y)}\right) \text{ si } \eta(x)q(x,y) > 0 \\ 1 \text{ si } \eta(x)q(x,y) = 0 \end{cases} . \tag{1.1}$$

Partant d'une variable aléatoire X_0 à valeurs dans E, le principe de l'algorithme de Metropolis-Hastings est de construire $(X_n)_{n\in\mathbb{N}}$ par la récurrence suivante :

— sachant $(X_0, \dots, X_n)$, on génère une proposition Y_{n+1} suivant la probabilité $q(X_n, y)\lambda(dy)$ et indépendamment une variable aléatoire U_{n+1} uniforme sur $[0, 1]$,

— on pose $X_{n+1} = Y_{n+1}1_{\{U_{n+1}\leq\alpha(X_n,Y_{n+1})\}} + X_n 1_{\{U_{n+1}>\alpha(X_n,Y_{n+1})\}}$, c'est-à-dire que la proposition est acceptée avec probabilité $\alpha(X_n, Y_{n+1})$ et la position X_n est conservée sinon.

Ainsi pour $f : E \to \mathbb{R}$ mesurable bornée,

$$\begin{aligned}
&\mathbb{E}[f(X_{n+1})|X_0, X_1, \dots, X_n] \\
&= \mathbb{E}[\mathbb{E}[f(Y_{n+1})1_{\{U_{n+1}\leq\alpha(X_n,Y_{n+1})\}} + f(X_n)1_{\{U_{n+1}>\alpha(X_n,Y_{n+1})\}}|X_0, X_1, \dots, X_n, Y_{n+1}|X_0, X_1, \dots, X_n] \\
&= \mathbb{E}[f(Y_{n+1})\alpha(X_n, Y_{n+1}) + f(X_n)(1 - \alpha(X_n, Y_{n+1}))|X_0, X_1, \dots, X_n] \\
&= \int_E f(y)\alpha(X_n, y)q(X_n, y)\lambda(dy) + f(X_n)\int_E (1 - \alpha(X_n, y))q(X_n, y)\lambda(dy) \\
&= \int_E f(y)P(X_n, dy)
\end{aligned}$$

où

$$P(x,dy) = \alpha(x,y)q(x,y)\lambda(dy) + R(x)\delta_x(dy) \text{ avec } R(x) = \int_E (1-\alpha(x,z))q(x,z)\lambda(dz). \quad (1.2)$$

Ainsi $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de noyau de transition P. Pour $y \neq x$,

$$\begin{aligned}\eta(x)q(x,y)\alpha(x,y) &= \begin{cases}\eta(x)q(x,y)\min\left(1, \frac{\eta(y)q(y,x)}{\eta(x)q(x,y)}\right) \text{ si } \eta(x)q(x,y) > 0\\ \eta(x)q(x,y)\times 1 \text{ si } \eta(x)q(x,y) = 0\end{cases}\\ &= \min(\eta(x)q(x,y), \eta(y)q(y,x)).\end{aligned}$$

est une fonction symétrique de (x,y) i.e. $\eta(x)q(x,y)\alpha(x,y) = \eta(y)q(y,x)\alpha(y,x)$. Cela entraîne que

$$\eta(x)\lambda(dx)\alpha(x,y)q(x,y)\lambda(dy) = \eta(y)\lambda(dy)\alpha(y,x)q(y,x)\lambda(dx).$$

Comme clairement, $\eta(x)\lambda(dx)R(x)\delta_x(dy) = \eta(y)\lambda(dy)R(y)\delta_y(dx)$, on en déduit que $\pi(dx)P(x,dy) = \pi(dy)P(y,dx)$. On dit que la probabilité π est réversible pour le noyau P. Comme $P(y,E) = 1$, par intégration en x sur E, cela implique que $\pi P = \pi$: la probabilité π est invariante par le noyau P. Notons que dans le cas particulier où $E = \mathbb{R}^d$ et $q(x,y) = \psi(y-x)$ pour une densité de probabilité paire ψ par rapport à la mesure de Lebesgue λ (exemple : $\psi(z) = e^{-\frac{|z|^2}{2\sigma^2}}/(2\pi\sigma^2)^{d/2}$), alors le rapport $\frac{\eta(y)q(y,x)}{\eta(x)q(x,y)}$ se simplifie en $\frac{\eta(y)}{\eta(x)}$. Comme les variables aléatoires $Y_{n+1} - X_n$ sont alors i.i.d. suivant la densité ψ, on parle d'algorithme de Metropolis-Hastings par marche aléatoire.

Remarque 1.2.1. — *Notons que la réversibilité de π pour le noyau P est préservée pour*

$$\alpha(x,y) = \begin{cases} a\left(\frac{\eta(y)q(y,x)}{\eta(x)q(x,y)}\right) \text{ si } \eta(x)q(x,y) > 0\\ 1 \text{ si } \eta(x)q(x,y) = 0\end{cases},$$

où la fonction $a : \mathbb{R}_+ \to [0,1]$ vérifie $a(0) = 0$ et $\forall u > 0,\ a(u) = ua(1/u)$ si bien que $a(u) \leq \min(1,u)$. Le choix $a(u) = \min(1,u)$ fait dans l'algorithme de Metropolis-Hastings maximise la probabilité d'accepter les propositions. Il est privilégié en pratique pour ses meilleures propriétés asymptotiques (voir le paragraphe 1.6) mais d'autres sont possibles comme $a(u) = \frac{\sum_{k=1}^K u^k}{1+\sum_{k=1}^K u^k}$ pour $K \in \mathbb{N}^*$.

— *Notons enfin que le choix de la valeur de $\alpha(x,y)$ lorsque $q(x,y) = 0$ ne change rien au noyau P. En revanche, choisir $\alpha(x,y) = 1$ lorsque $\eta(x) = 0$ et $q(x,y) > 0$ permet de maximiser la probabilité de bouger de l'état x.*

Exercice 1.2.2. *On suppose que π est invariante pour le noyau P. Soient $f, g : E \to \mathbb{R}$ mesurables et telles que $\pi(f^2 + g^2) < \infty$.*

1. *Montrer que $(Pf)^2 \leq Pf^2$ et en déduire à l'aide de l'inégalité de Cauchy-Schwarz que $\pi(|gPf|) < \infty$.*
2. *Lorsque π est réversible pour P, montrer que $\pi(gPf) = \pi(fPg)$.*

Dans la suite de ce chapître, nous allons donner des conditions sur le noyau de transition P d'une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ à valeurs dans E assurant que

— la chaîne admet une unique probabilité invariante π et, lorsque $n \to \infty$, la loi de X_n converge à vitesse géométrique vers π dans une norme bien choisie,

— pour toute fonction $f : E \to \mathbb{R}$ mesurable et telle que $\pi(|f|) < \infty$, les moyennes ergodiques $\frac{1}{n}\sum_{k=0}^{n} f(X_k)$ convergent p.s. vers $\pi(f)$,

— le comportement asymptotique de l'erreur dans la loi des grands nombres ergodique précédente est régi par un théorème de la limite centrale.

1.3 Ergodicité d'une chaîne de Markov à espace d'état général

Soit P un noyau markovien sur $(E, \mathcal{E})$. Nous allons nous intéresser à des conditions assurant que pour toute probabilité $\mu \in \mathcal{P}(E)$, μP^n converge lorsque $n \to \infty$ vers π où $\pi \in \mathcal{P}(E)$ est invariante c'est-à-dire vérifie $\pi P = \pi$. Pour cela, nous commençons par introduire une distance permettant de quantifier cette convergence.

1.3.1 distance en variation totale

Définition 1.3.1. *On appelle distance en variation totale sur $\mathcal{P}(E)$ la distance $d_{\mathrm{TV}}(\mu, \sigma) := \sup_{A \in \mathcal{E}} |\mu(A) - \sigma(A)|$.*

Pour obtenir une expression alternative de cette distance, nous introduisons, la décomposition de Jordan-Hahn de la mesure signée $\mu - \sigma$: $\mu - \sigma = \xi^+ - \xi^-$ où ξ^+ et ξ^- sont deux mesures positives sur $(E, \mathcal{E})$ telles qu'il existe $B \in \mathcal{E}$ vérifiant $\xi^-(B) = \xi^+(B^c) = 0$.

Exercice 1.3.2. *Montrer l'existence et l'unicité de la décomposition de Jordan-Hahn lorsque l'espace d'états E est fini ou dénombrable et $\mathcal{E}$ l'ensemble des parties de E (tribu discrète).*

Proposition 1.3.3. *La mesure $\mu \wedge \sigma$ sur $(E, \mathcal{E})$ définie par $\forall A \in \mathcal{E},\ \mu \wedge \sigma(A) = \sigma(A \cap B) + \mu(A \cap B^c)$ est la plus grande mesure positive majorée à la fois par μ et par σ. En outre,*

$$d_{\mathrm{TV}}(\mu, \sigma) = 1 - \mu \wedge \sigma(E) = \sup_{\varphi : |\varphi| \leq 1/2} |\mu(\varphi) - \sigma(\varphi)|,$$

où le supremum porte sur l'ensemble des fonctions mesurables $\varphi : E \to \mathbb{R}$ telles que $\sup_{x \in E} |\varphi(x)| \leq 1/2$.

Attention, généralement, en dehors des cas $A \subset B$ et $A \subset B^c$, $\mu \wedge \sigma(A) \neq \mu(A) \wedge \sigma(A)$.

Exercice 1.3.4. *Montrer que pour tout $\gamma > 0$, $d_{\mathrm{TV}}(\mu, \sigma) = \frac{1}{2\gamma} \sup_{\varphi : |\varphi| \leq \gamma} |\mu(\varphi) - \sigma(\varphi)|$.*

Exercice 1.3.5. *On suppose que $\mu(dx) = f(x)\lambda(dx)$, $\sigma(dx) = g(x)\lambda(dx)$ avec $f, g : E \to \mathbb{R}_+$ deux fonctions mesurables telles que $\int_E f(x)\lambda(dx) = \int_E g(x)\lambda(dx) = 1$.*

1. *Calculer $\mu(A) - \sigma(A)$ pour $A \in \mathcal{E}$ et en déduire que*

$$\begin{aligned} d_{\mathrm{TV}}(\mu, \sigma) &= \int_E \max(f(x) - g(x), 0)\lambda(dx) = \int_E \max(g(x) - f(x), 0)\lambda(dx) \\ &= \frac{1}{2}\int_E |g(x) - f(x)|\lambda(dx). \end{aligned}$$

2. *Vérifier que* $\xi^+(dx) = 1_{\{f(x)\geq g(x)\}}(f(x)-g(x))\lambda(dx)$ *et* $\xi^-(dx) = 1_{\{g(x)\geq f(x)\}}(g(x)-f(x))\lambda(dx)$ *est la décomposition de Jordan-Hahn de* $\mu-\sigma$ *et préciser les ensembles* B *et* B^c.
3. *Vérifier que* $\mu\wedge\sigma(dx) = (f(x)\wedge g(x))\lambda(dx)$.

Pour μ et σ quelconques, le théorème de Radon-Nikodym assure que μ et σ admettent des densités par rapport à $\lambda=\mu+\sigma$ et l'exercice précédent permet d'en déduire l'existence de la décomposition de Jordan-Hahn.

Démonstration : Pour $A\in\mathcal{E}$, comme, $\mu(A\cap B)-\sigma(A\cap B)=\xi^+(A\cap B)\geq 0$ et $\sigma(A\cap B^c)-\mu(A\cap B^c)=\xi^-(A\cap B^c)\geq 0$, $\mu\wedge\sigma(A)\leq\min(\mu(A),\sigma(A))$.

Par ailleurs, si η est une mesure positive telle que $\forall C\in\mathcal{E}$, $\eta(C)\leq\min(\mu(C),\sigma(C))$, pour $A\in\mathcal{E}$, $\eta(A)=\eta(A\cap B)+\eta(A\cap B^c)\leq\sigma(A\cap B)+\mu(A\cap B^c)=\mu\wedge\sigma(A)$.

Notons que comme $0=\mu(E)-\sigma(E)=\xi^+(E\cap B)-\xi^-(E\cap B^c)=\xi^+(B)-\xi^-(B^c)$, $\xi^+(B)=\xi^-(B^c)$. Pour $A\in\mathcal{E}$, comme $\mu(A)-\sigma(A)=\xi^+(A\cap B)-\xi^-(A\cap B^c)$,

$$-\xi^-(B^c)\leq\mu(A)-\sigma(A)\leq\xi^+(B),$$

où la première inégalité est une égalité pour $A=B^c$ et la seconde pour $A=B$. On en déduit que

$$\begin{aligned}d_{\mathrm{TV}}(\mu,\sigma)&=\xi^+(B)=\mu(B)-\sigma(B)=(1-\mu(B^c))-\sigma(B)=1-(\mu(E\cap B^c)+\sigma(E\cap B))\\&=1-\mu\wedge\sigma(E).\end{aligned}\tag{1.3}$$

Pour $\varphi:E\to\mathbb{R}$ mesurable telle que $\sup_{x\in E}|\varphi(x)|\leq 1/2$, $\xi^+(\varphi)$ (resp. $\xi^-(\varphi)$) est maximum (resp. minimum) pour φ égale à $1/2$ sur B (resp. à $-1/2$ sur B^c) et ne dépend pas des valeurs prises par φ sur B^c (resp. B). Ainsi,

$$\begin{aligned}\mu(\varphi)-\sigma(\varphi)&=\xi^+(\varphi)-\xi^-(\varphi)\leq\xi^+(1_B-1/2)-\xi^-(1_B-1/2)\\&=\mu(1_B-1/2)-\sigma(1_B-1/2)=\mu(B)-\sigma(B)=d_{\mathrm{TV}}(\mu,\sigma).\end{aligned}$$

Par symétrie, $\sigma(\varphi)-\mu(\varphi)\leq\sigma(B^c)-\mu(B^c)=d_{\mathrm{TV}}(\mu,\sigma)$, ce qui achève la démonstration. □

Remarque 1.3.6. *Pour* $A\in\mathcal{E}$,

$$\mu(A)-\mu\wedge\sigma(A)=\mu(A)-(\sigma(A\cap B)+\mu(A\cap B^c))=\mu(A\cap B)-\sigma(A\cap B)=\xi_+(A)$$

si bien que $\xi_+=\mu-\mu\wedge\sigma$, *ce qui entraîne* $\xi_-=\sigma-\mu+\xi_+=\sigma-\mu\wedge\sigma$.

Lemme 1.3.7. *Pour tous* $\mu,\sigma\in\mathcal{P}(E)$, *on a* $d_{\mathrm{TV}}(\mu,\sigma)=\inf_{X\sim\mu,Y\sim\sigma}\mathbb{P}(X\neq Y)$ *où l'infimum porte sur tous les couples de variables aléatoires* (X,Y) *avec* X *de loi* μ *et* Y *de loi* σ.

Démonstration : Pour $X\sim\mu$ et $Y\sim\sigma$ définies sur le même espace de probabilité et $A\in\mathcal{E}$, on a

$$\begin{aligned}\mu(A)-\sigma(A)&=\mathbb{P}(X\in A,X=Y)+\mathbb{P}(X\in A,X\neq Y)-\mathbb{P}(Y\in A,X=Y)-\mathbb{P}(Y\in A,X\neq Y)\\&=\mathbb{P}(X\in A,X\neq Y)-\mathbb{P}(Y\in A,X\neq Y)\end{aligned}$$

si bien que $-\mathbb{P}(X \neq Y) \leq -\mathbb{P}(Y \in A, X \neq Y) \leq \mu(A) - \sigma(A) \leq \mathbb{P}(X \in A, X \neq Y) \leq \mathbb{P}(X \neq Y)$. On en déduit que $\mathbb{P}(X \neq Y) \geq \sup_{A\in\mathcal{E}} |\mu(A) - \sigma(A)| = d_{\mathrm{TV}}(\mu, \sigma)$. Lorsque $d_{\mathrm{TV}}(\mu, \sigma) = 1$, l'inégalité est une égalité pour tout couple (X, Y) avec $X \sim \mu$ et $Y \sim \sigma$. Lorsque $d_{\mathrm{TV}}(\mu, \sigma) = 0$, c'est une égalité pour $Y = X \sim \mu$. Nous allons maintenant exhiber un couple pour lequel c'est une égalité lorsque $0 < d_{\mathrm{TV}}(\mu, \sigma) < 1$ i.e. $0 < \mu \wedge \sigma(E) < 1$. Soit $U \sim \mathcal{U}[0, 1]$ et $(Z, \tilde{X}, \tilde{Y})$ un triplet indépendant avec $Z \sim \frac{\mu\wedge\sigma}{\mu\wedge\sigma(E)}$, $\tilde{X} \sim \frac{\mu-\mu\wedge\sigma}{1-\mu\wedge\sigma(E)}$ et $\tilde{Y} \sim \frac{\sigma-\mu\wedge\sigma}{1-\mu\wedge\sigma(E)}$ (on peut par exemple prendre Z, $\tilde{X}$ et $\tilde{Y}$ indépendantes). On pose

$$(X, Y) = 1_{\{U\leq\mu\wedge\sigma(E)\}}(Z, Z) + 1_{\{U>\mu\wedge\sigma(E)\}}(\tilde{X}, \tilde{Y}).$$

Alors pour $f : E \to \mathbb{R}$ mesurable bornée, par indépendance de U et $(Z, \tilde{X})$ on a

$$\begin{aligned}
\mathbb{E}[f(X)] &= \mathbb{E}[1_{\{U\leq\mu\wedge\sigma(E)\}}f(Z)] + \mathbb{E}[1_{\{U>\mu\wedge\sigma(E)\}}f(\tilde{X})] \\
&= \mathbb{P}(U \leq \mu \wedge \sigma(E))\mathbb{E}[f(Z)] + \mathbb{P}(U > \mu \wedge \sigma(E))\mathbb{E}[f(\tilde{X})] \\
&= \mu \wedge \sigma(E) \int_E f(x)\frac{\mu \wedge \sigma(dx)}{\mu \wedge \sigma(E)} + (1 - \mu \wedge \sigma(E)) \int_E f(x)\frac{\mu(dx) - \mu \wedge \sigma(dx)}{1 - \mu \wedge \sigma(E)} \\
&= \int_E f(x)\mu(dx)
\end{aligned}$$

si bien que $X \sim \mu$. Un calcul analogue de $\mathbb{E}[f(Y)]$ assure que $Y \sim \sigma$. Par ailleurs, $\mathbb{P}(X \neq Y) \leq \mathbb{P}(U > \mu \wedge \sigma(E)) = 1 - \mu \wedge \sigma(E) = d_{\mathrm{TV}}(\mu, \sigma)$. □

Exercice 1.3.8. *Soit $A \in \mathcal{E}$ t.q. $\mu(A) = 1$.*

1. *Comment comparer $\mu(A^c \cap B)$ et $\sigma(A^c \cap B)$? En déduire que $\sigma(A \cap B) = \sigma(B)$.*
2. *Montrer que $\mu(A \cap B^c) = \mu(B^c)$ et conclure que $\mu \wedge \sigma(A) = 1 - d_{\mathrm{TV}}(\mu, \sigma)$.*

Proposition 1.3.9. *L'espace $\mathcal{P}(E)$ muni de la distance en variation totale est complet.*

Démonstration : Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de Cauchy. Pour tout $n \in \mathbb{N}$, μ_n est absolument continue par rapport à la probabilité $\mu = \sum_{n\in\mathbb{N}} \frac{\mu_n}{2^{n+1}}$ et admet donc une densité $p_n : \mathbb{E} \to \mathbb{R}_+$ par rapport à μ :

$$\forall A \in \mathcal{E},\ \mu_n(A) = \int_E 1_A(x)p_n(x)\mu(dx).$$

Comme le supremum à la seconde ligne est atteint pour $\varphi(x) = \frac{1}{2}\mathrm{signe}(p_n(x) - p_m(x))$,

$$\begin{aligned}
d_{\mathrm{TV}}(\mu_n, \mu_m) &= \sup_{\varphi:|\varphi|\leq 1/2} \left| \int_E \varphi(x)(p_n(x) - p_m(x))\mu(dx) \right| \\
&= \int_E \frac{1}{2}|p_n(x) - p_m(x)|\mu(dx) = \frac{1}{2}\|p_n - p_m\|_{L^1(\mu)},
\end{aligned}$$

la suite $(p_n)_{n\in\mathbb{N}}$ est de Cauchy dans $L^1(\mu)$ qui est complet. Elle converge donc vers p_∞. Comme $\mu(dx)$ p.p. $p_n(x) \geq 0$, $\mu(dx)$ p.p. $\max(-p_\infty(x), 0) \leq |p_n(x) - p_\infty(x)|$. Ainsi $\int_E \max(-p_\infty(x), 0)\mu(dx) \leq \int_E |p_n(x) - p_\infty(x)|\mu(dx)$, puis par passage à la limite $n \to \infty$, $\int_E \max(-p_\infty(x), 0)\mu(dx) = 0$ i.e. $\mu(dx)$ p.p. $p_\infty(x) \geq 0$. Par ailleurs, le passage à la limite $n \to \infty$ dans l'inégalité $| \int_E p_\infty(x)\mu(dx) - \int_E p_n(x)\mu(dx)| \leq \int_E |p_n(x) - p_\infty(x)|\mu(dx)$ assure que $\int_E p_\infty(x)\mu(dx) = 1$. Ainsi p_∞ est une densité de probabilité par rapport à μ. Notons μ_∞ la probabilité qui admet cette densité. Comme précédemment $d_{\mathrm{TV}}(\mu_n, \mu_\infty) = \frac{1}{2}\|p_n - p_\infty\|_{L^1(\mu)}$, où le second membre tend vers 0 lorsque $n \to \infty$. □

1.3.2 Conditions de Doeblin et ergodicité uniforme

Définition 1.3.10. *On dit que le noyau markovien P satisfait*

— *la condition de Doeblin si*

$$\exists \alpha \in]0,1],\ \forall x, y \in E,\ P(x,.) \wedge P(y,.)(E) \geq \alpha.$$

— *la condition de Doeblin uniforme si*

$$\exists \nu \in \mathcal{P}(E),\ \exists \alpha \in]0,1],\ \forall A \in \mathcal{E},\ \inf_{x \in E} P(x, A) \geq \alpha \nu(A).$$

La condition de Doeblin uniforme est plus forte que la condition de Doeblin en ce que pour tous $x, y \in E$, la mesure $P(x,.) \wedge P(y,.)$ domine la même mesure $\alpha\nu$ de masse totale α.

Théorème 1.3.11. *Sous la condition de Doeblin, le noyau markovien P admet une unique probabilité invariante π et*

$$\forall \mu \in \mathcal{P}(E),\ \forall n \in \mathbb{N},\ d_{\mathrm{TV}}(\mu P^n, \pi) \leq (1-\alpha)^n d_{\mathrm{TV}}(\mu, \pi).$$

Remarque 1.3.12. *Comme $\sup_{x \in E} |\varphi(x)| \leq 1/2$ implique que $\forall x \in E$, $|P\varphi(x)| \leq \int_E |\varphi(y)| P(x, dy) \leq 1/2$, on a toujours*

$$d_{\mathrm{TV}}(\mu P, \sigma P) = \sup_{\varphi : |\varphi| \leq 1/2} |\mu(P\varphi) - \sigma(P\varphi)| \leq \sup_{\psi : |\psi| \leq 1/2} |\mu(\psi) - \sigma(\psi)| = d_{\mathrm{TV}}(\mu, \sigma), \qquad (1.4)$$

même sans la condition de Doeblin.

La preuve du théorème repose sur le résultat suivant

Lemme 1.3.13. *Sous la condition de Doeblin, $\forall x, y \in E$, $|P\varphi(x) - P\varphi(y)| \leq 2(1-\alpha) \sup_{z \in E} |\varphi(z)|$ et*

$$\forall \mu, \sigma \in \mathcal{P}(E),\ d_{\mathrm{TV}}(\mu P, \sigma P) \leq (1-\alpha) d_{\mathrm{TV}}(\mu, \sigma). \qquad (1.5)$$

Nous allons en déduire le théorème avant de démontrer le lemme.
Démonstration : D'après (1.5), l'application $\mathcal{P}(E) \ni \mu \mapsto \mu P \in \mathcal{P}(E)$ est contractante pour d_{TV} distance qui rend l'espace $\mathcal{P}(E)$ complet d'après la proposition 1.3.9. D'après le théorème de point fixe de Picard, elle admet un unique point fixe π qui est l'unique probabilité invariante par P. Enfin, en itérant l'inégalité (1.5), on obtient

$$\forall \mu, \sigma \in \mathcal{P}(E),\ \forall n \in \mathbb{N},\ d_{\mathrm{TV}}(\mu P^n, \sigma P^n) \leq (1-\alpha)^n d_{\mathrm{TV}}(\mu, \sigma)$$

et l'inégalité énoncée dans le théorème en découle pour le choix $\sigma = \pi$. $\square$

Démonstration : Pour $\varphi : E \to \mathbb{R}$ une fonction mesurable bornée et $x, y \in E$, on a

$$\begin{aligned}|P\varphi(x) - P\varphi(y)| &= \Big| \int_E \varphi(z)P(x,dz) - \int_E \varphi(z)P(x,.)\wedge P(y,.)(dz) \\ &\qquad + \int_E \varphi(z)P(x,.)\wedge P(y,.)(dz) - \int_E \varphi(z)P(y,dz)\Big| \\ &\le \left|\int_E \varphi(z)(P(x,dz) - P(x,.)\wedge P(y,.)(dz))\right| + \left|\int_E \varphi(z)(P(y,dz) - P(x,.)\wedge P(y,.)(dz))\right| \\ &\le \int_E |\varphi(z)|(P(x,dz) - P(x,.)\wedge P(y,.)(dz)) + \int_E |\varphi(z)|(P(y,dz) - P(x,.)\wedge P(y,.)(dz)) \\ &\le \sup_{z\in E}|\varphi(z)| \left(\int_E (P(x,dz) - P(x,.)\wedge P(y,.)(dz)) + \int_E (P(y,dz) - P(x,.)\wedge P(y,.)(dz))\right) \\ &\le 2(1-\alpha)\sup_{z\in E}|\varphi(z)|,\end{aligned}$$

où on a utilisé la positivité des mesures $P(x,.)-P(x,.)\wedge P(y,.)$ et $P(y,.)-P(x,.)\wedge P(y,.)$ pour la seconde inégalité et la condition de Doeblin pour la quatrième inégalité. En posant $c_\varphi = (1-\alpha)\sup_{z\in E}|\varphi(z)| - \sup_{y\in E} P\varphi(y)$, on en déduit que pour $x \in E$,

$$P\varphi(x) + c_\varphi = (1-\alpha)\sup_{z\in E}|\varphi(z)| + \inf_{y\in E}(P\varphi(x) - P\varphi(y)) \ge -(1-\alpha)\sup_{z\in E}|\varphi(z)|,$$

D'autre part,

$$P\varphi(x) + c_\varphi \le P\varphi(x) + (1-\alpha)\sup_{z\in E}|\varphi(z)| - P\varphi(x) = (1-\alpha)\sup_{z\in E}|\varphi(z)|,$$

si bien que

$$\sup_{x\in E}|P\varphi(x) + c_\varphi| \le (1-\alpha)\sup_{z\in E}|\varphi(z)|.$$

Ainsi

$$\begin{aligned}d_{\mathrm{TV}}(\mu P, \sigma P) &= \sup_{\varphi:|\varphi|\le 1/2}|\mu P(\varphi) - \sigma P(\varphi)| = \sup_{\varphi:|\varphi|\le 1/2}|\mu(P\varphi) - \sigma(P\varphi)| \\ &= \sup_{\varphi:|\varphi|\le 1/2}|\mu(P\varphi + c_\varphi) - \sigma(P\varphi + c_\varphi)| \\ &\le \sup_{\psi:|\psi|\le(1-\alpha)/2}|\mu(\psi) - \sigma(\psi)| = (1-\alpha)d_{\mathrm{TV}}(\mu,\sigma),\end{aligned}$$

où on a utilisé $\mu(c_\varphi) = c_\varphi = \sigma(c_\varphi)$ pour la troisième égalité et le résultat de l'exercice 1.3.4 pour la dernière. □

Proposition 1.3.14. *La condition de Doeblin pour une itérée P^m de P est équivalente à l'ergodicité uniforme, à savoir l'existence de $\pi \in \mathcal{P}(E)$ telle que $\lim_{n\to\infty}\sup_{x\in E} d_{\mathrm{TV}}(P^n(x,.),\pi) = 0$.*

Démonstration : Si P est uniformément ergodique, choisir m assez grand pour que $\sup_{x\in E} d_{\mathrm{TV}}(P^m(x,.),\pi) \le \frac{1-\alpha}{2}$ assure par inégalité triangulaire que

$$\forall x, y \in E,\ 1-\alpha \ge d_{\mathrm{TV}}(P^m(x,.), P^m(y,.)) = 1 - P^m(x,.)\wedge P^m(y,.)(E).$$

Réciproquement, s'il existe $m \geq 1$ t.q. le noyau P^m satisfait la condition de Doeblin, alors le théorème 1.3.11 entraîne que P^m admet une unique probabilité invariante π. Comme $\pi P = \pi P^m P = \pi P P^m$, la probabilité πP est invariante pour P^m. Par unicité, $\pi P = \pi$ et π est invariante pour P. C'est la seule probabilité invariante pour P puisque toute probabilité invariante pour P l'est pour P^m. En combinant l'égalité $\mu P^n = \mu P^{n-\lfloor \frac{n}{m} \rfloor m} P^m \cdots P^m$ où P^m apparaît $\lfloor \frac{n}{m} \rfloor$ fois, l'estimation du théorème appliqué au noyau P^m et (1.4), on obtient que

$$\forall n \in \mathbb{N},\ d_{\mathrm{TV}}(\mu P^n, \pi) = d_{\mathrm{TV}}(\mu P^n, \pi P^n) \leq (1-\alpha)^{\lfloor n/m \rfloor} d_{\mathrm{TV}}(\mu, \pi).$$

Puisque la distance en variation totale est majorée par 1 et que, si δ_x désigne la masse de Dirac en $x \in E$, $\delta_x P^n = P^n(x, .)$, on conclut que $\lim_{n\to\infty} \sup_{x\in E} d_{\mathrm{TV}}(P^n(x, .), \pi) = 0$. $\square$

1.3.3 Algorithme de recuit simulé

Cet algorithme a pour objectif de minimiser une fonction $V : E \to \mathbb{R}$ mesurable. Pour $\beta \geq 0$, on note $Z_\beta = \int_E e^{-\beta V(x)} \lambda(dx)$ où λ est une mesure positive de référence sur $(E, \mathcal{E})$. Dans le cas d'une probabilité de Boltzmann-Gibbs, $\beta = \frac{1}{k_B T}$, c'est-à-dire que β est l'inverse de la température au facteur multiplicatif près $1/k_B$ avec k_B la constante de Boltzmann. On note également $v_\star = \inf\{v \in \mathbb{R} : \lambda(\{x : V(x) \leq v\}) > 0\}$ l'essentiel infimum de V pour la mesure λ et on suppose que $v_\star > -\infty$. On suppose enfin que

$$B := \{\beta \geq 0 : Z_\beta < \infty\} \neq \emptyset.$$

Comme $Z_\beta = e^{-\beta v_\star} \int_E e^{-\beta(V(x)-v_\star)} \lambda(dx)$ où l'intégrale est une fonction décroissante de β, B est un intervalle de la forme $[\underline{\beta}, +\infty[$ ou $]\underline{\beta}, +\infty[$ avec $\underline{\beta} \in [0, +\infty[$. Pour tout $\beta \in B$, $\pi_\beta(dx) = \frac{1}{Z_\beta} e^{-\beta V(x)} \lambda(dx)$ est une probabilité sur $(E, \mathcal{E})$. Notons que pour $\beta \geq 0$, $Z_\beta \leq e^{-\beta v_\star} \lambda(E)$ si bien que lorsque $\lambda(E) < \infty$, alors $B = [0, +\infty[$. Le lemme suivant assure que lorsque $\beta \to \infty$ (i.e. la température tend vers 0), π_β se concentre sur les minimas de V.

Lemme 1.3.15. *Pour tout $\varepsilon > 0$, $\lim_{\beta\to\infty} \pi_\beta\left(\{x \in E : V(x) \geq v_\star + \varepsilon\}\right) = 0$.*

Démonstration : On remarque que

$$\pi_\beta(dx) = \frac{1}{\hat{Z}_\beta} e^{-\beta(V(x)-v_\star)} \lambda(dx) \text{ où } \hat{Z}_\beta = \int_E e^{-\beta(V(x)-v_\star)} \lambda(dx). \tag{1.6}$$

Donc

$$\begin{aligned}
\pi_\beta\left(\{x \in E : V(x) \geq v_\star + \varepsilon\}\right) &= \frac{\int_E 1_{\{V(x) \geq v_\star + \varepsilon\}} e^{-\beta(V(x)-v_\star)} \lambda(dx)}{\int_E e^{-\beta(V(x)-v_\star)} \lambda(dx)} \\
&\leq \frac{\int_E 1_{\{V(x) \geq v_\star + \varepsilon\}} e^{-\beta(V(x)-v_\star)} \lambda(dx)}{e^{-\frac{\beta\varepsilon}{2}} \lambda(\{x : V(x) \leq v_\star + \varepsilon/2\})} \\
&\leq \frac{\int_E 1_{\{V(x) \geq v_\star + \varepsilon\}} e^{-\frac{\beta}{2}(V(x)-v_\star)} \lambda(dx)}{\lambda(\{x : V(x) \leq v_\star + \varepsilon/2\})}.
\end{aligned}$$

où $\lambda(\{x : V(x) \leq v_\star + \varepsilon/2\}) > 0$ par définition de $v_\star$ et, pour la dernière inégalité, on a utilisé que $V(x) - v_\star \geq \varepsilon$ est équivalent à $V(x) - v_\star \geq \frac{\varepsilon}{2} + \frac{V(x)-v_\star}{2}$. Lorsque $\beta \geq 2(\underline{\beta}+1)$, $e^{-(\underline{\beta}+1)(V(x)-v_\star)} \geq 1_{\{V(x)\geq v_\star+\varepsilon\}} e^{-\frac{\beta}{2}(V(x)-v_\star)}$ avec le membre de droite qui tend vers 0 lorsque $\beta \to \infty$. Le théorème de convergence dominée assure que $\lim_{\beta\to\infty} \int_E 1_{\{V(x)\geq v_\star+\varepsilon\}} e^{-\frac{\beta}{2}(V(x)-v_\star)}\lambda(dx) = 0$, ce qui conclut la démonstration. □

Le lemme suivant contrôle la dépendance en le paramètre β de la mesure de Gibbs π_β lorsque l'essentiel supremum $\bar{v} = \sup\{v \in \mathbb{R} : \lambda(\{x \in E : V(x) \geq v\}) > 0\}$ de V pour la mesure λ est fini. Notons que comme $Z_\beta \geq e^{-\beta\bar{v}}\lambda(E)$, si $\bar{v} < \infty$, alors le fait que $Z_\beta < \infty$ pour $\beta > \underline{\beta}$ entraîne que $\lambda(E) < \infty$ si bien que $B = [0, +\infty[$.

Lemme 1.3.16. *On suppose que $\bar{v} < \infty$. Pour tous $\tilde{\beta}, \beta \geq 0$, on a*

$$d_{\mathrm{TV}}(\pi_{\tilde{\beta}}, \pi_\beta) \leq (\bar{v} - v_\star)|\tilde{\beta} - \beta|.$$

Démonstration : Pour fixer les idées, on suppose que $\tilde{\beta} > \beta$ ce qui entraîne que $\hat{Z}_{\tilde{\beta}} \leq \hat{Z}_\beta$. D'après l'exercice 1.3.5 et (1.6),

$$\begin{aligned}
2d_{\mathrm{TV}}(\pi_{\tilde{\beta}}, \pi_\beta) &= \int_E \left| \frac{e^{-\tilde{\beta}(V(x)-v_\star)}}{\hat{Z}_{\tilde{\beta}}} - \frac{e^{-\beta(V(x)-v_\star)}}{\hat{Z}_\beta} \right| \lambda(dx) \\
&= \int_E \left| e^{-\tilde{\beta}(V(x)-v_\star)} \left(\frac{1}{\hat{Z}_{\tilde{\beta}}} - \frac{1}{\hat{Z}_\beta} \right) + \frac{e^{-\tilde{\beta}(V(x)-v_\star)} - e^{-\beta(V(x)-v_\star)}}{\hat{Z}_\beta} \right| \lambda(dx) \\
&\leq \left(\frac{1}{\hat{Z}_{\tilde{\beta}}} - \frac{1}{\hat{Z}_\beta} \right) \int_E e^{-\tilde{\beta}(V(x)-v_\star)}\lambda(dx) + \frac{1}{\hat{Z}_\beta} \int_E \left(e^{-\beta(V(x)-v_\star)} - e^{-\tilde{\beta}(V(x)-v_\star)} \right) \lambda(dx) \\
&= 2\left(1 - \frac{\hat{Z}_{\tilde{\beta}}}{\hat{Z}_\beta} \right).
\end{aligned}$$

Comme $1 - e^{-y} \leq y$ pour $y \geq 0$, on obtient que

$$\begin{aligned}
\hat{Z}_\beta - \hat{Z}_{\tilde{\beta}} &= \int_E (1 - e^{-(\tilde{\beta}-\beta)(V(x)-v_\star)}) e^{-\beta(V(x)-v_\star)}\lambda(dx) \\
&\leq (\tilde{\beta} - \beta) \int_E (V(x) - v_\star) e^{-\beta(V(x)-v_\star)}\lambda(dx) \leq (\tilde{\beta} - \beta)(\bar{v} - v_\star)\hat{Z}_\beta.
\end{aligned}$$

Cette inégalité se récrit $1 - \frac{\hat{Z}_{\tilde{\beta}}}{\hat{Z}_\beta} \leq (\tilde{\beta} - \beta)(\bar{v} - v_\star)$, ce qui permet de conclure. □

On suppose que la densité de proposition q par rapport à la mesure de référence λ de l'algorithme de Metropolis-Hastings est symétrique : $\boxed{\forall x, y \in E,\ q(x,y) = q(y,x)}$. Pour la probabilité cible π_β, on a $\alpha(x,y) = \min(1, e^{-\beta(V(y)-V(x))}) = e^{-\beta(V(y)-V(x))^+}$ (où, par rapport à (1.1), le remplacement de 1 par $e^{-\beta(V(y)-V(x))}$ pour $q(x,y) = 0$ et $V(y) > V(x)$ est sans conséquence d'après le second point dans la remarque 1.2.1). Notons $Q(x,dy) = q(x,y)\lambda(dy)$ le noyau de proposition et pour tout $\beta \geq 0$

$$\begin{aligned}
&P_\beta(x,dy) = e^{-\beta(V(y)-V(x))^+} q(x,y)\lambda(dy) + R_\beta(x)\delta_x(dy) \\
&\text{où } R_\beta(x) = \int_E (1 - e^{-\beta(V(z)-V(x))^+}) q(x,z)\lambda(dz).
\end{aligned}$$

On a $P_0 = Q$ et pour tout $\beta \in B$, P_β est le noyau de Metropolis-Hastings associé à π_β.

Proposition 1.3.17. *On suppose que $\bar{v} < \infty$ (si bien que $B = [0, +\infty[$) et qu'il existe $m \in \mathbb{N}^*$ tel que Q^m satisfait la condition de Doeblin pour la constante $\alpha > 0$. Alors*

$$\kappa := \sup \left\{ v : \int_{E \times E} 1_{\{(V(y)-V(x))^+ \geq v\}} q(x,y) \lambda(dx) \lambda(dy) > 0 \right\} \leq \bar{v} - v_\star,$$

et pour toute constante $h \in]0, \frac{1}{m\kappa}[$ et toute probabilité μ_0 sur E la suite $\mu_n = \mu_0 P_{\beta_1} P_{\beta_2} \dots P_{\beta_n}$ où $(\beta_n = h \ln(n))_{n \geq 1}$ vérifie $\lim_{n \to \infty} d_{\mathrm{TV}}(\mu_n, \pi_{\beta_n}) = 0$ si bien que $\lim_{n \to \infty} \mu_n(\{x \in E : V(x) \geq v_\star + \varepsilon\}) = 0$ pour tout $\varepsilon > 0$.

Exercice 1.3.18. *En remarquant que pour $\varepsilon > 0$, $(V(y) - V(x))^+ \geq \bar{v} - v_\star + \varepsilon \Rightarrow V(y) \geq \bar{v} + \frac{\varepsilon}{2}$ ou $V(x) \leq v_\star - \frac{\varepsilon}{2}$, vérifier que $\kappa \leq \bar{v} - v_\star$.*

Remarque 1.3.19. *Si on note $(X_n)_n$ une chaîne de Markov inhomogène à valeurs dans E telle que $X_0 \sim \mu_0$ et*

$$\forall n \in \mathbb{N},\ \mathbb{P}(X_{n+1} \in A | (X_0, \dots, X_n)) = P_{\beta_{n+1}}(X_n, A)$$

alors μ_n est la loi de X_n. Sous les hypothèses de la proposition, pour $h \in]0, \frac{1}{m\kappa}[$, $\lim_{n \to \infty} \mathbb{P}(V(X_n) \geq v_\star + \varepsilon) = 0$ pour tout $\varepsilon > 0$, c'est-à-dire que la suite $(V(X_n))_n$ converge en probabilité vers $v_\star$.

Notons que $\mu_1(dy) = \mu_0 P_0(dy) = \mu_0 Q(dy) = \int_{x \in E} q(x,y) \mu_0(dx) \lambda(dy)$, si bien que μ_1 possède la densité $f_1(y) = \int_{x \in E} q(x,y) \mu_0(dx)$ par rapport à la mesure de référence λ. Supposons que pour $n \in \mathbb{N}^*$, μ_n possède la densité f_n par rapport à λ, alors $\mu_{n+1}(dy) = \int_{x \in E} e^{-\beta_{n+1}(V(y)-V(x))^+} q(x,y) \mu_n(dx) \lambda(dy) + R_{\beta_{n+1}}(y) f_n(y) \lambda(dy) = f_{n+1}(y) \lambda(dy)$ pour $f_{n+1}(y) = \int_{x \in E} e^{-\beta_{n+1}(V(y)-V(x))^+} q(x,y) \mu_n(dx) + R_{\beta_{n+1}}(y) f_n(y)$. Ainsi, nous avons montré par récurrence que pour tout $n \in \mathbb{N}^*$, μ_n possède une densité f_n par rapport à la mesure de référence λ. Soit $A = \left\{ x \in E : \int_E 1_{\{(V(y)-V(x))^+ > \kappa\}} q(x,y) \lambda(dy) = 0 \right\}$. Par définition de κ, $\lambda(A^c) = 0$. Modifions désormais la définition de P_β en

$$P_\beta(x, dy) = e^{-\beta((V(y)-V(x))^+ \wedge \kappa)} q(x,y) \lambda(dy) + \int_E (1 - e^{-\beta((V(z)-V(x))^+ \wedge \kappa)}) q(x,z) \lambda(dz) \delta_x(dy).$$

Pour tout $x \in A$, la probabilité $P_\beta(x, .)$ est inchangée. Comme $\lambda(A^c) = 0$, pour tous $\beta > 0$ et $\mu \in \mathcal{P}(E)$ qui possède une densité par rapport à λ, la valeur de μP_β n'est pas affectée par cette modification. Ainsi par récurrence sur $n \in \mathbb{N}^*$, l'égalité $\mu_n = \mu_1 P_{\beta_2} \dots P_{\beta_n}$ reste vraie malgré la modification. Comme $P_{\beta_1} = P_0$ est inchangée, on a $\mu_1 = \mu_0 P_{\beta_1}$ si bien que l'égalité $\mu_n = \mu_0 P_{\beta_1} P_{\beta_2} \dots P_{\beta_n}$ reste également vraie. Comme π_β possède une densité par rapport à λ, la modification ne change rien à la valeur de $\pi_\beta P_\beta$ si bien que π_β reste invariante par P_β. L'intérêt de la modification est d'assurer le résultat suivant.

Lemme 1.3.20. *Si pour $m \in \mathbb{N}^*$, Q^m satisfait la condition de Doeblin pour la constante $\alpha \in]0, 1]$, alors pour $\beta_1, \dots, \beta_m \geq 0$, $P_{\beta_1} P_{\beta_2} \dots P_{\beta_m}$ satisfait la condition de Doeblin pour la constante $\alpha e^{-\kappa \sum_{n=1}^m \beta_n}$.*

Démonstration du lemme 1.3.20 : La définition modifiée de P_β entraîne que pour $\beta \geq 0$, $P_\beta \geq e^{-\beta\kappa} Q$ au sens où pour tout $x \in E$, la mesure $e^{-\beta\kappa} Q(x, dy)$ est majorée par la probabilité $P_\beta(x, dy)$. En itérant, on en déduit que $P_{\beta_1} P_{\beta_2} \dots P_{\beta_m} \geq$

$e^{-\kappa\sum_{n=1}^m \beta_n} Q^m$ si bien que pour tous $x, y \in E$, $P_{\beta_1}P_{\beta_2}\dots P_{\beta_m}(x,.) \wedge P_{\beta_1}P_{\beta_2}\dots P_{\beta_m}(y,.) \geq e^{-\kappa\sum_{n=1}^m \beta_n}\left(Q^m(x,.) \wedge Q^m(y,.)\right)$. $\square$

Démonstration de la proposition 1.3.17 : La majoration de κ découle de l'exercice 1.3.18. Comme β_n tend vers $+\infty$ avec n, pour $\varepsilon > 0$ et $A_\varepsilon = \{x \in E : V(x) \geq v_\star + \varepsilon\}$, le lemme 1.3.16 entraîne que $\lim_{n\to\infty} \pi_{\beta_n}(A_\varepsilon) = 0$. Comme $|\pi_{\beta_n}(A_\varepsilon) - \mu_n(A_\varepsilon)| \leq d_{\mathrm{TV}}(\pi_{\beta_n}, \mu_n)$, il suffit de montrer que $d_{\mathrm{TV}}(\pi_{\beta_n}, \mu_n)$ tend vers 0 lorsque $n \to \infty$ pour conclure que $\mu_n(A_\varepsilon)$ aussi.

Pour $n, k \in \mathbb{N}^*$, comme par invariance de π_β par P_β,

$$\forall \ell \in \{1,\dots,k\},\ (\pi_{\beta_{n+\ell}} - \pi_{\beta_{n+\ell-1}})P_{\beta_{n+\ell}}\dots P_{\beta_{n+k}} = \pi_{\beta_{n+\ell}}P_{\beta_{n+\ell+1}}\dots P_{\beta_{n+k}} - \pi_{\beta_{n+\ell-1}}P_{\beta_{n+\ell}}\dots P_{\beta_{n+k}},$$

on a $\pi_{\beta_{n+k}} = \sum_{\ell=1}^k (\pi_{\beta_{n+\ell}} - \pi_{\beta_{n+\ell-1}})P_{\beta_{n+\ell}}\dots P_{\beta_{n+k}} + \pi_{\beta_n}P_{\beta_{n+1}}\dots P_{\beta_{n+k}}$. Comme $\mu_{n+k} = \mu_n P_{\beta_{n+1}}\dots P_{\beta_{n+k}}$, avec la définition de d_{TV}, on en déduit que

$$\begin{aligned}
d_{\mathrm{TV}}(\pi_{\beta_{n+k}}, \mu_{n+k}) \leq& \sum_{\ell=1}^k d_{\mathrm{TV}}(\pi_{\beta_{n+\ell}}P_{\beta_{n+\ell}}\dots P_{\beta_{n+k}}, \pi_{\beta_{n+\ell-1}}P_{\beta_{n+\ell}}\dots P_{\beta_{n+k}}) \\
&+ d_{\mathrm{TV}}(\pi_{\beta_n}P_{\beta_{n+1}}\dots P_{\beta_{n+k}}, \mu_n P_{\beta_{n+1}}\dots P_{\beta_{n+k}}) \\
\leq& \sum_{\ell=1}^k d_{\mathrm{TV}}(\pi_{\beta_{n+\ell}}, \pi_{\beta_{n+\ell-1}}) + d_{\mathrm{TV}}(\mu_n, \pi_{\beta_n}) \qquad (1.7) \\
\leq& (\bar{v} - v_\star)h\sum_{\ell=1}^k (\ln(n+\ell) - \ln(n+\ell-1)) + d_{\mathrm{TV}}(\mu_n, \pi_{\beta_n}) \\
=& (\bar{v} - v_\star)h\ln\left(\frac{n+k}{n}\right) + d_{\mathrm{TV}}(\mu_n, \pi_{\beta_n})
\end{aligned}$$

où on a utilisé (1.4) pour la seconde inégalité puis le lemme 1.3.16 pour la troisième. On pose $\beta_0 = 0$ si bien que l'inégalité reste vraie pour $n = 0$ avec $\ln\left(\frac{n+k}{n}\right)$ remplacé par $\ln(k)$ au second membre. Pour $n = \ell m$ avec $\ell \in \mathbb{N}$ et $k \in \{1,\dots,m-1\}$, cela implique que $d_{\mathrm{TV}}(\pi_{\beta_{\ell m+k}}, \mu_{\ell m+k}) \leq (\bar{v} - v_\star)h\ln\left(\frac{\ell m+m-1}{(\ell m)\vee 1}\right) + d_{\mathrm{TV}}(\pi_{\beta_{\ell m}}, \mu_{\ell m})$. Comme $\lim_{\ell\to\infty} \ln\left(\frac{\ell m+m-1}{\ell m}\right) = 0$, il suffit donc de montrer que $d_{\mathrm{TV}}(\pi_{\beta_{\ell m}}, \mu_{\ell m})$ tend vers 0 lorsque $\ell \to \infty$ pour conclure. Pour $k = m$ on peut améliorer l'inégalité (1.7) en remarquant que d'après les lemmes 1.3.20 et 1.3.13,

$$\begin{aligned}
d_{\mathrm{TV}}(\pi_{\beta_n}P_{\beta_{n+1}}\dots P_{\beta_{n+m}}, \mu_n P_{\beta_{n+1}}\dots P_{\beta_{n+m}}) &\leq (1 - \alpha e^{-\kappa\sum_{\ell=1}^m \beta_{n+\ell}})d_{\mathrm{TV}}(\pi_{\beta_n}, \mu_n) \\
&\leq (1 - \alpha e^{-m\kappa\beta_{n+m}})d_{\mathrm{TV}}(\pi_{\beta_n}, \mu_n)
\end{aligned}$$

si bien que

$$d_{\mathrm{TV}}(\pi_{\beta_{n+m}}, \mu_{n+m}) \leq (\bar{v} - v_\star)h\ln\left(\frac{n+m}{n\vee 1}\right) + (1 - \alpha e^{-m\kappa\beta_{n+m}})d_{\mathrm{TV}}(\mu_n, \pi_{\beta_n}).$$

En posant pour $\ell \in \mathbb{N}$, $z_\ell = d_{\mathrm{TV}}(\pi_{\beta_{\ell m}}, \mu_{\ell m})$, $\alpha_{\ell+1} = \alpha e^{-m\kappa\beta_{(\ell+1)m}} = \alpha((\ell+1)m)^{-m\kappa h}$ et $b_{\ell+1} = (\bar{v} - v_\star)h\ln\left(\frac{(\ell+1)m}{(\ell m)\vee 1}\right)$ on en déduit que

$$z_{\ell+1} \leq (1 - \alpha_{\ell+1})z_\ell + b_{\ell+1}.$$

Pour $\ell \to \infty$, on a $\alpha_{\ell+1} \sim \alpha(m\ell)^{-m\kappa h}$ et comme pour $\ell \geq 1$, $b_{\ell+1} = (\bar{v} - v_\star)h \ln\left(1 + \frac{1}{\ell}\right)$, $b_{\ell+1} \sim (\bar{v} - v_\star)h/\ell$. Si $m\kappa h \leq 1$, alors $\sum_{\ell \geq 1} \alpha_\ell = \infty$ et si $m\kappa h < 1$, alors $\frac{b_{\ell+1}}{\alpha_{\ell+1}} \sim \frac{1}{\alpha}(\bar{v} - v_\star)hm^{m\kappa h}\ell^{m\kappa h - 1}$ si bien que $\lim_{\ell\to\infty} \frac{b_\ell}{\alpha_\ell} = 0$. Avec le lemme suivant on conclut que si $h < \frac{1}{m\kappa}$ alors $z_\ell = d_{\mathrm{TV}}(\mu_{\ell m}, \pi_{\beta_{\ell m}})$ tend vers 0 lorsque $\ell \to \infty$ et donc $d_{\mathrm{TV}}(\mu_n, \pi_{\beta_n})$ tend également vers 0 lorsque $n \to \infty$.

□

Lemma 1.3.21. *Soit $(z_\ell)_{\ell\in\mathbb{N}}$ une suite de $\mathbb{R}_+$ vérifiant*

$$\forall \ell \in \mathbb{N},\ z_{\ell+1} \leq (1 - \alpha_{\ell+1})z_\ell + b_{\ell+1}. \tag{1.8}$$

avec $(\alpha_\ell)_{\ell\geq1}$ suite de $]0,1[$ telle que $\sum_{\ell\geq1}\alpha_\ell = \infty$ et $(b_\ell)_{\ell\geq1}$ suite de $\mathbb{R}_+$ telle que $\lim_{\ell\to\infty} \frac{b_\ell}{\alpha_\ell} = 0$. Alors $\lim_{\ell\to\infty} z_\ell = 0$.

Démonstration : Posons $A_0 = 1$ et pour $\ell \geq 1$, $A_\ell = \prod_{k=1}^{\ell} \frac{1}{1-\alpha_k}$. Notons que $(A_\ell)_{\ell\in\mathbb{N}}$ est une suite croissante. Comme $-\ln(A_\ell) = \sum_{k=1}^{\ell} \ln(1-\alpha_k) \leq -\sum_{k=1}^{\ell}\alpha_k$ et $\sum_{k\geq1}\alpha_k = \infty$, A_ℓ tend vers $+\infty$ avec ℓ. En multipliant (1.8) par $A_{\ell+1}$ et en utilisant que $A_{\ell+1}(1-\alpha_{\ell+1}) = A_\ell$, on obtient que $A_{\ell+1}z_{\ell+1} \leq A_\ell z_\ell + A_{\ell+1}b_{\ell+1}$. Par récurrence, on en déduit que

$$A_\ell z_\ell \leq A_0 z_0 + \sum_{k=1}^{\ell} A_k b_k.$$

Comme $A_0 = 1$ et $\alpha_k A_k = A_k - A_{k-1}$, on obtient que

$$\forall \ell \in \mathbb{N},\ z_\ell \leq \frac{z_0}{A_\ell} + \frac{1}{A_\ell}\sum_{k=1}^{\ell}(A_k - A_{k-1})\frac{b_k}{\alpha_k}.$$

Avec la croissance de la suite $(A_k)_{k\in\mathbb{N}}$ on en déduit que pour $\ell_0 \in \mathbb{N}^*$,

$$\forall \ell \geq \ell_0,\ z_\ell \leq \frac{1}{A_\ell}\left(z_0 + \sum_{k=1}^{\ell_0}(A_k - A_{k-1})\frac{b_k}{\alpha_k}\right) + \max_{k>\ell_0}\frac{b_k}{\alpha_k}.$$

Comme le second terme est arbitrairement petit pour ℓ_0 assez grand tandis qu'à ℓ_0 fixé le premier terme tend vers 0 lorsque $\ell \to \infty$, on conclut que $\lim_{\ell\to\infty} z_\ell = 0$. □

En pratique, on travaille avec un nombre fini N d'itérations de l'algorithme et on utilise des suites $(\beta_n)_{1\leq n\leq N}$ qui croissent beaucoup plus vite que la suite $h\ln(n)$ (typiquement des puissances positives de n).

Notons que l'hypothèse que Q satisfait la condition de Doeblin est assez restrictive sur le couple (E, λ).

Lemme 1.3.22. *On suppose que $E = \mathbb{R}^d$ muni de la mesure de Lebesgue λ. Soit $Q(x,dz) = q(x,z)dz$ pour $q : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ mesurable, symétrique et telle que $\int_{\mathbb{R}^d} q(x,z)dz = 1$ pour tout $x \in \mathbb{R}^d$. Alors $\sup_{m\in\mathbb{N}^*}\inf_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}(Q^m(x,.)\wedge Q^m(y,.))(\mathbb{R}^d) = 0$ i.e. aucune itérée Q^m de Q ne satisfait la condition de Doeblin.*

Remarque 1.3.23. — *Notons que sous les hypothèses du lemme, pour tout* $\beta \geq 0$, P_β *ne satisfait pas non plus la condition de Doeblin, même si, comme* $B \neq \emptyset$ *et* $\bar{v} < \infty$ *impliquent* $\lambda(E) < \infty$, *ce n'est pas restrictif en vue de la proposition 1.3.17. En effet, pour* $x \neq y$, $P_\beta(x, .)$ *et* $P_\beta(y, .)$ *admettent respectivement les densités* $R_\beta(x)1_{\{z=x\}} + 1_{\mathbb{R}^d \setminus \{x,y\}}(z)e^{-\beta(V(z)-V(x))^+}q(x, z)$ *et* $R_\beta(y)1_{\{z=y\}} + 1_{\mathbb{R}^d \setminus \{x,y\}}(z)e^{-\beta(V(z)-V(y))^+}q(y, z)$ *par rapport à la mesure de référence* $\lambda(dz) + \delta_x(dz) + \delta_y(dz)$. *Et, d'après l'exercice 1.3.5,*

$$
\begin{aligned}
(P_\beta(x, .) \wedge P_\beta(y, .))(\mathbb{R}^d) &= \int_{\mathbb{R}^d} (e^{-\beta(V(z)-V(x))^+}q(x, z)) \wedge (e^{-\beta(V(z)-V(y))^+}q(y, z))dz \\
&\leq \int_{\mathbb{R}^d} q(x, z) \wedge q(y, z)dz = (Q(x, .) \wedge Q(y, .))(\mathbb{R}^d).
\end{aligned}
$$

— *Toujours sous les hypothèses du lemme, pour le noyau* P *de l'algorithme de Metropolis-Hastings sur* $\mathbb{R}^d$ *défini par* (1.2), *un raisonnement analogue entraîne que pour* $y \neq x$,

$$
\begin{aligned}
(P(x, .) \wedge P(y, .))(\mathbb{R}^d) &= \int_{\mathbb{R}^d} (\alpha(x, z)q(x, z)) \wedge (\alpha(y, z)q(y, z))dz \\
&\leq \int_{\mathbb{R}^d} q(x, z) \wedge q(y, z)dz = (Q(x, .) \wedge Q(y, .))(\mathbb{R}^d)
\end{aligned}
$$

si bien que la condition de Doeblin n'est pas satisfaite par le noyau P.

— *Si, en revanche,* E *est un compact de* $\mathbb{R}^d$ *muni de la mesure de Lebesgue* λ, *pour la fonction symétrique* $q(x, y) = \frac{1}{\lambda(E)}$, *le noyau* Q *satisfait la condition de Doeblin.*

Démonstration : Notons que pour $m \geq 2$, $Q^m(x, dy) = q_m(x, y)dy$ où, en utilisant la symétrie de q pour la seconde égalité,

$$
\begin{aligned}
q_m(x, y) &= \int_{\mathbb{R}^{(m-1)d}} q(x, z_1)q(z_1, z_2) \dots q(z_{m-2}, z_{m-1})q(z_{m-1}, y)dz_1 \dots dz_{m-1} \\
&= \int_{\mathbb{R}^{(m-1)d}} q(y, z_{m-1})q(z_{m-1}, z_{m-2}) \dots q(z_2, z_1)q(z_1, x)dz_{m-1} \dots dz_1 = q_m(y, x).
\end{aligned}
$$

Le noyau Q^m possèdant lui aussi une densité symétrique, il suffit de montrer que $\inf_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}(Q(x, .) \wedge Q(y, .))(\mathbb{R}^d) = 0$.

Supposons que $\forall x, y \in \mathbb{R}^d$, $(Q(x, .) \wedge Q(y, .))(\mathbb{R}^d) \geq \alpha > 0$ et obtenons une contradiction. D'après l'exercice 1.3.5, on a $(Q(x, .) \wedge Q(y, .))(\mathbb{R}^d) = \int_{\mathbb{R}^d} q(x, z) \wedge q(y, z)dz$. Fixons $x \in \mathbb{R}^d$. Comme, par convergence dominée, $\lim_{M\to\infty} \int 1_{\{|z-x|\geq M\}}q(x, z)dz = 0$, il existe $M_x \in]0, +\infty[$ tel que $\int 1_{\{|z-x|\geq M_x\}}q(x, z)dz \leq \frac{\alpha}{2}$. Alors

$$
\begin{aligned}
\forall y \in \mathbb{R}^d, \int_{\mathbb{R}^d} 1_{\{|z-x|<M_x\}}q(y, z)dz &\geq \int_{\mathbb{R}^d} 1_{\{|z-x|<M_x\}}q(x, z) \wedge q(y, z)dz \\
&\geq \int_{\mathbb{R}^d} q(x, z) \wedge q(y, z)dz - \int_{\mathbb{R}^d} 1_{\{|z-x|\geq M_x\}}q(x, z)dz \geq \frac{\alpha}{2}.
\end{aligned}
$$

Donc $\int_{\mathbb{R}^d\times\mathbb{R}^d} 1_{\{|z-x|<M_x\}}q(y, z)dzdy = \infty$ alors que, par la symétrie de q et le théorème de Fubini, cette intégrale est aussi égale à $\int_{\mathbb{R}^d} 1_{\{|z-x|<M_x\}} \int_{\mathbb{R}^d} q(z, y)dydz = \int_{\mathbb{R}^d} 1_{\{|z-x|<M_x\}}dz < \infty$. □

Cela a motivé l'étude de conditions plus locales qui font l'objet du paragraphe suivant.

1.3.4 Ergodicité géométrique sous condition de dérive

Ce chapître est inspiré de [35]. Nous allons travailler sous la condition de dérive suivante :

$$\text{(D1) } \exists V : E \to \mathbb{R}_+ \text{ mesurable, } \exists K \in \mathbb{R}_+,\ \exists \gamma \in]0,1[,\ \forall x \in E,\ PV(x) \leq \gamma V(x) + K$$

$$\text{(D2) } \exists R > \frac{2K}{1-\gamma},\ \exists \alpha \in]0,1],\ \forall x,y \text{ t.q. } V(x)+V(y) \leq R,\ P(x,.) \wedge P(y,.)(E) \geq \alpha.$$

Notons que la condition de Doeblin implique la condition de dérive pour le choix $V \equiv 0$.

Notons également que (D2) est une conséquence de

$$\exists R > \frac{2K}{1-\gamma},\ \exists \nu \in \mathcal{P}(E),\ \exists \alpha \in]0,1],\ \forall A \in \mathcal{E},\ \inf_{x:V(x)\leq R} P(x,A) \geq \alpha\nu(A)$$

où l'ensemble $\{x \in E : V(x) \leq R\}$ est appelé “small set” dans la littérature (attention “petite set” désigne une notion reliée mais légèrement plus faible).

La fonction V qui porte le nom de fonction de Lyapunov tend en général vers l'infini à l'infini lorsque $E = \mathbb{R}^d$ (dans l'application à l'algorithme de Metropolis-Hastings par marche aléatoire du paragraphe 1.3.5, $V(x) = \eta^{-1/2}(x)$ où η est la densité cible). La condition (D1) signifie alors qu'en moyenne le noyau P rapproche de l'origine et donc du compact où une condition de Doeblin locale est satisfaite d'après (D2). L'inégalité $R > 2K/(1-\gamma)$ signifie que ce compact (d'autant plus grand que R l'est) doit être suffisamment grand par rapport à la force de rappel en moyenne vers l'origine (d'autant plus petite que K et γ sont grands). Plus généralement la condition de Doeblin est satisfaite localement sur l'ensemble de niveau $\{x : V(x) \leq \frac{R}{2}\}$ tandis que si $V(x) \geq \frac{R}{2}$ alors $K < (1-\gamma)V(x)$ si bien que $PV(x) \leq \gamma V(x) + K < V(x)$ et en moyenne V diminue par application du noyau P i.e. le noyau P rapproche de cet ensemble de niveau.

Définition 1.3.24. — *Pour $\beta \geq 0$, on note $\|\|_\beta$ la norme sur l'espace vectoriel des fonctions $\varphi : E \to \mathbb{R}$ mesurables et telles que $\sup_{x\in E} \frac{|\varphi(x)|}{1+\beta V(x)} < \infty$ définie par $\|\varphi\|_\beta = \sup_{x\in E} \frac{|\varphi(x)|}{1+\beta V(x)}$.*
- *Pour $\beta > 0$, on lui associe la distance d_β sur $\mathcal{P}_V(E) := \{\mu \in \mathcal{P}(E) : \mu(V) < \infty\}$ définie par $d_\beta(\mu,\sigma) = \sup_{\varphi:\|\varphi\|_\beta \leq 1} |\mu(\varphi) - \sigma(\varphi)|$. Pour $\mu, \sigma \in \mathcal{P}(E)$, on pose toujours $d_\beta(\mu,\sigma) = \sup_{\varphi:\|\varphi\|_\beta \leq 1} |\mu(\varphi) - \sigma(\varphi)|$ où rien ne garantit que le supremum soit fini lorsque V n'est pas bornée.*
- *Pour $\beta = 0$, on lui associe la distance d_0 sur $\mathcal{P}(E)$ définie par $d_0(\mu,\sigma) = \sup_{\varphi:|\varphi|\leq 1} |\mu(\varphi) - \sigma(\varphi)|$.*

Exercice 1.3.25. *Montrer que pour $\beta > 0$, d_β est une distance sur $\mathcal{P}_V(E)$.*

Notons que $\|\|_0$ est la norme du supremum sur l'espace des fonctions mesurables bornées sur E et que, d'après l'exercice 1.3.4, $d_0(\mu,\sigma) = 2d_{\mathrm{TV}}(\mu,\sigma)$. Pour $\beta > 0$, $\|\|_\beta$ est une norme sur l'espace plus grand lorsque $\sup_{x\in E} V(x) = \infty$

$$\mathcal{V} = \left\{\varphi : E \to \mathbb{R} \text{ mesurable t.q. } \sup_{x\in E} \frac{|\varphi(x)|}{1+V(x)} < \infty\right\}.$$

Pour $0 \leq \beta' \leq \beta$, comme $\|\varphi\|_\beta \leq \|\varphi\|_{\beta'}$, $d_{\beta'} \leq d_\beta$.

En reprenant la preuve de la proposition 1.3.9, et en remarquant que

$$\begin{aligned} d_\beta(\mu_n,\mu_m) &= \sup_{\varphi:|\varphi|\leq 1+\beta V}\left|\int_E \varphi(x)(p_n(x)-p_m(x))\mu(dx)\right| \\ &= \int_E (1+\beta V(x))|p_n(x)-p_m(x)|\mu(dx) = \|p_n-p_m\|_{L^1((1+\beta V)\mu)}, \end{aligned}$$

où $(1+\beta V)\mu$ désigne la mesure de densité $(1+\beta V(x))$ par rapport à $\mu(dx)$, on vérifie que

Proposition 1.3.26. *Pour $\beta > 0$, l'espace $\mathcal{P}_V(E)$ muni de d_β est complet.*

Théorème 1.3.27. *On suppose que le noyau markovien P satisfait (D1) et (D2). Alors il admet une unique probabilité invariante π. En outre, $\pi(V) \leq \frac{K}{1-\gamma}$ et*

$$\forall \beta > 0,\ \forall \mu \in \mathcal{P}(E),\ \forall n \in \mathbb{N},\ d_\beta(\mu P^n, \pi) \leq (\chi(\alpha,\beta,\gamma,K,R))^n d_\beta(\mu,\pi). \tag{1.9}$$

où $\chi(\alpha,\beta,\gamma,K,R) = (1-\alpha+\beta K) \vee \frac{2+\beta\gamma R+2\beta K}{2+\beta R} \in]0,1[$ si $\beta \in]0,\frac{\alpha}{K}[$.

Remarque 1.3.28. — *On a*

$$\begin{aligned} (2+\beta R)^2 \frac{d}{d\beta}\frac{2+\beta\gamma R+2\beta K}{2+\beta R} &= (\gamma R+2K)(2+\beta R) - R(2+\beta\gamma R+2\beta K) \\ &= 2(2K-(1-\gamma)R) < 0. \end{aligned}$$

Ainsi $\mathbb{R}_+ \ni \beta \mapsto g(\beta) = (1-\alpha+\beta K) - \frac{2+\beta\gamma R+2\beta K}{2+\beta R}$ est une fonction strictement croissante telle que $g(0) = -\alpha$ et $\lim_{\beta\to\infty} g(\beta) = +\infty$. Donc $\exists! \beta_\star > 0$ tel que $g(\beta_\star) = 0$. Pour $\beta \in [0,\beta_\star[$, $\chi(\alpha,\beta,\gamma,K,R) = \frac{2+\beta\gamma R+2\beta K}{2+\beta R} > \frac{2+\beta_\star\gamma R+2\beta_\star K}{2+\beta_\star R} = 1-\alpha+\beta_\star K$. Pour $\beta \in]\beta_\star,+\infty[$, $\chi(\alpha,\beta,\gamma,K,R) = 1-\alpha+\beta K > 1-\alpha+\beta_\star K$. On conclut que $\inf_{\beta\geq 0}\chi(\alpha,\beta,\gamma,K,R) = \chi(\alpha,\beta_\star,\gamma,K,R)$.

— *Si pour $m \geq 2$, P^m satisfait (D1) et (D2), en raisonnant comme dans la Proposition 1.3.14, on vérifie que P admet une unique probabilité invariante π. Si, en outre, le noyau P lui-même vérifie la condition (D1), alors pour $\varphi \in \mathcal{V}$, la définition de $\|\|_\beta$ entraîne que*

$$\forall x \in E,\ |P\varphi(x)| \leq P|\varphi|(x) \leq \|\varphi\|_\beta P(1+\beta V)(x) \leq \|\varphi\|_\beta(1+\gamma\beta V(x)+\beta K), \tag{1.10}$$

si bien que $\|P\varphi\|_\beta \leq (1+\beta K)\|\varphi\|_\beta$. On en déduit que $d_\beta(\mu P,\sigma P) \leq (1+\beta K)d_\beta(\mu,\sigma)$ et avec l'estimation du théorème pour le noyau P^m que

$$\forall \mu \in \mathcal{P}(E),\ \forall n \in \mathbb{N},\ d_\beta(\mu P^n,\pi) \leq (1+\beta K)^{m-1}(\chi(\alpha,\beta,\gamma,K,R))^{\lfloor n/m\rfloor} d_\beta(\mu,\pi).$$

La preuve du théorème repose sur le résultat suivant

Lemme 1.3.29. *Sous (D1) et (D2) pour toute fonction $\varphi \in \mathcal{V}$,*

$$\forall \beta > 0, \forall x,y \in E,\ |P\varphi(x)-P\varphi(y)| \leq \chi(\alpha,\beta,\gamma,K,R)(2+\beta(V(x)+V(y)))\|\varphi\|_\beta \tag{1.11}$$

$$\forall \mu,\sigma \in \mathcal{P}(E),\ d_\beta(\mu P,\sigma P) \leq \chi(\alpha,\beta,\gamma,K,R)d_\beta(\mu,\sigma). \tag{1.12}$$

Démontrons ce lemme avant de revenir à la preuve du théorème.
Démonstration : Soit $\varphi \in \mathcal{V}$ et $x, y \in E$. On distingue deux cas.

$\boxed{\text{Si } V(x) + V(y) \geq R}$, en utilisant (1.10) à la seconde inégalité, on obtient

$$|P\varphi(x) - P\varphi(y)| \leq |P\varphi(x)| + |P\varphi(y)| \leq (2 + \beta\gamma(V(x) + V(y)) + 2\beta K)\|\varphi\|_\beta.$$

La décroissance de $\mathbb{R}_+ \ni r \to f(r) := \frac{2+\beta\gamma r+2\beta K}{2+\beta r} = \gamma + \frac{2(1-\gamma+\beta K)}{2+\beta r}$ entraîne que

$$\begin{aligned} 2 + \beta\gamma(V(x) + V(y)) + 2\beta K &= \frac{2 + \beta\gamma(V(x) + V(y)) + 2\beta K}{2 + \beta(V(x) + V(y))}(2 + \beta(V(x) + V(y))) \\ &\leq \frac{2 + \beta\gamma R + 2\beta K}{2 + \beta R}(2 + \beta(V(x) + V(y))). \end{aligned}$$

On conclut donc que

$$|P\varphi(x) - P\varphi(y)| \leq \frac{2 + \beta\gamma R + 2\beta K}{2 + \beta R}(2 + \beta(V(x) + V(y)))\|\varphi\|_\beta.$$

$\boxed{\text{Si } V(x) + V(y) \leq R}$, alors en utilisant (D2) à la quatrième inégalité puis (D1) à la cinquième, on obtient

$$\begin{aligned} &|P\varphi(x) - P\varphi(y)| \\ &\leq \left| \int_E \varphi(z)(P(x, dz) - P(x, .) \wedge P(y, .)(dz)) \right| + \left| \int_E \varphi(z)(P(y, dz) - P(x, .) \wedge P(y, .)(dz)) \right| \\ &\leq \|\varphi\|_\beta \Big(\int_E (1 + \beta V(z))(P(x, dz) - P(x, .) \wedge P(y, .)(dz)) \\ &\qquad + \int_E (1 + \beta V(z))(P(y, dz) - P(x, .) \wedge P(y, .)(dz)) \Big) \\ &\leq \|\varphi\|_\beta \Big(P(x, E) - P(x, .) \wedge P(y, .)(E) + \beta PV(x) + P(y, E) - P(x, .) \wedge P(y, .)(E) + \beta PV(y) \Big) \\ &\leq \|\varphi\|_\beta (2(1 - \alpha) + \beta PV(x) + \beta PV(y)) \leq \|\varphi\|_\beta (2(1 - \alpha) + \gamma\beta(V(x) + V(y)) + 2\beta K) \\ &\leq ((1 - \alpha + \beta K) \vee \gamma)(2 + \beta(V(x) + V(y)))\|\varphi\|_\beta. \end{aligned}$$

Comme $\gamma = \lim_{r\to+\infty} f(r) < f(R) = \frac{2+\beta\gamma R+2\beta K}{2+\beta R}$, on a démontré (1.11).

Supposons maintenant $\|\varphi\|_\beta \leq 1$, notons χ à la place de $\chi(\alpha, \beta, \gamma, K, R)$ et posons $c_\varphi = \inf_{y\in E}(\chi(1 + \beta V(y)) - P\varphi(y))$. Pour $x \in E$, on a d'une part, d'après (1.11),

$$P\varphi(x)+c_\varphi = \inf_{y\in E}(P\varphi(x)-P\varphi(y)+\chi(2+\beta(V(x)+V(y)))-\chi(1+\beta V(x)) \geq -\chi(1+\beta V(x)),$$

ce qui assure que $c_\varphi > -\infty$. D'autre part, d'après la définition de c_φ,

$$P\varphi(x) + c_\varphi \leq P\varphi(x) + \chi(1 + \beta V(x)) - P\varphi(x) = \chi(1 + \beta V(x)).$$

Ainsi $\|P\varphi + c_\varphi\|_\beta \leq \chi$. On conclut que

$$\begin{aligned} d_\beta(\mu P, \sigma P) = \sup_{\varphi:\|\varphi\|_\beta\leq 1} |\mu P(\varphi) - \sigma P(\varphi)| &= \sup_{\varphi:\|\varphi\|_\beta\leq 1} |\mu(P\varphi + c_\varphi) - \sigma(P\varphi + c_\varphi)| \\ &\leq \sup_{\psi:\|\psi\|_\beta\leq\chi} |\mu(\psi) - \sigma(\psi)| = \chi d_\beta(\mu, \sigma). \end{aligned}$$

□

Démontrons maintenant le théorème 1.3.27.

Démonstration : Soit $\beta \in]0, \frac{\alpha}{K}[$. Pour alléger les notations, nous noterons χ à la place de $\chi(\alpha, \beta, \gamma, K, R)$. Montrons que $\chi \in]0, 1[$. La condition $\beta < \alpha/K$ assure que $1-\alpha+\beta K < 1$. Par ailleurs, les conditions $\gamma \in]0, 1[$, $\beta > 0$ et $R > \frac{2K}{1-\gamma}$ assurent que $\beta R > \beta\gamma R + 2\beta K$ si bien que $\frac{2+\beta\gamma R+2\beta K}{2+\beta R} < 1$. Ainsi, sous (D1) et (D2), $\chi \in]0, 1[$.

D'après (D1), pour $\mu \in \mathcal{P}_V(E)$,

$$\mu P(V) = \mu(PV) \leq \mu(\gamma V + K) = \gamma\mu(V) + K \tag{1.13}$$

si bien que $\mu P \in \mathcal{P}_V(E)$. D'après (1.12), l'application $\mathcal{P}_V(E) \ni \mu \mapsto \mu P \in \mathcal{P}_V(E)$ est contractante pour d_β distance qui rend l'espace $\mathcal{P}_V(E)$ complet d'après la proposition 1.3.26. D'après le théorème de point fixe de Picard, elle admet un unique point fixe π qui est l'unique élément de $\mathcal{P}_V(E)$ invariant par P. Pour le choix $\mu = \pi$, (1.13) s'écrit $\pi(V) \leq \gamma\pi(V) + K$ si bien que $\pi(V) \leq \frac{K}{1-\gamma}$.

L'inégalité $d_\beta(\mu P^n, \pi) \leq \chi^n d_\beta(\mu, \pi)$ s'obtient en itérant (1.12) pour le choix $\sigma = \pi$. Pour $x \in E$, comme $d_\beta(\delta_x, \pi) \leq \delta_x(1 + \beta V) + \pi(1 + \beta V) \leq 2 + \beta V(x) + \frac{\beta K}{1-\gamma} < \infty$, on en déduit que $\lim_{n\to\infty} d_\beta(P^n(x, .), \pi) = 0$.

Soit $\tilde{\pi}$ une autre probabilité invariante. On a $d_\beta(\pi, \tilde{\pi}) = d_\beta(\pi P, \tilde{\pi} P) \leq \chi d_\beta(\pi, \tilde{\pi})$ d'après (1.12). On en déduit que $d_\beta(\pi, \tilde{\pi}) = 0$ dès lors que $d_\beta(\pi, \tilde{\pi}) < \infty$, condition assurée par $\tilde{\pi} \in \mathcal{P}_V(E)$. Sans cette condition, il faut trouver un autre argument. Comme $d_\beta \geq d_0 = 2d_{\mathrm{TV}}$, pour tout $x \in E$, $P^n(x, .)$ converge en variation totale vers π et pour $A \in \mathcal{E}$, la suite $(P^n(x, A))_{n\geq 1}$ à valeurs dans $[0, 1]$ converge vers $\pi(A)$. Par convergence dominée, on en déduit que $\tilde{\pi} P^n(A) = \int_E P^n(x, A)\tilde{\pi}(dx)$ converge vers $\int_E \pi(A)\tilde{\pi}(dx) = \pi(A)$ lorsque $n \to \infty$. Comme par invariance de $\tilde{\pi}$, $\tilde{\pi}(A) = \tilde{\pi} P^n(A)$, on conclut que $\tilde{\pi} = \pi$.

□

1.3.5 Application à l'algorithme de Metropolis-Hastings

On se place dans le cas où

- $E = \mathbb{R}^d$ muni de sa tribu borélienne $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$,
- $\lambda(dx) = dx$ est la mesure de Lebesgue sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$,
- $\pi(dx) = \frac{\eta(x)dx}{\int_{\mathbb{R}^d} \eta(y)dy}$ où η est une fonction mesurable de $\mathbb{R}^d$ dans $\mathbb{R}_+$ telle que $\int_{\mathbb{R}^d} \eta(y)dy \in]0, \infty[$.

On note $q(x, y)$ la densité du noyau de proposition par rapport à la mesure de Lebesgue. Notons que le noyau P s'écrit alors

$$P(x, dy) = \alpha(x, y)q(x, y)dy + \left(\int_{\mathbb{R}^d} (1 - \alpha(x, z))q(x, z)dz\right) \delta_x(dy)$$

$$\text{avec } \alpha(x, y) = 1_{\{\eta(x)q(x,y)>0\}} \min\left(\frac{\eta(y)q(y, x)}{\eta(x)q(x, y)}, 1\right) + 1_{\{\eta(x)q(x,y)=0\}}.$$

Le résultat principal de ce paragraphe, énoncé dans la proposition 1.3.35, est que si η est C^1 strictement positive et telle que $\lim_{|x|\to\infty} \frac{x}{|x|} \cdot \nabla \ln \eta(x) = -\infty$, $\limsup_{|x|\to\infty} \frac{x}{|x|} \cdot \frac{\nabla\eta(x)}{|\nabla\eta(x)|} < 0$

et $q(x,y) = \psi(y-x)$ pour une densité ψ paire et telle que $\forall M > 0$, $\inf_{|x|\leq M} \psi(x) > 0$, alors l'algorithme de Metropolis-Hastings par marche aléatoire, noté MHMA dans la suite, est géométriquement ergodique puisqu'il vérifie (D1) et (D2) pour $V(x) = \eta^{-1/2}(x)$. Notons que, d'après la remarque 1.3.23, pour cet algorithme, la condition de Doeblin n'est jamais satisfaite sur $\mathbb{R}^d$. Comme $q(x,y) = \psi(y-x)$ on peut même simplifier la preuve de ce résultat en remarquant que pour $x \neq y \in \mathbb{R}^d$, pour tout $z \in \mathbb{R}^d$, $|x-y| \leq |x-z| + |z-y|$ si bien que $|z-x| \vee |z-y| \geq \frac{|x-y|}{2}$ et $(\alpha(x,z)\psi(z-x)) \wedge (\alpha(y,z)\psi(z-y)) \leq 1_{\{|z-x| \geq \frac{|x-y|}{2}\}}\psi(z-x) + 1_{\{|z-y| \geq \frac{|x-y|}{2}\}}\psi(z-y)$. Ainsi,

$$P(x,.) \wedge P(y,.)(\mathbb{R}^d) = \int_{\mathbb{R}^d} (\alpha(x,z)\psi(z-x)) \wedge (\alpha(y,z)\psi(z-y))dz \leq 2 \int_{|w| \geq \frac{|x-y|}{2}} \psi(w)dw$$

où, par convergence dominée, le membre de droite tend vers 0 lorsque $|x-y| \to \infty$.

Le lemme suivant donne une condition suffisante pour que la condition de Doeblin soit satisfaite lorsque le point de départ x est restreint à un compact.

Lemme 1.3.30. *Si* $\forall M \in]0,+\infty[$, $\underline{q}_M := \inf_{|x| \vee |y| \leq M} q(x,y) > 0$ *et* $\bar{\eta}_M := \sup_{|x| \leq M} \eta(x) < \infty$*, alors pour tout* M *assez grand pour que* $\int_{|z| \leq M} \eta(z)dz > 0$*,*

$$\forall |x| \leq M, \ \forall A \in \mathcal{B}(\mathbb{R}^d), \ P(x,A) \geq \alpha_M \nu_M(A),$$

où $\alpha_M := \frac{\underline{q}_M}{\bar{\eta}_M} \int_{|z| \leq M} \eta(z)dz > 0$ *et* $\nu_M(dy) := \frac{1_{\{|y| \leq M\}}\eta(y)dy}{\int_{|z| \leq M} \eta(z)dz}$ *est une probabilité.*

Remarque 1.3.31. *Les hypothèses sont satisfaites si* η *est continue et* q *est strictement positive et continue.*

Démonstration : En ne tenant pas compte de la contribution liée à la possibilité que l'algorithme partant de x reste en x si la proposition y n'est pas acceptée, on obtient que pour $|x| \leq M$ et $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned} P(x,A) &\geq \int_A 1_{\{\eta(y)>0\}} \left(\frac{\eta(y)q(y,x)}{\eta(x)q(x,y)} \wedge 1 \right) q(x,y)dy = \int_A 1_{\{\eta(y)>0\}} \left(\frac{q(y,x)}{\eta(x)} \wedge \frac{q(x,y)}{\eta(y)} \right) \eta(y)dy \\ &\geq \int_A 1_{\{\eta(y)>0\}} \left(\frac{q(y,x)}{\eta(x)} \wedge \frac{q(x,y)}{\eta(y)} \right) 1_{\{|y| \leq M\}}\eta(y)dy \geq \frac{\underline{q}_M}{\bar{\eta}_M} \int_A 1_{\{y \leq M\}}\eta(y)dy. \end{aligned}$$

Sous nos hypothèses, q ne s'annule pas et on utilise les conventions $\frac{\eta(y)q(y,x)}{\eta(x)q(x,y)} = \infty = \frac{q(y,x)}{\eta(x)}$ si $\eta(x) = 0$. □

Il est plus difficile de vérifier (D1) et nous nous contenterons d'énoncer des conditions obtenues par Jarner et Hansen [39] pour que (D1) soit satisfaite par $V(x) = \eta^{-1/2}(x)$ dans le cas particulier de l'algorithme de Metropolis-Hastings par marche aléatoire où $q(x,y) = \psi(y-x)$ pour une densité de probabilité ψ paire sur $\mathbb{R}^d$.

Vérifions tout d'abord que pour le choix $V(x) = \eta^{-1/2}(x)$, $x \mapsto \frac{PV(x)}{V(x)}$ est borné. Ce résultat est vrai pour un espace d'état E quelconque, dès lors que la densité de proposition est symétrique.

Lemme 1.3.32. *Tout algorithme de Metropolis-Hastings avec une densité de proposition* q *symétrique* ($\forall x,y \in E,\ q(x,y) = q(y,x)$) *est tel que*

$$\forall x \in E \ t.q. \ \eta(x) > 0, \ \frac{P\eta^{-1/2}(x)}{\eta^{-1/2}(x)} \leq \frac{5}{4}.$$

On en déduit que si la fonction $\eta^{-1/2}$ est localement bornée, il en va de même pour la fonction $P\eta^{-1/2}$.

Démonstration : Soit $x \in E$ t.q. $\eta(x) > 0$. Rappelons que partant de x, une proposition y est acceptée avec probabilité $\alpha(x,y) = \min(1, \frac{\eta(y)q(y,x)}{\eta(x)q(x,y)}) = \min(1, \frac{\eta(y)}{\eta(x)})$. Notons $\mathcal{A}_x = \{y \in E : \eta(y) \geq \eta(x)\}$ l'ensemble des propositions qui sont acceptées avec probabilité 1 et $\mathcal{R}_x = \{y \in E : \eta(y) < \eta(x)\}$ son complémentaire. La proposition $y \in \mathcal{R}_x$ est acceptée avec probabilité $\frac{\eta(y)}{\eta(x)}$ et sinon, l'algorithme reste au point x. Donc

$$\begin{aligned}\frac{P\eta^{-1/2}(x)}{\eta^{-1/2}(x)} &= \int_{\mathcal{A}_x} \frac{\eta^{-1/2}(y)}{\eta^{-1/2}(x)} q(x,y)\lambda(dy) + \int_{\mathcal{R}_x} \frac{\eta^{-1/2}(y)}{\eta^{-1/2}(x)} \frac{\eta(y)}{\eta(x)} q(x,y)\lambda(dy) \\ &\quad + \int_{\mathcal{R}_x} \frac{\eta^{-1/2}(x)}{\eta^{-1/2}(x)} \left(1 - \frac{\eta(y)}{\eta(x)}\right) q(x,y)\lambda(dy) \\ &= \int_{\mathcal{A}_x} \frac{\eta^{1/2}(x)}{\eta^{1/2}(y)} q(x,y)\lambda(dy) + \int_{\mathcal{R}_x} \left(\frac{\eta^{1/2}(y)}{\eta^{1/2}(x)} + 1 - \frac{\eta(y)}{\eta(x)}\right) q(x,y)\lambda(dy) \quad (1.14)\end{aligned}$$

Sur $\mathcal{A}_x$, $\frac{\eta^{1/2}(x)}{\eta^{1/2}(y)} \leq 1$ et sur $\mathcal{R}_x$, $\frac{\eta(y)}{\eta(x)} < 1$. Comme $\sup_{u\in[0,1]}(\sqrt{u} + 1 - u) = \frac{5}{4}$, on en déduit que

$$\frac{P\eta^{-1/2}(x)}{\eta^{-1/2}(x)} \leq \int_{\mathcal{A}_x} q(x,y)\lambda(dy) + \frac{5}{4}\int_{\mathcal{R}_x} q(x,y)\lambda(dy) \leq \frac{5}{4}\int_{\mathbb{R}^d} q(x,y)\lambda(dy) = \frac{5}{4}.$$

□

Le résultat de Jarner et Hansen porte sur la classe suivante de lois cibles.

Définition 1.3.33. *La loi π est dite sous-exponentielle si la fonction η est C^1, strictement positive et vérifie $\lim_{|x|\to\infty} \frac{x}{|x|} \cdot \nabla \ln \eta(x) = -\infty$.*

Si π est sous-exponentielle, pour $\beta > 0$, en choisissant M_β t.q. $\sup_{|x|\geq M_\beta} \frac{x}{|x|}.\nabla \ln \eta(x) \leq -\beta$, on obtient que pour $|y| \geq M_\beta$,

$$\ln \eta(y) - \ln \eta\left(\frac{M_\beta y}{|y|}\right) = \int_{M_\beta}^{|y|} \frac{y}{|y|}.\nabla \ln \eta\left(\frac{uy}{|y|}\right) du \leq -\beta(|y| - M_\beta).$$

Donc

$$\forall |y| \geq M_\beta,\ \eta(y) \leq \eta\left(\frac{M_\beta y}{|y|}\right) e^{-\beta(|y|-M_\beta)} \leq e^{\beta M_\beta} \sup_{|x|\leq M_\beta} \eta(x) e^{-\beta|y|}. \quad (1.15)$$

Ainsi la densité $\frac{\eta(y)}{\int_{\mathbb{R}^d} \eta(x)dx}$ de π tend vers 0 plus vite que n'importe quelle exponentielle de $|y|$ lorsque $|y| \to \infty$, ce qui justifie la terminologie sous-exponentielle.

Exercice 1.3.34. *Vérifier que si les fonctions η_1 et η_2 sont proportionnelles aux densités de deux lois sous-exponentielles, alors les lois de densités proportionnelles aux fonctions $\eta_1\eta_2$ et $a_1\eta_1 + a_2\eta_2$ où $a_1, a_2 > 0$ sont sous-exponentielles.*

Dans la suite, nous noterons $\mathcal{S}_{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$ la sphère unité de $\mathbb{R}^d$ et ω_{d-1} la mesure de surface sur cette sphère.

Proposition 1.3.35. *Soit π sous-exponentielle et telle que $\limsup_{|x|\to\infty} \frac{x}{|x|} \cdot \frac{\nabla\eta(x)}{|\nabla\eta(x)|} < 0$.*

(i) *Tout algorithme MHMA avec la densité ψ paire et telle que* $\inf_{\zeta\in\mathcal{S}_{d-1}}\int_0^\infty \psi(r\zeta)r^{d-1}dr > 0$ *vérifie*

$$\exists K \in \mathbb{R}_+,\ \exists\gamma\in]0,1[,\ \forall x\in\mathbb{R}^d,\ P\eta^{-1/2}(x)\le\gamma\eta^{-1/2}(x)+K. \tag{1.16}$$

(ii) *Tout algorithme MHMA avec la densité ψ paire et telle que $\forall M \in]0,+\infty[$, $\inf_{|x|\le M}\psi(x)>0$ est géométriquement ergodique.*

Exemple 1.3.36. *Les hypothèses de la proposition portant sur π sont satisfaites par la loi gaussienne $\mathcal{N}_d(\mu,\Gamma)$ dont la matrice de covariance $\Gamma\in\mathbb{R}^{d\times d}$ est supposée définie positive. En effet, $\nabla\ln\eta(x) = -\Gamma^{-1}(x-\mu)$ si bien que*

$$\frac{x}{|x|}\cdot\nabla\ln\eta(x) = -\frac{x}{|x|}\cdot\Gamma^{-1}x+\frac{x}{|x|}\cdot\Gamma^{-1}\mu\le|\Gamma^{-1}\mu|-|x|\inf_{\xi\in\mathcal{S}_{d-1}}\xi\cdot\Gamma^{-1}\xi$$

tend vers $-\infty$ lorsque $|x|\to\infty$. Ainsi π est sous-exponentielle. Par ailleurs, $\limsup_{|x|\to\infty}\frac{x}{|x|}\cdot\frac{\nabla\eta(x)}{|\nabla\eta(x)|} = -\inf_{\xi\in\mathcal{S}_{d-1}}\frac{\xi\cdot\Gamma^{-1}\xi}{|\xi||\Gamma^{-1}\xi|}$. L'exercice qui suit permet de montrer que $\inf_{\xi\in\mathcal{S}_{d-1}}\frac{\xi\cdot\Gamma^{-1}\xi}{|\Gamma^{-1}\xi|} = \frac{2\sqrt{\underline{\lambda}\bar{\lambda}}}{\underline{\lambda}+\bar{\lambda}}$ où $\underline{\lambda}$ et $\bar{\lambda}$ désignent la plus petite et la plus grande valeurs propres de Γ^{-1}.

Exercice 1.3.37. *On suppose que $M\in\mathbb{R}^{d\times d}$ est une matrice symétrique définie positive. On note $0<\lambda_1\le\lambda_2\le\ldots\le\lambda_d$ ses valeurs propres et on considère une base orthonormée de vecteurs propres de $\mathbb{R}^d$ associés à ces valeurs propres. Pour $\xi\in\mathcal{S}_{d-1}$, on note $a_1,\ldots,a_d$ les coordonnées de ξ dans cette base orthonormée.*

1. *Vérifier que $\left(\frac{|M\xi|}{\xi\cdot M\xi}\right)^2 = \frac{\sum_{i=1}^d\lambda_i^2p_i}{(\sum_{i=1}^d\lambda_ip_i)^2}$ où $(p_i=a_i^2)_{1\le i\le d}$ est une probabilité. Que se passe-t-il si $\lambda_d=\lambda_1$?*
2. *Soit X est une variable aléatoire à valeurs dans $[\lambda_1,\lambda_d]$. Dans le cas où $\lambda_d>\lambda_1$, vérifier que si $q=\mathbb{E}[\frac{X-\lambda_1}{\lambda_d-\lambda_1}]$, on a $q\in[0,1]$, $\lambda_1(1-q)+\lambda_dq=\mathbb{E}[X]$ et $\lambda_1^2(1-q)+\lambda_d^2q=\mathbb{E}[X^2]+\mathbb{E}[(\lambda_d-X)(X-\lambda_1)]$. Conclure que $\frac{\mathbb{E}[X^2]}{\mathbb{E}^2[X]}\le\frac{\lambda_1^2(1-q)+\lambda_d^2q}{(\lambda_1(1-q)+\lambda_dq)^2}$.*
3. *Vérifier que $\sup_{q\in[0,1]}\frac{\lambda_1^2+q(\lambda_d^2-\lambda_1^2)}{(\lambda_1+q(\lambda_d-\lambda_1))^2}=\frac{(\lambda_1+\lambda_d)^2}{4\lambda_1\lambda_d}$.*
4. *Conclure que $\sup_{\xi\in\mathcal{S}_{d-1}}\frac{|M\xi|}{\xi\cdot M\xi}=\frac{(\lambda_1+\lambda_d)}{2\sqrt{\lambda_1\lambda_d}}$.*

Remarque 1.3.38. — *S'il existe $M\in]0,+\infty[$ t.q. $\inf_{|x|\le M}\psi(x)>0$, alors $\inf_{\zeta\in\mathcal{S}_{d-1}}\int_0^\infty\psi(r\zeta)r^{d-1}dr>0$. Ainsi l'hypothèse faite sur ψ dans* **(ii)** *implique celle faite dans* (**i**).

— *Les hypothèses que nous formulons dans la proposition et plus généralement dans tout ce paragraphe ne dépendent pas de la constante de proportionalité entre η et la densité de π : elles sont satisfaites par η si et seulement si elles sont satisfaites par $c\eta$ pour toute constante multiplicative $c\in]0,+\infty[$.*

— *Sous les hypothèses de la proposition, la condition* (1.16) *reste valable avec la fonction $\eta^{-1/2}$ remplacée par η^{-s} où $s\in]0,1[$ est arbitraire.*

La preuve repose sur les deux lemmes suivants.

Lemme 1.3.39. *Soit π telle que la fonction η est C^1, telle que $\nabla\eta$ ne s'annule pas en dehors d'un compact et que $\ell := -\limsup_{|x|\to\infty}\frac{x}{|x|}\cdot\frac{\nabla\eta(x)}{|\nabla\eta(x)|}>0$. On suppose*

$$\exists\gamma\in]0,\ell[\ \text{t.q.}\ \inf_{\xi\in\mathcal{S}_{d-1}}\int_{r=0}^\infty\int_{\zeta\in\mathcal{S}_{d-1}}1_{\{|\zeta-\xi|\le\gamma\}}\psi(r\zeta)\omega_{d-1}(d\zeta)r^{d-1}dr>0. \tag{1.17}$$

alors l'ensemble $\mathcal{A}_x=\{y\in\mathbb{R}^d:\eta(y)\ge\eta(x)\}$ est tel que $\liminf_{|x|\to\infty}\int_{\mathcal{A}_x}\psi(y-x)dy>0$.

Remarque 1.3.40. *Une condition suffisante pour que* (1.17) *soit satisfaite est que* $\inf_{\zeta\in\mathcal{S}_{d-1}}\int_0^\infty \psi(r\zeta)r^{d-1}dr > 0$. *Notons que la fonction* $\zeta \in \mathcal{S}_{d-1} \mapsto \int_0^\infty \psi(r\zeta)r^{d-1}dr$ *s'interprète comme la densité de* $\frac{Z}{|Z|}$ *par rapport à* ω_{d-1} *lorsque* Z *possède la densité* ψ.

Lemme 1.3.41. *Pour* $\varepsilon \in]0,1]$ *et* $x \in \mathbb{R}^d$, *on note*

$$\mathcal{C}_x^\varepsilon = \{y \in \mathbb{R}^d : \varepsilon \leq \frac{\eta(y)}{\eta(x)} \leq \frac{1}{\varepsilon}\}.$$

Si π *est sous-exponentielle, alors* $\limsup_{|x|\to\infty}\int_{\mathcal{C}_x^\varepsilon}\psi(y-x)dy = 0$.

Démontrons la proposition 1.3.35 avant de prouver les deux lemmes.
Démonstration : Montrons d'abord **(ii)** en supposant **(i)**. Comme $q(x,y) = \psi(y-x)$, pour $M \in]0,+\infty[$, $\inf_{|x|\vee|y|\leq M} q(x,y) \geq \inf_{|z|\leq 2M}\psi(z)$. La continuité et la stricte positivité de η permettent d'appliquer le lemme 1.3.30 et d'obtenir que pour tout $M>0$, il existe une constante $\alpha_M > 0$ telle que

$$\forall x,y \in \mathbb{R}^d \text{ t.q. } |x|\vee|y| \leq M,\ P(x,.)\wedge P(y,.) \geq \alpha_M.$$

Comme (1.15) assure que $\lim_{|x|\to\infty}\eta(x) = 0$, on peut choisir M telle que $\sup_{|x|>M}\eta(x) < \frac{(1-\gamma)^2}{4K^2}$ où γ et K sont les constantes qui apparaissent dans (1.16). Alors $R := \inf_{|x|>M}\eta^{-1/2}(x) > \frac{2K}{1-\gamma}$ et $\eta^{-1/2}(x) \leq R \Rightarrow |x| \leq M$. Ainsi la condition de dérive (D1) et (D2) est satisfaite pour $V(x) = \eta^{-1/2}(x)$ et l'algorithme MHMA est géométriquement ergodique.

Il nous reste à démontrer **(i)**. Soit $\varepsilon > 0$. D'après (1.14) et comme $\int_{\mathcal{R}_x}\psi(y-x)dy = 1 - \int_{\mathcal{A}_x}\psi(y-x)dy$,

$$\frac{P\eta^{-1/2}(x)}{\eta^{-1/2}(x)} - 1 = \int_{\mathcal{C}_x^\varepsilon}\left(1_{\mathcal{R}_x}(y)\left(\frac{\eta^{1/2}(y)}{\eta^{1/2}(x)} - \frac{\eta(y)}{\eta(x)}\right) + 1_{\mathcal{A}_x}(y)\frac{\eta^{1/2}(x)}{\eta^{1/2}(y)}\right)\psi(y-x)dy$$
$$+ \int_{\mathcal{R}_x\cap(\mathcal{C}_x^\varepsilon)^c}\left(\frac{\eta^{1/2}(y)}{\eta^{1/2}(x)} - \frac{\eta(y)}{\eta(x)}\right)\psi(y-x)dy + \int_{\mathcal{A}_x\cap(\mathcal{C}_x^\varepsilon)^c}\frac{\eta^{1/2}(x)}{\eta^{1/2}(y)}\psi(y-x)dy - \int_{\mathcal{A}_x}\psi(y-x)dy$$

Pour $y \in \mathcal{R}_x$, $\eta(y) < \eta(x)$ et donc $0 \leq \frac{\eta^{1/2}(y)}{\eta^{1/2}(x)} - \frac{\eta(y)}{\eta(x)} \leq \frac{1}{4}1_{\mathcal{C}_x^\varepsilon}(y) + \sqrt{\varepsilon}1_{(\mathcal{C}_x^\varepsilon)^c}(y)$ (en utilisant $\max_{u\in[0,1]}\sqrt{u}-u = \frac{1}{4}$) tandis que pour $y \in \mathcal{A}_x$, $\eta(y) \geq \eta(x)$ si bien que $0 \leq \frac{\eta^{1/2}(x)}{\eta^{1/2}(y)} \leq 1_{\mathcal{C}_x^\varepsilon}(y) + \sqrt{\varepsilon}1_{(\mathcal{C}_x^\varepsilon)^c}(y)$. Ainsi

$$\begin{aligned}\frac{P\eta^{-1/2}(x)}{\eta^{-1/2}(x)} - 1 &\leq \int_{\mathcal{C}_x^\varepsilon}\psi(y-x)dx + \sqrt{\varepsilon}\int_{(\mathcal{R}_x\cup\mathcal{A}_x)\cap(\mathcal{C}_x^\varepsilon)^c}\psi(y-x)dy - \int_{\mathcal{A}_x}\psi(y-x)dy \\ &\leq \int_{\mathcal{C}_x^\varepsilon}\psi(y-x)dx + \sqrt{\varepsilon} - \int_{\mathcal{A}_x}\psi(y-x)dy.\end{aligned}$$

D'après le lemme 1.3.39 et la remarque 1.3.40, $\liminf_{|x|\to\infty}\int_{\mathcal{A}_x}\psi(y-x)dy > 0$. Pour le choix $\varepsilon = \left(\frac{1}{3}\liminf_{|x|\to\infty}\int_{\mathcal{A}_x}\psi(y-x)dy\right)^2$, le lemme 1.3.41 assure

$$\exists M < \infty,\ \forall |x| \geq M,\ P\eta^{-1/2}(x) \leq \left(1 - \frac{1}{3}\liminf_{|x|\to\infty}\int_{\mathcal{A}_x}\psi(y-x)dy\right)\eta^{-1/2}(x).$$

Posons $\gamma = \left(1 - \frac{1}{3}\liminf_{|x|\to\infty}\int_{\mathcal{A}_x}\psi(y-x)dy\right)$. Pour $|x| \leq M$, d'après le lemme 1.3.32,

$$P\eta^{-1/2}(x) \leq \frac{5}{4}\eta^{-1/2}(x) \leq \gamma\eta^{-1/2}(x) + \left(\frac{5}{4} - \gamma\right)\sup_{|y|\leq M}\eta^{-1/2}(y),$$

où le second membre est fini car η est continue et strictement positive. Donc (1.16) est satisfaite pour $K = \left(\frac{5}{4} - \gamma\right)\sup_{|y|\leq M}\eta^{-1/2}(y)$. □

Remarque 1.3.42. *Soit π sous-exponentielle et ψ paire et t.q. $\forall M \in]0,+\infty[$, $\inf_{|x|\leq M}\psi(x) > 0$. La preuve qui précède assure que l'algorithme MHMA est géométriquement ergodique dès lors que $\liminf_{|x|\to\infty}\int_{\mathcal{A}_x}\psi(y-x)dy > 0$.*
Par ailleurs, comme $P(x,\{x\}) = \int_{\mathcal{R}_x}\left(1 - \frac{\eta(y)}{\eta(x)}\right)\psi(y-x)dy$, on a

$$\begin{aligned}
P(x,\{x\}) - 1 + \int_{\mathcal{A}_x}\psi(y-x)dy &= -\int_{\mathcal{R}_x\cap\mathcal{C}_x^\varepsilon}\frac{\eta(y)}{\eta(x)}\psi(y-x)dy \\
&\geq -\int_{\mathcal{R}_x\cap\mathcal{C}_x^\varepsilon}\psi(y-x)dy - \varepsilon\int_{\mathcal{R}_x\cap(\mathcal{C}_x^\varepsilon)^c}\psi(y-x)dy \\
&\geq -\int_{\mathcal{C}_x^\varepsilon}\psi(y-x)dy - \varepsilon.
\end{aligned}$$

Avec le lemme 1.3.41 qui s'applique comme π est sous-exponentielle, on en déduit que $\lim_{|x|\to\infty}\left(P(x,\{x\}) - 1 + \int_{\mathcal{A}_x}\psi(y-x)dy\right) = 0$. Le théorème 5.1 p103 [67], assure que si $\limsup_{|x|\to\infty}P(x,\{x\}) = 1$, l'algorithme MHMA n'est pas géométrique ergodique. Donc $\liminf_{|x|\to\infty}\int_{\mathcal{A}_x}\psi(y-x)dy > 0$ est une condition nécessaire et suffisante d'ergodicité géométrique dans ce cadre (voir le théorème 4.1 p349 [39]).

Démontrons maintenant le lemme 1.3.39.
Démonstration : Soit $\gamma \in]0,\ell[$ t.q. $\inf_{\xi\in\mathcal{S}_{d-1}}\int_{r=0}^{\infty}\int_{\zeta\in\mathcal{S}_{d-1}}1_{\{|\zeta-\xi|\leq\gamma\}}\psi(r\zeta)\omega_{d-1}(d\zeta)r^{d-1}dr > 0$. Comme par convergence dominée, $\lim_{K\to\infty}\int_{r=K}^{\infty}\int_{\zeta\in\mathcal{S}_{d-1}}\psi(r\zeta)\omega_{d-1}(d\zeta)r^{d-1}dr = 0$ on peut choisir $K > 0$ tel que $\inf_{\xi\in\mathcal{S}_{d-1}}\int_{r=0}^{K}\int_{\zeta\in\mathcal{S}_{d-1}}1_{\{|\zeta-\xi|\leq\gamma\}}\psi(r\zeta)\omega_{d-1}(d\zeta)r^{d-1}dr > 0$.

Pour $x \in \mathbb{R}^d$ t.q. $|x| > 0$, soit $\mathcal{W}_x = \{x + r\zeta :\ 0 < r \leq K$ et $|\zeta - \frac{x}{|x|}| \leq \gamma\}$. Comme $\int_{\mathcal{W}_x}\psi(y-x)dy = \int_{r=0}^{K}\int_{\zeta\in\mathcal{S}_{d-1}}1_{\{|\zeta-\frac{x}{|x|}|\leq\gamma\}}\psi(r\zeta)\omega_{d-1}(d\zeta)r^{d-1}dr$, il suffit de montrer que $\mathcal{W}_x \in \mathcal{A}_x$ pour $|x|$ assez grand pour conclure.

D'après l'hypothèse sur π,

$$\exists M > 0,\ \forall |y| \geq M,\ \frac{y}{|y|}.\frac{\nabla\eta(y)}{|\nabla\eta(y)|} < -\frac{\ell+\gamma}{2}. \tag{1.18}$$

Par ailleurs, pour $x \in \mathbb{R}^d$ tel que $|x| > K$ et y dans la boule $B(x,K)$ fermée centrée en x et de rayon K ,

$$\left|\frac{y}{|y|} - \frac{x}{|x|}\right| \leq \left|\frac{|x|-|y|}{|x||y|}y\right| + \left|\frac{y-x}{|x|}\right| \leq 2\frac{|y-x|}{|x|} \leq \frac{2K}{|x|}.$$

Soit $|x| \geq \frac{4K}{\ell-\gamma} \vee (M+K)$. Pour $y \in \mathcal{W}_x$, en utilisant la définition de $\mathcal{W}_x$ et les deux inégalités précédentes, on a donc

$$\frac{x-y}{|x-y|}.\frac{\nabla\eta(y)}{|\nabla\eta(y)|} \leq \left|\frac{x-y}{|x-y|} - \frac{x}{|x|}\right| + \left|\frac{x}{|x|} - \frac{y}{|y|}\right| + \frac{y}{|y|}.\frac{\nabla\eta(y)}{|\nabla\eta(y)|} < \gamma + \frac{\ell-\gamma}{2} - \frac{\ell+\gamma}{2} < 0.$$

Pour $z \neq x$ t.q. $\eta(z) = \eta(x)$, comme la fonction $f(t) = \eta(x + t(z - x)) - \eta(x)$ s'annule en $t = 0$ et $t = 1$, le théorème de Rolle assure l'existence de $t \in]0, 1[$ tel que $f'(t) = 0$ i.e. $(z - x).\nabla\eta(x + t(z - x)) = 0$. Ainsi $y = x + t(z - x)$ vérifie $\frac{x-y}{|x-y|}.\frac{\nabla\eta(y)}{|\nabla\eta(y)|} = 0$ si bien que $y \notin \mathcal{W}_x$. Comme, pour $z \in \mathcal{W}_x$, $x + t(z - x) \in \mathcal{W}_x$ pour tout $t \in]0, 1[$, on conclut que $\forall z \in \mathcal{W}_x$, $\eta(z) \neq \eta(x)$. L'inégalité (1.18) écrite au point x assure l'existence de $\varepsilon \in]0, \frac{K}{|x|}]$ tel que $\eta((1 - \varepsilon)x) > \eta(x)$. Soit $y \in \mathcal{W}_x$. Comme $(1 - \varepsilon)x \in \mathcal{W}_x$ et que $\mathcal{W}_x$ est convexe, le segment $[(1 - \varepsilon)x, y]$ est inclus dans $\mathcal{W}_x$. La fonction continue $z \mapsto \eta(z) - \eta(x)$ ne s'annulant pas sur ce segment, $\eta(y) - \eta(x) > 0$, c'est-à-dire que $y \in \mathcal{A}_x$. $\square$

Nous terminerons ce paragraphe par la démonstration du lemme 1.3.41

Démonstration : Soit $\alpha > 0$. On veut montrer que pour $|x|$ grand, $\int_{\mathcal{C}_x^\varepsilon} \psi(y - x)dy \leq \alpha$. On commence par choisir $K \in]0, +\infty[$ t.q. $\int_{|z|>K} \psi(z) \leq \frac{\alpha}{3}$ et on note $B(x, K)$ la boule fermée de $\mathbb{R}^d$ centrée en x et de rayon K. Notons que $\int 1_{\{|y-x|>K\}}\psi(y-x)dy \leq \frac{\alpha}{3}$. Ensuite le théorème de convergence dominée assure que $\int \psi(z)1_{\{\psi(z)>m\}}dz$ converge vers 0 lorsque m tend vers l'infini si bien que l'on peut choisir $m_\alpha > 0$ tel que $\int \psi(z)1_{\{\psi(z)>m_\alpha\}}dz \leq \frac{\alpha}{3}$.

$$\begin{aligned}\int_{\mathcal{C}_x^\varepsilon} \psi(y - x)dy \leq &\int 1_{\{|y-x|>K\}}\psi(y - x)dy + \int 1_{\{\psi(y-x)>m_\alpha\}}\psi(y - x)dy \\ &+ \int_{\mathcal{C}_x^\varepsilon \cap B(x,K)} 1_{\{\psi(y-x)\leq m_\alpha\}}\psi(y - x)dy \\ \leq &\frac{\alpha}{3} + \frac{\alpha}{3} + m_\alpha\lambda(\mathcal{C}_x^\varepsilon \cap B(x, K)). \end{aligned} \tag{1.19}$$

Il suffit donc de majorer $\lambda(\mathcal{C}_x^\varepsilon \cap B(x, K))$ pour $|x|$ grand. Soit $x \in \mathbb{R}^d$ t.q. $|x| > K$. On pose $T_x = \{\xi \in \mathcal{S}_{d-1} : \exists r \in [0, +\infty),\ r\xi \in B(x, K)\}$. En passant en coordonnées polaires puis en utilisant que pour $y \in B(x, K)$, $|x| - K \leq |y| \leq |x| + K$, on obtient

$$\begin{aligned}\lambda(\mathcal{C}_x^\varepsilon \cap B(x, K)) &= \int_{\mathcal{S}_{d-1}} \int_0^\infty 1_{\mathcal{C}_x^\varepsilon \cap B(x,K)}(r\xi)r^{d-1}dr\omega_{d-1}(d\xi) \\ &\leq \int_{T_x} \int_{|x|-K}^{+\infty} 1_{\mathcal{C}_x^\varepsilon}(r\xi)(|x| + K)^{d-1}dr\omega_{d-1}(d\xi)\end{aligned}$$

Pour $\beta > 0$, en raisonnant comme pour obtenir (1.15), on vérifie que pour tout $\xi \in \mathcal{S}_{d-1}$, $\forall r \geq \rho \geq M_\beta$, $\eta(r\xi) \leq e^{-\beta(r-\rho)}\eta(\rho\xi)$. Donc si $|x| - K \geq M_\beta$ et $\xi \in T_x$, en notant $\underline{r}_\xi = \inf\{r \in [|x|-K, |x|+K] : r\xi \in \mathcal{C}_x^\varepsilon\}$, on a $\eta(\underline{r}_\xi\xi) \leq \frac{\eta(x)}{\varepsilon}$ par définition de $\mathcal{C}_x^\varepsilon$ et $\forall r > \underline{r}_\xi - \frac{2\ln\varepsilon}{\beta}$, $\eta(r\xi) \leq e^{-\beta(r-\underline{r}_\xi)}\eta(\underline{r}_\xi\xi) < \varepsilon\eta(x)$ si bien que $r\xi \notin \mathcal{C}_x^\varepsilon$. Donc $\int_{|x|-K}^{+\infty} 1_{\mathcal{C}_x^\varepsilon}(r\xi)dr \leq -\frac{2\ln\varepsilon}{\beta}$. Ainsi,

$$\forall |x| > K + M_\beta,\ \lambda(\mathcal{C}_x^\varepsilon \cap B(x, K)) \leq -\frac{2\ln\varepsilon}{\beta}(|x| + K)^{d-1}\omega_{d-1}(T_x). \tag{1.20}$$

Pour majorer $\omega_{d-1}(T_x)$, on remarque que pour $\xi \in T_x$ et $r_\xi \in [0, +\infty[$ tel que $r_\xi\xi \in B(x, K)$, $|r_\xi\xi| \in [|x| - K, |x| + K]$ si bien que $r_\xi \in [|x| - K, |x| + K]$. Donc pour $r \in [|x| - K, |x| + K]$, $|r - r_\xi| \leq 2K$ et $|x - r\xi| \leq |x - r_\xi\xi| + |(r_\xi - r)\xi| \leq K + 2K \leq 3K$. Ainsi $\{r\xi :\ (r, \xi) \in [|x| - K, |x| + K] \times T_x\} \subset B(x, 3K)$ et

$$\lambda(B(x, 3K)) \geq \omega_{d-1}(T_x) \int_{|x|-K}^{|x|+K} r^{d-1}dr \geq 2K(|x| - K)^{d-1}\omega_{d-1}(T_x).$$

Avec (1.20), on conclut que

$$\forall|x| > K + M_\beta,\ \lambda(\mathcal{C}_x^\varepsilon \cap B(x,K)) \leq -\frac{\ln\varepsilon}{K\beta}\left(\frac{|x|+K}{|x|-K}\right)^{d-1}\lambda(B(0,3K)).$$

Pour $|x| \geq K+1$, $\left(\frac{|x|+K}{|x|-K}\right)^{d-1} \leq (2K+1)^{d-1}$. D'après (1.19), le choix $\beta = -\frac{3m_\alpha \ln\varepsilon}{\alpha K}(2K+1)^{d-1}\lambda(B(0,3K))$ assure que

$$\forall|x| \geq K + (M_\beta \vee 1),\ \int_{\mathcal{C}_x^\varepsilon} \psi(y-x)dy \leq \alpha.$$

□

1.4 Loi forte des grands nombres pour les moyennes ergodiques

La loi des grands nombres que nous allons énoncer repose sur le théorème ergodique de Birkhoff (voir par exemple [16] corollary 6.23 p115) :

Théorème 1.4.1. *Soit $(\Omega,\mathcal{A},\mathbb{P})$ un espace de probabilité et $T:\Omega\to\Omega$ une application mesurable qui préserve la probabilité $\mathbb{P}$ ($\forall A\in\mathcal{A}$, $\mathbb{P}(T^{-1}(A)) = \mathbb{P}(A)$) et qui est ergodique : $\forall A\in\mathcal{A}$, $T^{-1}(A) = A \Rightarrow \mathbb{P}(A)\in\{0,1\}$. Alors, pour toute variable aléatoire Y intégrable, $\mathbb{P}\left(\{\omega : \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}Y(T^k(\omega)) = \mathbb{E}[Y]\}\right) = 1$ où on pose $T^0(\omega) = \omega$ et pour $k\in\mathbb{N}$, $T^{k+1}(\omega) = T(T^k(\omega))$ (c'est-à-dire que T^k est l'itérée k fois de T).*

Soit $(X_n)_{n\in\mathbb{N}}$ chaîne de Markov de noyau de transition P sur $(E,\mathcal{E})$. Pour mettre en valeur le rôle joué par la loi μ de X_0, on note $\mathbb{P}_\mu$ la probabilité sur l'espace Ω sous-jacent. Dans le cas particulier $\mu = \delta_x$ où $x\in E$, on adopte la notation simplifiée $\mathbb{P}_x$ au lieu de $\mathbb{P}_{\delta_x}$. On note $\mathbb{E}_\mu$ (resp. $\mathbb{E}_x$) l'espérance sous $\mathbb{P}_\mu$ (resp. $\mathbb{P}_x$).

Définition 1.4.2. *Une probabilité π invariante par P (i.e. qui vérifie $\pi P = \pi$) est dite extrémale s'il n'existe pas deux probabilités invariantes distinctes π_1 et π_2 et $t\in]0,1[$ tels que $\pi = t\pi_1 + (1-t)\pi_2$.*

Remarque 1.4.3.
- *Si π et σ sont deux probabilités invariantes pour P, alors pour tout $t\in[0,1]$, $(t\pi+(1-t)\sigma)P = t\pi P + (1-t)\sigma P = t\pi + (1-t)\sigma$ si bien que la probabilité $t\pi+(1-t)\sigma$ est invariante. Ainsi l'ensemble des probabilités invariantes pour P est un sous-ensemble convexe de $\mathcal{P}(E)$.*
- *Si π et σ sont deux probabilités invariantes extrémales distinctes pour P, alors $\pi\wedge\sigma(E) = 1 - d_{\mathrm{TV}}(\pi,\sigma) < 1$. Comme π et σ majorent toutes deux $\pi\wedge\sigma$, on obtient que $\pi\wedge\sigma = \pi P\wedge\sigma P \geq (\pi\wedge\sigma)P$. Comme l'égalité $P(x,E) = 1$ valable pour tout $x\in E$ entraîne que $\pi\wedge\sigma(E) = (\pi\wedge\sigma)P(E)$, on en déduit que $\pi\wedge\sigma = (\pi\wedge\sigma)P$. Si $\pi\wedge\sigma(E) > 0$, la probabilité $\frac{\pi\wedge\sigma}{\pi\wedge\sigma(E)}$ est alors invariante par P et elle est distincte soit de π soit de σ. Dans le premier cas, comme, $\pi = \pi - \pi\wedge\sigma + \pi\wedge\sigma = (1-\pi\wedge\sigma(E))\frac{\pi-\pi\wedge\sigma}{1-\pi\wedge\sigma(E)} + \pi\wedge\sigma(E)\frac{\pi\wedge\sigma}{\pi\wedge\sigma(E)}$, la probabilité $\frac{\pi-\pi\wedge\sigma}{1-\pi\wedge\sigma(E)}$ est également invariante par linéarité et distincte de $\frac{\pi\wedge\sigma}{\pi\wedge\sigma(E)}$ car $\pi\neq\frac{\pi\wedge\sigma}{\pi\wedge\sigma(E)}$, ce qui contredit l'extrémalité de π. Dans le second cas, par un raisonnement analogue, on contredit l'extrémalité de σ. Donc $\pi\wedge\sigma(E) = 0$ et il existe $B\in\mathcal{E}$ tel que $\pi(B) = 1 = \sigma(B^c)$, d'après (1.3).*

Proposition 1.4.4. *Soit π une probabilité invariante extrémale pour P. Alors $\mathbb{P}_\pi((X_n)_{n\in\mathbb{N}} \in A) \in \{0,1\}$ pour tout ensemble $A \in \mathcal{E}^{\otimes\mathbb{N}}$ invariant par l'opérateur de décalage $T((x_k)_{k\in\mathbb{N}}) = (x_k)_{k\geq 1}$ au sens où $T^{-1}(A) = A$. En outre, pour toute fonction $f : E \to \mathbb{R}$ mesurable et telle que $\pi(|f|) < \infty$, $\mathbb{P}_\pi\left(\frac{1}{n}\sum_{j=0}^{n-1} f(X_j) \to_{n\to\infty} \pi(f)\right) = 1$.*

Démonstration : Soit A invariant par l'opérateur de décalage et pour $n \in \mathbb{N}$, $Y_n = \mathbb{E}_\pi[1_A((X_k)_{k\in\mathbb{N}})|\mathcal{F}_n]$ où $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$. La martingale $(Y_n)_{n\in\mathbb{N}}$ est fermée par la variable aléatoire $1_A((X_k)_{k\in\mathbb{N}})$ qui est mesurable par rapport à la plus petite tribu contenant $\bigcup_{k\in\mathbb{N}} \mathcal{F}_k$. Elle converge donc $\mathbb{P}_\pi$ p.s. vers $1_A((X_k)_{k\in\mathbb{N}})$.

Comme

$$T^{-1}(A) = \{(x_k)_{k\in\mathbb{N}} \in E^{\mathbb{N}} : T((x_k)_{k\in\mathbb{N}}) \in A\} = \{(x_k)_{k\in\mathbb{N}} \in E^{\mathbb{N}} : (x_k)_{k\geq 1} \in A\},$$

par invariance de A par T, $1_A((x_k)_{k\in\mathbb{N}}) = 1_A((x_k)_{k\geq 1})$ et par récurrence $1_A((x_k)_{k\in\mathbb{N}}) = 1_A((x_k)_{k\geq n})$ pour tout $n \in \mathbb{N}$. Ainsi $Y_n = \mathbb{E}_\pi[1_A((X_k)_{k\geq n})|\mathcal{F}_n]$ et la propriété de Markov assure que $Y_n = \varphi(X_n)$ où la fonction $\varphi : E \to \mathbb{R}$ mesurable est définie par $\varphi(x) = \mathbb{E}_x[1_A((X_k)_{k\geq 0})]$. La suite $(\varphi(X_n))_{n\in\mathbb{N}}$ converge donc $\mathbb{P}_\pi$ p.s. vers $1_A((X_k)_{k\in\mathbb{N}})$ de même que $(Z_n := \sup_{k\geq n} \varphi(X_k))_{n\in\mathbb{N}}$. Comme π est invariante par P, sous $\mathbb{P}_\pi$, $(X_k)_{k\geq n}$ a même loi que $(X_k)_{k\geq 0}$, ce qui assure que la loi de Z_n et celle de $\varphi(X_n)$ ne dépendent pas de n. Ainsi, pour tout $n \in \mathbb{N}$, $\varphi(X_n)$ et Z_n ont même loi que $1_A((X_k)_{k\in\mathbb{N}})$. Comme $\mathbb{P}_\pi(\varphi(X_n) \leq Z_n) = 1 = \mathbb{P}_\pi(Z_{n+1} \leq Z_n)$, cela implique que $\mathbb{P}_\pi(\varphi(X_n) = \sup_{k\geq n} \varphi(X_k)) = 1 = \mathbb{P}_\pi(\sup_{k\geq n} \varphi(X_k) = \sup_{k\geq n+1} \varphi(X_k))$. On en déduit que $\mathbb{P}_\pi$ p.s., la suite $(\varphi(X_n))_{n\in\mathbb{N}}$ est constante et égale à $1_A((X_k)_{k\in\mathbb{N}})$. Pour $B = \varphi^{-1}(\{1\})$, on en déduit que $\mathbb{P}_\pi$ p.s.,

$$1_B(X_0) = 1_B(X_1) = 1_A((X_k)_{k\in\mathbb{N}}),$$

et donc $\pi(B) = \mathbb{P}_\pi((X_n)_{n\in\mathbb{N}} \in A)$.

Supposons que $0 < \pi(B) < 1$. En notant $\pi(.|B)$ et $\pi(.|B^c)$ les probabilités définies sur $(E, \mathcal{E})$ par $\pi(C|B) = \frac{\pi(C\cap B)}{\pi(B)}$ et $\pi(C|B^c) = \frac{\pi(C\cap B^c)}{\pi(B^c)}$ pour tout $C \in \mathcal{E}$, on a

$$\begin{aligned}\pi(C|B) &= \frac{\mathbb{E}_\pi[1_{C\cap B}(X_1)]}{\pi(B)} = \frac{\mathbb{E}_\pi[1_C(X_1)1_B(X_1)]}{\pi(B)} = \frac{\mathbb{E}_\pi[1_C(X_1)1_B(X_0)]}{\pi(B)} \\ &= \int_{E\times E} 1_C(y)P(x,dy)\frac{1_B(x)\pi(dx)}{\pi(B)} = \pi(.|B)P(C).\end{aligned}$$

Ainsi $\pi(.|B)$ est invariante par P et, de manière analogue, $\pi(.|B^c)$ l'est également. Comme $\pi(.|B)$ et $\pi(.|B^c)$ sont distinctes ($\pi(B|B) = 1 = \pi(B^c|B^c)$) et comme $\pi = \pi(B)\pi(.|B) + \pi(B^c)\pi(.|B^c)$, on contredit l'extrémalité de π. Donc $\mathbb{P}_\pi((X_n)_{n\in\mathbb{N}} \in A) = \pi(B) \in \{0,1\}$.

Sur $(E^{\mathbb{N}}, \mathcal{E}^{\otimes\mathbb{N}})$ muni de la loi de $(X_n)_{n\geq 0}$ sous $\mathbb{P}_\pi$, l'opérateur de décalage $T : (x_n)_{n\geq 0} \to (x_n)_{n\geq 1}$ est donc ergodique. Il préserve cette loi par invariance de π pour P. Le théorème ergodique de Birkhoff entraîne alors que pour toute fonction $F : E^{\mathbb{N}} \to \mathbb{R}$ mesurable et telle que $\mathbb{E}_\pi[|F((X_k)_{k\in\mathbb{N}})|] < \infty$, $(\frac{1}{n}\sum_{j=0}^{n-1} F((X_k)_{k\geq j})_{n\geq 1}$ converge $\mathbb{P}_\pi$ p.s. vers $\mathbb{E}_\pi[F((X_k)_{k\in\mathbb{N}})]$ lorsque $n \to \infty$. Dans le cas particulier où $F((x_k)_{k\in\mathbb{N}}) = f(x_0)$ avec $f : E \to \mathbb{R}$ mesurable et telle que $\pi(|f|) < \infty$, on en déduit que $\mathbb{P}_\pi$ p.s., $\frac{1}{n}\sum_{j=0}^{n-1} f(X_j)$ converge vers $\pi(f)$ lorsque $n \to \infty$. □

Exercice 1.4.5. *Soit $(\Omega, \mathcal{A})$ un espace mesurable et $T : \Omega \to \Omega$ une application mesurable.*

1. *Montrer que l'ensemble $\mathcal{P}$ des probabilités $\mathbb{P}$ sur $(\Omega, \mathcal{A})$ qui sont préservées par T ($\forall A \in \mathcal{A}$, $\mathbb{P}(T^{-1}(A)) = \mathbb{P}(A)$) est convexe.*
2. *Montrer que si $\mathbb{P} \in \mathcal{P}$ est extrêmale i.e. ne s'écrit pas $\mathbb{P} = t\mathbb{P}_1 + (1-t)\mathbb{P}_2$ pour deux éléments distincts $\mathbb{P}_1$ et $\mathbb{P}_2$ de $\mathcal{P}$ et $t \in]0,1[$, alors T est ergodique pour $\mathbb{P}$.*

Nous sommes maintenant prêts à énoncer la loi forte des grands nombres sous condition de dérive.

Théorème 1.4.6. *On suppose que le noyau markovien P satisfait (D1) et (D2) et on note π sa probabilité invariante dont l'existence et l'unicité sont assurées par le théorème 1.3.27. Alors pour toute fonction $f : E \to \mathbb{R}$ mesurable et telle que $\pi(|f|) < \infty$,*

$$\forall \mu \in \mathcal{P}(E),\ \mathbb{P}_\mu \left(\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \longrightarrow_{n\to\infty} \pi(f) \right) = 1.$$

Remarque 1.4.7. — *Comme, d'après le théorème 1.3.27, $\pi(V) < \infty$, $\sup_{x\in E} \frac{|f(x)|}{1+V(x)} < \infty$ (i.e. $f \in \mathcal{V}$) est une condition suffisante pour que $\pi(|f|) < \infty$.*
— *Notons que si $f : E \to \mathbb{R}_+$ est mesurable et vérifie $\pi(f) = +\infty$, alors appliquer le théorème avec $f \wedge m$ où $m \in \mathbb{N}$ entraîne que pour tout $x \in E$, $\mathbb{P}_x \left(\liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \geq \pi(f \wedge m)\right) = 1$. Comme, par convergence monotone, $\lim_{m\to\infty} \pi(f \wedge m) = \pi(f) = \infty$, $\mathbb{P}_x \left(\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = +\infty\right) = 1$.*

Démonstration : Comme P admet π comme unique probabilité invariante, celle-ci est extrémale. D'après la proposition 1.4.4, pour $f : E \to \mathbb{R}$ mesurable et telle que $\pi(|f|) < \infty$, $\mathbb{P}_\pi \left(\frac{1}{n} \sum_{j=0}^{n-1} f(X_j) \to_{n\to\infty} \pi(f)\right) = 1$.

Reste à démontrer que la convergence a aussi lieu $\mathbb{P}_\mu$ p.s. où $\mu \in \mathcal{P}(E)$ est quelconque. Comme

$$\mathbb{P}_\mu \left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) = \pi(f) \right) = \int_E \mathbb{P}_x \left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) = \pi(f) \right) \mu(dx),$$

il suffit de montrer qu'elle a lieu $\mathbb{P}_x$ p.s. pour tout $x \in E$. Soit $k \in \mathbb{N}$, comme $\lim_{n\to\infty} \frac{n+k}{n} = 1$ et $\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{k-1} f(X_j) = 0$, $\frac{1}{n} \sum_{j=0}^{n-1} f(X_j)$ et $\frac{1}{n} \sum_{j=k}^{k+n-1} f(X_j) = \frac{n+k}{n} \times \frac{1}{n+k} \sum_{j=0}^{k+n-1} f(X_j) - \frac{1}{n} \sum_{j=0}^{k-1} f(X_j)$ ont mêmes valeurs d'adhérence lorsque $n \to \infty$. Avec la propriété de Markov et en posant $\varphi(y) = \mathbb{P}_y \left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) = \pi(f)\right)$, on en déduit que

$$\begin{aligned}
\mathbb{P}_x\left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) = \pi(f) \right) &= \mathbb{P}_x \left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=k}^{k+n-1} f(X_j) = \pi(f) \right) \\
&= P^k \varphi(x) \geq \pi(\varphi) - |\pi(\varphi) - P^k \varphi(x)| \\
&= \mathbb{P}_\pi \left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) = \pi(f) \right) - |\pi(\varphi - 1/2) - P^k(\varphi - 1/2)(x)| \\
&\geq 1 - \sup_{\psi : |\psi| \leq 1/2} |\pi(\psi) - P^k(x, \psi)| = 1 - d_{\mathrm{TV}}(\pi, P^k(x, .)).
\end{aligned}$$

D'après le théorème 1.3.27,

$$d_{\text{TV}}(P^k(x,.),\pi) = d_{\text{TV}}(\delta_x P^k,\pi) \leq \frac{\chi^k}{2} d_\beta(\delta_x,\pi) \leq \frac{\chi^k}{2}(2+\beta V(x) + \frac{\beta K}{1-\gamma})$$

tend vers 0 lorsque $n \to \infty$ pour $\beta \in]0, \frac{\alpha}{K}[$, ce qui achève la preuve. □

Dans le cas où P satisfait la condition de Doeblin uniforme, la probabilité invariante π est explicite et l'exercice suivant explique comment obtenir la conclusion du théorème 1.4.6 à l'aide de la loi forte des grands nombres usuelle.

Exercice 1.4.8. *On suppose que P satisfait la condition de Doeblin uniforme de la définition 1.3.10. Sous cette condition, $\alpha\nu(dy) = g(x,y)P(x,dy)$ pour une densité g définie sur $E \times E$, mesurable et à valeurs dans $[0,1]$. On se donne $(U_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. suivant la loi uniforme sur $[0,1]$ indépendante de la chaîne de Markov $(X_n)_{n\geq 0}$ de noyau P. On note $\tau_0 = 0$ et pour $k \geq 1$,*

$$\tau_k = \inf\{n > \tau_{k-1} : U_n \leq g(X_{n-1}, X_n)\}.$$

On note Q le noyau markovien défini par $Q(x,dy) = \frac{P(x,dy)-\alpha\nu(dy)}{1-\alpha}$.

1. *Montrer que $\pi = \alpha \sum_{n\in\mathbb{N}}(1-\alpha)^n \nu Q^n$ est l'unique probabilité invariante de P.*

Soit $f : E \to \mathbb{R}$ mesurable et telle que $\pi(|f|) < \infty$.

2. *Pour tout $x \in E$, montrer que pour $\varphi : E \to \mathbb{R}$ mesurable bornée, $\int_E \varphi(y)(1 - g(x,y))P(x,dy) = (1-\alpha)\int_E \varphi(y)Q(x,dy)$.*
3. *Soient $\psi_1 : \mathbb{N}^* \to \mathbb{R}$ et $\psi_2, \psi_3 : \bigcup_{n\in\mathbb{N}^*}\{n\} \times E^n \to \mathbb{R}$ mesurables bornée. En décomposant sur les valeurs de (τ_1, τ_2, τ_3), montrer que*

$$\mathbb{E}\left[\psi_1(\tau_1)\psi_2(\tau_2-\tau_1, X_{\tau_1}, \ldots, X_{\tau_2-1})\psi_3(\tau_3-\tau_2, X_{\tau_2}, \ldots, X_{\tau_3-1})\right]$$
$$= \sum_{n\in\mathbb{N}^*} \alpha(1-\alpha)^{n-1}\psi_1(n) \prod_{i=2}^{3} \sum_{n\in\mathbb{N}^*} \alpha(1-\alpha)^{n-1} \int_{E^n} \psi_i(n, x_1, \ldots, x_n)\nu(dx_1) \prod_{k=1}^{n-1} Q(x_k, dx_{k+1}).$$

   En déduire que les variables aléatoires $\sum_{j=\tau_1}^{\tau_2-1} f(X_j)$ et $\sum_{j=\tau_2}^{\tau_3-1} f(X_j)$ sont i.i.d. intégrables d'espérance commune $\frac{\pi(f)}{\alpha}$. Comment ce résultat se généralise-t-il ?
4. *En déduire que p.s., $\frac{1}{k}\sum_{j=0}^{\tau_k-1} f(X_j) \longrightarrow_{k\to\infty} \frac{\pi(f)}{\alpha}$ et $\frac{\tau_k}{k} \longrightarrow_{k\to\infty} \frac{1}{\alpha}$.*
5. *Pour $n \in \mathbb{N}$, on note $\kappa(n) = \max\{k \in \mathbb{N} : \tau_k \leq n\}$. Montrer que p.s. $\lim_{n\to\infty} \kappa(n) = +\infty$. Lorsque $f \geq 0$, remarquer que*

$$\frac{1}{\tau_{\kappa(n)+1}} \sum_{j=0}^{\tau_{\kappa(n)}-1} f(X_j) \leq \frac{1}{n}\sum_{j=0}^{n-1} f(X_j) \leq \frac{1}{\tau_{\kappa(n)}} \sum_{j=0}^{\tau_{\kappa(n)+1}-1} f(X_j)$$

   et en déduire que p.s., $\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f(X_j) = \pi(f)$. Comment généraliser ce résultat sans supposer f de signe constant ?

1.5 Théorème de la limite centrale pour les moyennes ergodiques

Pour contrôler l'erreur commise dans la loi forte ergodique étudiée au paragraphe précédent, on s'intéresse maintenant au comportement asymptotique lorsque $n \to \infty$

de

$$\xi_n = \sqrt{n}\left(\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) - \pi(f)\right).$$

S'il existe une fonction $F : E \to \mathbb{R}$ mesurable et telle que $\forall x \in E,\ P|F|(x) < \infty$ solution de l'équation de Poisson

$$\forall x \in E,\ F(x) - PF(x) = f(x) - \pi(f), \tag{1.21}$$

alors, pour $n \geq 1$, on peut récrire $\xi_n = \frac{F(X_0)-PF(X_{n-1})}{\sqrt{n}} + \frac{1}{\sqrt{n}} M^F_{n-1}$ où $M^F_{n-1} = \sum_{k=1}^{n-1}(F(X_k) - PF(X_{k-1}))$. Introduisons la filtration $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$. Comme d'après la propriété de Markov, $\mathbb{E}[F(X_k)|\mathcal{F}_{k-1}] = PF(X_{k-1})$, $(M^F_n)_{n\in\mathbb{N}}$ est une martingale locale nulle en 0. On peut alors étudier le comportement asymptotique de $(\xi_n)_{n\geq 1}$ à l'aide du théorème de la limite centrale pour les martingales. C'est la motivation de l'étude qui suit de l'équation de Poisson (1.21). Notons que si F est solution, alors pour toute constante $c \in \mathbb{R}$, $F + c$ est également solution.

Proposition 1.5.1. *On suppose que le noyau Markovien P satisfait (D1) et (D2). Pour $f \in \mathcal{V}$, la série de terme général $(P^n f - \pi(f))_{n\in\mathbb{N}}$ converge normalement dans $\mathcal{V}$ muni de $\|.\|_1$, $F = \sum_{n\in\mathbb{N}}(P^n f - \pi(f))$ est solution de l'équation de Poisson* (1.21) *et si $\tilde{F}$ désigne une autre solution dans $\mathcal{V}$, $F - \tilde{F} = \pi(F - \tilde{F})$.*

Remarque 1.5.2. — *Notons que, pour tout $\beta > 0$, comme pour tout $x \in E$, $\frac{\min(\beta,1)}{1+\beta V(x)} \leq \frac{1}{1+V(x)} \leq \frac{\max(\beta,1)}{1+\beta V(x)}$,*

$$\min(\beta, 1)\|\|_\beta \leq \|\|_1 \leq \max(\beta, 1)\|\|_\beta.$$

La convergence normale a donc lieu simultanément pour toutes les normes $\|\|_\beta$.

— *Comme $F \in \mathcal{V}$, (D1) implique que $\forall x \in E$, $P|F|(x) < \infty$.*

Démonstration : Soit $f \in \mathcal{V}$, $\beta \in]0, \frac{\alpha}{K}[$ et $x \in E$. Avec la définition de d_β et le théorème 1.3.27, on obtient que

$$|P^n f(x) - \pi(f)| \leq d_\beta(P^n(x, .), \pi)\|f\|_\beta \leq \chi^n d_\beta(\delta_x, \pi)\|f\|_\beta, \tag{1.22}$$

avec $\chi \in]0, 1[$ ne dépendant ni de n ni de x. Comme

$$d_\beta(\delta_x, \pi) \leq 1 + \beta V(x) + \pi(1 + \beta V) \leq 2 + \frac{\beta K}{1-\gamma} + \beta V(x) \leq \max(2 + \frac{\beta K}{1-\gamma}, \beta)(1 + V(x)),$$

on obtient que

$$\|P^n f - \pi(f)\|_1 \leq \max(2 + \frac{\beta K}{1-\gamma}, \beta)\chi^n\|f\|_\beta, \tag{1.23}$$

d'où le caractère normalement convergent de la série. Soit $F = \sum_{n\in\mathbb{N}}(P^n f - \pi(f))$. Pour $\varphi \in \mathcal{V}$, $\|P\varphi\|_1 \leq \|\varphi\|_1\|P(1 + V)\|_1 \leq \|\varphi\|_1\|1 + K + \gamma V\|_1 \leq (1 + K)\|\varphi\|_1$ où on a utilisé (D1) pour la seconde inégalité. Ainsi l'application linéaire $\varphi \to P\varphi$ de $\mathcal{V}$ muni de $\|\|_1$ dans lui-même est continue. On en déduit que $PF = \sum_{n\in\mathbb{N}} P(P^n f - \pi(f)) = \sum_{n\in\mathbb{N}}(P^{n+1} f - \pi(f)) = \sum_{n\geq 1}(P^n f - \pi(f))$, d'où $F - PF = P^0 f - \pi(f) = f - \pi(f)$.

Soit $\tilde{F}$ une autre solution de l'équation de Poisson dans $\mathcal{V}$. Alors $G := F - \tilde{F} \in \mathcal{V}$ vérifie $G = PG$ si bien que $G = P^n G$ et $G - \pi(G) = P^n G - \pi(G)$. L'estimation (1.23) reste valable pour f remplacée par $G - \pi(G)$, si bien qu'avec le premier point de la remarque précédente, on obtient

$$\begin{aligned}\|G - \pi(G)\|_1 = \|P^n G - \pi(G)\|_1 &\leq \max(2 + \frac{\beta K}{1-\gamma}, \beta)\chi^n \|G - \pi(G)\|_\beta \\ &\leq \max(2 + \frac{\beta K}{1-\gamma}, \beta)\chi^n \max(\beta^{-1}, 1)\|G - \pi(G)\|_1.\end{aligned}$$

En choisissant n assez grand pour que $\max(2 + \frac{\beta K}{1-\gamma}, \beta)\chi^n \max(\beta^{-1}, 1) < 1$ on conclut que $\|G - \pi(G)\|_1 = 0$, c'est-à-dire que $G = \pi(G)$.

□

L'inégalité de Jensen et (D1) entraînent que

$$P\sqrt{V}(x) \leq \sqrt{PV(x)} \leq \sqrt{\gamma V(x) + K} \leq \sqrt{\gamma}\sqrt{V(x)} + \sqrt{K}.$$

Comme $\sqrt{V(x)} + \sqrt{V(y)} \leq \sqrt{R}$ implique $V(x) + V(y) \leq R$ et $\sqrt{R} > \frac{2\sqrt{K}}{1-\sqrt{\gamma}} \Leftrightarrow R > \frac{4K}{(1-\sqrt{\gamma})^2}$, en remplaçant (V, K, γ) par $(\sqrt{V}, \sqrt{K}, \sqrt{\gamma})$ dans la proposition 1.5.1, on obtient le corollaire suivant :

Corollaire 1.5.3. *Supposons que le noyau markovien P satisfait (D1) et*

$$\text{(D2}'\text{)}\ \exists R > \frac{4K}{(1-\sqrt{\gamma})^2},\ \exists \alpha \in]0, 1],\ \forall x, y\ t.q.\ V(x) + V(y) \leq R,\ P(x, .) \wedge P(y, .)(E) \geq \alpha.$$

Alors pour $f : E \to \mathbb{R}$ mesurable et telle que $f^2 \in \mathcal{V}$, $F = \sum_{n\in\mathbb{N}}(P^n f - \pi(f))$ est solution de l'équation de Poisson (1.21) et vérifie $F^2 \in \mathcal{V}$.

Remarque 1.5.4.
- *Comme $\gamma \in]0, 1[$, $0 < \gamma < \sqrt{\gamma} < 1$ si bien que $(1 - \sqrt{\gamma})^2 < 1 - \sqrt{\gamma} < 1 - \gamma$, $\frac{2K}{1-\gamma} < \frac{2K}{(1-\sqrt{\gamma})^2}$ et (D2') implique (D2).*
- *La convergence de la série de terme général $(P^n f)_{n\in\mathbb{N}}$ est normale pour toute norme de la forme $\|g\| = \sup_{x\in E} \frac{|g(x)|}{1+\beta\sqrt{V(x)}}$ avec $\beta > 0$.*
- *D'après la preuve de la proposition 1.5.1 (voir en particulier (1.22)), pour tout $\beta \in]0, \frac{\alpha}{\sqrt{K}}[$ et tout $n \in \mathbb{N}$,*

$$\sup_{x\in E} \frac{|P^n f(x) - \pi(f)|}{1 + \sqrt{V(x)}} \leq \max(2 + \frac{\beta\sqrt{K}}{1-\sqrt{\gamma}}, \beta)\chi^n \sup_{x\in E} \frac{|f(x)|}{1 + \beta\sqrt{V(x)}} \tag{1.24}$$

où $\chi = \chi(\alpha, \beta, \sqrt{\gamma}, \sqrt{K}, \sqrt{R}) \in]0, 1[$ d'après le théorème 1.3.27

Exercice 1.5.5. *On suppose que P satisfait la condition de Doeblin uniforme de la définition 1.3.10 et on note Q le noyau markovien défini par $Q(x, dy) = \frac{P(x,dy)-\alpha\nu(dy)}{1-\alpha}$. Montrer que pour $f : E \to \mathbb{R}$ mesurable bornée, $F = \sum_{n\in\mathbb{N}}(1 - \alpha)^n Q^n f$ est solution de l'équation de Poisson $F - PF = f - \pi(f)$.*

Théorème 1.5.6. *Soit $(X_n)_{n\geq 0}$ chaîne de Markov dont le noyau P satisfait (D1) et (D2'). Alors pour toute fonction $f : E \to \mathbb{R}$ mesurable et telle que $\sup_{x\in E} \frac{f^2(x)}{1+V(x)} < \infty$ et quelle que soit la loi $\mu \in \mathcal{P}(E)$ de X_0, la suite $\left(\xi_n = \sqrt{n}\left(\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) - \pi(f)\right)\right)_{n\geq 1}$ converge en loi vers $\mathcal{N}_1(0, \sigma^2(f))$ où $\sigma^2(f) = \pi(F^2) - \pi((PF)^2)$ avec F solution de l'équation de Poisson (1.21) dont l'existence est assurée par le corollaire 1.5.3.*

Remarque 1.5.7. — *Notons que comme $F^2 \in \mathcal{V}$ et $\pi(1+V) \leq 1 + \frac{K}{1-\gamma}$ (d'après le théorème 1.3.11), $\pi(F^2) < \infty$. Par ailleurs l'inégalité de Jensen et l'invariance de π par P assurent que $\pi((PF)^2) \leq \pi(PF^2) = \pi(F^2)$ si bien que $0 \leq \sigma^2(f) < \infty$.*

— *Notons que $F^2 - (PF)^2 = (F - PF)(F - PF + 2PF) = (f - \pi(f))^2 + 2(f - \pi(f)) \sum_{k\geq 1}(P^k f - \pi(f))$ avec $\sum_{k\in\mathbb{N}} \sup_{x\in E} \frac{|P^k f(x) - \pi(f)|}{1+\sqrt{V(x)}} < \infty$ si bien que $\sum_{k\in\mathbb{N}} \sup_{x\in E} \frac{|(f(x)-\pi(f))(P^k f(x)-\pi(f))|}{1+V(x)} < \infty$. Ainsi la variance asymptotique admet l'expression alternative*

$$\sigma^2(f) = \pi((f-\pi(f))^2) + 2\sum_{k\geq 1} \mathbb{E}_\pi[(f(X_0) - \pi(f))(f(X_k) - \pi(f))],$$

où le premier terme du membre de droite est la variance de $f(X_0)$ et le terme d'indice $k \geq 1$ de la somme est la covariance entre $f(X_k)$ et $f(X_0)$, lorsque X_0 est distribuée suivant la probibilité invariante π.

— *Par ailleurs, par l'inégalité de Jensen,*

$$(F(X_k) - PF(X_{k-1}))^2 \leq 2(F^2(X_k) + (PF(X_{k-1}))^2) \leq 2(F^2(X_k) + PF^2(X_{k-1}))).$$

Donc, pour $x \in E$, $\mathbb{E}_x\left[(F(X_k) - PF(X_{k-1}))^2\right] \leq 4P^k F^2(x)$ où le membre de droite est fini d'après (D1) puisque $F^2 \in \mathcal{V}$. Ainsi, sous $\mathbb{P}_x$, $(M_n^F)_{n\geq 1}$ est une martingale de carré intégrable.

Notons que $< M^F >_{n-1} = \sum_{k=0}^{n-2} \mathbb{E}[(F(X_{k+1}) - PF(X_k))^2|\mathcal{F}_k] = \sum_{k=0}^{n-2}(PF^2(X_k) - (PF(X_k))^2)$ si bien que, par la loi forte des grands nombres ergodique, $\frac{<M^F>_{n-1}}{n}$ converge presque sûrement vers $\pi(PF^2 - (PF)^2) = \sigma^2(f)$ par invariance de π pour P.

Démonstration : Le fait que pour $\mu \in \mathcal{P}(E)$ et $\varphi : \mathbb{R} \to \mathbb{R}$ continue bornée, $\mathbb{E}_\mu[\varphi(\xi_n)] = \int_E \mathbb{E}_x[\varphi(\xi_n)]\mu(dx)$ et le théorème de convergence dominée assurent qu'il suffit de démontrer le résultat sous $\mathbb{P}_x$ avec $x \in E$ quelconque.

Avant de voir comment contrôler l'erreur introduite en approchant F par une fonction bornée, nous allons commencer par montrer le résultat en supposant F bornée par m

On a $\xi_n = \frac{F(X_0)-PF(X_{n-1})}{\sqrt{n}} + \frac{1}{\sqrt{n}} M_{n-1}^F$ où $M_{n-1}^F = \sum_{k=1}^{n-1}(F(X_k) - PF(X_{k-1}))$ et $\frac{F(X_0)-PF(X_{n-1})}{\sqrt{n}}$ converge p.s. vers 0 lorsque $n \to \infty$ car $|F| \leq m$ et $|PF| \leq P|F| \leq m$ entraîne que $\frac{|F(X_0)-PF(X_{n-1})|}{\sqrt{n}} \leq \frac{2m}{\sqrt{n}}$. Le théorème de Slutsky assure qu'il suffit de montrer que $\frac{M_{n-1}^F}{\sqrt{n}}$ converge en loi vers $\mathcal{N}_1(0, \sigma^2(f))$.

Pour cela, utilisons la fonction caractéristique et introduisons $a_k^n = \mathbb{E}_x\left[e^{i\frac{u}{\sqrt{n}}(F(X_k)-PF(X_{k-1}))}|\mathcal{F}_{k-1}\right]$. Comme $|F| \leq m$ et $|PF| \leq P|F| \leq m$,

$$\left|e^{i\frac{u}{\sqrt{n}}(F(X_k)-PF(X_{k-1}))} - 1 - \frac{iu}{\sqrt{n}}(F(X_k) - PF(X_{k-1})) + \frac{u^2}{2n}(F(X_k) - PF(X_{k-1}))^2\right| \leq \frac{4m^3|u|^3}{3n^{3/2}}.$$

En prenant l'espérance conditionnelle du terme entre valeurs absolues et en remarquant que $\mathbb{E}_x[F(X_k) - PF(X_{k-1})|\mathcal{F}_{k-1}] = 0$ et

$$\mathbb{E}_x[(F(X_k) - PF(X_{k-1}))^2|\mathcal{F}_{k-1}] = [PF^2 - (PF)^2](X_{k-1}),$$

on en déduit que

$$\left| a_k^n - 1 + \frac{u^2}{2n}[PF^2 - (PF)^2](X_{k-1}) \right| \leq \frac{4m^3|u|^3}{3n^{3/2}}. \tag{1.25}$$

En particulier, il existe $C \in]0, +\infty[$ telle que $\forall n \geq 1, \forall 1 \leq k \leq n-1,\ |a_k^n| \geq 1 - \frac{C}{n}$. On suppose maintenant $n > C$ si bien que $\left|\frac{1}{\prod_{k=1}^{n-1} a_k^n}\right| \leq \left(1 - \frac{C}{n}\right)^{-n}$.

Comme $\frac{e^{i\frac{u}{\sqrt{n}} M_{n-2}^F}}{\prod_{k=1}^{n-1} a_k^n}$ est $\mathcal{F}_{n-2}$-mesurable,

$$\begin{aligned}
\mathbb{E}_x \left[\frac{e^{i\frac{u}{\sqrt{n}} M_{n-1}^F}}{\prod_{k=1}^{n-1} a_k^n} \right] &= \mathbb{E}_x \left[e^{i\frac{u}{\sqrt{n}}(F(X_{n-1}) - PF(X_{n-2}))} \frac{e^{i\frac{u}{\sqrt{n}} M_{n-2}^F}}{\prod_{k=1}^{n-1} a_k^n} \right] \\
&= \mathbb{E}_x \left[a_{n-1}^n \frac{e^{i\frac{u}{\sqrt{n}} M_{n-2}^F}}{\prod_{k=1}^{n-1} a_k^n} \right] = \mathbb{E}_x \left[\frac{e^{i\frac{u}{\sqrt{n}} M_{n-2}^F}}{\prod_{k=1}^{n-2} a_k^n} \right],
\end{aligned}$$

et, en itérant ce raisonnement, $\mathbb{E}_x \left[\frac{e^{i\frac{u}{\sqrt{n}} M_{n-1}^F}}{\prod_{k=1}^{n-1} a_k^n} \right] = 1$. Ainsi,

$$\begin{aligned}
\left| \mathbb{E}_x \left[e^{i\frac{u}{\sqrt{n}} M_{n-1}^F} \right] - e^{-\frac{u^2 \sigma_F^2}{2}} \right| &= \left| \mathbb{E}_x \left[e^{i\frac{u}{\sqrt{n}} M_{n-1}^F} \left(1 - \frac{e^{-\frac{u^2\sigma_F^2}{2}}}{\prod_{k=1}^{n-1} a_k^n} \right) \right] \right| \leq \mathbb{E}_x \left| \frac{\prod_{k=1}^{n-1} a_k^n - e^{-\frac{u^2\sigma_F^2}{2}}}{\prod_{k=1}^{n-1} a_k^n} \right| \\
&\leq \left(1 - \frac{C}{n}\right)^{-n} \mathbb{E}_x \left| \prod_{k=1}^{n-1} a_k^n - e^{-\frac{u^2\sigma_F^2}{2}} \right|
\end{aligned}$$

où le premier facteur au second membre tend vers e^C lorsque $n \to \infty$. Pour montrer que le second facteur tend vers 0, on pose $b_k^n = e^{-\frac{u^2}{2n}[PF^2 - (PF)^2](X_{k-1})}$ et on écrit

$$\mathbb{E}_x \left| \prod_{k=1}^{n-1} a_k^n - e^{-\frac{u^2\sigma_F^2}{2}} \right| \leq \left| \prod_{k=1}^{n-1} a_k^n - \prod_{k=1}^{n-1} b_k^n \right| + \mathbb{E}_x \left| e^{-\frac{u^2}{2n} \sum_{k=0}^{n-2} [PF^2 - (PF)^2](X_k)} - e^{-\frac{u^2\sigma_F^2}{2}} \right|.$$

La positivité et la convergence presque sûre de $\frac{n-1}{n} \times \frac{1}{n-1} \sum_{k=0}^{n-2} [PF^2 - (PF)^2](X_k)$ vers $\sigma_F^2 := \pi(PF^2 - (PF)^2)$ déduite du théorème 1.4.6 assurent que le second terme au membre de droite de l'inégalité tend vers 0 lorsque $n \to \infty$ par convergence dominée. Montrons que

$$\prod_{k=1}^{n-1} a_k^n - \prod_{k=1}^{n-1} b_k^n = \sum_{k=1}^{n-1} \prod_{j=1}^{k-1} a_j^n (a_k^n - b_k^n) \prod_{j=k+1}^{n-1} b_j^n$$

tend également vers 0, ce qui terminera la preuve dans le cas où F est bornée.

Comme $0 \leq PF^2 - (PF)^2 \leq m^2$,

$$\left| 1 - \frac{u^2}{2n}[PF^2 - (PF)^2](X_{k-1}) - b_k^n \right| \leq \frac{u^4 m^4}{8n^2} \text{ et avec (1.25), } |a_k^n - b_k^n| \leq \frac{4m^3|u|^3}{3n^{3/2}} + \frac{u^4m^4}{8n^2}.$$

En outre, les définitions de a_k^n et b_k^n assurent que $\forall 1 \leq k \leq n-1,\ |a_k^n| \vee |b_k^n| \leq 1$. On en déduit que

$$\left| \prod_{k=1}^{n-1} a_k^n - \prod_{k=1}^{n-1} b_k^n \right| \leq \frac{4m^3|u|^3}{3\sqrt{n}} + \frac{u^4m^4}{8n},$$

où le second membre tend vers 0 lorsque $n \to \infty$.

Traitons maintenant le cas général où F n'est plus supposée bornée. Clairement $\frac{F(X_0)}{\sqrt{n}}$ tend p.s. vers 0 lorsque $n \to \infty$. Par ailleurs, en utilisant l'inégalité de Jensen puis (D1), on obtient

$$\mathbb{E}_x\left[\left(\frac{PF(X_{n-1})}{\sqrt{n}}\right)^2\right] \leq \frac{1}{n}\mathbb{E}_x[PF^2(X_{n-1})] = \frac{P^nF^2(x)}{n} \leq \frac{\|F^2\|_1}{n}\left(1+\gamma^nV(x)+K\sum_{k=0}^{n-1}\gamma^k\right),$$

ce qui assure que $\frac{PF(X_{n-1})}{\sqrt{n}}$ converge vers 0 en moyenne quadratique lorsque $n \to \infty$. Le théorème de Slutsky assure qu'il suffit de montrer que $\frac{M^F_{n-1}}{\sqrt{n}}$ converge en loi vers $\mathcal{N}_1(0,\sigma^2(f))$ pour conclure la démonstration. Pour cela, nous allons remplacer F par la fonction bornée $F_m = -m \vee F \wedge m$ où $m \in \mathbb{N}$. En introduisant la notation $G_m = F - F_m$, nous avons $M^F_{n-1} - M^{F_m}_{n-1} = M^{G_m}_{n-1}$ si bien qu'avec la propriété de martingale de $(M^{G_m}_k)_{k\geq 1}$,

$$\begin{aligned}\mathbb{E}_x\left[\left(\frac{M^F_{n-1}-M^{F_m}_{n-1}}{\sqrt{n}}\right)^2\right] &= \frac{1}{n}\sum_{k=1}^{n-1}\mathbb{E}_x\left[(G_m(X_k)-PG_m(X_{k-1}))^2\right]\\ &= \frac{1}{n}\sum_{k=1}^{n-1}\mathbb{E}_x\left[G_m^2(X_k)-(PG_m(X_{k-1}))^2\right] \leq \frac{1}{n}\sum_{k=1}^{n-1}P^kG_m^2(x).\end{aligned}$$

Comme $G_m^2 \leq F^2$, la fonction G_m^2 est dans $\mathcal{V}$ et, en remplaçant f par G_m^2 dans (1.22), on obtient que $\lim_{k\to\infty} P^kG_m^2(x) = \pi(G_m^2)$ puis, par passage à la moyenne de Cesaro, que $\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}P^kG_m^2(x) = \pi(G_m^2)$. Comme $G_m(y) = 0$ pour $m \geq |F(y)|$ et $G_m^2 \leq F^2$ avec $\pi(F^2) < \infty$, le théorème de convergence dominée assure que $\lim_{m\to\infty}\pi(G_m^2) = 0$ et donc que

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{E}_x\left[\left(\frac{M^F_{n-1}-M^{F_m}_{n-1}}{\sqrt{n}}\right)^2\right] = 0.$$

Montrons d'autre part que $\lim_{m\to\infty}\sigma^2_{F_m} = \sigma^2(f)$. Comme $F^2 \in \mathcal{V}$, $F \in \mathcal{V}$. La condition (D1), l'inégalité $|F_m| \leq |F|$ ainsi que la convergence de F_m vers F assurent que pour tout $y \in \mathbb{E}$, $PF_m(y)$ converge vers $PF(y)$ par convergence dominée lorsque $m \to \infty$. Comme $(PF_m)^2 \leq P(F_m)^2 \leq PF^2$ et que $\pi(PF^2) = \pi(F^2) < \infty$, le théorème de convergence dominée assure que $\lim_{m\to\infty}\pi((PF_m)^2) = \pi((PF)^2)$. De même $\lim_{m\to\infty}\pi(F_m^2) = \pi(F^2)$ et donc $\lim_{m\to\infty}\sigma^2_{F_m} = \sigma^2(f)$.

Pour $u \in \mathbb{R}^*$, comme $\mathbb{R} \ni x \to e^{iux}$ est lipschitzienne de constante $|u|$,

$$\begin{aligned}&\left|\mathbb{E}_x\left[e^{iu\frac{M^F_{n-1}}{\sqrt{n}}}\right] - e^{-\frac{\sigma_F^2u^2}{2}}\right|\\ &\leq \left|\mathbb{E}_x\left[e^{iu\frac{M^F_{n-1}}{\sqrt{n}}} - e^{iu\frac{M^{F_m}_{n-1}}{\sqrt{n}}}\right]\right| + \left|\mathbb{E}_x\left[e^{iu\frac{M^{F_m}_{n-1}}{\sqrt{n}}}\right] - e^{-\frac{\sigma^2_{F_m}u^2}{2}}\right| + \left|e^{-\frac{\sigma^2_{F_m}u^2}{2}} - e^{-\frac{\sigma^2_Fu^2}{2}}\right|\\ &\leq |u|\mathbb{E}_x^{1/2}\left[\left(\frac{M^F_{n-1}-M^{F_m}_{n-1}}{\sqrt{n}}\right)^2\right] + \left|\mathbb{E}_x\left[e^{iu\frac{M^{F_m}_{n-1}}{\sqrt{n}}}\right] - e^{-\frac{\sigma^2_{F_m}u^2}{2}}\right| + \left|e^{-\frac{\sigma^2_{F_m}u^2}{2}} - e^{-\frac{\sigma^2_Fu^2}{2}}\right|.\end{aligned} \tag{1.26}$$

Pour $\varepsilon > 0$ fixé, on peut choisir m_ε tel que $\left|e^{-\frac{\sigma^2_{F_{m_\varepsilon}}u^2}{2}} - e^{-\frac{\sigma^2_F u^2}{2}}\right| \leq \frac{\varepsilon}{3}$ et $\limsup_{n\to\infty} \mathbb{E}_x^{1/2}\left[\left(\frac{M^F_{n-1}-M^{F_{m_\varepsilon}}_{n-1}}{\sqrt{n}}\right)^2\right] \leq \frac{\varepsilon}{6|u|}$. Alors il existe n^1_ε tel que pour $n \geq n^1_\varepsilon$, $\mathbb{E}_x^{1/2}\left[\left(\frac{M^F_{n-1}-M^{F_{m_\varepsilon}}_{n-1}}{\sqrt{n}}\right)^2\right] \leq \frac{\varepsilon}{3|u|}$. Par ailleurs, la première étape de la démonstration, appliquée à la fonction bornée F_{m_ε}, assure l'existence de n^2_ε t.q. $\forall n \geq n^2_\varepsilon$, $\left|\mathbb{E}_x\left[e^{iu\frac{M^{F_{m_\varepsilon}}_{n-1}}{\sqrt{n}}}\right] - e^{-\frac{\sigma^2_{F_{m_\varepsilon}}u^2}{2}}\right| \leq \frac{\varepsilon}{3}$. En choisissant $m = m_\varepsilon$ dans (1.26), on conclut que

$$\forall n \geq n^1_\varepsilon \vee n^2_\varepsilon,\ \left|\mathbb{E}_x\left[e^{iu\frac{M^F_{n-1}}{\sqrt{n}}}\right] - e^{-\frac{\sigma^2_F u^2}{2}}\right| \leq \varepsilon,$$

ce qui achève la démonstration.

□

1.6 Comparaison des variances asymptotiques

Nous allons donner une condition suffisante due Tierney [73] permettant de comparer les variances asymptotiques $\sigma_0^2(f) = \pi(F_0^2) - \pi((P_0F_0)^2)$ et $\sigma_1^2(f) = \pi(F_1^2) - \pi((P_1F_1)^2)$ associées à deux noyaux markoviens P_0 et P_1 dans le théorème 1.5.6.

Proposition 1.6.1. *Supposons que la probabilité π est réversible pour les deux noyaux markoviens P_0 et P_1 et que $\pi(g(P_1 - P_0)g) \geq 0$ pour toute fonction $g : E \to \mathbb{R}$ mesurable vérifiant $\pi(g^2) < \infty$. Supposons également que*

$$\exists V : E \to \mathbb{R}_+ \text{ mesurable}, \exists K \in \mathbb{R}_+,\ \exists \gamma \in]0,1[,\ \forall x \in E,\ P_0V \vee P_1V(x) \leq \gamma V(x) + K,$$

$$\exists R > \frac{4K}{(1-\sqrt{\gamma})^2},\ \exists \alpha \in]0,1], \forall x, y \ t.q.\ V(x) + V(y) \leq R,$$
$$(P_0(x,.) \wedge P_0(y,.)(E)) \wedge (P_1(x,.) \wedge P_1(y,.)(E)) \geq \alpha.$$

Alors $\sigma_0^2(f) \leq \sigma_1^2(f)$ pour toute fonction $f : E \to \mathbb{R}$ mesurable et t.q. $\sup_{x\in E} \frac{|f(x)|}{1+\sqrt{V(x)}} < \infty$.

Remarque 1.6.2. *Pour $g : E \to \mathbb{R}$ mesurable et telle que $\pi(g^2) < \infty$ et P un noyau markovien admettant π comme probabilité invariante, $\pi(P|g|) = \pi(|g|) \leq \sqrt{\pi(g^2)}$ ce qui assure que $Pg(x)$ est bien défini $\pi(dx)$ p.p.. Par l'inégalité de Cauchy-Schwarz, celle de Jensen qui assure que $(Pg)^2 \leq Pg^2$ puis l'invariance de π par P,*

$$(\pi(|gPg|))^2 \leq \pi(g^2)\pi((Pg)^2) \leq \pi(g^2)\pi(Pg^2) = \left(\pi(g^2)\right)^2 < \infty.$$

Comme $|g(P_1 - P_0)g| \leq |gP_1g| + |gP_0g|$, on en déduit que $\pi(|g(P_1 - P_0)g|) < \infty$.

Le lemme suivant donne une condition suffisante pour que l'hypothèse de la proposition soit satisfaite.

Lemme 1.6.3. *Soient P_0 et P_1 deux noyaux markoviens admettant π comme probabilité invariante. Si pour tout $A \in \mathcal{E}$ et tout $x \in E \setminus A$, $P_0(x, A) \geq P_1(x, A)$, alors $\pi(g(P_1 - P_0)g) \geq 0$ pour toute fonction $g : E \to \mathbb{R}$ mesurable vérifiant $\pi(g^2) < \infty$.*

Démonstration du lemme 1.6.3 : Soit $P(x, dy) = (\delta_x(dy)+P_0(x, dy)-P_1(x, dy))$. Pour $x \in E$ et $A \in \mathcal{E}$, $\delta_x(A) + P_0(x, A) - P_1(x, A)$ est égal à $P_0(x, A) - P_1(x, A) \geq 0$ si $x \notin A$ et à $1-P_1(x, A)+P_0(x, A) \geq P_0(x, A) \geq 0$ si $x \in A$. Comme $\delta_x(E)+P_0(x, E)-P_1(x, E) = 1$ on en déduit que P est un noyau markovien. Comme la probabilité π est invariante par les noyaux $\delta_x(dy)$, P_0 et P_1, elle est invariante par P, ce qui implique que $\int_E g^2(x)\pi(dx) = \frac{1}{2}\int_{E\times E}(g^2(x) + g^2(y))\pi(dx)P(x, dy)$. Ainsi

$$\begin{aligned}
\pi(g(P_1g - P_0g)) &= \int_{E\times E} g(x)g(y)\pi(dx)(P_1(x, dy) - P_0(x, dy)) \\
&= \int_{E\times E} g(x)g(y)\pi(dx)(\delta_x(dy) - P(x, dy)) \\
&= \int_E g^2(x)\pi(dx) - \int_{E\times E} g(x)g(y)\pi(dx)P(x, dy) \\
&= \frac{1}{2}\int_{E\times E}(g^2(x) + g^2(y) - 2g(x)g(y))\pi(dx)P(x, dy) \\
&= \frac{1}{2}\int_{E\times E}(g(x) - g(y))^2\pi(dx)P(x, dy) \geq 0.
\end{aligned}$$

□

Pour un espace d'état E fini, Peskun [65] a montré que $\sigma_0^2(f) \leq \sigma_1^2(f)$ lorsque π est réversible par rapport aux deux noyaux P_0 et P_1 qui vérifient la condition de comparaison en hypothèse de ce lemme. Cette condition permet de vérifier l'optimalité en terme de variance asymptotique du choix de Metropolis-Hastings pour la probabilité d'accepter la proposition dans l'algorithme du même nom.

Example 1.6.4. *Pour l'algorithme de Metropolis-Hastings décrit au paragraphe 1.2, soit*

$$P_1(x, dy) = \alpha(x, y)q(x, y)\lambda(dy) + \left(\int_E(1 - \alpha(x, z))q(x, z)\lambda(dz)\right)\delta_x(dy),$$

avec

$$\alpha(x, y) = \begin{cases} a\left(\frac{\eta(y)q(y,x)}{\eta(x)q(x,y)}\right) & si\ \eta(x)q(x, y) > 0 \\ 1\ si\ \eta(x)q(x, y) = 0 \end{cases},$$

pour une fonction $a : \mathbb{R}_+ \to [0, 1]$ vérifiant $a(0) = 0$ et $\forall u > 0,\ a(u) = ua(1/u)$. On note P_0 et α_0 le noyau et la probabilité d'acceptation obtenus pour le choix $a_0(u) = \min(u, 1)$ de Metropolis-Hastings. Pour $u > 0$, comme a est à valeurs dans $[0, 1]$, $a(u) \leq 1$ et $a(u) = ua(1/u) \leq u$ si bien que $a(u) \leq a_0(u)$ pour tout $u \geq 0$. Ainsi $\alpha(x, y) \leq \alpha_0(x, y)$ pour tous $x, y \in E$ et pour tout $A \in \mathcal{E}$ t.q. $x \notin A$,

$$P_1(x, A) = \int_A \alpha(x, y)q(x, y)\lambda(dy) \leq \int_A \alpha_0(x, y)q(x, y)\lambda(dy) = P_0(x, A).$$

Comme $\pi(dx) = \frac{\eta(x)\lambda(dx)}{\int_E \eta(y)\lambda(dy)}$ est réversible pour P_0 et P_1, la proposition et le lemme assurent, sous réserve que les conditions de dérive en hypothèse de la proposition soient *satisfaites, que la variance asymptotique correspondant au choix de Metropolis-Hastings est minimale parmi tous les choix de fonction* $a : \mathbb{R}_+ \to [0,1]$ *vérifiant* $a(0) = 0$ *et* $\forall u > 0,\ a(u) = ua(1/u)$.

Démontrons maintenant la proposition 1.6.1.

Démonstration : Soit $f : E \to \mathbb{R}$ mesurable et t.q. $\sup_{x\in E} \frac{|f(x)|}{1+\sqrt{V(x)}} < \infty$. Comme les variances asymptotiques pour les fonctions f et $f - \pi(f)$ sont égales, nous supposerons que $\pi(f) = 0$ sans que cela ne soit restrictif.

Pour $t \in [0,1]$, notons $P_t = tP_1 + (1-t)P_0$. Comme $P_tV \leq P_0V \vee P_1V$ et $P_t(x,.) \wedge P_t(y,.) \geq t(P_1(x,.) \wedge P_1(y,.)) + (1-t)(P_0(x,.) \wedge P_0(y,.))$, les hypothèses du théorème 1.3.27, du corollaire 1.5.3 et du théorème 1.5.6 sont satisfaites pour le noyau P_t. La probabilité π étant réversible pour P_0 et P_1, elle l'est également pour P_t et c'est l'unique probabilité invariante pour ce noyau d'après le théorème 1.3.27. D'après le théorème 1.5.6, la variance asymptotique de l'estimateur ergodique de $\pi(f)$ sous le noyau P_t est $\sigma_t^2(f) = \pi(F_t^2) - \pi((P_tF_t)^2)$ où $F_t = \sum_{n\in\mathbb{N}} P_t^n f$ est solution de l'équation de Poisson

$$F_t - P_tF_t = f. \tag{1.27}$$

Une relecture attentive des preuves du théorème 1.3.27 et de la proposition 1.5.1 montre que (1.24) se généralise en : $\forall \beta \in]0, \frac{\alpha}{\sqrt{K}}[,\ \exists \chi \in]0,1[,\ \forall n \in \mathbb{N},\ \forall t_1,\ldots,t_n \in [0,1]$,

$$\sup_{x\in E} \frac{|P_{t_1}P_{t_2}\ldots P_{t_n}f(x)|}{1+\sqrt{V(x)}} \leq \max\left(2 + \frac{\beta\sqrt{K}}{1-\sqrt{\gamma}}, \beta\right)\chi^n \sup_{x\in E} \frac{|f(x)|}{1+\beta\sqrt{V(x)}}. \tag{1.28}$$

Choisissons désormais $\beta = \frac{\alpha}{2\sqrt{K}}$ et posons $C = \max\left(2 + \frac{\beta\sqrt{K}}{1-\sqrt{\gamma}}, \beta\right) \sup_{x\in E} \frac{|f(x)|}{1+\beta\sqrt{V(x)}} < \infty$.

Nous allons vérifier que $t \mapsto \sigma_t^2(f)$ est dérivable sur $[0,1]$ de dérivée $\frac{d\sigma_t^2}{dt}(f) = 2\pi(F_t(P_1F_t - P_0F_t))$. Comme l'hypothèse appliquée à $g = F_t$ assure que cette dérivée est positive, cela entraîne que $\sigma_1^2(f) \geq \sigma_0^2(f)$.

Pour $g : E \to \mathbb{R}$ mesurable telle que $\sup_{y\in E} \frac{|g(y)|}{1+\sqrt{V(y)}} < \infty$ et $x \in E$, $[0,1] \ni t \mapsto P_t^n g(x)$ est un polynôme de degré n donc une fonction C^∞. Pour identifier $\frac{d}{dt}P_t^n f(x)$ remarquons que pour $t \in [0,1]$ et $h \neq 0$ tel que $t+h \in [0,1]$,

$$\frac{1}{h}(P_{t+h}^n f(x) - P_t^n f(x)) = \sum_{m=0}^{n-1} P_{t+h}^m \frac{P_{t+h} - P_t}{h} P_t^{n-1-m} f(x) = \sum_{m=0}^{n-1} P_{t+h}^m (P_1 - P_0) P_t^{n-1-m} f(x)$$

converge vers $\frac{d}{dt}P_t^n f(x) = \sum_{m=0}^{n-1} P_t^m (P_1 - P_0) P_t^{n-1-m} f(x)$ lorsque $h \to 0$ puisque $\sup_{y\in E} \frac{|(P_1-P_0)P_t^{n-1-m}f(y)|}{1+\sqrt{V(y)}} \leq \sup_{y\in E} \frac{|P_1P_t^{n-1-m}f(y)|}{1+\sqrt{V(y)}} + \sup_{y\in E} \frac{|P_0P_t^{n-1-m}f(y)|}{1+\sqrt{V(y)}} \leq 2C\chi^{n-m}$, d'après (1.28). En outre, de manière analogue, $\sup_{t\in[0,1],y\in E} \frac{|\frac{d}{dt}P_t^n f(y)|}{1+\sqrt{V(y)}} \leq 2nC\chi^n$ où le second membre est sommable sur $n \in \mathbb{N}$. Le théorème des accroissements finis et le théorème de convergence dominée entraînent alors que $t \mapsto F_t(x) = \sum_{n\in\mathbb{N}} P_t^n f(x)$ est dérivable sur $[0,1]$ de dérivée $\frac{dF_t(x)}{dt} = \sum_{n\in\mathbb{N}} \frac{d}{dt}P_t^n f(x) = \sum_{n\in\mathbb{N}^*} \sum_{m=0}^{n-1} P_t^m(P_1 - P_0)P_t^{n-1-m} f(x)$ avec $\sup_{t\in[0,1],y\in E} \frac{|\frac{dF_t(y)}{dt}|}{1+\sqrt{V(y)}} \leq 2C\chi \sum_{n\in\mathbb{N}^*} n\chi^{n-1} = \frac{2C\chi}{(1-\chi)^2}$. Avec l'hypothèse de croissance faite sur f, cela implique que $\sup_{t\in[0,1],y\in E} \frac{|f(y)\frac{dF_t(y)}{dt}|}{1+V(y)} < \infty$. Comme $\pi(1+V) < \infty$, on en déduit que $\pi(\sup_{t\in[0,1]}|f\frac{dF_t(y)}{dt}|)<\infty$. D'après l'équation de Poisson,

$$F_t^2-(P_tF_t)^2=(F_t-P_tF_t)(2F_t-(F_t-P_tF_t))=f(2F_t-f),$$

si bien que $\sigma_t^2(f)=\pi(f(2F_t-f))$. Une nouvelle application du théorème des accroissements finis et du théorème de convergence dominée assure que $t\mapsto\sigma_t^2(f)$ est dérivable sur $[0,1]$ de dérivée

$$\frac{d\sigma_t^2(f)}{dt}=2\pi\left(f\frac{dF_t}{dt}\right)=2\pi\left((F_t-P_tF_t)\frac{dF_t}{dt}\right).$$

La réversibilité de π pour P_t s'écrit $\pi(dx)P_t(x,dy)=\pi(dy)P_t(y,dx)$ et implique que

$$\begin{aligned}\pi\left(\frac{dF_t}{dt}P_tF_t\right)&=\int_{E\times E}F_t(y)\frac{dF_t}{dt}(x)\pi(dx)P_t(x,dy)=\int_{E\times E}F_t(y)\frac{dF_t}{dt}(x)\pi(dy)P_t(y,dx)\\&=\pi\left(F_tP_t\frac{dF_t}{dt}\right),\end{aligned}$$

si bien que $\frac{d\sigma_t^2(f)}{dt}=2\pi\left(F_t\left(\frac{dF_t}{dt}-P_t\frac{dF_t}{dt}\right)\right)$. Par ailleurs, en dérivant formellement l'équation de Poisson (1.27), on obtient que $\frac{dF_t}{dt}-P_t\frac{dF_t}{dt}$ doit être égal à $\frac{dP_t}{dt}F_t=(P_1-P_0)F_t$. Pour le montrer rigoureusement, on écrit

$$\begin{aligned}P_t\frac{dF_t}{dt}&=\sum_{n\in\mathbb{N}^*}\sum_{m=0}^{n-1}P_t^{m+1}(P_1-P_0)P_t^{n-1-m}f=\sum_{n\in\mathbb{N}^*}\sum_{m=1}^{n}P_t^{m}(P_1-P_0)P_t^{n-m}f\\&=\sum_{n\geq2}\sum_{m=1}^{n-1}P_t^{m}(P_1-P_0)P_t^{n-1-m}f=\sum_{n\geq2}\sum_{m=0}^{n-1}P_t^{m}(P_1-P_0)P_t^{n-1-m}f-\sum_{n\geq2}(P_1-P_0)P_t^{n-1}f\\ \text{et }\frac{dF_t}{dt}-P_t\frac{dF_t}{dt}&=(P_1-P_0)f+\sum_{n\geq2}(P_1-P_0)P_t^{n-1}f=(P_1-P_0)\sum_{n\in\mathbb{N}^*}P_t^{n-1}f\\&=(P_1-P_0)F_t.\end{aligned}$$

On conclut que $\frac{d\sigma_t^2(f)}{dt}=2\pi(F_t(P_1-P_0)F_t)$. $\square$

Chapitre 2

Systèmes de particules en interaction

Lorsque l'on veut calculer $\mathbb{E}[f_n(X_{0:n})]$ où $X_{0:n} = (X_0, \ldots, X_n)$ avec $(X_k)_{k\in\mathbb{N}}$ chaîne de Markov à valeurs dans E issue de X_0 distribuée suivant la probabilité η_0 et $f_n : E^{n+1} \to \mathbb{R}$ telle que $\mathbb{P}(f_n(X_{0:n}) \neq 0)$ est très faible (événement rare), alors la précision de l'estimateur de Monte-Carlo standard va être mauvaise. L'échantillonnage préférentiel consiste à effectuer un changement de probabilité de manière à privilégier les trajectoires de la chaîne de Markov qui se dirigent vers les zones où f_n est non nulle. À cet effet, on peut introduire des fonctions $g_p : E^{p+1} \mapsto]0, +\infty[$ bornées et définir une nouvelle probabilité $\mathbb{Q}_n$ par $\frac{d\mathbb{Q}_n}{d\mathbb{P}} = \frac{\prod_{p=0}^{n-1} g_p(X_{0:p})}{\mathbb{E}\left[\prod_{p=0}^{n-1} g_p(X_{0:p})\right]}$.

Exemple 2.0.1. *Si $E = \mathbb{R}$ et que $f_n(x_{0:n}) = \varphi(x_n)$ avec $\varphi(x)$ non nulle seulement pour de très grandes valeurs de x, on peut choisir $g_p(x_{0:p}) = \psi(x_p)$ avec ψ strictement positive et croissante pour privilégier les trajectoires qui sont déjà hautes à l'instant p ou bien $g_p(x_{0:p}) = \psi(x_p - x_{p-1})$ pour privilégier les trajectoires qui montent entre l'instant $p-1$ et l'instant p supposé supérieur ou égal à 1.*

Alors on a

$$\mathbb{E}[f_n(X_{0:n})] = \mathbb{E}^{\mathbb{Q}_n}\left[\tilde{f}_n(X_{0:n})\right]\mathbb{E}\left[\prod_{p=0}^{n-1} g_p(X_{0:p})\right] \quad \text{où} \quad \tilde{f}_n(x_{0:n}) := f_n(x_{0:n})\prod_{p=0}^{n-1} g_p^{-1}(x_{0:p}). \tag{2.1}$$

Pour implémenter l'échantillonnage préférentiel correspondant au changement de $\mathbb{P}$ à $\mathbb{Q}_n$, il faut à la fois être capable de simuler $(X_k)_{0\leq k\leq n}$ sous la probabilité $\mathbb{Q}_n$ et de calculer la constante de normalisation $\mathbb{E}[\prod_{p=0}^{n-1} g_p(X_{0:p})]$. Notons que cette constante de normalisation est une espérance du même type que celle de départ mais portant sur la trajectoire $X_{0:n-1}$ raccourcie d'une unité de temps, ce qui suggère d'itérer la transformation effectuée sur l'espérance de départ. Posons $\boxed{\gamma_0 = \eta_0}$ et pour $p \in \{1, \ldots, n\}$, introduisons la mesure positive γ_p sur E^{p+1} définie par

$$\forall h_p : E^{p+1} \to \mathbb{R} \text{ mesurable bornée }, \boxed{\gamma_p(h_p) = \mathbb{E}\left[h_p(X_{0:p})\prod_{k=0}^{p-1} g_k(X_{0:k})\right]}$$

ainsi que la probabilité η_p sur E^{p+1} obtenue par normalisation de γ_p :

$$\boxed{\eta_p(h_p) = \frac{\gamma_p(h_p)}{\gamma_p(1)} = \frac{\mathbb{E}\left[h_p(X_{0:p})\prod_{k=0}^{p-1} g_k(X_{0:k})\right]}{\mathbb{E}\left[\prod_{k=0}^{p-1} g_k(X_{0:k})\right]}.} \tag{2.2}$$

En particulier $\mathbb{E}[f_n(X_{0:n})] = \mathbb{E}\left[\tilde{f}_n(X_{0:n})\prod_{k=0}^{n-1} g_k(X_{0:k})\right] = \gamma_n(\tilde{f}_n)$ et $\mathbb{E}^{\mathbb{Q}_n}[h_n(X_{0:n})] = \eta_n(h_n)$.

Comme $\gamma_p(1) = \gamma_{p-1}(g_{p-1}) = \eta_{p-1}(g_{p-1})\gamma_{p-1}(1)$ et $\gamma_0(1) = 1$, par récurrence sur p on obtient que

$$\text{pour } p \geq 1, \quad \boxed{\gamma_p(1) = \prod_{k=0}^{p-1} \eta_k(g_k).}$$

On en déduit que

$$\mathbb{E}[f_n(X_{0:n})] = \gamma_n(\tilde{f}_n) = \eta_n(\tilde{f}_n)\gamma_n(1) = \eta_n(\tilde{f}_n)\prod_{k=0}^{n-1}\eta_k(g_k). \tag{2.3}$$

Plus généralement, pour tout $p \in \{1, \ldots, n-1\}$,

$$\boxed{\gamma_p(h_p) = \eta_p(h_p)\gamma_p(1) = \eta_p(h_p)\prod_{k=0}^{p-1}\eta_k(g_k).} \tag{2.4}$$

L'exercice suivant fournit un example où remplacer un calcul d'espérance par un calcul de produit d'espérances comme dans la formule (2.3) conduit à une réduction de variance significative.

Exercice 2.0.2. *On souhaite estimer une probabilité de la forme $p = \mathbb{P}(X \in A)$ avec p très petite (X désigne par exemple le vecteur des paramètres de fonctionnement d'une centrale nucléaire et A l'ensemble des valeurs de ces paramètres conduisant à la fusion du cœur de la centrale). On se donne pour cela une suite $(X_i)_{i\geq 1}$ de variables aléatoires indépendantes et identiquement distribuées suivant la loi de X et on note $\hat{p}_n = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i\in A\}}$.*

1. *Quel est le comportement asymptotique de $\hat{p}_n$ pour $n \to +\infty$? Donner la variance de $\hat{p}_n$.*

On suppose maintenant que l'on sait simuler indépendamment des X_i une suite $(Y_i)_{i\geq 1}$ de variables aléatoires indépendantes et identiquement distribuées suivant la loi conditionnelle de X sachant $X \in B$ où B contient A. En particulier $\mathbb{P}(Y_i \in A) = \mathbb{P}(X \in A|X \in B)$. On note $\hat{q}_n = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i\in B\}}$ et $\hat{r}_n = \frac{1}{n}\sum_{i=1}^n 1_{\{Y_i\in A\}}$.

2. *Quel est le comportement asymptotique de $\hat{q}_n \times \hat{r}_n$ pour $n \to +\infty$? Calculer l'espérance et la variance de $\hat{q}_n \times \hat{r}_n$. Expliquer pourquoi, lorsque $\mathbb{P}(X \in B)$ et $\mathbb{P}(X \in A|X \in B)$ sont du même ordre, $\hat{q}_n \times \hat{r}_n$ approche p beaucoup plus précisément que $\hat{p}_n$.*

3. *On suppose plus généralement que $p = \prod_{j=1}^k p_j$ et que les estimateurs indépendants $(\hat{p}_n^j)_{1\leq j\leq k}$ sont respectivement les moyennes empiriques de n variables aléatoires de Bernoulli de paramètre p_j indépendantes.*
   (a) *Montrer que $\prod_{j=1}^k \hat{p}_n^j$ converge presque sûrement vers p lorsque $n \to \infty$.*
   (b) *Vérifier que la variance asymptotique $\mathcal{V}_k(p_1, \ldots, p_k) = \lim_{n\to\infty} n\mathrm{Var}\left(\frac{\prod_{j=1}^k \hat{p}_n^j}{\prod_{j=1}^k p_j}\right)$ qui mesure la précision de cet estimateur est donnée par $\mathcal{V}_k(p_1, \ldots, p_k) = \sum_{j=1}^k \frac{1-p_j}{p_j} = \sum_{j=1}^k \frac{1}{p_j} - k$.*

(c) *Montrer que* $\frac{1}{k}\sum_{j=1}^{k}\frac{1}{p_j} \geq e^{-\frac{1}{k}\sum_{j=1}^{k}\ln(p_j)}$ *et en déduire que* $\mathcal{V}_k(p^{1/k},\ldots,p^{1/k}) = \inf_{(p_1,\ldots,p_k)\in(\mathbb{R}_+^*)^k:\prod_{j=1}^{k}p_j=p}\mathcal{V}_k(p_1,\ldots,p_k)$.

(d) *Comparer* $\lim_{k\to\infty}\mathcal{V}_k(p^{1/k},\ldots,p^{1/k})$ *à* $\mathcal{V}_1(p)$.

Le choix sous forme produit du changement de probabilité $\frac{d\mathbb{Q}_n}{d\mathbb{P}} = \frac{\prod_{p=0}^{n-1} g_p(X_{0:p})}{\mathbb{E}\left[\prod_{p=0}^{n-1} g_p(X_{0:p})\right]} = \frac{\prod_{p=0}^{n-1} g_p(X_{0:p})}{\prod_{p=0}^{n-1}\eta_p(g_p)}$ qui peut sembler étrange de prime abord, va permettre de construire dynamiquement une suite de mesures empiriques $(\eta_p^M = \frac{1}{M}\sum_{m=1}^{M}\delta_{X_{0:p}^{p,m}})_{0\leq p\leq n}$ qui approchent les probabilités $(\eta_p)_{0\leq p\leq n}$ dans la limite $M\to\infty$ d'un grand nombre de particules. Il n'y a pas besoin de savoir simuler $(X_k)_{0\leq k\leq n}$ sous $\mathbb{Q}_n$ pour bénéficier en partie de la réduction de variance par échantillonnage préférentiel. En suivant (2.3), on approchera $\mathbb{E}\left[f_n(X_{0:n})\right]$ par $\eta_n^M(\tilde{f}_n)\prod_{p=0}^{n-1}\eta_p^M(g_p)$.

Avant d'expliquer la construction des mesures empiriques η_p^M, montrons que les mesures η_p apparaissent également naturellement dans des problèmes de filtrage. On considère un signal (par exemple la position d'un aéronef dans le ciel) modélisé par une chaîne de Markov $(X_k)_{k\in\mathbb{N}}$ avec X_0 distribuée suivant la probabilité η_0. On n'observe pas directement le signal mais des informations $Y_k = \varphi_k(X_k) + V_k$ (écho radar par exemple) où $\varphi_k : E\to\mathbb{R}^d$ et les bruits d'observation $(V_k)_{k\in\mathbb{N}}$ sont supposés indépendants à valeurs $\mathbb{R}^d$ de densités respectives $(q_k)_{k\in\mathbb{N}}$ par rapport à la mesure de Lebesgue et indépendants de la chaîne de Markov $(X_k)_{k\in\mathbb{N}}$. La densité conditionnelle de $Y_{0:p}$ sachant $X_{0:p+1} = x_{0:p+1}$ est $\prod_{k=0}^{p} q_k(y_k - \varphi_k(x_k))$. La formule de Bayes assure alors que

$$\mathbb{P}(X_{0:p+1}\in dx_{0:p+1}|Y_{0:p} = y_{0:p}) \propto \prod_{k=0}^{p} q_k(y_k-\varphi_k(x_k))\mathbb{P}(X_{0:p+1}\in dx_{0:p+1}).$$

Ainsi, pour le choix $g_k(x_{0:k}) = q_k(y_k-\varphi_k(x_k))$ pour tout $k\in\{0,\ldots,p\}$, la loi conditionnelle de $X_{0:p+1}$ sachant $Y_{0:p} = y_{0:p}$ est η_{p+1}. La loi conditionnelle $\hat{\eta}_p$ de $X_{0:p}$ sachant $Y_{0:p} = y_{0:p}$ s'obtient par la formule de la loi marginale :

$$\hat{\eta}_p(h_p) = \int_{E^{p+2}} h_p(x_{0:p})\eta_{p+1}(dx_{0:p+1}) = \frac{\mathbb{E}[h_p(X_{0:p})\prod_{k=0}^{p} g_k(X_{0:k})]}{\gamma_{p+1}(1)} = \frac{\gamma_p(h_p g_p)}{\gamma_p(g_p)} = \frac{\eta_p(h_p g_p)}{\eta_p(g_p)}.$$

La construction des probabilités $(\eta_p^M)_{0\leq p\leq n}$ se fait par récurrence sur p. La chaîne de Markov $(X_k)_{k\geq 0}$ est supposée inhomogène au sens où, pour $h : E\to\mathbb{R}$ mesurable bornée, $\mathbb{E}[h(X_k)|X_{0:k-1}] = P_k h(X_{k-1})$ pour un noyau Markovien P_k pouvant dépendre de l'instant k. Notons que le cas où l'espace E_k où X_k prend ses valeurs dépend de k peut rentrer dans le cadre introduit ici en posant $E = \bigcup_{k\in\mathbb{N}} E_k$.

Initialisation : On génère un échantillon $(X_0^{0,m})_{1\leq m\leq M}$ i.i.d. suivant la loi η_0 de X_0.

Passage de $\eta_p^M = \frac{1}{M}\sum_{m=1}^{M}\delta_{X_{0:p}^{p,m}}$ **à** $\eta_{p+1}^M = \frac{1}{M}\sum_{m=1}^{M}\delta_{X_{0:p+1}^{p+1,m}}$ **:** pour $p\in\{0,\ldots,n-1\}$ ce passage s'effectue en deux étapes :

Sélection : on génère des vecteurs aléatoires $(X_{0:p}^{p+1,m})_{1\leq m\leq M}$ conditionnellement indépendants sachant $\mathcal{F}_p = \sigma((X_{0:k}^{k,m})_{1\leq m\leq M, 0\leq k\leq p})$ et vérifiant

$$\mathbb{E}\left[\frac{1}{M}\sum_{m=1}^{M}\delta_{X_{0:p}^{p+1,m}}\middle|\mathcal{F}_p\right] = \frac{1}{M\eta_p^M(g_p)}\sum_{m=1}^{M} g_p(X_{0:p}^{p,m})\delta_{X_{0:p}^{p,m}}. \tag{2.5}$$

Après avoir décrit la seconde étape, nous allons revenir sur plusieurs approches permettant d'effectuer cette étape de sélection. Notons que $\hat{\eta}_p^M = \frac{1}{M}\sum_{m=1}^M \delta_{X_{0:p}^{p+1,m}}$ (égale à la loi marginale des $p+1$ premières coordonnées sous η_{p+1}^M) constitue une approximation de la mesure $\hat{\eta}_p$ introduite dans le problème de filtrage présenté plus haut.

Mutation : sachant $\mathcal{G}_{p+1} = \sigma((X_{0:k}^{k,m})_{1\leq m\leq M, 0\leq k\leq p}, (X_{0:p}^{p+1,m})_{1\leq m\leq M})$, on génère des variables aléatoires $(X_{p+1}^{p+1,m})_{1\leq m\leq M}$ conditionnellement indépendantes et respectivement distribuées suivant $P_{p+1}(X_p^{p+1,m}, .)$. Pour tout $m \in \{1,\ldots,M\}$, on pose $X_{0:p+1}^{p+1,m} = (X_{0:p}^{p+1,m}, X_{p+1}^{p+1,m})$ et $\eta_{p+1}^M = \frac{1}{M}\sum_{m=1}^M \delta_{X_{0:p+1}^{p+1,m}}$.

Intuitivement, l'étape de sélection permet de se débarrasser du poids multiplicatif individuel $g_p(X_{0:p}^{p,m})$ affecté à chaque trajectoire entre l'instant p et l'instant $p+1$ en répliquant les trajectoires pour lesquelles ce poids est élevé et en supprimant celles pour lesquelles ce poids est faible. L'étape de mutation consiste juste à ajouter une nouvelle position tirée suivant le noyau markovien P_{p+1} à chaque trajectoire issue de l'étape de sélection.

Pour l'étape de sélection, partant de M valeurs $(y^m)_{1\leq m\leq M}$ (les trajectoires $(X_{0:p}^{p,m})_{1\leq m\leq M}$) dans un espace d'états $\mathcal{Y}$ quelconque (l'espace produit E^{p+1}) et d'une probabilité $(p^m)_{1\leq m\leq M}$ $((g_p(X_{0:p}^{p,m})/\sum_{\ell=1}^M g_p(X_{0:p}^{p,l}))_{1\leq m\leq M})$, il faut être capable de générer des variables aléatoires indépendantes $(Y^m)_{1\leq m\leq M}$ à valeurs dans $\{y^1,\ldots,y^M\}$,

$$\mathbb{E}\left[\frac{1}{M}\sum_{m=1}^M \delta_{Y^m}\right] = \sum_{m=1}^M p^m \delta_{y^m}, \tag{2.6}$$

i.e pour toute fonction $f : \mathcal{Y} \to \mathbb{R}$ mesurable bornée, $\mathbb{E}\left[\frac{1}{M}\sum_{m=1}^M f(Y^m)\right] = \sum_{m=1}^M p^m f(y^m)$. Notons que, lorsque les y^ℓ sont distincts, $\delta_{Y^m} = \sum_{\ell=1}^M 1_{\{Y^m=y^\ell\}}\delta_{y^\ell}$ et (2.6) est équivalent à $\mathbb{E}\left[\sum_{m=1}^M 1_{\{Y^m=y^\ell\}}\right] = Mp^\ell$ pour tout $\ell \in \{1,\ldots,M\}$. Sans supposer les y^ℓ distincts, (2.6) est équivalent à $\mathbb{E}\left[\sum_{m=1}^M 1_{\{Y^m=y^\ell\}}\right] = M\sum_{j=1}^M 1_{\{y^j=y^\ell\}}p^j$ pour tout $\ell \in \{1,\ldots,M\}$. Il y a bien sûr de multiples façons de générer de telles variables aléatoires :

Échantillonnage multinomial : on génére les Y^m i.i.d. suivant la probabilité $\sum_{\ell=1}^M p^\ell \delta_{y^\ell}$ ($\forall \ell, m \in \{1,\ldots,M\}$, $\mathbb{P}(Y^m = y^\ell) = \sum_{j=1}^M 1_{\{y^j=y^\ell\}}p^j$). Notons que $\mathbb{E}[1_{\{Y^m=y^\ell\}}] = \sum_{j=1}^M 1_{\{y^j=y^\ell\}}p^j$ pour tous $m,\ell \in \{1,\ldots,M\}$.

Échantillonnage résiduel : Pour $x \in \mathbb{R}$, on note $\lfloor x\rfloor$ et $\lceil x\rceil$ les parties entières inférieure et supérieure de x, c'est-à-dire les entiers relatifs vérifiant $\lfloor x\rfloor \leq x < \lfloor x\rfloor + 1$ et $\lceil x\rceil - 1 < x \leq \lceil x\rceil$ et $\{x\} = x - \lfloor x\rfloor \in [0,1[$ sa partie fractionnaire. L'échantillonnage résiduel consiste à choisir $Y^m = y^\ell$ pour $\sum_{j=1}^{\ell-1}\lfloor Mp^j\rfloor < m \leq \sum_{j=1}^{\ell}\lfloor Mp^j\rfloor$ (on réplique donc $\lfloor Mp^\ell\rfloor$ fois la valeur y^ℓ) et, lorsque $\sum_{j=1}^M\lfloor Mp^j\rfloor < M$ à générer les $r = M - \sum_{j=1}^M\lfloor Mp^j\rfloor = \sum_{j=1}^M\{Mp^j\}$ variables aléatoires $Y_{\sum_{j=1}^M\lfloor Mp^j\rfloor+1},\ldots,Y_M$ i.i.d. suivant $\frac{1}{M-\sum_{j=1}^M\lfloor Mp^j\rfloor}\sum_{\ell=1}^M\{Mp^\ell\}\delta_{y^\ell}$. Alors, si les y^ℓ sont distincts, pour tout $\ell \in \{1,\ldots,M\}$, $\sum_{m=1}^{M-r} 1_{\{Y^m=y^\ell\}} = \lfloor Mp^\ell\rfloor$ et, lorsque $r > 0$, pour tout $m \in \{M-r+1,\ldots,M\}$, $\mathbb{E}[1_{\{Y^m=y^\ell\}}] = \frac{\{Mp^\ell\}}{r}$.

Échantillonnage stratifié : à partir de M variables aléatoires $(U^m)_{1\leq m\leq M}$ i.i.d. suivant la loi uniforme sur $[0,1]$ on construit

$$Y^m = \sum_{\ell=1}^{M} 1_{\{\sum_{j=1}^{\ell-1} p^j < (m-U^m)/M \leq \sum_{j=1}^{\ell} p^j\}} y^\ell \text{ pour } m \in \{1,\dots,M\}.$$

Échantillonnage systématique : on prend les variables U^m toutes égales à une seule variable U uniforme sur $[0,1]$ dans la formule précédente. Notons que si y^ℓ est distinct de y^j pour $j \neq \ell$, alors cette valeur est tirée $\lfloor M \sum_{j=1}^{\ell} p^j + U \rfloor - \lfloor M \sum_{j=1}^{\ell-1} p^j + U \rfloor$ fois.

Avec l'échantillonnage multinomial, l'échantillonnage résiduel et l'échantillonnage stratifié, les variables aléatoires Y^m sont indépendantes (pour l'échantillonnage résiduel, on pourra remarquer qu'une variable aléatoire constante est indépendant de toute variable aléatoire), alors qu'avec l'échantillonnage systématique, on perd cette propriété d'indépendance (sauf si pour tout $m \in \{1,\dots,M\}$, Mp^m est entier, auquel cas les échantillonnages résiduel, stratifié et systématique coincident). C'est pourquoi seules les trois premières techniques d'échantillonnage sont éligibles pour l'étape de sélection de la méthode particulaire décrite plus haut.

Notons que lorsque les y^ℓ sont distincts, alors la mesure empirique $\frac{1}{M}\sum_{m=1}^{M} \delta_{Y^m}$ s'écrit $\frac{1}{M}\sum_{\ell=1}^{M}\sum_{m=1}^{M} 1_{\{Y^m=y^\ell\}}\delta_{y^\ell}$. L'objectif du problème suivant est de comparer les variances des variables aléatoires $\sum_{m=1}^{M} 1_{\{Y^m=y^\ell\}}$ pour les différentes techniques d'échantillonnage.

Problème 2.0.3. *On suppose les y^ℓ distincts. Pour $\ell \in \{1,\dots,M\}$, la variable aléatoire $\sum_{m=1}^{M} 1_{\{Y^m=y^\ell\}}$ est respectivement notée*

— *N_ℓ dans l'échantillonnage multinomial,*
— *$\tilde{N}_\ell$ dans l'échantillonnage résiduel,*
— *$\bar{N}_\ell$ dans l'échantillonnage stratifié,*
— *$\hat{N}_\ell$ dans l'échantillonnage systématique.*

Condition (2.6) **:** *Remarquer que $\frac{1}{M}\sum_{m=1}^{M} f(Y^m) = \sum_{\ell=1}^{M}\left(\frac{1}{M}\sum_{m=1}^{M} 1_{\{Y^m=y^\ell\}} f(y^\ell)\right)$ et en déduire que* (2.6) *est équivalente à $\forall \ell \in \{1,\dots,M\}$, $\mathbb{E}\left[\sum_{m=1}^{M} 1_{\{Y^m=y^\ell\}}\right] = Mp^\ell$.*

Échantillonnage multinomial *:*

1. *Pour $\ell \in \{1,\dots,M\}$ vérifier que $\mathbb{E}(N_\ell) = Mp^\ell$ et calculer $\mathrm{Var}(N_\ell)$.*
2. *Vérifier que le vecteur $(N_1,\dots,N_M)$ suit la loi multinomiale de paramètre $(M,p^1,\dots,p^M)$: $\forall (n_1,\dots,n_M) \in \mathbb{N}^M$*

$$\mathbb{P}(N_1 = n_1,\dots,N_M = n_M) = 1_{\{\sum_{m=1}^{M} n_m = M\}} M! \prod_{m=1}^{M} \frac{(p^m)^{n_m}}{n_m!}.$$

Les variables aléatoires N_m sont-elles indépendantes ?

Variance optimale : *Soit N une variable aléatoire à valeurs dans $\mathbb{Z}$ intégrable d'espérance ν. On note $x_+ = \max(x,0)$ et $x_- = \max(-x,0)$ pour $x \in \mathbb{R}$.*

1. *Vérifier que $\mathbb{E}[(N-\nu)_+] = \mathbb{E}[(N-\nu)_-]$. Montrer que $(N-\nu)_+ \geq (1-\{\nu\})1_{\{N>\nu\}}$ et $(N-\nu)_- \geq \{\nu\}1_{\{N\leq\nu\}}$.*
2. *En déduire que $\mathrm{Var}(N) = \mathbb{E}[(N-\nu)_+^2 + (N-\nu)_-^2] \geq \mathbb{E}[(N-\nu)_+] \geq \{\nu\}(1-\{\nu\})$ et exhiber une loi sous laquelle les deux dernières inégalités sont des égalités.*

Échantillonnage résiduel :

3. *Que se passe-t-il si* $r = 0$ *?*

On suppose désormais $r > 0$.

4. *Pour* $\ell \in \{1, \dots, M\}$ *vérifier que* $\mathbb{E}(\tilde{N}_\ell) = Mp^\ell$ *et calculer* $\mathrm{Var}(\tilde{N}_\ell)$. *En déduire que la répartition résiduelle est optimale en termes de variance si* $r = 1$ *mais pas si* $r \geq 2$.
5. *Vérifier que* $r \leq M(1-p^\ell)+\{Mp^\ell\}$ *et en déduire que* $\mathrm{Var}(\tilde{N}_\ell) \leq \frac{M\{Mp^\ell\}(1-p^\ell)}{M(1-p^\ell)+\{Mp^\ell\}}$. *Conclure que* $\mathrm{Var}(\tilde{N}_\ell) \leq \mathrm{Var}(N_\ell)$. *Quel intérêt voyez-vous à l'échantillonnage résiduel par rapport à l'échantillonnage multinomial ?*

Échantillonnage stratifié :

6. *Vérifier que pour* $\ell, m \in \{1, \dots, M\}$,

$$\mathbb{E}\left[1_{\{p^1+\ldots+p^{\ell-1}<\frac{m-U_m}{M}\leq p^1+\ldots+p^\ell\}}\right] = M\int_{\frac{m-1}{M}}^{\frac{m}{M}} 1_{\{p^1+\ldots+p^{\ell-1}<u\leq p^1+\ldots+p^\ell\}}du$$

et en déduire que $\mathbb{E}[\bar{N}_\ell] = Mp^\ell = \mathbb{E}[\hat{N}_\ell]$.

7. *Pour* x *et* y *deux réels tels que* $x < y$ *et* $j \in \mathbb{Z}$, *vérifier*

$$]x,y]\cap]j-1,j] = \begin{cases}]j-1,j] \text{ si et seulement si } \lceil x\rceil + 1 \leq j \leq \lfloor y\rfloor \\ \emptyset \text{ si } j \leq \lfloor x\rfloor \text{ ou } j \geq \lceil y\rceil + 1 \end{cases}.$$

En déduire que $\bar{N}_\ell$ *prend ses valeurs dans* $\{\lfloor n(p_1+\ldots+p_\ell)\rfloor - \lceil n(p_1+\ldots+p_{\ell-1})\rceil, \lfloor n(p_1+\ldots+p_\ell)\rfloor - \lfloor n(p_1+\ldots+p_{\ell-1})\rfloor, \lceil n(p_1+\ldots+p_\ell)\rceil - \lfloor n(p_1+\ldots+p_{\ell-1})\rfloor\}$. *Vérifier que pour* X *variable aléatoire à valeurs dans* $[a,b]$, $\mathrm{Var}(X) \leq \frac{(b-a)^2}{4}$ *et en déduire que* $\mathrm{Var}(\bar{N}_\ell) \leq 1$.

8. *Soit* x *et* y *deux réels. Vérifier que* $y \geq \lfloor y-x\rfloor + \lfloor x\rfloor$ *et* $x \geq \lfloor y\rfloor - \lceil y-x\rceil$. *En déduire que* $\lfloor y-x\rfloor \leq \lfloor y\rfloor - \lfloor x\rfloor \leq \lceil y-x\rceil \leq \lfloor y-x\rfloor + 1$. *Lorsque* $x < y$, *vérifier que* $]x,y]\cap\mathbb{N}$ *est égal à* $\emptyset$ *si* $\lfloor y\rfloor = \lfloor x\rfloor$ *et comprend les* $\lfloor y\rfloor - \lfloor x\rfloor$ *éléments* $\{\lfloor x\rfloor+1, \dots, \lfloor y\rfloor\}$ *sinon. En déduire que l'intervalle*

$$]M(p_1+\ldots+p_{\ell-1})+U, M(p_1+\ldots+p_\ell)+U]$$

contient soit $\lfloor Mp_\ell\rfloor$ *soit* $\lfloor Mp_\ell\rfloor + 1$ *éléments de* $\{1, \dots, M\}$. *En déduire que* $\hat{N}_\ell - \lfloor Mp_i\rfloor$ *suit une loi de Bernoulli de paramètre à préciser et conclure que* $\mathrm{Var}(\hat{N}_\ell) = \{Mp_\ell\}(1-\{Mp_\ell\})$, *c'est-à-dire que la répartition systématique est optimale en termes de variance. Vérifier que ce résultat reste vrai pour* $\bar{N}_\ell$ *lorsque* $M(p^1+\ldots+p^{\ell-1})$ *ou* $M(p^1+\ldots+p^\ell)$ *sont entiers, propriété toujours satisfaite pour* $\ell \in \{1, M\}$.

9. *Vérifier que*

$$\mathrm{Var}(\bar{N}_\ell) = M\left(p^\ell - M\sum_{m=1}^{M}\left(\int_{\frac{m-1}{M}}^{\frac{m}{M}} 1_{\{p^1+\ldots+p^{\ell-1}<u\leq p^1+\ldots+p^\ell\}}du\right)^2\right)$$

et en déduire par l'inégalité de Cauchy-Schwarz que $\mathrm{Var}(\bar{N}_\ell) \leq \mathrm{Var}(N_\ell)$.

10. *À l'aide de la question 7, montrer que le cardinal de*

$$\{m \in \{1,\dots,M\} : \int_{\frac{m-1}{M}}^{\frac{m}{M}} 1_{\{p^1+\dots+p^{\ell-1}<u\leq p^1+\dots+p^\ell\}}du = 1/M\}$$

est égal soit à $\lfloor Mp^\ell \rfloor$ *soit à* $\lfloor Mp^\ell \rfloor - 1$. *En remarquant que dans le premier cas, si* $a, b \in [0, 1/M]$ *avec* $a + b = \{Mp^\ell\}/M$, $a^2 + b^2 \geq \frac{\{Mp^\ell\}^2}{2M^2}$ *et dans le second, si* $a, b \in [0, 1/M]$ *avec* $a + b = (1 + \{Mp^\ell\})/M$, $a^2 + b^2 \geq \frac{(1+\{Mp^\ell\})^2}{2M^2}$, *en déduire que*

$$\sum_{m=1}^{M} \left(\int_{\frac{m-1}{M}}^{\frac{m}{M}} 1_{\{p^1+\dots+p^{\ell-1}<u\leq p^1+\dots+p^\ell\}}du \right)^2$$
$$\geq \min \left(\frac{\lfloor Mp^\ell \rfloor + \{Mp^\ell\}^2/2}{M^2}, \frac{Mp^\ell + (\{Mp^\ell\}^2 - 1)/2}{M^2} \right).$$

Conclure que $\mathrm{Var}(\bar{N}_\ell) \leq \max\left(\{Mp^\ell\}(1 - \frac{\{Mp^\ell\}}{2}), \frac{1-\{Mp^\ell\}^2}{2}\right)$.
Pour $M = 3$ *et* $p^1 = \frac{1}{12}$, $p^2 = \frac{1}{2}$ *et* $p^3 = \frac{5}{12}$, *vérifier que* $\mathrm{Var}(\bar{N}_2) = \frac{3}{8} > \frac{1}{4} = \mathrm{Var}(\tilde{N}_2)$.

Exercice 2.0.4. *On suppose que* $\underline{p} := \min_{1\leq m\leq M} p^m > 0$ *(on peut restreindre l'ensemble des valeurs* $(y^m)_{1\leq m\leq M}$ *pour s'y ramener).*

1. *Montrer que* $M\underline{p} \leq 1$. *Que vaut* $(p^1,\dots,p^M)$ *en cas d'égalité ? Que donnent alors l'échantillonnage résiduel et l'échantillonnage stratifié ?*
2. *On suppose* $M\underline{p} < 1$, *on se donne des variables aléatoires* $(Z^m)_{1\leq m\leq M}$ *indépendantes et telles que* $\mathbb{E}\left[\frac{1}{M}\sum_{m=1}^M \delta_{Z^m}\right] = \frac{1}{1-M\underline{p}} \sum_{m=1}^M (p^m - \underline{p})\delta_{y^m}$ *et indépendamment une suite* $(U^m)_{1\leq m\leq M}$ *de variables aléatoires indépendantes et uniformes sur* $[0, 1]$.
   (a) *Vérifier que si pour* $m \in \{1,\dots,M\}$, $Y^m = 1_{\{U_m\leq M\underline{p}\}}y^m + 1_{\{U_m>M\underline{p}\}}Z^m$ *alors*

$$\mathbb{E}\left[\frac{1}{M}\sum_{m=1}^M \delta_{Y^m}\right] = \sum_{m=1}^M p^m \delta_{y^m} \tag{2.7}$$

   et les variables aléatoires $(Y^m)_{1\leq m\leq M}$ *sont indépendantes.*
   (b) *Si* $(Y^1,\dots,Y^M) = 1_{\{U_1\leq M\underline{p}\}}(y^1,\dots,y^M) + 1_{\{U_1>M\underline{p}\}}(Z^1,\dots,Z^M)$, *vérifier que* (2.7) *reste vraie mais que les variables aléatoires* $(Y^m)_{1\leq m\leq M}$ *ne sont pas indépendantes.*

En suivant la formule (2.4), on introduit l'approximation suivante de la mesure non normalisée γ_p : $\gamma_0^M = \eta_0^M$ et

$$\text{pour } p \geq 1, \boxed{\gamma_p^M(h_p) = \eta_p^M(h_p) \prod_{k=0}^{p-1} \eta_k^M(g_k).} \tag{2.8}$$

Notons que $\eta_p^M(h_p) = \frac{\gamma_p^M(h_p)}{\gamma_p^M(1)}$ si bien que la probabilité empirique η_p^M s'obtient par normalisation de la mesure ponctuelle γ_p^M.

Proposition 2.0.5. *On suppose que pour tout $p \in \{0,\dots,n-1\}$, la fonction $g_p : E^{p+1} \to]0,+\infty[$ est mesurable et bornée. Pour tout $p \in \{0,\dots,n\}$ et toute fonction $h_p : E^{p+1} \to \mathbb{R}$ mesurable bornée,*

$$\mathbb{E}\left[\gamma_p^M(h_p)\right] = \gamma_p(h_p).$$

En particulier, si $\tilde{f}_n$ est bornée, l'estimateur $\eta_n^M(\tilde{f}_n)\prod_{p=0}^{n-1}\eta_p^M(g_p) = \gamma_n^M(\tilde{f}_n)$ de $\mathbb{E}[f_n(X_{0:n})]$ est sans biais.

Démonstration : On raisonne par récurrence sur p. On a $\gamma_0^M = \eta_0^M = \frac{1}{M}\sum_{m=1}^M \delta_{X_0^{0,m}}$ où les variables $(X_0^{0,m})_{1\le m\le M}$ sont toutes distribuées suivant $\eta_0 = \gamma_0$, ce qui assure que $\mathbb{E}[\gamma_0^M(h_0)] = \gamma_0(h_0)$. Ainsi, l'hypothèse de récurrence est satisfaite au rang $p = 0$. Supposons-la satisfaite au rang p. Rappelons que $\mathcal{F}_p = \sigma((X_{0:k}^{k,m})_{1\le m\le M, 0\le k\le p})$ et $\mathcal{G}_{p+1} = \sigma((X_{0:k}^{k,m})_{1\le m\le M,0\le k\le p}, (X_{0:p}^{p+1,m})_{1\le m\le M})$. En notant $P_{p+1}h_{p+1}(x_{0:p}) = \int_E h_{p+1}(x_{0:p+1})P_{p+1}(x_p, dx_{p+1})$, on a d'après l'étape de mutation

$$\mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{G}_{p+1}] = \frac{1}{M}\sum_{m=1}^M P_{p+1}h_{p+1}(X_{0:p}^{p+1,m}).$$

Ensuite comme $\mathcal{F}_p \subset \mathcal{G}_{p+1}$,

$$\begin{aligned}\mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{F}_p] = \mathbb{E}[\mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{G}_{p+1}]|\mathcal{F}_p] &= \mathbb{E}\left[\frac{1}{M}\sum_{m=1}^M P_{p+1}h_{p+1}(X_{0:p}^{p+1,m})\middle|\mathcal{F}_p\right]\\ &= \frac{\eta_p^M(g_pP_{p+1}h_{p+1})}{\eta_p^M(g_p)},\end{aligned} \tag{2.9}$$

où on a utilisé (2.5) pour la dernière égalité. Comme $\prod_{k=0}^p \eta_k^M(g_k)$ est $\mathcal{F}_p$ mesurable, on en déduit que

$$\begin{aligned}\mathbb{E}\left[\eta_{p+1}^M(h_{p+1})\prod_{k=0}^p\eta_k^M(g_k)\right] &= \mathbb{E}\left[\mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{F}_p]\prod_{k=0}^p\eta_k^M(g_k)\right]\\ &= \mathbb{E}\left[\eta_p^M(g_pP_{p+1}h_{p+1})\prod_{k=0}^{p-1}\eta_k^M(g_k)\right] = \mathbb{E}\left[\gamma_p^M(g_pP_{p+1}h_{p+1})\right].\end{aligned}$$

L'hypothèse de récurrence au rang p et la définition de γ_p assurent alors que

$$\mathbb{E}\left[\eta_{p+1}^M(h_{p+1})\prod_{k=0}^p\eta_k^M(g_k)\right] = \gamma_p(g_pP_{p+1}h_{p+1}) = \mathbb{E}\left[P_{p+1}h_{p+1}(X_{0:p})\prod_{k=0}^p g_k(X_{0:k})\right].$$

Comme la propriété de Markov assure que $P_{p+1}h_{p+1}(X_{0:p}) = \mathbb{E}[h_{p+1}(X_{0:p+1})|X_{0:p}]$ et que $\prod_{k=0}^p g_k(X_{0:k})$ est une fonction de $X_{0:p}$,

$$\begin{aligned}\mathbb{E}\left[P_{p+1}h_{p+1}(X_{0:p})\prod_{k=0}^p g_k(X_{0:k})\right] &= \mathbb{E}\left[\mathbb{E}[h_{p+1}(X_{0:p+1})|X_{0:p}]\prod_{k=0}^p g_k(X_{0:k})\right]\\ &= \mathbb{E}\left[h_{p+1}(X_{0:p+1})\prod_{k=0}^p g_k(X_{0:k})\right] = \gamma_{p+1}(h_{p+1}),\end{aligned}$$

ce qui achève la démonstration. □

Remarque 2.0.6. — *Notons que nous venons d'établir que* $\gamma_{p+1}(h_{p+1}) = \gamma_p(g_p P_{p+1} h_{p+1})$. *Comme* $\gamma_{p+1}(1) = \gamma_p(g_p) = \eta_p(g_p)\gamma_p(1)$, *avec* (2.2), *on en déduit que*

$$\eta_{p+1}(h_{p+1}) = \frac{\gamma_{p+1}(h_{p+1})}{\gamma_{p+1}(1)} = \frac{\gamma_p(g_p P_{p+1} h_{p+1})}{\eta_p(g_p)\gamma_p(1)} = \frac{\eta_p(g_p P_{p+1} h_{p+1})}{\eta_p(g_p)}. \tag{2.10}$$

— *En regardant attentivement la preuve précédente on constate que la conclusion de la proposition 2.0.5 reste vraie sans indépendance dans l'étape d'initialisation et sans indépendance conditionnelle dans les étapes de sélection et de mutation. En particulier, elle reste vraie si l'étape de sélection repose sur l'échantillonnage systématique.*

Théorème 2.0.7. *On suppose que pour tout* $p \in \{0, \ldots, n-1\}$, $0 < \underline{g}_p := \inf_{x_{0:p} \in E^{p+1}} g_p(x_{0:p}) \leq \bar{g}_p := \sup_{x_{0:p} \in E^{p+1}} g_p(x_{0:p}) < \infty$ *et que les propriétés d'indépendance et d'indépendance conditionnelle sont satisfaites à l'initialisation et lors des sélections et des mutations. Alors pour tout* $p \in \{0, \ldots, n\}$ *et toute fonction* $h_p : E^{p+1} \to \mathbb{R}$ *mesurable bornée (par* $\overline{|h_p|} < \infty$*),*

$$\sup_{M \geq 1} \sqrt{M} \mathbb{E}^{1/4} \left[(\eta_p^M(h_p) - \eta_p(h_p))^4 \right] < \infty \tag{2.11}$$

$$\sup_{M \geq 1} \sqrt{M} \mathbb{E}^{1/4} \left[(\gamma_p^M(h_p) - \gamma_p(h_p))^4 \right] < \infty. \tag{2.12}$$

En outre $\mathbb{P}\left(\lim_{M\to\infty} \eta_p^M(h_p) = \eta_p(h_p)\right) = 1 = \mathbb{P}\left(\lim_{M\to\infty} \gamma_p^M(h_p) = \gamma_p(h_p)\right)$.

Pour des résultats plus fins (convergence presque sûre sous des hypothèses plus faibles, théorème de la limite centrale associé, inégalités de concentration), nous renvoyons au livre de Del Moral [22] et au problème 2.0.9 pour un théorème de la limite centrale dans le cas particulier de l'échantillonnage multinomial.

Comme $\mathbb{E}[f_n(X_{0:n})] = \gamma_n(\tilde{f}_n)$ avec $\tilde{f}_n(x_{0:n}) = f_n(x_{0:n}) \prod_{p=0}^{n-1} g_p^{-1}(x_{0:p})$ bornée si f_n est bornée et les fonctions g_p minorées par une constante positive, le théorème implique le corollaire suivant.

Corollaire 2.0.8. *Sous les hypothèses d'indépendance et de minoration/majoration des fonctions* g_p *du théorème 2.0.7, si* $f_n : E^{n+1} \to \mathbb{R}$ *est bornée,*

$$\sup_{M \geq 1} \sqrt{M} \mathbb{E}^{1/4} \left[(\gamma_n^M(\tilde{f}_n) - \mathbb{E}[f_n(X_{0:n})])^4 \right] < \infty$$

et $\mathbb{P}\left(\lim_{M\to\infty} \gamma_n^M(\tilde{f}_n) = \mathbb{E}[f_n(X_{0:n})]\right) = 1$.

Cette fois, la propriété d'indépendance conditionnelle à l'étape de sélection va jouer un rôle clé dans la preuve du théorème, si bien que celle-ci ne s'applique pas à l'échantillonnage systématique.

Démonstration : La condition (2.11) implique que $\mathbb{E}\left[\sum_{M\geq 1} (\eta_p^M(h_p) - \eta_p(h_p))^4\right] < \infty$ d'où $\mathbb{P}\left(\sum_{M\geq 1} (\eta_p^M(h_p) - \eta_p(h_p))^4 < \infty\right) = 1 = \mathbb{P}(\lim_{M\to\infty} (\eta_p^M(h_p) - \eta_p(h_p))^4 = 0)$. Par le même raisonnement, (2.12) implique que $\mathbb{P}(\lim_{M\to\infty} (\gamma_p^M(h_p) - \gamma_p(h_p))^4 = 0) = 1$.

Montrons l'estimation (2.11) par récurrence sur p. Pour $p=0$, on a

$$\mathbb{E}[(\eta_0^M(h_0)-\eta_0(h_0))^4] = \frac{1}{M^4}\sum_{m_1,m_2,m_3,m_4=1}^{M} \mathbb{E}\left[\prod_{i=1}^4 (h_0(X_0^{0,m_i})-\eta_0(h_0))\right].$$

Parmi les M^4 espérances qui figurent dans la somme au second membre beaucoup sont nulles : en effet dès que l'un des quatre indices m_i est différent des trois autres, par indépendance des $(X_0^{0,m})_{1\leq m\leq M}$, l'espérance est égale à l'espérance du produit des termes correspondant aux trois autres indices par $\mathbb{E}[(h_0(X_0^{0,m_i})-\eta_0(h_0)]=0$.
Il reste donc seulement les contributions des cas où

— les 4 indices sont égaux soit M termes en $\mathbb{E}[(h_0(X_0^{0,1})-\eta_0(h_0))^4]$

— 2 des indices prennent une valeur et les 2 autres une autre valeur soit $3M(M-1)$ termes en $\mathbb{E}^2[(h_0(X_0^{0,1})-\eta_0(h_0))^2]$ (3 pour le choix de l'indice $m_k,\ k=2,3$ ou 4, égal au premier indice m_1, $M(M-1)$ pour le choix de deux valeurs différentes parmi M).

Ainsi

$$\mathbb{E}[(\eta_0^M(h_0)-\eta_0(h_0))^4] = \frac{1}{M^3}\mathbb{E}[(h_0(X_0^{0,1})-\eta_0(h_0))^4] + \frac{3(M-1)}{M^3}(\text{Var }(h_0(X_0^{0,1})))^2.$$

L'hypothèse de récurrence est donc satisfaite au rang $p=0$. Supposons la satisfaite au rang p et déduisons la au rang $p+1$. Reprenons les notations de la preuve de la proposition 2.0.5. Comme, d'après (2.9), $\mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{F}_p] = \frac{\eta_p^M(g_pP_{p+1}h_{p+1})}{\eta_p^M(g_p)}$ et d'après (2.10), $\eta_{p+1}(h_{p+1}) = \frac{\eta_p(g_pP_{p+1}h_{p+1})}{\eta_p(g_p)}$,

$$\eta_{p+1}^M(h_{p+1})-\eta_{p+1}(h_{p+1}) = \eta_{p+1}^M(h_{p+1})-\mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{F}_p]+\frac{\eta_p^M(g_pP_{p+1}h_{p+1})}{\eta_p^M(g_p)}-\frac{\eta_p(g_pP_{p+1}h_{p+1})}{\eta_p(g_p)}.$$

Donc

$$\begin{aligned}\mathbb{E}[(\eta_{p+1}^M(h_{p+1})-\eta_{p+1}(h_{p+1}))^4] &\leq 8\mathbb{E}\left[\left(\eta_{p+1}^M(h_{p+1})-\mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{F}_p]\right)^4\right]\\ &+8\mathbb{E}\left[\left(\frac{\eta_p^M(g_pP_{p+1}h_{p+1})}{\eta_p^M(g_p)}-\frac{\eta_p(g_pP_{p+1}h_{p+1})}{\eta_p(g_p)}\right)^4\right].\end{aligned} \tag{2.13}$$

Pour le second terme au second membre, on utilise la décomposition suivante

$$\begin{aligned}\frac{\eta_p^M(g_pP_{p+1}h_{p+1})}{\eta_p^M(g_p)}-\frac{\eta_p(g_pP_{p+1}h_{p+1})}{\eta_p(g_p)} =& \frac{\eta_p^M(g_pP_{p+1}h_{p+1})-\eta_p(g_pP_{p+1}h_{p+1})}{\eta_p^M(g_p)}\\ &+\frac{\eta_p(g_pP_{p+1}h_{p+1})(\eta_p(g_p)-\eta_p^M(g_p))}{\eta_p^M(g_p)\eta_p(g_p)}.\end{aligned}$$

Comme $0<\frac{1}{\eta_p^M(g_p)}\leq\frac{1}{\underline{g}_p}$ et $\left|\frac{\eta_p(g_pP_{p+1}h_{p+1})}{\eta_p(g_p)}\right|\leq\overline{|h_{p+1}|}$, on en déduit que

$$\begin{aligned}\mathbb{E}\left[\left(\frac{\eta_p^M(g_pP_{p+1}h_{p+1})}{\eta_p^M(g_p)}-\frac{\eta_p(g_pP_{p+1}h_{p+1})}{\eta_p(g_p)}\right)^4\right] \leq& \frac{8}{\underline{g}_p^4}\mathbb{E}[(\eta_p^M(g_pP_{p+1}h_{p+1})-\eta_p(g_pP_{p+1}h_{p+1}))^4]\\ &+\frac{8\overline{|h_{p+1}|}^4}{\underline{g}_p^4}\mathbb{E}[(\eta_p(g_p)-\eta_p^M(g_p))^4].\end{aligned}$$

L'hypothèse de récurrence au rang p appliquée avec $h_p = g_p P_{p+1} h_{p+1}$ et avec $h_p = g_p$ assure donc que

$$\sup_{M\geq 1} M^2 \mathbb{E}\left[\left(\frac{\eta_p^M(g_p P_{p+1} h_{p+1})}{\eta_p^M(g_p)} - \frac{\eta_p(g_p P_{p+1} h_{p+1})}{\eta_p(g_p)}\right)^4\right] < \infty. \tag{2.14}$$

Pour traiter le premier terme au second membre de (2.13), on utilise les propriétés de l'étape de mutation puis de l'étape de sélection qui assurent que les variables aléatoires $(X_{0:p+1}^{p+1,m})_{1\leq m\leq M}$ sont conditionnellement indépendantes sachant $\mathcal{F}_p$. Comme

$$\begin{aligned}
&\mathbb{E}\left[\left(\eta_{p+1}^M(h_{p+1}) - \mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{F}_p]\right)^4\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\left(\frac{1}{M}\sum_{m=1}^M (h_{p+1}(X_{0:p+1}^{p+1,m}) - \mathbb{E}[h_{p+1}(X_{0:p+1}^{p+1,m})|\mathcal{F}_p])\right)^4 \middle| \mathcal{F}_p\right]\right] \\
&= \frac{1}{M^4}\sum_{m_1,m_2,m_3,m_4=1}^M \mathbb{E}\left[\mathbb{E}\left[\prod_{i=1}^4 (h_{p+1}(X_{0:p+1}^{p+1,m_i}) - \mathbb{E}[h_{p+1}(X_{0:p+1}^{p+1,m_i})|\mathcal{F}_p])\middle|\mathcal{F}_p\right]\right].
\end{aligned}$$

L'indépendance conditionnelle assure que dès que l'un des quatre indices m_i est distinct des trois autres $\mathbb{E}\left[\prod_{i=1}^4 (h_{p+1}(X_{0:p+1}^{p+1,m_i}) - \mathbb{E}[h_{p+1}(X_{0:p+1}^{p+1,m_i})|\mathcal{F}_p])\middle|\mathcal{F}_p\right] = 0$. Comme dans le cas $p = 0$, on en déduit avec la bornitude de h_{p+1} que

$$\sup_{M\geq 1} M^2 \mathbb{E}\left[\left(\eta_{p+1}^M(h_{p+1}) - \mathbb{E}[\eta_{p+1}^M(h_{p+1})|\mathcal{F}_p]\right)^4\right] < \infty.$$

En combinant cette inégalité, (2.14) et (2.13), on conclut que l'hypothèse de récurrence est vraie au rang p, ce qui achève la démonstration de (2.11). Pour en déduire (2.12), on note $h_k = g_k$ pour $k \in \{0, \ldots, p-1\}$ de manière à ce que

$$\gamma_p^M(h_p) - \gamma_p(h_p) = \prod_{k=0}^p \eta_k^M(h_k) - \prod_{k=0}^p \eta_k(h_k) = \sum_{k=0}^p \left(\prod_{j=0}^{k-1}\eta_j^M(h_j)\right)(\eta_k^M(h_k) - \eta_k(h_k))\left(\prod_{j=k+1}^p \eta_j(h_j)\right).$$

Par conséquent,

$$\begin{aligned}
(\gamma_p^M(h_p) - \gamma_p(h_p))^4 &\leq (p+1)^3 \sum_{k=0}^p \left(\prod_{j=0}^{k-1}\eta_j^M(h_j) \prod_{j=k+1}^p \eta_j(h_j)\right)^4 (\eta_k^M(h_k) - \eta_k(h_k))^4 \\
&\leq (p+1)^3 \sum_{k=0}^p \left(\prod_{\substack{j=0\\ j\neq k}}^p \overline{|h_j|}\right)^4 (\eta_k^M(h_k) - \eta_k(h_k))^4.
\end{aligned}$$

En prenant l'espérance dans cette inégalité, on déduit (2.12) de (2.11). □

Problème 2.0.9. *On se place dans le cadre et les notations de ce chapitre. On suppose que pour tout $p \in \mathbb{N}$ la fonction $g_p : E^{p+1} \to \mathbb{R}$ est mesurable et telle que*

$$0 < \underline{g}_p := \inf_{x_{0:p}\in E^{p+1}} g_p(x_{0:p}) \leq \bar{g}_p := \sup_{x_{0:p}\in E^{p+1}} g_p(x_{0:p}) < \infty.$$

On suppose que les positions initiales des particules $(X_0^{0,m})_{m\geq 1}$ *sont i.i.d. suivant la probabilité* η_0 *et que la sélection est effectuée par échantillonnage multinomial, résiduel ou stratifié. Pour* $h_p : E^{p+1} \to \mathbb{R}$ *mesurable bornée, on pose* $\hat{\eta}_p^M(h_p) = \frac{\eta_p^M(g_p h_p)}{\eta_p^M(g_p)}$ *et* $\hat{\eta}_p(h_p) = \frac{\eta_p(g_p h_p)}{\eta_p(g_p)}$. *Pour* $h_{p+1} : E^{p+2} \to \mathbb{R}$ *mesurable bornée, on note* $P_{p+1}h_{p+1} : E^{p+1} \to \mathbb{R}$ *la fonction définie par* $P_{p+1}h_{p+1}(x_{0:p}) = \int_{x_{p+1}\in E} h_{p+1}(x_{0:p+1})P_{p+1}(x_p, dx_{p+1})$.

1. *Que peut-on dire de* $\mathbb{P}\left(\lim_{M\to\infty}(\eta_p^M(g_p h_p), \eta_p^M(g_p)) = (\eta_p(g_p h_p), \eta_p(g_p))\right)$ *? En déduire que* $\mathbb{P}\left(\lim_{M\to\infty}\hat{\eta}_p^M(h_p) = \hat{\eta}_p(h_p)\right) = 1$.
2. *Vérifier que* $\eta_{p+1}(h_{p+1}) = \hat{\eta}_p(P_{p+1}h_{p+1})$.
3. *Pour quelle fonction* $h_{p+1} : E^{p+2} \to \mathbb{R}$, *a-t-on* $\frac{1}{M}\sum_{m=1}^M h_p(X_{0:p}^{p+1,m}) = \eta_{p+1}^M(h_{p+1})$ *? Vérifier qu'alors* $\eta_{p+1}(h_{p+1}) = \hat{\eta}_p(h_p)$ *et en déduire que* $\mathbb{P}\left(\lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^M h_p(X_{0:p}^{p+1,m}) = \hat{\eta}_p(h_p)\right) = 1$.
4. *Montrer que lorsque* $M \to \infty$, *pour* $h_0 : E \to \mathbb{R}$ *mesurable bornée,* $\sqrt{M}(\eta_0^M(h_0) - \eta_0(h_0))$ *converge en loi vers* $\mathcal{N}_1(0, \mathbb{V}_0(h_0))$ *en explicitant* $\mathbb{V}_0(h_0)$.

On suppose désormais que l'étape de sélection est effectuée suivant l'échantillonnage multinomial. L'objectif de la suite du problème est de montrer par récurrence sur p que pour $h_p : E^{p+1} \to \mathbb{R}$ *mesurable bornée,* $\sqrt{M}(\eta_p^M(h_p) - \eta_p(h_p))$ *converge en loi lorsque* $M \to \infty$ *vers* $\mathcal{N}_1(0, \mathbb{V}_p(h_p))$. *On suppose l'hypothèse de récurrence satisfaite au rang p et on se donne* $h_{p+1} : E^{p+2} \to \mathbb{R}$ *mesurable bornée.*

5. *Pour* $\hat{h}_p : E^{p+1} \to \mathbb{R}$ *mesurable bornée, on pose* $\tilde{h}_p = \frac{g_p}{\eta_p(g_p)}(\hat{h}_p - \hat{\eta}_p(\hat{h}_p))$. *Remarquer que* $\eta_p(\tilde{h}_p) = 0$ *et* $\hat{\eta}_p^M(\hat{h}_p) - \hat{\eta}_p(\hat{h}_p) = \frac{\eta_p(g_p)}{\eta_p^M(g_p)}\eta_p^M(\tilde{h}_p)$. *Vérifier que* $\sqrt{M}\left(\hat{\eta}_p^M(\hat{h}_p) - \hat{\eta}_p(\hat{h}_p) - \eta_p^M(\tilde{h}_p)\right)$ *converge en probabilité vers 0 lorsque* $M \to \infty$ *et en déduire que* $\sqrt{M}(\hat{\eta}_p^M(\hat{h}_p) - \hat{\eta}_p(\hat{h}_p))$ *converge en loi vers* $\mathcal{N}_1\left(0, \mathbb{V}_p(\tilde{h}_p)\right)$.
6. *Vérifier que pour* $v \in \mathbb{R}$ *et* $\mathcal{F}_p = \sigma((X_{0:k}^{k,m})_{1\leq m\leq M, 0\leq k\leq p})$,
$$\mathbb{E}\left[e^{iv\sqrt{M}(\eta_{p+1}^M(h_{p+1})-\eta_{p+1}(h_{p+1}))}\right] = \mathbb{E}\Bigg[e^{iv\sqrt{M}(\hat{\eta}_p^M(P_{p+1}h_{p+1})-\hat{\eta}_p(P_{p+1}h_{p+1}))} \times \mathbb{E}\left[e^{iv\sqrt{M}\left(\eta_{p+1}^M(h_{p+1})-\hat{\eta}_p^M(P_{p+1}h_{p+1})\right)}|\mathcal{F}_p\right]\Bigg].$$
7. *On pose* $\hat{\eta}_p^M P_{p+1}(dx_{0:p+1}) := \hat{\eta}_p^M(dx_{0:p})P_{p+1}(x_p, dx_{p+1})$. *Vérifier que* $\hat{\eta}_p^M P_{p+1}(h_{p+1}) = \hat{\eta}_p^M(P_{p+1}h_{p+1})$ *et montrer que conditionnellement à* $\mathcal{F}_p$, *les trajectoires* $(X_{0:p+1}^{p+1,m})_{1\leq m\leq M}$ *sont i.i.d. suivant* $\hat{\eta}_p^M P_{p+1}(dx_{0:p+1})$.
8. *Vérifier que pour* $x \in \mathbb{R}$, $\left|e^{ix} - 1 - ix + \frac{x^2}{2}\right| \leq \frac{|x|^3}{6}$ *puis, en notant* $\overline{|h_{p+1}|} = \sup_{x_{0:p+1}\in E^{p+2}}|h_{p+1}(x_{0:p+1})|$, *montrer que pour* $v \in \mathbb{R}$,
$$\Bigg|\mathbb{E}\left[e^{i\frac{v}{\sqrt{M}}(h_{p+1}(X_{0:p+1}^{p+1,1})-\hat{\eta}_p^M(P_{p+1}h_{p+1}))}|\mathcal{F}_p\right] - 1 + \frac{v^2}{2M}\left(\hat{\eta}_p^M(P_{p+1}h_{p+1}^2) - (\hat{\eta}_p^M(P_{p+1}h_{p+1}))^2\right)\Bigg| \leq \frac{4|v|^3\overline{|h_{p+1}|}^3}{3M^{3/2}}.$$
9. *On pose* $\mathbb{W}_{p+1}(h_{p+1}) := \hat{\eta}_p(P_{p+1}h_{p+1}^2) - (\hat{\eta}_p(P_{p+1}h_{p+1}))^2$ *et*
$$\varepsilon_M = M\left(\mathbb{E}\left[e^{i\frac{v}{\sqrt{M}}(h_{p+1}(X_{0:p+1}^{p+1,1})-\hat{\eta}_p^M(P_{p+1}h_{p+1}))}|\mathcal{F}_p\right] - 1 + \frac{v^2}{2M}\mathbb{W}_{p+1}(h_{p+1})\right)$$

Vérifier que

$$|\varepsilon_M| \leq \frac{4|v|^3\overline{|h_{p+1}|}^3}{3\sqrt{M}} + \frac{v^2}{2}|\hat{\eta}_p^M(P_{p+1}h_{p+1}^2) - \hat{\eta}_p(P_{p+1}h_{p+1}^2)|$$
$$+ \frac{v^2}{2}|(\hat{\eta}_p^M(P_{p+1}h_{p+1}))^2 - (\hat{\eta}_p(P_{p+1}h_{p+1}))^2|$$

et en déduire le comportement de la suite $(\varepsilon_M)_{M\geq 1}$ *lorsque* $M \to \infty$.

10. *Comment comparer* $P_{p+1}h_{p+1}^2$ *et* $(P_{p+1}h_{p+1})^2$ *? En déduire que* $\mathbb{W}_{p+1}(h_{p+1}) \geq 0$.
11. *À l'aide de la formule du binôme, vérifier que pour* $M \geq \frac{v^2\mathbb{W}_{p+1}(h_{p+1})}{4}$,

$$\left|\left(1 - \frac{v^2\mathbb{W}_{p+1}(h_{p+1})}{2M} + \frac{\varepsilon_M}{M}\right)^M - \left(1 - \frac{v^2\mathbb{W}_{p+1}(h_{p+1})}{2M}\right)^M\right| \leq \sum_{k=1}^{M} \frac{|\varepsilon_M|^k}{k!} \leq e^{|\varepsilon_M|} - 1$$

et en déduire que $\mathbb{E}\left[e^{iv\sqrt{M}\left(\eta_{p+1}^M(h_{p+1}) - \hat{\eta}_p^M(P_{p+1}h_{p+1})\right)}|\mathcal{F}_p\right]$ *converge presque sûrement vers* $e^{-\frac{v^2\mathbb{W}_{p+1}(h_{p+1})}{2}}$ *lorsque* $M \to \infty$*. puis que*

$$\lim_{M\to\infty} \mathbb{E}\left[\left|\mathbb{E}\left[e^{iv\sqrt{M}\left(\eta_{p+1}^M(h_{p+1}) - \hat{\eta}_p^M(P_{p+1}h_{p+1})\right)}|\mathcal{F}_p\right] - e^{-\frac{v^2\mathbb{W}_{p+1}(h_{p+1})}{2}}\right|\right] = 0.$$

12. *Vérifier que*

$$\left|\mathbb{E}\left[e^{iv\sqrt{M}(\eta_{p+1}^M(h_{p+1}) - \eta_{p+1}(h_{p+1}))}\right] - e^{-\frac{v^2\mathbb{W}_{p+1}(h_{p+1})}{2}}\mathbb{E}\left[e^{iv\sqrt{M}(\hat{\eta}_p^M(P_{p+1}h_{p+1}) - \hat{\eta}_p(P_{p+1}h_{p+1}))}\right]\right|$$
$$\leq \mathbb{E}\left[\left|\mathbb{E}\left[e^{iv\sqrt{M}\left(\eta_{p+1}^M(h_{p+1}) - \hat{\eta}_p^M(P_{p+1}h_{p+1})\right)}|\mathcal{F}_p\right] - e^{-\frac{v^2\mathbb{W}_{p+1}(h_{p+1})}{2}}\right|\right]$$

et conclure que, lorsque $M \to \infty$, $\sqrt{M}(\eta_{p+1}^M(h_{p+1}) - \eta_{p+1}(h_{p+1}))$ *converge en loi vers* $\mathcal{N}_1(0, \mathbb{V}_{p+1}(h_{p+1}))$ *avec*

$$\mathbb{V}_{p+1}(h_{p+1}) = \mathbb{W}_{p+1}(h_{p+1}) + \mathbb{V}_p(\mathbb{L}_p h_{p+1}) \text{ où } \mathbb{L}_p h_{p+1} = \frac{g_p}{\eta_p(g_p)}(P_{p+1}h_{p+1} - \hat{\eta}_p(P_{p+1}h_{p+1})).$$

Chapitre 3

Discrétisation des EDS

3.1 Équations différentielles stochastiques

3.1.1 Existence et unicité, applications en finance

Soit $T > 0$ un horizon de temps (la maturité d'une option par exemple). Sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ muni d'une filtration $(\mathcal{F}_t)_{t\geq 0}$, on considère un $\mathcal{F}_t$-mouvement brownien $W_t = \begin{pmatrix} W_t^1 \\ \vdots \\ W_t^d \end{pmatrix}$ à valeurs $\mathbb{R}^d$ et Y une variable aléatoire $\mathcal{F}_0$ mesurable à valeurs $\mathbb{R}^n$. On se donne également des coefficients $\sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ et $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ mesurables et on s'intéresse à l'Equation Différentielle Stochastique

$$\begin{cases} dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt \\ X_0 = Y \end{cases} \tag{3.1}$$

Définition 3.1.1. *On appelle solution de l'EDS* (3.1) *un processus* $(X_t)_{t\in[0,T]}$ *continu* $\mathcal{F}_t$*-adapté à valeurs* $\mathbb{R}^n$ *tel que*
— *p.s.,* $\int_0^T |b(s, X_s)| + |\sigma(s, X_s)|^2 ds < +\infty$,
— *p.s.,* $\forall t \in [0,T]$, $X_t = Y + \int_0^t \sigma(s, X_s)dW_s + \int_0^t b(s, X_s)ds$.

Le principal résultat d'existence et d'unicité pour les EDS est l'analogue du théorème de Cauchy-Lipschitz pour les équations différentielles ordinaires et s'obtient sous des hypothèses similaires sur les coefficients de l'équation.

Théorème 3.1.2 (d'Itô)**.** *On suppose que*

$$\text{(Lip) } \exists \text{K} > 0,\ \forall \text{t} \in [0, \text{T}],\ \begin{cases} \forall x \in \mathbb{R}^n,\ |\sigma(t,x)| + |b(t,x)| \leq K(1 + |x|) \\ \forall x, y \in \mathbb{R}^n,\ |\sigma(t,x) - \sigma(t,y)| + |b(t,x) - b(t,y)| \leq K|x - y| \end{cases} .$$

Alors l'EDS (3.1) *admet une unique solution* $(X_t)_{t\in[0,T]}$ *(si* X'_t *est une autre solution, alors p.s.,* $\forall t \in [0,T]$, $X_t = X'_t$*). En outre, si* $\mathbb{E}(|Y|^2) < +\infty$*, alors*

$$\mathbb{E}\left(\sup_{t\leq T} |X_t|^2\right) \leq C(1 + \mathbb{E}\left(|Y|^2\right)) \text{ où } C \text{ ne dépend pas de } Y. \tag{3.2}$$

Démonstration : La démonstration repose sur une technique de point fixe. L'espace

$$\mathcal{E} = \left\{ (X_t)_{t\in[0,T]} \text{ processus } \mathcal{F}_t\text{-adapté continu à valeurs } \mathbb{R}^n \text{ t.q. } \mathbb{E}\left(\sup_{t\le T}|X_t|^2\right) < +\infty \right\}$$

muni de la norme $\sqrt{\mathbb{E}\left(\sup_{t\le T}|X_t|^2\right)}$ est un espace de Banach.

• Dans le cas où $\mathbb{E}(|Y|^2) < +\infty$ commençons par montrer que toute solution X de (3.1) vérifie (3.2) et appartient donc à $\mathcal{E}$. Pour $\nu_m = \inf\{t \ge 0 : |X_t| \ge m\}$ (convention $\inf \emptyset = T$), on a

$$\sup_{s\le t}|X_{s\wedge\nu_m}|^2 \le 3\left(|Y|^2 + \sup_{s\le t}\left|\int_0^s 1_{\{r\le\nu_m\}}\sigma(r,X_r)dW_r\right|^2 + \left(\int_0^t |b(r\wedge\nu_m, X_{r\wedge\nu_m})|dr\right)^2\right).$$

En utilisant l'inégalité de Doob pour l'intégrale stochastique, l'inégalité de Cauchy-Schwarz pour l'intégrale classique, puis l'hypothèse (Lip), on en déduit que pour t dans $[0,T]$

$$\begin{aligned}\mathbb{E}\left(\sup_{s\le t}|X_{s\wedge\nu_m}|^2\right) &\le C\left(\mathbb{E}(|Y|^2) + \int_0^t \mathbb{E}\left(|\sigma(r\wedge\nu_m, X_{r\wedge\nu_m})|^2 + t|b(r\wedge\nu_m, X_{r\wedge\nu_m})|^2\right)dr\right)\\ &\le C\left(\mathbb{E}(|Y|^2) + \int_0^t 1 + \mathbb{E}\left(\sup_{s\le r}|X_{s\wedge\nu_m}|^2\right)dr\right)\end{aligned}$$

où la constante C peut changer d'une ligne à l'autre mais ne dépend ni de t dans $[0,T]$ ni de m (on a par exemple majoré le facteur t par T). Par le Lemme de Gronwall, on en déduit que

$$\mathbb{E}\left(\sup_{s\le T}|X_{s\wedge\nu_m}|^2\right) \le C(\mathbb{E}(|Y|^2) + T)e^{CT}.$$

Comme par continuité des trajectoires de X, p.s. $\nu_m = T$ pour m assez grand, on a p.s. $\sup_{s\le T}|X_{s\wedge\nu_m}| = \sup_{s\le T\wedge\nu_m}|X_s| \to \sup_{s\le T}|X_s|$ lorsque $m\to+\infty$. On déduit alors (3.2) du lemme de Fatou.

On introduit ensuite Φ qui à un processus $X = (X_t)_{t\in[0,T]}$ de $\mathcal{E}$ associe le processus $(\Phi(X)_t)_{t\in[0,T]}$ défini par

$$\Phi(X)_t = Y + \int_0^t \sigma(s,X_s)dW_s + \int_0^t b(s,X_s)ds,$$

dont on vérifie facilement qu'il appartient à $\mathcal{E}$ en reprenant le calcul ci-dessus mais sans la technique de localisation. Comme nous avons vérifié que toute solution de (3.1) est dans $\mathcal{E}$, un processus X est solution de (3.1) si et seulement si c'est un point fixe de Φ.
Pour $X, X' \in \mathcal{E}$ et $t \le T$, en utilisant l'inégalité de Cauchy-Schwarz puis l'inégalité de Doob et enfin l'hypothèse (Lip), il vient

$$\begin{aligned}\mathbb{E}\left(\sup_{s\le t}|\Phi(X)_s - \Phi(X')_s|^2\right) &\le 2\mathbb{E}\Bigg(\sup_{s\le t}\left|\int_0^s(\sigma(r,X_r)-\sigma(r,X'_r))dW_r\right|^2\\ &\qquad + t\int_0^t |b(r,X_r)-b(r,X'_r)|^2dr\Bigg)\\ &\le C\int_0^t \mathbb{E}\left(|\sigma(r,X_r)-\sigma(r,X'_r)|^2 + |b(r,X_r)-b(r,X'_r)|^2\right)dr\\ &\le C\int_0^t \mathbb{E}\left(\sup_{s\le r}|X_s - X'_s|^2\right)dr\end{aligned}$$

où la constante C ne dépend pas de t ni de X et X'. En itérant cette inégalité, on obtient que pour $k \in \mathbb{N}^*$, si Φ^k désigne la composée k-ième de Φ,

$$\begin{aligned}\|\Phi^k(X) - \Phi^k(X')\|^2 &\leq C^k \int_0^T \int_0^{r_1} \dots \int_0^{r_{k-1}} \mathbb{E}\left(\sup_{s \leq r_k} |X_s - X'_s|^2\right) dr_k \dots dr_1 \\ &\leq C^k \|X - X'\|^2 \int_0^T \int_0^{r_1} \dots \int_0^{r_{k-1}} dr_k \dots dr_1 = \frac{C^k T^k}{k!} \|X - X'\|^2.\end{aligned}$$

Donc pour $\bar{k}$ assez grand, $\Phi^{\bar{k}}$ est une contraction sur $\mathcal{E}$. Soit X son unique point fixe. Tout point fixe de Φ est point fixe de $\Phi^{\bar{k}}$ et donc égal à X. Par ailleurs,

$$\Phi^{\bar{k}}(\Phi(X)) = \Phi^{\bar{k}+1}(X) = \Phi(\Phi^{\bar{k}}(X)) = \Phi(X)$$

assure que $\Phi(X)$ est point fixe de $\Phi^{\bar{k}}$ et donc égal à X. On conclut donc que (3.1) admet X comme unique solution.

• Dans le cas où $\mathbb{E}(|Y|^2) = +\infty$, pour $m \in \mathbb{N}^*$, on note $(X_t^m)_{t \in [0,T]}$ la solution de l'EDS issue de la condition initiale $Y 1_{\{|Y| < m\}}$. On vérifie ensuite que

$$X_t = \sum_{m \in \mathbb{N}^*} 1_{\{m-1 \leq |Y| < m\}} X_t^m$$

est solution de (3.1) et que toute autre solution lui est égale. □

Exemple. *La plupart des modèles d'actifs financiers reposent sur des EDS.*

• **le processus d'Ornstein-Uhlenbeck :** *pour $n = d = 1$ et $c \in \mathbb{R}$, l'unique solution de*

$$dX_t = dW_t + cX_t dt, \; X_0 = Y$$

est $X_t = Ye^{ct} + \int_0^t e^{c(t-s)} dW_s$.

• **le modèle de Black-Scholes :** *pour $n = d = 1$ et $\sigma, r \in \mathbb{R}$, l'unique solution de*

$$dX_t = \sigma X_t dW_t + rX_t dt, \; X_0 = Y$$

est $X_t = Ye^{\sigma W_t + (r - \frac{\sigma^2}{2})t}$.

• **le modèle à volatilité locale (ou modèle de Dupire) :** *pour $n = d = 1$, $r \in \mathbb{R}$, et $\eta : [0,T] \times \mathbb{R} \to \mathbb{R}$ vérifiant*

$$\exists K > 0, \; \forall (t,x) \in [0,T] \times \mathbb{R}, \; |\eta(t,x)| + |x||\partial_x \eta(t,x)| \leq K,$$

l'équation différentielle stochastique

$$dX_t = \eta(t, X_t) X_t dW_t + rX_t dt, \; X_0 = Y$$

admet une unique solution. On vérifie que $X_t = Ye^{\int_0^t \eta(s,X_s)dW_s + (rt - \frac{1}{2}\int_0^t \eta^2(s,X_s)ds)}$.

• **le modèle à volatilité stochastique :** *pour $n = d = 2$, on peut considérer un modèle de la forme*

$$\begin{cases} dX_t = f(Y_t) X_t (\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2) + rX_t dt \\ dY_t = \eta(Y_t) dW_t^1 + b(Y_t) dt \end{cases}, (X_0, Y_0) = (x_0, y_0) \in \mathbb{R}_+^* \times \mathbb{R}$$

où $\eta, b : \mathbb{R} \to \mathbb{R}$ sont supposées lipschitziennes, $r \in \mathbb{R}$ et $\rho \in [-1, 1]$ est le coefficient de corrélation entre le brownien W_t^1 qui dirige le processus de volatilité et celui $B_t = \rho W_t^1 + \sqrt{1-\rho^2} W_t^2$ qui dirige l'évolution de l'actif. À cause du terme quadratique $f(y)x$, la fonction

$$\sigma(x, y) = \begin{pmatrix} \rho f(y)x & \sqrt{1-\rho^2} f(y)x \\ \eta(x) & 0 \end{pmatrix}$$

n'est globalement lipschitzienne que si et seulement si f est constante. On ne peut donc pas appliquer le théorème d'Itô au couple (X_t, Y_t). Néanmoins, l'équation différentielle stochastique autonome satisfaite par $(Y_t)_{t\in[0,T]}$ admet unique solution d'après le théorème d'Itô. On peut ensuite utiliser le caractère linéaire de l'équation satisfaite par X_t. Si f est localement bornée, alors la continuité de $(Y_t)_{t\in[0,T]}$ entraîne que $\mathbb{P}(\int_0^T f^2(Y_t)dt < \infty) = 1$ et donc que l'on peut définir pour tout t dans $[0, T]$, $\xi_t = x_0 e^{\int_0^t f(Y_s)dB_s + rt - \frac{1}{2}\int_0^t f^2(Y_s)ds}$. La formule d'Itô assure que $d\xi_t = f(Y_t)\xi_t dB_t + r\xi_t dt$ avec $\xi_0 = x_0$ si bien que le couple $(\xi_t, Y_t)_{t\in[0,T]}$ est solution de l'EDS de départ. Pour l'unicité, on remarque que $\zeta_t = e^{-\int_0^t f(Y_s)dB_s - rt + \frac{1}{2}\int_0^t f^2(Y_s)ds}$ satisfait $d\zeta_t = \zeta_t(-f(Y_t)dB_t + (f^2(Y_t) - r)dt)$. Ainsi, par la formule d'intégration par partie, toute solution $(X_t)_{t\in[0,T]}$ vérifie

$$\begin{aligned} d(\zeta_t X_t) &= \zeta_t dX_t + X_t d\zeta_t + d\langle \zeta, X\rangle_t \\ &= X_t \zeta_t \big(f(Y_t)dB_t + rdt - f(Y_t)dB_t + (f^2(Y_t) - r)dt - f^2(Y_t)dt\big) = 0. \end{aligned}$$

Le prix d'un Call d'échéance T et de Strike K portant sur le sous-jacent X est donné par $C = \mathbb{E}\left(e^{-rT}(X_T - K)^+\right)$. Dans le cas du modèle de Black-Scholes, on dispose à la fois d'une formule fermée pour ce prix et de la possibilité de simuler exactement $X_T = ye^{\sigma W_T + (r - \frac{\sigma^2}{2})T}$ ($y > 0$ est le cours initial du sous-jacent) pour calculer C par la méthode de Monte Carlo. Mais dès que l'on complique un peu le modèle en supposant que la volatilité est soit une fonction locale du temps et du sous-jacent, soit stochastique, ces deux propriétés très agréables disparaissent en général. Pour pouvoir calculer le prix, on peut néanmoins discrétiser l'équation différentielle stochastique qui donne le modèle de façon à générer $\bar{X}_T$ qui approche X_T. La méthode de Monte Carlo consiste alors à approcher C par $\frac{1}{M}\sum_{i=1}^M e^{-rT}(\bar{X}_T^i - K)^+$ où les variables $\bar{X}_T^i$, générées par le même schéma mais pour des accroissements browniens independants sont i.i.d.. L'erreur se décompose alors classiquement en un terme de biais plus un terme d'erreur statistique :

$$\begin{aligned} C - \frac{1}{M}\sum_{i=1}^M e^{-rT}(\bar{X}_T^i - K)^+ =& \mathbb{E}\left(e^{-rT}(X_T - K)^+\right) - \mathbb{E}\left(e^{-rT}(\bar{X}_T - K)^+\right) \\ &+ \mathbb{E}\left(e^{-rT}(\bar{X}_T - K)^+\right) - \frac{1}{M}\sum_{i=1}^M e^{-rT}(\bar{X}_T^i - K)^+. \end{aligned}$$

Le comportement du terme d'erreur statistique est bien connu et découle du théorème de la limite centrale : pour peu que $\mathbb{E}\left([(\bar{X}_T - K)^+]^2\right) < +\infty$, ce terme se comporte comme $e^{-rT}\sqrt{\frac{\mathrm{Var}\,((\bar{X}_T - K)^+)}{M}}\, G$ où $G \sim \mathcal{N}_1(0, 1)$ lorsque M est grand. Dans le prochain chapitre, nous présenterons différentes techniques de réduction de variance permettant de diminuer ce terme.

Dans le présent chapitre, nous allons nous attacher à analyser le terme de biais qui dépend bien sûr du schéma de discrétisation retenu pour approcher X_T.

Remarque 3.1.3. — *En dimension $n = d = 1$, l'unicité énoncée dans le théorème 3.1.2 reste vraie sans supposer que le coefficient de diffusion σ est lipschitzien dans la variable d'espace x mais sous l'hypothèse plus faible d'existence d'une fonction $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ telle que $\int_{0+} \frac{du}{\rho(u)} = +\infty$ et que*

$$\forall t \in [0,T],\ \forall x, y \in \mathbb{R},\ |\sigma(t,x) - \sigma(t,y)|^2 \leq \rho(|x-y|).$$

C'est ce qui assure l'unicité pour l'EDS de Cox-Ingersoll-Ross

$$dX_t = \eta\sqrt{X_t}dW_t + (a - bX_t)dt,\ X_0 = Y \geq 0$$

où $\eta, a \geq 0$ et $b \in \mathbb{R}$. On peut également montrer l'existence pour cette EDS et on en déduit facilement l'existence et l'unicité pour le modèle à volatilité stochastique d'Heston où le carré de la volatilité est un processus de Cox-Ingersoll-Ross :

$$\begin{cases} dS_t = \sqrt{V_t}S_t dW_t^1 + rS_t dt \\ dV_t = \eta\sqrt{V_t}(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2) + (a - bV_t)dt \end{cases}.$$

— *Lorsque l'on ne fixe pas le mouvement Brownien mais que l'on cherche un couple $(X_t, W_t)_{t\in[0,T]}$ tel que $(W_t)_t$ est un mouvement brownien et que $(X_t)_t$ vérifie* (3.1) *on définit une notion de solution plus faible que celle étudiée plus haut.*
Soit par exemple $(B_t)_t$ est un mouvement brownien à valeurs $\mathbb{R}^d$ independant de Y à valeurs $\mathbb{R}^d$ et $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ une fonction bornée et mesurable (mais pas plus régulière). Par le théorème de Girsanov, sous la probabilité $\mathbb{Q}$ de densité $\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^T b(s, Y + B_s)dB_s - \frac{1}{2}\int_0^T b^2(s, Y + B_s)ds\right)$ le processus $(W_t = B_t - \int_0^t b(s, Y + B_s)ds)_{t\in[0,T]}$ est un mouvement Brownien. Si on pose $X_t = Y + B_t$, on a alors

$$X_t = Y + W_t + \int_0^t b(s, Y + B_s)ds = Y + W_t + \int_0^t b(s, X_s)ds.$$

Ainsi le couple (X_t, W_t) est solution faible de (3.1) *avec le coefficient de diffusion σ constant égal à la matrice identité $d \times d$.*

3.1.2 Propriétés des solutions

Sous l'hypothèse (Lip), il est possible de contrôler les moments de la solution de l'équation

$$dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt,\ X_0 = y \in \mathbb{R}^n \tag{3.3}$$

lorsque la condition initiale y est déterministe.

Lemme 3.1.4. *Soit $p \geq 1$. Il existe une constante C dépendant de p, de T, des coefficients σ et b mais pas de la condition initiale y telle que*

$$\sup_{t\leq T} \mathbb{E}(|X_t|^{2p}) \leq C(1 + |y|^{2p}). \tag{3.4}$$

Démonstration : La fonction $f : x \in \mathbb{R}^n \mapsto |x|^{2p}$ a pour gradient $\nabla f(x) = 2p|x|^{2p-2}x$ et pour hessienne $\nabla^2 f(x) = 2p|x|^{2p-4}(|x|^2 I_n + (2p-2)xx^*)$ où $I_n, xx^* \in \mathbb{R}^{n\times n}$ désignent respectivement la matrice identité et la matrice $(x_i x_j)_{ij}$. On note $a(t,x) = \sigma\sigma^*(t,x) \in \mathbb{R}^{n\times n}$. La formule d'Itô assure

$$\begin{aligned}|X_t|^{2p} =& |y|^{2p} + \int_0^t 2p|X_s|^{2p-2}X_s.b(s,X_s) + p|X_s|^{2p-4}\mathrm{tr}\left[(|X_s|^2 I_n + (2p-2)X_sX_s^*)a(s,X_s)\right]ds \\ &+ \int_0^t 2p|X_s|^{2p-2}X_s.\sigma(s,X_s)dW_s.\end{aligned}$$

Pour pouvoir assurer que l'espérance de l'intégrale stochastique est nulle, on utilise une procédure de localisation. Pour $\nu_m = \inf\{t \geq 0 : \ |X_t| \geq m\}$ (convention $\inf\emptyset = T$), $\mathbb{E}\left(\int_0^{\nu_m\wedge t} 2p|X_s|^{2p-2}X_s.\sigma(s,X_s)dW_s\right) = 0$, ce qui assure, en utilisant (Lip) pour la seconde inégalité, puis $|x|^{2p-2} + |x|^{2p-1} + |x|^{2p} \leq 3(1+|x|^{2p})$ pour la troisième

$$\begin{aligned}\mathbb{E}\left(|X_{\nu_m\wedge t}|^{2p}\right) = |y|^{2p} + \mathbb{E}\Bigg(\int_0^{\nu_m\wedge t} & 2p|X_s|^{2p-2}X_s.b(s,X_s) \\ &+ p|X_s|^{2p-4}\mathrm{tr}\left[(|X_s|^2 I_n + (2p-2)X_sX_s^*)a(s,X_s)\right]ds\Bigg) \\ \leq |y|^{2p} + C\int_0^t & \mathbb{E}\left(|X_{\nu_m\wedge s}|^{2p-1}(1+|X_{\nu_m\wedge s}|) + |X_{\nu_m\wedge s}|^{2p-2}(1+|X_{\nu_m\wedge s}|)^2\right)ds \\ \leq |y|^{2p} + C\int_0^t & 1 + \mathbb{E}\left(|X_{\nu_m\wedge s}|^{2p}\right)ds \\ \leq |y|^{2p} + Ct &+ C\int_0^t \mathbb{E}\left(|X_{\nu_m\wedge s}|^{2p}\right)ds,\end{aligned}$$

Comme $|X_{\nu_m\wedge s}| \leq 1_{\{\nu_m=0\}}|y| + 1_{\{\nu_m>0\}}m \leq |y|\vee m$, la fonction $s \mapsto \mathbb{E}[|X_{\nu_m\wedge s}|^{2p}]$ est localement intégrable. Le lemme de Gronwall implique alors que

$$\forall t \leq T,\ \mathbb{E}\left(|X_{\nu_m\wedge t}|^{2p}\right) \leq (|y|^{2p} + CT)e^{Ct} \leq (|y|^{2p} + CT)e^{CT}.$$

Comme lorsque $m \to +\infty$, $\nu_m \to T$ p.s., $X_{\nu_m\wedge t} \to X_t$ p.s. et le Lemme de Fatou implique que $\mathbb{E}\left(|X_t|^{2p}\right) \leq \liminf_{m\to+\infty}\mathbb{E}\left(|X_{\nu_m\wedge t}|^{2p}\right)$. On en déduit que

$$\forall t \leq T,\ \mathbb{E}\left(|X_t|^{2p}\right) \leq (|y|^{2p} + CT)e^{CT}.$$

□

Nous allons également contrôler les moments des accroissements de la solution de (3.3), en utilisant l'inégalité de Burkholder Davis Gundy qui généralise l'inégalité de Doob ($p = 1$).

Lemme 3.1.5. *Soit $p \geq \frac{1}{2}$. Il existe une constante $C_p > 0$ telle que pour tout processus $(H_s)_{s\geq 0}$ à valeurs $\mathbb{R}^{n\times d}$ $\mathcal{F}_s$-adapté et tout $t \geq 0$ tel que $\mathbb{P}\left(\int_0^t |H_s|^2 ds < +\infty\right) = 1$,*

$$\mathbb{E}\left(\sup_{s\leq t}\left|\int_0^s H_r dW_r\right|^{2p}\right) \leq C_p\mathbb{E}\left(\left(\int_0^t |H_r|^2 dr\right)^p\right). \tag{3.5}$$

Avec le lemme 3.1.4, nous allons en déduire la majoration suivante :

Proposition 3.1.6. *Soit $p \geq 1$. Il existe une constante C finie dépendant de p, des coefficients σ et b et de T mais pas de la condition initiale y telle que*

$$\forall 0 \leq s \leq t \leq T,\ \mathbb{E}(|X_t - X_s|^{2p}) \leq C(1+|y|^{2p})(t-s)^p. \tag{3.6}$$

Dans la preuve nous aurons besoin du résultat suivant : pour $q \geq 1$,

$$\left|\sum_{k=1}^{K} a_k\right|^q \leq K^{q-1}\sum_{k=1}^{K}|a_k|^q \text{ et } \left|\int_A f(x)dx\right|^q \leq |A|^{q-1}\int_A |f(x)|^q dx, \tag{3.7}$$

où $f : \mathbb{R}^\ell \to \mathbb{R}^n$ est supposée mesurable et $|A|$ désigne la mesure de Lebesgue du sous-ensemble borélien A de $\mathbb{R}^\ell$. Ces inégalités évidentes si $|A| \in \{0, +\infty\}$ découlent sinon de l'inégalité de Jensen qui assure :

$$\left(\frac{1}{K}\sum_{k=1}^{K}|a_k|\right)^q \leq \frac{1}{K}\sum_{k=1}^{K}|a_k|^q \text{ et } \left(\frac{1}{|A|}\int_A |f(x)|dx\right)^q \leq \frac{1}{|A|}\int_A |f(x)|^q dx.$$

Démonstration : Par l'inégalité triangulaire

$$|X_t - X_s| \leq \left|\int_s^t \sigma(r, X_r)dW_r\right| + \left|\int_s^t b(r, X_r)dr\right|.$$

Avec (3.7) et (3.5), on en déduit

$$\begin{aligned}
\mathbb{E}\left(|X_t - X_s|^{2p}\right) &\leq 2^{2p-1}\left[\mathbb{E}\left(\left|\int_s^t \sigma(r, X_r)dW_r\right|^{2p}\right) + \mathbb{E}\left(\left|\int_s^t b(r, X_r)dr\right|^{2p}\right)\right] \\
&\leq C\left[\mathbb{E}\left(\left(\int_s^t |\sigma(r, X_r)|^2 dr\right)^p\right) + (t-s)^{2p-1}\mathbb{E}\left(\int_s^t |b(r, X_r)|^{2p}dr\right)\right] \\
&\leq C\left[(t-s)^{p-1}\int_s^t \mathbb{E}\left(|\sigma(r, X_r)|^{2p}\right)dr + T^p(t-s)^{p-1}\int_s^t \mathbb{E}\left(|b(r, X_r)|^{2p}\right)dr\right].
\end{aligned}$$

On conclut facilement en remarquant que d'après (Lip), (3.7) et le Lemme 3.1.4 sur les moments de X_r,

$$\mathbb{E}\left(|\sigma(r, X_r)|^{2p} + |b(r, X_r)|^{2p}\right) \leq C\mathbb{E}((1 + |X_r|)^{2p}) \leq C2^{2p-1}(1 + \mathbb{E}(|X_r|^{2p}) \leq C(1 + |y|^{2p}).$$

□

3.2 Le schéma d'Euler

On subdivise l'intervalle $[0, T]$ en N sous-intervalles de même longueur et pour $k \in \{0, \ldots, N\}$, on pose $t_k = \frac{kT}{N}$. Le schéma d'Euler consiste à discrétiser l'EDS

$$dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt,\ X_0 = y \in \mathbb{R}^n$$

suivant la grille $(t_k)_k$ en posant

$$\begin{cases} \bar{X}_0 = y \\ \forall 0 \leq k \leq N-1,\ \bar{X}_{t_{k+1}} = \bar{X}_{t_k} + \sigma(t_k, \bar{X}_{t_k})(W_{t_{k+1}} - W_{t_k}) + b(t_k, \bar{X}_{t_k})(t_{k+1} - t_k) \end{cases} \tag{3.8}$$

Pour passer d'un instant de discrétisation au suivant on fige les coefficients de diffusion et de dérive à leur valeur au début de l'intervalle. Afin d'implémenter ce schéma, il suffit de savoir simuler les accroissements $(W_{t_{k+1}} - W_{t_k})_{0\leq k\leq N-1}$ qui sont i.i.d. suivant la loi gaussienne $\mathcal{N}_d(0, \frac{T}{N}I_d)$ où I_d désigne la matrice identité de dimension d. Pour effectuer les preuves de convergence, il est commode d'introduire le schéma d'Euler en temps continu défini par

$$\forall 0 \leq k \leq N-1,\ \forall t \in [t_k, t_{k+1}],\ \bar{X}_t = \bar{X}_{t_k} + \sigma(t_k, \bar{X}_{t_k})(W_t - W_{t_k}) + b(t_k, \bar{X}_{t_k})(t - t_k).$$

Bien sûr, même si pour $t \in [t_k, t_{k+1}]$ et $G \sim \mathcal{N}_d(0, I_d)$ vecteur gaussien centré indépendant de l'accroissement $W_{t_{k+1}} - W_{t_k}$ et admettant la matrice identité $d \times d$ comme matrice de covariance, le vecteur aléatoire $\frac{t-t_k}{t_{k+1}-t_k}(W_{t_{k+1}} - W_{t_k}) + \sqrt{\frac{(t_{k+1}-t)(t-t_k)}{t_{k+1}-t_k}}G$ a même loi conditionnelle sachant $W_{t_{k+1}} - W_{t_k}$ que $W_t - W_{t_k}$, il n'est pas possible de générer la trajectoire du schéma d'Euler en temps continu en plus d'un nombre fini d'instants.
Si pour $s \in [0, T]$, on note $\tau_s = \lfloor \frac{Ns}{T} \rfloor \times \frac{T}{N}$ le dernier instant de discrétisation avant s, on a

$$d\bar{X}_t = \sigma(\tau_t, \bar{X}_{\tau_t})dW_t + b(\tau_t, \bar{X}_{\tau_t})dt,\ \bar{X}_0 = y. \tag{3.9}$$

Lorsqu'il sera important de préciser le nombre N de pas de temps utilisés pour la discrétisation, nous utiliserons la notation $\bar{X}_t^N$.

3.2.1 Vitesse forte

Comme dans le schéma, les coefficients b et σ sont figés à leur valeur au début de chaque pas de temps, pour obtenir un résultat de convergence, il faut introduire une hypothèse de régularité en temps pour ces fonctions.

Théorème 3.2.1. *On suppose que les coefficient σ, b satisfont l'hypothèse (Lip) (voir théorème 3.1.2) et que*

$$\exists \alpha, K > 0,\ \forall x \in \mathbb{R}^n,\ \forall (s,t) \in [0,T],\ |\sigma(t,x)-\sigma(s,x)|+|b(t,x)-b(s,x)| \leq K(1+|x|)(t-s)^\alpha.$$

Alors pour $\beta = \min(\alpha, 1/2)$,

$$\forall p \geq 1,\ \exists C_p > 0, \forall y \in \mathbb{R}^n,\ \forall N \in \mathbb{N}^*,\ \mathbb{E}\left(\sup_{t\leq T} |X_t - \bar{X}_t^N|^{2p}\right) \leq \frac{C_p(1+|y|^{2p})}{N^{2\beta p}}.$$

En outre si $\gamma < \beta$, $N^\gamma \sup_{t\leq T} |X_t - \bar{X}_t^N|$ converge presque sûrement vers 0 lorsque N tend vers l'infini.

Remarque 3.2.2.
— *On a $\| \sup_{t\leq T} |X_t - \bar{X}_t^N| \|_{2p} = \left(\mathbb{E}\left(\sup_{t\leq T} |X_t - \bar{X}_t^N|^{2p}\right)\right)^{\frac{1}{2p}} \leq \frac{C}{N^\beta}$ et on dit que la vitesse forte du schéma d'Euler est en $\frac{1}{N^\beta}$. En particulier lorsque les coefficients de l'EDS ne dépendent pas du temps ou que $\alpha \geq \frac{1}{2}$, la vitesse forte est en $\frac{1}{\sqrt{N}}$.*
— *Sous des hypothèses de régularité sur $\sigma, b : \mathbb{R} \to \mathbb{R}$, Kurtz et Protter [46] ont montré que si $(X_t)_{t\in[0,T]}$ est solution de l'équation différentielle stochastique $dX_t = \sigma(X_t)dW_t + b(X_t)dt$ en dimension $n = d = 1$ avec coefficients homogènes en temps, alors le processus $(\sqrt{N}(X_t - \bar{X}_t^N))_{t\in[0,T]}$ d'erreur renormalisée du schéma d'Euler converge en loi vers $(Y_t)_{t\in[0,T]}$ solution de*

$$Y_t = \int_0^t Y_s(\sigma'(X_s)dW_s + b'(X_s)ds) + \sqrt{\frac{T}{2}} \int_0^t \sigma\sigma'(X_s)dB_s$$

où $(B_t)_{t\in[0,T]}$ est un mouvement brownien indépendant de $(W_t)_{t\in[0,T]}$ (résultat qui se généralise en dimension supérieure). Nous renvoyons au problème du paragraphe 3.8.5 pour une preuve de ce résultat dans le cas particulier du modèle de Black-Scholes. Notons qu'en dehors du cas où le coefficient de diffusion σ est constant, le processus $(Y_t)_{t\in[0,T]}$ est non nul, ce qui signifie que la vitesse de convergence obtenue dans le théorème 3.2.1 est la bonne.

— *D'après le problème 3.8.3, le résultat de convergence forte du théorème 3.2.1 reste vrai si le coefficient de dérive b est simplement mesurable en la variable temporelle (mais on garde l'hypothèse de régularité höldérienne en la variable temporelle pour le coefficient de diffusion σ) lorsque, dans la définition* (3.8) *du schéma d'Euler, on remplace $b(t_k, X^N_{t_k})$ par $b(\eta_k, X^N_{t_k})$ avec les variables aléatoires $(\eta_k)_{0\leq k\leq N-1}$ indépendantes de $(W_t)_{t\geq 0}$ et indépendantes et respectivement distribuées suivant les lois uniformes sur $[t_k, t_{k+1}]$.*

— *Lorsque pour $n = d = 1$ on souhaite calculer le prix $\mathbb{E}((\frac{1}{T}\int_0^T X_t dt - K)^+)$ d'un Call asiatique, on peut introduire $Y_t = \int_0^t X_s ds$ et vérifier qu'appliquer le schéma d'Euler au couple (X, Y) revient à appliquer le schéma d'Euler à X et à poser $\bar{Y}^N_{t_k} = \frac{T}{N}\sum_{j=0}^{k-1}\bar{X}^N_{t_j}$. Comme la fonction $z \to (z-K)^+$ est lipschitzienne de constante 1, on a alors*

$$\left|\mathbb{E}\left(\left(\frac{1}{T}\int_0^T X_t dt - K\right)^+\right) - \mathbb{E}\left(\left(\frac{1}{N}\sum_{j=0}^{N-1}\bar{X}^N_{t_j} - K\right)^+\right)\right| \leq \frac{1}{T}\mathbb{E}|Y_T - \bar{Y}^N_T| \leq \frac{C}{N^\beta}.$$

En fait, il est bien préférable numériquement d'approcher le prix de l'option asiatique par $\mathbb{E}\left(\left(\frac{1}{N}\left(\frac{\bar{X}^N_0+\bar{X}^N_T}{2} + \sum_{j=1}^{N-1}\bar{X}_{t_j}\right) - K\right)^+\right)$ ce qui revient à approcher l'intégrale $\int_0^T X_t dt$ par la méthode des trapèzes plutôt que par celle des rectangles (voir le paragraphe 3.7).

— *Dans le cas $\alpha \geq \frac{1}{2}$, on peut montrer que les vitesses fortes du schéma d'Euler constant par morceaux (pour $t \in [t_k, t_{k+1}[$, $\bar{X}^N_t = \bar{X}^N_{t_k}$) et du schéma interpolé linéairement (pour $t \in [t_k, t_{k+1}]$, $\bar{X}^N_t = [(t_{k+1}-t)\bar{X}^N_{t_k} + (t-t_k)\bar{X}^N_{t_{k+1}}]/(t_{k+1}-t_k)$) sont en $\sqrt{\frac{\log(N)}{N}}$.*

Démonstration : On a pour tout $u \in [0, T]$,

$$X_u - \bar{X}_u = \int_0^u \sigma(s, X_s) - \sigma(\tau_s, \bar{X}_{\tau_s})dW_s + \int_0^u b(s, X_s) - b(\tau_s, \bar{X}_{\tau_s})ds.$$

Donc pour $t \in [0, T]$, en utilisant (3.7) et l'inégalité de Burkolder Davis Gundy pour l'intégrale stochastique, on obtient comme dans la preuve de la proposition 3.1.6

$$\begin{aligned}\mathbb{E}\left(\sup_{u\le t}|X_u-\bar{X}_u|^{2p}\right) &\le 2^{2p-1}\mathbb{E}\left(\sup_{u\le t}\left|\int_0^u \sigma(s,X_s)-\sigma(\tau_s,\bar{X}_{\tau_s})dW_s\right|^{2p}\right)\\ &\quad+2^{2p-1}\mathbb{E}\left(\sup_{u\le t}\left|\int_0^u b(s,X_s)-b(\tau_s,\bar{X}_{\tau_s})ds\right|^{2p}\right)\\ &\le Ct^{p-1}\int_0^t \mathbb{E}\left(|\sigma(s,X_s)-\sigma(\tau_s,\bar{X}_{\tau_s})|^{2p}\right)ds\\ &\quad+2^{2p-1}t^{2p-1}\int_0^t \mathbb{E}\left(|b(s,X_s)-b(\tau_s,\bar{X}_{\tau_s})|^{2p}\right)ds. \end{aligned}\tag{3.10}$$

Avec (3.7) et les hypothèses de régularité faites sur les coefficients de l'EDS, on a

$$\begin{aligned}|\sigma(s,X_s)-\sigma(\tau_s,\bar{X}_{\tau_s})|^{2p} &\le 3^{2p-1}\Big(|\sigma(s,X_s)-\sigma(s,X_{\tau_s})|^{2p}+|\sigma(s,X_{\tau_s})-\sigma(\tau_s,X_{\tau_s})|^{2p}\\ &\qquad+|\sigma(\tau_s,X_{\tau_s})-\sigma(\tau_s,\bar{X}_{\tau_s})|^{2p}\Big)\\ &\le C\left(|X_s-X_{\tau_s}|^{2p}+(1+|X_{\tau_s}|^{2p})(s-\tau_s)^{2p\alpha}+|X_{\tau_s}-\bar{X}_{\tau_s}|^{2p}\right).\end{aligned}$$

En utilisant (3.6), (3.4) et $0\le s-\tau_s\le \frac{T}{N}$, on en déduit que

$$\mathbb{E}\left(|\sigma(s,X_s)-\sigma(\tau_s,\bar{X}_{\tau_s})|^{2p}\right)\le C\left(\frac{1+|y|^{2p}}{N^p}+\frac{1+|y|^{2p}}{N^{2\alpha p}}+\mathbb{E}\left(\sup_{u\le s}|X_u-\bar{X}_u|^{2p}\right)\right).$$

Cette inégalité reste vraie en remplaçant σ par b au membre de gauche et avec (3.10), on en déduit

$$\mathbb{E}\left(\sup_{u\le t}|X_u-\bar{X}_u|^{2p}\right)\le C\left(\frac{1+|y|^{2p}}{N^{2(\alpha\wedge\frac{1}{2})p}}+\int_0^t\mathbb{E}\left(\sup_{u\le s}|X_u-\bar{X}_u|^{2p}\right)ds\right).$$

La première assertion du théorème découle alors du lemme de Gronwall. Pour rendre l'argument qui précède rigoureux, il aurait fallu utiliser une technique de localisation, par exemple avec les temps d'arrêt $\nu_m=\inf\{t\ge 0 : |X_t-\bar{X}_t|\ge m\}$ qui sont tels que $\mathbb{E}\left(\sup_{u\le t}|X_{u\wedge\nu_m}-\bar{X}_{u\wedge\nu_m}|^{2p}\right)<+\infty$.

Pour $\gamma<\beta$,

$$\mathbb{E}\left(\left(N^\gamma\sup_{u\le T}|X_u-\bar{X}_u|\right)^{2p}\right)\le\frac{C}{N^{2(\beta-\gamma)p}}.$$

On choisit $p>\frac{1}{2(\beta-\gamma)}$, ce qui assure que

$$\mathbb{E}\left(\sum_{N\ge1}\left(N^\gamma\sup_{u\le T}|X_u-\bar{X}_u|\right)^{2p}\right)<+\infty \text{ puis } \mathbb{P}\left(\sum_{N\ge1}\left(N^\gamma\sup_{u\le T}|X_u-\bar{X}_u|\right)^{2p}<+\infty\right)=1.$$

Comme le terme général d'une série convergente tend vers 0, on conclut que $N^\gamma\sup_{u\le T}|X_u-\bar{X}_u|$ converge presque sûrement vers 0 lorsque N tend vers l'infini. $\square$

3.3 Vitesse faible :

Lorsque l'on souhaite calculer $\mathbb{E}(f(X_T))$ pour f une fonction lipschitzienne de constante de Lipschitz K, le théorème 3.2.1 assure que

$$\begin{aligned}|\mathbb{E}(f(X_T)) - \mathbb{E}(f(\bar{X}_T^N))| &\leq \mathbb{E}(|f(X_T) - f(\bar{X}_T^N)|) \leq \sqrt{\mathbb{E}(|f(X_T) - f(\bar{X}_T^N)|^2)} \\ &\leq K\sqrt{\mathbb{E}(|X_T - \bar{X}_T^N|^2)} \leq \frac{C}{N^\beta}. \end{aligned} \tag{3.11}$$

Mais la première inégalité qui consiste à majorer la valeur absolue de la différence des espérances par l'espérance de la valeur absolue de la différence est très grossière. En fait, la question de savoir si $\mathbb{E}(f(\bar{X}_T^N))$ est proche de $\mathbb{E}(f(X_T))$ revient à se demander si la loi de $\bar{X}_T^N$ est proche de celle de X_T. On va donc s'intéresser au problème de la convergence en loi du schéma d'Euler. La convergence en loi est une convergence contre des fonctions tests comme la fonction f introduite plus haut. C'est pourquoi on parle de convergence faible. Le résultat suivant a été démontré par Talay et Tubaro [70].

Théorème 3.3.1. *On suppose que b, σ sont des fonctions C^∞ sur $[0,T] \times \mathbb{R}^n$ avec des dérivées de tous ordres bornées et que $f : \mathbb{R}^n \to \mathbb{R}$ est C^∞ avec des dérivées à croissance polynomiale :*

$$\forall a = (a_1, \dots, a_n) \in \mathbb{N}^n,\ \exists p, C > 0,\ \forall x \in \mathbb{R}^n,\ \left|\frac{\partial^{a_1+\ldots+a_n} f}{\partial x_1^{a_1} \ldots \partial x_n^{a_n}}(x)\right| \leq C(1 + |x|^p).$$

Alors il existe une suite $(C_l)_{l\geq 1}$ de réels telle que pour tout $L \in \mathbb{N}^*$, on a un développement de l'erreur de la forme

$$\mathbb{E}(f(\bar{X}_T^N)) - \mathbb{E}(f(X_T)) = \frac{C_1}{N} + \frac{C_2}{N^2} + \ldots + \frac{C_L}{N^L} + \mathcal{O}\left(\frac{1}{N^{L+1}}\right).$$

Remarque 3.3.2.

— *En particulier cela implique que $|\mathbb{E}(f(\bar{X}_T^N)) - \mathbb{E}(f(X_T))| \leq \frac{C}{N}$ alors qu'utiliser la vitesse forte et le caractère lipschitzien de f conduit à majorer $|\mathbb{E}(f(\bar{X}_T^N)) - \mathbb{E}(f(X_T))|$ par $\frac{C}{\sqrt{N}}$ ($\alpha = 1$ dans le théorème 3.2.1 sous les hypothèses faites ici sur b et σ).*

— *On peut obtenir la majoration $|\mathbb{E}(f(\bar{X}_T^N)) - \mathbb{E}(f(X_T))| \leq \frac{C}{N}$ (sans développement de l'erreur) en supposant simplement σ, b C^4 avec des dérivées bornées et f C^4 avec des dérivées à croissance polynomiales.*

— **Extrapolation de Romberg (Richardson) :** *On a*

$$\begin{aligned}\mathbb{E}\left(2f(\bar{X}_T^{2N}) - f(\bar{X}_T^N)\right) - \mathbb{E}(f(X_T)) &\\ &= 2\left(\mathbb{E}(f(\bar{X}_T^{2N}) - \mathbb{E}(f(X_T))\right) - \left(\mathbb{E}(f(\bar{X}_T^N) - \mathbb{E}(f(X_T))\right) \\ &= 2\left(\frac{C_1}{2N} + \frac{C_2}{4N^2} + \mathcal{O}\left(\frac{1}{N^3}\right)\right) - \left(\frac{C_1}{N} + \frac{C_2}{N^2} + \mathcal{O}\left(\frac{1}{N^3}\right)\right) \\ &= -\frac{C_2}{2N^2} + \mathcal{O}\left(\frac{1}{N^3}\right).\end{aligned}$$

En calculant l'espérance de la combinaison linéaire $2f(\bar{X}_T^{2N}) - f(\bar{X}_T^N)$, on fait disparaître le terme principal de l'erreur et on obtient une vitesse faible en $\frac{1}{N^2}$. Notons que comme $f(\bar{X}_T^{2N})$ est proche de $f(\bar{X}_T^N)$ pour N grand, on peut penser que

Cov $(f(\bar{X}_T^{2N}), f(\bar{X}_T^N)) \geq 0$. *Il est préférable en termes de variance d'utiliser les mêmes accroissements browniens pour générer $\bar{X}_T^{2N}$ et $\bar{X}_T^N$ plutôt que d'utiliser des accroissements indépendants pour les deux schémas. En fait, on déduit facilement du théorème 3.2.1 que* Var $(2f(\bar{X}_T^{2N}) - f(\bar{X}_T^N)) \to$ Var $(f(X_T))$ *lorsque $N \to +\infty$.*
Cette technique d'extrapolation de Romberg se généralise de la façon suivante (voir [64]) : pour tout $l \geq 2$,

$$\mathbb{E}\left(\sum_{m=1}^{l} \frac{(-1)^{l-m} m^l}{m!(l-m)!} f(\bar{X}_T^{mN})\right) - \mathbb{E}(f(X_T)) = \frac{(-1)^{l-1} C_l}{l! N^l} + \mathcal{O}\left(\frac{1}{N^{l+1}}\right).$$

De même que précédemment, pour éviter que la variance n'explose, il faut utiliser les mêmes accroissements browniens pour générer $(\bar{X}_T^{mN})_{1\leq m\leq l}$. Il suffit d'itérer N fois la procédure utilisée à cet effet sur l'intervalle de temps $[0, T/N]$. Plutôt que de générer $(W_{kT/N\mathrm{ppcm}(1,\ldots,l)}, 0 \leq k \leq \mathrm{ppcm}(1,\ldots,l))$, il s'avère plus économique surtout lorsque l est grand de générer $(W_{kT/mN}, 1 \leq k \leq m, 1 \leq m \leq l, \mathrm{pgcd}(k,m) = 1)$. Par exemple, si $l = 4$, on se donne $G_1, \ldots, G_6$ six vecteurs gaussiens centrés réduits indépendants à valeurs dans $\mathbb{R}^d$ et on pose

$$\begin{aligned}
W_{\frac{T}{4N}} &= \sqrt{\frac{T}{4N}} G_1 \\
W_{\frac{T}{3N}} &= \sqrt{\frac{T}{4N}} G_1 + \sqrt{\frac{T}{12N}} G_2 \\
W_{\frac{T}{2N}} &= \sqrt{\frac{T}{4N}} G_1 + \sqrt{\frac{T}{12N}} G_2 + \sqrt{\frac{T}{6N}} G_3 \\
W_{\frac{2T}{3N}} &= \sqrt{\frac{T}{4N}} G_1 + \sqrt{\frac{T}{12N}} G_2 + \sqrt{\frac{T}{6N}} (G_3 + G_4) \\
W_{\frac{3T}{4N}} &= \sqrt{\frac{T}{4N}} G_1 + \sqrt{\frac{T}{12N}} (G_2 + G_5) + \sqrt{\frac{T}{6N}} (G_3 + G_4) \\
W_{\frac{T}{N}} &= \sqrt{\frac{T}{4N}} (G_1 + G_6) + \sqrt{\frac{T}{12N}} (G_2 + G_5) + \sqrt{\frac{T}{6N}} (G_3 + G_4).
\end{aligned}$$

La preuve du théorème utilise la régularité de la solution de l'Équation aux Dérivées Partielles

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) + \frac{1}{2}\sum_{i,j=1}^n a_{ij}(t,x) \frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^n b_i(t,x) \frac{\partial u}{\partial x_i}(t,x) = 0, \ (t,x) \in [0,T] \times \mathbb{R}^n \\ u(T,x) = f(x), \ x \in \mathbb{R}^n \end{cases} \tag{3.12}$$

où $a(t,x) = \sigma(t,x)\sigma^*(t,x)$ $(\forall i,j \in \{1,\ldots,n\},\ a_{ij}(t,x) = \sum_{l=1}^d \sigma_{il}(t,x)\sigma_{jl}(t,x))$. Cette régularité est assurée par le résultat suivant.

Proposition 3.3.3. *Sous les hypothèses du théorème,* (3.12) *admet une unique solution. En outre, cette solution est C^∞ avec des dérivées à croissance polynomiale. Enfin, pour $(X_t)_{t\in[0,T]}$ solution de* (3.3), *$(u(t,X_t))_{t\in[0,T]}$ est une $\mathcal{F}_t$-martingale de carré intégrable de valeur terminale $f(X_T)$ et*

$$\boxed{u(0,y) = \mathbb{E}(f(X_T)).} \tag{3.13}$$

Démonstration : Nous ne démonterrons pas le résultat d'existence, d'unicité et de régularité pour l'EDP (3.12). Par la formule d'Itô, pour $t \in [0,T]$,

$$u(t,X_t) = u(0,X_0) + \int_0^t \nabla_x u(s,X_s).\sigma(s,X_s)dW_s$$
$$+ \int_0^t \left(\frac{\partial u}{\partial t} + \frac{1}{2}\sum_{i,j=1}^n a_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} + b.\nabla_x u \right)(s,X_s)ds. \quad (3.14)$$

D'après (3.12), $u(T,X_T) = f(X_T)$ et l'intégrale en ds au second membre est nulle. Comme $X_0 = y$, $u(0,X_0) = u(0,y)$. Enfin la régularité de σ et u assure que la fonction $\sigma^*(s,x)\nabla_x u(s,x)$ est à croissance polynomiale en x. Avec le contrôle des moments de la solution de l'EDS donné par le lemme 3.1.4, on en déduit que $\mathbb{E}\left(\int_0^T |\sigma^*(s,X_s)\nabla_x u(s,X_s)|^2 ds\right) < +\infty$, ce qui entraîne que $\left(u(t,X_t) = u(0,y) + \int_0^t \nabla_x u(s,X_s).\sigma(s,X_s)dW_s\right)_{t\in[0,T]}$ est une martingale de carré intégrable. L'égalité des espérances des valeurs en $t=0$ et en $t=T$ de cette martingale entraîne (3.13). □

Nous pouvons maintenant démontrer le théorème 3.3.1.

Démonstration : Nous nous plaçons en dimension $n = d = 1$ pour simplifier et nous allons seulement faire apparaître les deux premiers termes du développement de l'erreur. D'après la condition terminale dans (3.12) et d'après (3.13),

$$\mathbb{E}(f(\bar{X}_T)) - \mathbb{E}(f(X_T)) = \mathbb{E}(u(T,\bar{X}_T)) - u(0,y) = \mathbb{E}\left(u(T,\bar{X}_T) - u(0,\bar{X}_0)\right) = \sum_{k=0}^{N-1}\mathcal{E}_k \quad (3.15)$$

avec $\mathcal{E}_k = \mathbb{E}\left(u(t_{k+1},\bar{X}_{t_{k+1}}) - u(t_k,\bar{X}_{t_k})\right)$.
Comme sur $[t_k,t_{k+1}]$, $d\bar{X}_t = \sigma(t_k,\bar{X}_{t_k})dW_t + b(t_k,\bar{X}_{t_k})dt$, la formule d'Itô assure que

$$u(t_{k+1},\bar{X}_{t_{k+1}}) - u(t_k,\bar{X}_{t_k}) = \int_{t_k}^{t_{k+1}} \sigma(t_k,\bar{X}_{t_k})\frac{\partial u}{\partial x}(t,\bar{X}_t)dW_t$$
$$+ \int_{t_k}^{t_{k+1}} \frac{\partial u}{\partial t}(t,\bar{X}_t) + \frac{1}{2}\sigma^2(t_k,\bar{X}_{t_k})\frac{\partial^2 u}{\partial x^2}(t,\bar{X}_t) + b(t_k,\bar{X}_{t_k})\frac{\partial u}{\partial x}(t,\bar{X}_t)dt$$

En combinant la croissance polynomiale en x des fonctions $\sigma(t_k,x)$ et $\frac{\partial u}{\partial x}(t,x)$ avec le contrôle des moments du schéma d'Euler déduit du lemme 3.1.4 et du théorème 3.2.1 sur la vitesse forte, on obtient que l'espérance de l'intégrale stochastique est nulle. On en déduit qu'en notant $\sigma_k = \sigma(t_k,\bar{X}_{t_k})$ et $b_k = b(t_k,\bar{X}_{t_k})$, on a

$$\mathcal{E}_k = \int_{t_k}^{t_{k+1}} \mathbb{E}\left(v_k(t,\bar{X}_t)\right)dt \text{ où } v_k(t,x) = \frac{\partial u}{\partial t}(t,x) + \frac{1}{2}\sigma_k^2\frac{\partial^2 u}{\partial x^2}(t,x) + b_k\frac{\partial u}{\partial x}(t,x).$$

Comme en écrivant la première ligne de (3.12) au point $(t,x) = (t_k,\bar{X}_{t_k})$, on a $v_k(t_k,\bar{X}_{t_k}) = 0$, la formule d'Itô assure comme précédemment que pour $t \in [t_k,t_{k+1}]$,

$$\mathbb{E}\left(v_k(t,\bar{X}_t)\right) = \int_{t_k}^t \mathbb{E}\left(w_k(s,\bar{X}_s)\right)ds, \text{ où}$$

$$w_k(t,x) = \left(\frac{\partial v_k}{\partial t} + \frac{1}{2}\sigma_k^2\frac{\partial^2 v_k}{\partial x^2} + b_k\frac{\partial v_k}{\partial x}\right)(t,x)$$
$$= \left(\frac{\partial^2 u}{\partial t^2} + \sigma_k^2\frac{\partial^3 u}{\partial t\partial x^2} + 2b_k\frac{\partial^2 u}{\partial t\partial x} + \frac{\sigma_k^4}{4}\frac{\partial^4 u}{\partial x^4} + \sigma_k^2 b_k\frac{\partial^3 u}{\partial x^3} + b_k^2\frac{\partial^2 u}{\partial x^2}\right)(t,x).$$

Notons que lorsque l'on dérive la fonction v_k par rapport à t ou x, on dérive les facteurs qui s'écrivent comme dérivées partielles de u mais pas les coefficients σ_k^2 et b_k.

Ainsi, en remarquant que pour toute fonction g intégrable, $\int_{t_k}^{t_{k+1}} \int_{t_k}^{t} g(s)dsdt = \int_{t_k}^{t_{k+1}} (t_{k+1} - s)g(s)ds$, il vient $\mathcal{E}_k = \int_{t_k}^{t_{k+1}} (t_{k+1} - s)\mathbb{E}\left(w_k(s, \bar{X}_s)\right) ds$. La régularité des fonctions u, σ et b et le contrôle des moments du schéma d'Euler déduit du lemme 3.1.4 et du théorème 3.2.1 assurent que l'espérance dans l'intégrale est bornée ce qui implique que

$$|\mathcal{E}_k| \leq C \int_{t_k}^{t_{k+1}} (t_{k+1} - s)ds = \frac{CT^2}{2N^2}$$

En reportant cette inégalité dans (3.15), on obtient que $|\mathbb{E}(f(\bar{X}_T)) - \mathbb{E}(f(X_T))| \leq \frac{CT^2}{2N}$.
Soit

$$w(t, x) = \left(\frac{\partial^2 u}{\partial t^2} + \sigma^2 \frac{\partial^3 u}{\partial t \partial x^2} + 2b \frac{\partial^2 u}{\partial t \partial x} + \frac{\sigma^4}{4} \frac{\partial^4 u}{\partial x^4} + \sigma^2 b \frac{\partial^3 u}{\partial x^3} + b^2 \frac{\partial^2 u}{\partial x^2}\right)(t, x).$$

La fonction $w_k(s, \bar{X}_s) - w(s, \bar{X}_s)$ est la somme des produits $\left(g(t_k, \bar{X}_{t_k}) - g(s, \bar{X}_s)\right) h(s, \bar{X}_s)$ pour $(g, h) \in \{(1, \frac{\partial^2 u}{\partial t^2}), (\sigma^2, \frac{\partial^3 u}{\partial t \partial x^2}), (2b, \frac{\partial^2 u}{\partial t \partial x}), (\frac{\sigma^4}{4}, \frac{\partial^4 u}{\partial x^4}), (\sigma^2 b, \frac{\partial^3 u}{\partial x^3}), (b^2, \frac{\partial^2 u}{\partial x^2})\}$. On a

$$\left(g(t_k, \bar{X}_{t_k}) - g(s, \bar{X}_s)\right) h(s, \bar{X}_s) = g(t_k, \bar{X}_{t_k}) \left(h(s, \bar{X}_s) - h(t_k, \bar{X}_{t_k})\right) - (gh(s, \bar{X}_s) - gh(t_k, \bar{X}_{t_k})),$$

où, par application de la formule d'Itô, l'espérance du second membre s'écrit

$$\int_{t_k}^{s} \mathbb{E}\left(g(t_k, \bar{X}_{t_k}) \left(\frac{\partial h}{\partial r} + \frac{\sigma_k^2}{2} \frac{\partial^2 h}{\partial x^2} + b_k \frac{\partial h}{\partial x}\right)(r, \bar{X}_r) - \left(\frac{\partial gh}{\partial r} + \frac{\sigma_k^2}{2} \frac{\partial^2 gh}{\partial x^2} + b_k \frac{\partial gh}{\partial x}\right)(r, \bar{X}_r)\right) dr.$$

Le terme dans l'espérance s'écrit comme somme des produits $\tilde{g}(t_k, \bar{X}_{t_k})\tilde{h}(r, \bar{X}_r)$ pour $(\tilde{g}, \tilde{h}) \in \{(g, \frac{\partial h}{\partial r}), (\frac{g\sigma^2}{2}, \frac{\partial^2 h}{\partial x^2}), (gb, \frac{\partial h}{\partial x}), (-1, \frac{\partial gh}{\partial r}), (-\frac{\sigma^2}{2}, \frac{\partial^2 gh}{\partial x^2}), (-b, \frac{\partial gh}{\partial x})\}$. Ainsi

$$\mathcal{E}_k = \int_{t_k}^{t_{k+1}} (t_{k+1} - s)\mathbb{E}(w(s, \bar{X}_s))ds + \int_{t_k}^{t_{k+1}} \frac{1}{2}(t_{k+1} - r)^2 \mathbb{E}\left(\sum_{i=1}^{I} \tilde{g}_i(t_k, \bar{X}_{t_k})\tilde{h}_i(r, \bar{X}_r)\right) dr$$

si bien qu'en notant $\bar{\tau}_t^N = \lceil \frac{Nt}{T} \rceil \frac{T}{N}$, l'instant de discrétisation juste après t,

$$\mathbb{E}(f(\bar{X}_T)) - \mathbb{E}(f(X_T)) = \int_0^T (\bar{\tau}_t^N - t)\mathbb{E}(w(t, \bar{X}_t))dt + \int_0^T \frac{1}{2}(\bar{\tau}_t^N - t)^2 \mathbb{E}\left(\sum_{i=1}^{I} \tilde{g}_i(\tau_t, \bar{X}_{\tau_t})\tilde{h}_i(t, \bar{X}_t)\right) dt.$$

En raison du facteur $(\bar{\tau}_t^N - t)^2$, le second terme du second membre se comporte en $\mathcal{O}(1/N^2)$. La convergence de $\mathbb{E}(w(t, \bar{X}_t))$ vers $\mathbb{E}(w(t, X_t))$ lorsque $N \to \infty$ se démontre comme celle de $\mathbb{E}(f(\bar{X}_t))$ vers $\mathbb{E}(f(X_t))$. Comme $\mathbb{E}(w(t, \bar{X}_t))$ est bornée uniformément en $N \in \mathbb{N}^*$ et $t \in [0, T]$, on en déduit que $[0, T] \ni t \mapsto \mathbb{E}(w(t, \bar{X}_t))$ converge dans $L^1([0, T])$ (pour la mesure de Lebesgue) vers $[0, T] \ni t \mapsto \mathbb{E}(w(t, X_t))$. D'autre part, $[0, T] \ni t \mapsto N(\bar{\tau}_t^N - t)$ converge faiblement vers sa moyenne $\frac{T}{2}$. On conclut que

$$\lim_{N \to \infty} N \left(\mathbb{E}(f(\bar{X}_T)) - \mathbb{E}(f(X_T))\right) = \frac{T}{2} \int_0^T \mathbb{E}(w(t, X_t))dt,$$

si bien que $C_1 = \frac{1}{2} \int_0^T \mathbb{E}(w(t, X_t))dt$. Notons $\tilde{C}_1(t)$ la constante qui apparaît de manière analogue dans le développement de $\mathbb{E}(w(t, \bar{X}_t)) - \mathbb{E}(w(t, X_t))$. On remarque maintenant que

$$\int_0^T (\bar{\tau}_t^N - t)\mathbb{E}(w(t, X_t))dt = \frac{T}{2N} \int_0^T \mathbb{E}(w(t, X_t))dt - \int_0^T \frac{1}{2}(\bar{\tau}_t^N - t)(t - \tau_t^N)\partial_t \mathbb{E}(w(t, X_t))dt$$

pour écrire

$$\begin{aligned}\mathbb{E}(f(\bar{X}_T)) - \mathbb{E}(f(X_T)) - \frac{C_1}{N} = &\int_0^T (\bar{\tau}_t^N - t)\left(\mathbb{E}(w(t,\bar{X}_t)) - \mathbb{E}(w(t,X_t))\right)dt \\ &- \int_0^T \frac{1}{2}(\bar{\tau}_t^N - t)(t - \tau_t^N)\partial_t \mathbb{E}(w(t,X_t))dt \\ &+ \int_0^T \frac{1}{2}(\bar{\tau}_t^N - t)^2 \mathbb{E}\left(\sum_{i=1}^I \tilde{g}_i(\tau_t, \bar{X}_{\tau_t})\tilde{h}_i(t,\bar{X}_t)\right)dt.\end{aligned}$$

On conclut que

$$C_2 = \frac{T}{2}\int_0^T \tilde{C}_1(t)dt - \frac{T^2}{12}\int_0^T \partial_t\mathbb{E}(w(t,X_t))dt + \frac{T^2}{6}\int_0^T \mathbb{E}\left(\sum_{i=1}^I \tilde{g}_i(t,X_t)\tilde{h}_i(t,X_t)\right)dt,$$

où le facteur $\int_0^T \partial_t\mathbb{E}(w(t,X_t))dt$ se récrit $\mathbb{E}(w(T,X_T)) - w(0,y)$.

□

En vue d'applications en finance, l'hypothèse de régularité faite sur f dans le théorème 3.3.1 n'est pas satisfaisante car les fonctions de payoff $x \to (x-K)^+$ et $x \to (K-x)^+$ du Call et du Put ne sont pas des fonctions C^1. Dans le cas de coefficients $\sigma(t,x) = \sigma(x)$ et $b(t,x) = b(x)$ homogènes en temps et C^∞ à dérivées bornées, Bally et Talay [6] [7] ont montré que l'on conserve un développement de l'erreur

$$\mathbb{E}(f(\bar{X}_T^N)) - \mathbb{E}(f(X_T)) = \frac{C_1}{N} + \mathcal{O}\left(\frac{1}{N^2}\right) \tag{3.16}$$

pour $f : \mathbb{R}^n \to \mathbb{R}$ mesurable bornée dès lors que σ et b satisfont une condition de non dégénérescence. L'outil utilisé, le calcul de Malliavin, dépasse le cadre de ce cours. La condition de non-dégénerescence assure que X_T admet une densité C^∞ par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$. Elle est impliquée par la condition d'uniforme ellipticité suivante :

$$\exists \varepsilon > 0,\ \forall x, \xi \in \mathbb{R}^n,\ \xi^*\sigma(x)\sigma^*(x)\xi \geq \varepsilon|\xi|^2.$$

Sous cette condition d'uniforme ellipticité, Guyon [34] a démontré que le développement (3.16) reste vrai lorsque f est une distribution tempérée, le terme $\mathbb{E}(f(\bar{X}_T^N))$ (resp. $\mathbb{E}(f(X_T))$) s'interprétant alors comme $< f, \rho_N >$ (resp. $< f, \rho >$) où ρ_N (resp. ρ) désigne la densité de $\bar{X}_T^N$ (resp. X_T). En particulier, f peut être choisie mesurable à croissance polynomiale, égale à une masse de Dirac ou à une dérivée de masse de Dirac.

Pour l'EDS $dX_t = dW_t + b(t,X_t)dt$ avec une fonction de dérive b mesurable bornée, la convergence faible à l'ordre 1/2 a été montrée dans [12] pour le schéma d'Euler avec randomisation de la variable temporelle :

$$X^N_{\frac{(k+1)T}{N}} = X^N_{\frac{kT}{N}} + W_{\frac{(k+1)T}{N}} - W_{\frac{kT}{N}} + b\left(\frac{(k+U_k)T}{N}, X^N_{\frac{kT}{N}}\right)\frac{T}{N},$$

où les variables aléatoires $(U_k)_{0\leq k\leq N-1}$ sont i.i.d. suivant la loi uniforme sur $[0,1]$ et indépendantes de $(W_t)_{t\geq 0}$. Plus précisément, $\sup_{N\geq 1}\sqrt{N}\sup_{\|f\|_\infty\leq 1}\left|\mathbb{E}(f(X_T)) - \mathbb{E}(f(X_T^N))\right| < \infty$ où le second supremum porte sur les fonctions $f : \mathbb{R}^n \to \mathbb{R}$ mesurables et bornées par 1.

Lorsque l'on s'intéresse à une option exotique dont le payoff dépend de la trajectoire du sous-jacent pas simplement au travers de sa valeur terminale X_T, l'analyse de l'erreur faible developpée dans ce paragraphe ne s'applique plus. Si $F : C([0,T],\mathbb{R}) \to \mathbb{R}$ est lispchitzienne lorsque $C([0,T],\mathbb{R})$ est muni de la norme sup, on peut généraliser (3.11) en

$$\left|\mathbb{E}[F((X_t)_{t\in[0,T]})] - \mathbb{E}[F((\bar{X}_t^N)_{t\in[0,T]})]\right| \leq \mathbb{E}\left|F((X_t)_{t\in[0,T]}) - F((\bar{X}_t^N)_{t\in[0,T]})\right| \leq \frac{C}{N^\beta}$$

mais il n'y aucune raison pour que cette inégalité donne la bonne vitesse de convergence. L'analyse d'erreur faible a été généralisée pour des options exotiques particulières dont le payoff se met sous la forme $F((X_t)_{t\in[0,T]}) = f(X_T, Y_T)$ avec Y_t une fonction de $(X_s)_{0\leq s\leq t}$ telle que le couple $((X_t, Y_t))_{0\leq t\leq T}$ reste Markovien, si bien que l'on dispose toujours d'une équation aux dérivées partielles de pricing. Les cas $Y_t = \int_0^t X_s ds$ et $Y_t = \max_{0\leq s\leq t} X_s$ correspondent respectivement aux options sur moyenne [72] et aux options barrières [31, 32, 33] ou lookback.
La théorie du transport optimal a permis d'obtenir un résultat qui s'applique à des payoffs lipschitziens (toujours avec $C([0,T],\mathbb{R})$ muni de la norme sup) quelconques [3]

$$\forall\varepsilon > 0,\ \exists C_\varepsilon < +\infty,\ \forall N \geq 1,\ \sup_{F:Lip(F)\leq 1} \left|\mathbb{E}[F((X_t)_{t\in[0,T]})] - \mathbb{E}[F((\bar{X}_t^N)_{t\in[0,T]})]\right| \leq \frac{C_\varepsilon}{N^{2/3-\varepsilon}}$$

dans le cas de l'équation différentielle stochastique $dX_t = \sigma(X_t)dW_t + b(X_t)dt$ en dimension $n = d = 1$ avec les fonctions σ et b respectivement C^4 et C^3 bornées ainsi que leurs dérivées et vérifiant $\inf_{x\in\mathbb{R}} \sigma^2(x) > 0$.

3.4 Le schéma de Milstein

Comme la possibilité de faire des extrapolations de Romberg permet de construire à partir du schéma d'Euler des schémas d'ordre de convergence (puissance de $1/N$ dans le terme principal de l'erreur) faible arbitrairement élevé, il semble plus intéressant d'essayer d'améliorer l'ordre de convergence forte du schéma d'Euler plutôt que l'ordre de convergence faible. Le schéma de Milstein est obtenu en rajoutant des termes au schéma d'Euler et sa vitesse de convergence forte est en $\frac{1}{N}$ au lieu de $\frac{1}{\sqrt{N}}$ pour le schéma d'Euler (cas $\alpha = \frac{1}{2}$ dans les hypothèses du théorème 3.2.1).

3.4.1 Le cas de la dimension 1

On suppose $n = d = 1$ et on se place dans le cas homogène en temps $\sigma(t,x) = \sigma(x)$ et $b(t,x) = b(x)$. On souhaite discrétiser l'EDS

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt,\ X_0 = y \in \mathbb{R}.$$

Sur l'intervalle de temps $[t_k, t_{k+1}]$,

$$X_t = X_{t_k} + \int_{t_k}^t \sigma(X_s)dW_s + \int_{t_k}^t b(X_s)ds.$$

Comme

$$\mathbb{E}\left[\left(\int_{t_k}^{t}\sigma(X_s)dW_s\right)^2\right] = \mathbb{E}\left(\int_{t_k}^{t}\sigma^2(X_s)ds\right) \sim C(t-t_k)$$
$$\mathbb{E}\left[\left(\int_{t_k}^{t}b(X_s)ds\right)^2\right] \leq (t-t_k)\mathbb{E}\left(\int_{t_k}^{t}b^2(X_s)ds\right) \sim C(t-t_k)^2$$

on voit que le terme d'intégrale stochastique est dominant lorsque le pas de temps est petit. Donc pour améliorer la convergence du schéma, il faut améliorer la discrétisation de ce terme d'intégrale stochastique. Pour cela on remarque que

$$\sigma(X_s) \simeq \sigma(X_{t_k}) + \sigma'(X_{t_k})(X_s - X_{t_k}) \simeq \sigma(X_{t_k}) + \sigma'(X_{t_k})\sigma(X_{t_k})(W_s - W_{t_k}).$$

On en déduit que

$$\begin{aligned}\int_{t_k}^{t}\sigma(X_s)dW_s &\simeq \sigma(X_{t_k})\int_{t_k}^{t}dW_s + \sigma'(X_{t_k})\sigma(X_{t_k})\int_{t_k}^{t}(W_s - W_{t_k})dW_s\\ &= \sigma(X_{t_k})(W_t - W_{t_k}) + \frac{1}{2}\sigma\sigma'(X_{t_k})\left((W_t - W_{t_k})^2 - (t-t_k)\right).\end{aligned}$$

Le schéma de Milstein repose sur ce développement :

$$\begin{cases}\tilde{X}_0 = y \text{ et pour } k \in \{0,\dots,N-1\},\ \forall t \in [t_k, t_{k+1}],\\ \tilde{X}_t = \tilde{X}_{t_k} + \sigma(\tilde{X}_{t_k})(W_t - W_{t_k}) + \frac{1}{2}\sigma\sigma'(\tilde{X}_{t_k})\left((W_t - W_{t_k})^2 - (t-t_k)\right) + b(\tilde{X}_{t_k})(t-t_k)\end{cases}. \tag{3.17}$$

Pour générer le schéma aux instants de discrétisation $(\tilde{X}_{t_k})_{0\leq k\leq N}$, il suffit à nouveau de générer les accroissements $(W_{t_{k+1}} - W_{t_k})_{0\leq k\leq N-1}$ qui sont i.i.d. suivant la loi gaussienne $\mathcal{N}_1(0, \frac{T}{N})$. La vitesse forte du schéma est donnée par le résultat dont la démonstration fait l'objet du problème du paragraphe 3.8.6.

Théorème 3.4.1. *On suppose que les coefficients σ et b sont C^2 à dérivées bornées. Alors*

$$\forall p \geq 1,\ \exists C_p > 0,\ \forall N \in \mathbb{N}^*,\ \mathbb{E}\left(\sup_{t\leq T}|X_t - \tilde{X}_t^N|^{2p}\right) \leq \frac{C_p}{N^{2p}}.$$

En outre $\forall \gamma < 1$, $N^\gamma \sup_{t\leq T}|X_t - \tilde{X}_t^N|$ converge presque sûrement vers 0 lorsque N tend vers l'infini.

Remarque 3.4.2. — *Lorsque la fonction σ est constante, le schéma de Milstein coïncide avec le schéma d'Euler et la vitesse forte de ce dernier est donc en $\frac{C}{N}$.*

— *Pour éviter de calculer analytiquement la dérivée de σ, Newton [58] a proposé le schéma suivant qui a également une vitesse forte en $\frac{C}{N}$:*

$$\begin{cases}\hat{X}_0 = y \text{ et pour } k \in \{0,\dots,N-1\},\\ \hat{X}_{t_{k+1}} = \hat{X}_{t_k} + b(\hat{X}_{t_k})(t_{k+1}-t_k) - \sigma(\hat{X}_{t_k})\sqrt{t_{k+1}-t_k} + \sigma_k(W_{t_{k+1}} - W_{t_k} + \sqrt{t_{k+1}-t_k})\\ \text{où } \sigma_k = \sigma\left(\hat{X}_{t_k} + \frac{1}{2}\sigma(\hat{X}_{t_k})(W_{t_{k+1}} - W_{t_k} - \sqrt{t_{k+1}-t_k})\right).\end{cases}$$

Comme $\sigma_k \simeq \sigma(\hat{X}_{t_k}) + \frac{1}{2}\sigma\sigma'(\hat{X}_{t_k})(W_{t_{k+1}} - W_{t_k} - \sqrt{t_{k+1}-t_k})$, ce schéma est très proche du schéma de Milstein.

3.4.2 Le cas général

On effectue la même correction qu'en dimension 1 mais les notations sont plus lourdes. Pour $s \in [t_k, t_{k+1}]$ et $(i,j) \in \{1,\dots,n\} \times \{1,\dots,d\}$, on a

$$\begin{aligned}\sigma_{ij}(X_s) &\simeq \sigma_{ij}(X_{t_k}) + \sum_{l=1}^{n} \frac{\partial \sigma_{ij}}{\partial x_l}(X_{t_k})(X_s^l - X_{t_k}^l) \\ &\simeq \sigma_{ij}(X_{t_k}) + \sum_{l=1}^{n} \frac{\partial \sigma_{ij}}{\partial x_l}(X_{t_k}) \sum_{m=1}^{d} \sigma_{lm}(X_{t_k})(W_s^m - W_{t_k}^m).\end{aligned}$$

On introduit les notations

$$\sigma_j = \begin{pmatrix} \sigma_{1j} \\ \vdots \\ \sigma_{nj} \end{pmatrix} \text{ et } \partial\sigma_j = \begin{pmatrix} \frac{\partial \sigma_{1j}}{\partial x_1} & \dots & \frac{\partial \sigma_{1j}}{\partial x_n} \\ \vdots & \dots & \vdots \\ \frac{\partial \sigma_{nj}}{\partial x_1} & \dots & \frac{\partial \sigma_{nj}}{\partial x_n} \end{pmatrix} \tag{3.18}$$

pour la j-ième colonne de σ et pour la matrice $\left(\frac{\partial \sigma_{ij}}{\partial x_l}\right)_{1\leq i,l \leq n}$. Cela permet de récrire le développement précédent sous forme vectorielle :

$$\sigma_j(X_s) \simeq \sigma_j(X_{t_k}) + \sum_{m=1}^{d} \partial\sigma_j \sigma_m(X_{t_k})(W_s^m - W_{t_k}^m).$$

Dans $\sigma(X_s)dW_s = \sum_{j=1}^{d} \sigma_j(X_s)dW_s^j$, la j-ème colonne de σ vient multiplier dW_s^j. Donc

$$\sigma(X_s)dW_s \simeq \sigma(X_{t_k})dW_s + \sum_{j=1}^{d}\left(\sum_{m=1}^{d} \partial\sigma_j\sigma_m(X_{t_k})(W_s^m - W_{t_k}^m)\right) dW_s^j.$$

Le schéma de Milstein s'écrit donc :

$$\begin{cases} X_0 = y \text{ et pour } k \in \{0,\dots,N-1\},\ \forall t \in [t_k, t_{k+1}], \\ \tilde{X}_t = \tilde{X}_{t_k} + \sigma(\tilde{X}_{t_k})(W_t - W_{t_k}) + \sum_{j,m=1}^{d} \partial\sigma_j\sigma_m(\tilde{X}_{t_k}) \int_{t_k}^{t} (W_s^m - W_{t_k}^m)dW_s^j \\ \qquad + b(\tilde{X}_{t_k})(t - t_k). \end{cases} \tag{3.19}$$

Pour implémenter en pratique le schéma, on rencontre le problème suivant : pour $j \neq m$, on ne sait pas simuler $\int_{t_k}^{t_{k+1}} (W_s^m - W_{t_k}^m)dW_s^j$. Le seul cas où on sait mettre en œuvre le schéma est celui où la condition de **commutativité** suivante est satisfaite :

$$\forall j,m \in \{1,\dots,d\},\ \partial\sigma_j\sigma_m = \partial\sigma_m\sigma_j$$

$$\text{i.e. } \forall i \in \{1,\dots,n\},\ \forall x \in \mathbb{R}^n,\ \sum_{l=1}^{n} \frac{\partial \sigma_{ij}}{\partial x_l}(x)\sigma_{lm}(x) = \sum_{l=1}^{n} \frac{\partial \sigma_{im}}{\partial x_l}(x)\sigma_{lj}(x).$$

En effet comme pour $m \neq j$, $\int_{t_k}^t (W_s^m - W_{t_k}^m) dW_s^j + \int_{t_k}^t (W_s^j - W_{t_k}^j) dW_s^m = (W_t^m - W_{t_k}^m)(W_t^j - W_{t_k}^j)$, sous la condition de commutativité,

$$\begin{aligned}\sum_{j,m=1}^d \partial\sigma_j\sigma_m(\tilde{X}_{t_k}) \int_{t_k}^t (W_s^m - W_{t_k}^m) dW_s^j &= \sum_{j=1}^d \partial\sigma_j\sigma_j(X_{t_k}) \frac{1}{2} \left((W_t^j - W_{t_k}^j)^2 - (t - t_k)\right) \\ &\quad + \sum_{j=1}^d \sum_{m=1}^{j-1} \partial\sigma_j\sigma_m(\tilde{X}_{t_k})(W_t^m - W_{t_k}^m)(W_t^j - W_{t_k}^j) \\ &= \frac{1}{2} \sum_{j,m=1}^d \partial\sigma_j\sigma_m(\tilde{X}_{t_k})(W_t^m - W_{t_k}^m)(W_t^j - W_{t_k}^j) \\ &\quad - \frac{1}{2} \sum_{j=1}^d \partial\sigma_j\sigma_j(X_{t_k})(t - t_k).\end{aligned}$$

Et le schéma ne fait plus intervenir que les accroissements $(W_t - W_{t_k})$.
Que la condition de commutativité soit satisfaite ou non, le résultat de convergence énoncé en dimension 1 dans le théorème 3.4.1 reste valable.

Remarque 3.4.3. — *Une condition suffisante de commutativité est que la dimension d du mouvement brownien soit égale à 1.*

— *Comme la condition de commutativité fait intervenir des dérivées des coefficients de la matrice de diffusion σ, elle est également vérifiée lorsque σ est constante. Dans ce cas, les schémas d'Euler et de Milstein coincident et la vitesse forte du schéma d'Euler est en $\frac{C}{N}$.*

— *Lorsque la dimension d du mouvement Brownien est égale à 2, on dispose d'une expression pour la fonction caractéristique du triplet $(W_t^1, W_t^2, \int_0^t W_s^1 dW_s^2 - \int_0^t W_s^2 dW_s^1)$. Gaines et Lyons [27] ont tiré parti de cette expression pour construire une technique de simulation du triplet. Comme*

$$\int_0^t W_s^1 dW_s^2 = \frac{1}{2}\left(W_t^1 W_t^2 + \int_0^t W_s^1 dW_s^2 - \int_0^t W_s^2 dW_s^1\right),$$

générer le triplet permet ensuite d'implémenter le schéma de Milstein.

— *Parmi les développements récents concernant la discrétisation des EDS, on peut également citer la technique de simulation exacte en dimension $n = d = 1$ proposée récemment par Beskos Papaspiliopoulos et Roberts [13] et qui fait l'objet du problème du paragraphe 3.8.11.*

— *La méthode de Monte Carlo multipas proposée par Giles [28] et étudiée dans le problème du paragraphe 3.8.7 permet de tirer partie des propriétés de convergence forte du schéma pour réduire la variance tout en contrôlant le biais grâce à l'analyse de l'erreur faible. Pour tirer pleinement partie de cette méthode, Giles et Szpruch [29] ont récemment proposé un schéma tel que la différence entre le schéma sur une grille grossière et la demi-somme, sur la grille de pas moitié, du schéma et du schéma dit antithétique où les accroissements browniens d'indices impairs et pairs sont échangés converge à l'ordre fort 1. Ce schéma est également étudié dans le paragraphe 3.8.7.*

3.5 Schémas d'ordre faible élevé

On a vu que pour obtenir un schéma d'ordre fort égal à 1, il faut faire intervenir des intégrales itérées d'ordre 2 du mouvement brownien $\int_{t_k}^t (W_s^m - W_{t_k}^m) dW_s^j$ que l'on ne sait pas simuler pour $m \neq j$. À l'aide de développements de Taylor stochastiques, on peut construire des schémas d'ordre fort plus élevé (voir [42]) qui font intervenir des intégrales itérées d'ordre supérieur que l'on sait encore moins simuler. En remplaçant ces intégrales itérées par des variables aléatoires définies sur un espace fini et qui ont mêmes moments jusqu'à un certain ordre, Kuzuoka [47] [48] a récemment introduit une nouvelle classe de schémas d'ordre faible élevé. Cette classe de schémas et son application en finance font l'objet d'une recherche active [51] [59] [60]. Ninomiya et Victoir [61] et Ninomiya et Ninomiya [62] ont récemment proposé une autre approche très intéressante : l'évolution de leurs schémas de discrétisation sur un pas de temps s'obtient en intégrant des équations différentielles ordinaires construites à partir des coefficients de l'EDS sur des horizons aléatoires bien choisis (voir le problème du paragraphe 3.8.14). Notons également les travaux d'Alfonsi [2] et de Tanaka et Kohatsu-Higa [71] sur la composition de schémas avec des applications respectives au modèle à volatilité stochastique d'Heston et à des modèles à sauts.

3.5.1 Développements de Taylor stochastiques

Pour effectuer ces développements, il est beaucoup plus pratique de récrire l'EDS $X_t = y + \int_0^t \sigma(X_s) dW_s + \int_0^t b(X_s) ds$ en remplaçant l'intégrale stochastique d'Itô par une intégrale de Stratonovich. L'intérêt est que les règles du calcul différentiel ordinaire s'appliquent lorsque l'on considère des intégrales de Stratonovich : il n'y a plus de terme faisant intervenir des dérivées secondes à ajouter. Nous rappelons que pour un processus unidimensionnel adapté régulier $(H_s)_{s\leq T}$ et pour $t \in [0,T]$, l'intégrale de Stratonovich $\int_0^t H_s \circ dW_s^j$ est égale à la limite en probabilité de $\sum_{k=0}^{N-1} \frac{1}{2}(H_{t_{k+1}\wedge t} + H_{t_k \wedge t})(W^j_{t_{k+1}\wedge t} - W^j_{t_k\wedge t})$ lorsque $N \to +\infty$. Ainsi

$$\int_0^t H_s \circ dW_s^j = \int_0^t H_s dW_s^j + \frac{1}{2} < H, W^j >_t .$$

L'EDS se récrit donc

$$X_t = y + \int_0^t \sigma_0(X_s) ds + \sum_{j=1}^d \int_0^t \sigma_j(X_s) \circ dW_s^j \tag{3.20}$$

où $\sigma_0 = b - \frac{1}{2} \sum_{j=1}^d \partial\sigma_j \sigma_j$ avec la matrice $\partial\sigma_j$ et le vecteur σ_j définis dans (3.18). Introduisons les opérateurs différentiels $V_j = \sum_{i=1}^n \sigma_{ij} \partial_{x_i}$ pour $0 \leq j \leq d$: pour une fonction $g : \mathbb{R}^n \to \mathbb{R}$ régulière, on note $V_j g$ la fonction définie sur $\mathbb{R}^n$ par $V_j g(x) = \sum_{i=1}^n \sigma_{ij}(x) \partial_{x_i} g(x)$. Comme les règles du calcul différentiel usuel s'appliquent pour les integrales de Stratonovich, pour f une fonction régulière sur $\mathbb{R}^n$, on a

$$f(X_t) = f(y) + \int_0^t V_0 f(X_s) ds + \sum_{j=1}^d \int_0^t V_j f(X_s) \circ dW_s^j = f(y) + \sum_{j=0}^d \int_0^t V_j f(X_s) \circ dW_s^j,$$

où on pose $W_s^0 = s$.
En remarquant que $V_j f(X_s) = V_j f(y) + \sum_{l=0}^d \int_0^s V_l V_j f(X_u) \circ dW_u^l$, on obtient

$$f(X_t) = f(y) + \sum_{j=0}^d V_j f(y) \int_0^t \circ dW_s^j + \sum_{j,l=0}^d \int_{0\le u\le s\le t} V_l V_j f(X_u) \circ dW_u^l \circ dW_s^j.$$

Par récurrence, on conclut que pour $m \in \mathbb{N}^*$,

$$\begin{aligned} f(X_t) =& f(y) + \sum_{k=1}^m \sum_{j_1,\dots,j_k=0}^d V_{j_1} V_{j_2} \dots V_{j_k} f(y) W_t^{(j_1,\dots,j_k)} \\ &+ \sum_{j_1,\dots,j_{m+1}=0}^d \int_{0\le s_1\le\dots\le s_{m+1}\le t} V_{j_1} \dots V_{j_{m+1}} f(X_{s_1}) \circ dW_{s_1}^{j_1} \circ \dots \circ dW_{s_{m+1}}^{j_{m+1}} \end{aligned}$$

où $W_t^{(j_1,\dots,j_k)} = \int_{0\le s_1\le\dots\le s_k\le t} \circ dW_{s_1}^{j_1} \circ \dots \circ dW_{s_k}^{j_k}$.
Pour $j \in \{1,\dots,d\}$, W_s^j est de l'ordre de $\sqrt{s}$ tandis que $W_s^0 = s$. En d'autres termes, pour $t > 0$, $(W_s^0, W_s^1, \dots, W_s^d)_{s\in[0,t]}$ a même loi que $(tW_{\frac{s}{t}}^0, \sqrt{t}W_{s/t}^1, \dots, W_{s/t}^d)_{s\in[0,t]}$ ce qui entraîne que

$$\begin{aligned} W_t^{(j_1,\dots,j_k)} &\overset{\mathcal{L}}{=} t^{(k+\mathrm{Card}\{1\le l\le k: j_l=0\})/2} \int_{0\le s_1\le\dots\le s_k\le t} \circ dW_{s_1/t}^{j_1} \circ \dots \circ dW_{s_k/t}^{j_k} \\ &= t^{(k+\mathrm{Card}\{1\le l\le k: j_l=0\})/2} \int_{0\le u_1\le\dots\le u_k\le 1} \circ dW_{u_1}^{j_1} \circ \dots \circ dW_{u_k}^{j_k} \\ &= t^{(k+Card\{1\le l\le k: j_l=0\})/2} W_1^{(j_1,\dots,j_k)}. \end{aligned}$$

Ainsi pour obtenir des termes du même ordre dans le développement de Taylor précédent, il faut compter les intégrales par rapport à W^0 deux fois.
C'est pourquoi pour $\alpha = (j_1,\dots,j_k) \in \mathcal{A} = \bigcup_{l\in\mathbb{N}^*} \{0,\dots,d\}^l$, on pose $|\alpha| = k$ et $\|\alpha\| = k + \mathrm{Card}\{1 \le l \le k : j_l = 0\}$. On écrit enfin

$$\begin{aligned} f(X_t) =& f(y) + \sum_{\alpha:1\le\|\alpha\|\le m} V_{j_1} \dots V_{j_k} f(y) W_t^\alpha + R_{m,f}^{W,y}(t) \text{ où} \qquad (3.21) \\ R_{m,f}^{W,y}(t) =& \sum_{\alpha:|\alpha|\le m, \|\alpha\|>m} V_{j_1} \dots V_{j_k} f(y) W_t^\alpha \\ &+ \sum_{j_1,\dots,j_{m+1}=0}^d \int_{0\le s_1\le\dots\le s_{m+1}\le t} V_{j_1} \dots V_{j_{m+1}} f(X_{s_1}) \circ dW_{s_1}^{j_1} \circ \dots \circ dW_{s_{m+1}}^{j_{m+1}}. \end{aligned}$$

Comme le reste $R_{m,f}^{W,y}(t)$ ne comporte que des termes qui se comportent comme t à une puissance supérieure ou égale à $(m+1)/2$, le résultat suivant dont la preuve se trouve dans [42] n'est pas surprenant.

Proposition 3.5.1. *Lorsque les fonctions f, b et σ sont suffisamment régulières, le reste $R_{m,f}^{W,y}(t)$ est tel que pour $p \ge 1$, $\mathbb{E}(|R_{m,f}^{W,y}(t)|^{2p})^{\frac{1}{2p}} \le Ct^{\frac{m+1}{2}}$ où la constante C dépend de m, p, f, b et σ mais pas de t.*

3.5.2 Schémas d'ordre faible élevé

La simulation des intégrales browniennes itérées W_t^α pose problème. Pour lever cette difficulté et obtenir des schémas implémentables, Kusuoka [47] [48] propose de remplacer les intégrales itérées qui apparaissent dans le développement de Taylor stochastique (3.21) par des variables aléatoires qui ont les mêmes moments jusqu'à l'ordre m.

Définition 3.5.2. *Soit $m \in \mathbb{N}^*$. On dit qu'une famille $(\zeta^\alpha)_{\|\alpha\| \leq m}$ de variables aléatoires avec des moments de tous ordres finis préserve les moments jusqu'à l'ordre m si*

$$\forall k \in \mathbb{N}^*,\ \forall \alpha_1, \ldots, \alpha_k \in \mathcal{A} \ t.q.\ \|\alpha_1\| + \ldots + \|\alpha_k\| \leq m,\ \mathbb{E}\left(\prod_{l=1}^{k} \zeta^{\alpha_l}\right) = \mathbb{E}\left(\prod_{l=1}^{k} W_1^{\alpha_l}\right). \tag{3.22}$$

L'exemple suivant est tiré de [60] où d'autres familles préservant les moments sont également données.

Exemple 3.5.3. *Soit η t.q. $\mathbb{P}(\eta = 0) = \frac{2}{3}$ et $\mathbb{P}(\eta = \pm\sqrt{3}) = \frac{1}{6}$.*
On obtient une famille qui préserve les moments jusqu'à l'ordre 5 en dimension $d = 1$ en posant

$$\zeta^0 = 1,\ \zeta^1 = \eta,\ \zeta^{(1,1)} = \frac{1}{2}\eta^2,\ \zeta^{(1,0)} = \zeta^{(0,1)} = \frac{1}{2}\eta,\ \zeta^{(1,1,1)} = \frac{1}{6}\eta^3,$$
$$\zeta^{(1,1,0)} = \zeta^{(0,1,1)} = \frac{1}{4},\ \zeta^{(0,0)} = \frac{1}{2},\ \zeta^{(1,1,1,1)} = \frac{1}{8} \text{ et } \zeta^\alpha = 0 \text{ sinon.}$$

Exercice 3.5.4. *Montrer que la famille précédente préserve bien les moments jusqu'à l'ordre 5.*

En écrivant le développement (3.21) pour $f(x) = I(x)$ où $I(x) = x$ est la fonction identité sur $\mathbb{R}^n$ et en remplaçant les intégrales browniennes itérées par les variables correspondantes de la famille qui préserve les moments jusqu'à l'ordre m on approche la loi de X_t par celle de

$$Y_t^y = y + \sum_{1 \leq \|\alpha\| \leq m} t^{\|\alpha\|/2} V_{j_1} \ldots V_{j_k} I(y) \zeta^\alpha.$$

Pour $g : \mathbb{R}^n \to \mathbb{R}$, on note $Q_t g$ la fonction définie sur $\mathbb{R}^n$ par $Q_t g(x) = \mathbb{E}(g(Y_t^x))$ où Y_t^x s'obtient en remplaçant y par x dans le second membre de l'équation précédente.
On peut approcher $\mathbb{E}(f(X_t))$ par $\mathbb{E}(f(Y_t^y))$. Le résultat suivant assure qu'en itérant N fois cette approximation, on obtient un schéma dont l'ordre de convergence faible est $(m-1)/2$.

Théorème 3.5.5. *Lorsque les fonctions f, b et σ sont suffisamment régulières,*

$$\left| \mathbb{E}(f(X_T)) - Q_{\frac{T}{N}}^N f(y) \right| \leq \frac{C}{N^{(m-1)/2}}$$

où la constante C dépend de m, f, b et σ mais pas de N.

Remarque 3.5.6. — *En posant $\zeta^j = W_1^j$ pour $0 \leq j \leq d$, $\zeta^{(j,j)} = \frac{1}{2}$ pour $1 \leq j \leq d$, $\zeta^{(j,l)} = 0$ pour $0 \leq j \neq l \leq d$ et $\zeta^{(j,k,l)} = 0$ pour $1 \leq j, k, l \leq d$, on obtient une famille qui préserve les moments jusqu'à l'ordre $m = 3$. Par la propriété de Markov, cette famille est telle que $Q_{\frac{T}{N}}^N f(y) = \mathbb{E}(f(\bar{X}_T^N))$. On retrouve ainsi que l'ordre de convergence faible du schéma d'Euler est $1 = (3-1)/2$.*

— *Bien entendu, pour approcher numériquement $\mathbb{E}(f(X_T))$, il est préférable d'utiliser une famille qui préserve les moments jusqu'à l'ordre m définie sur un espace de probabilité fini de cardinal aussi petit que possible. La famille qui préserve les moments jusqu'à l'ordre 5 de l'exemple 3.5.3 peut être générée sur un espace de probabilité de cardinal 3. Pour ce choix, le calcul exact de $Q^N_{\frac{T}{N}} f(y)$ est possible en utilisant un arbre non-recombinant avec 3^N feuilles. Le cardinal de l'espace de probabilité nécessaire pour générer une famille qui préserve les moments jusqu'à l'ordre m augmente bien sûr avec m et avec la dimension d du mouvement brownien. Si le calcul exact de $Q^N_{\frac{T}{N}} f(y)$ n'est plus possible, on doit recourir à une technique d'échantillonnage partiel comme la méthode de Monte-Carlo (voir [60]).*

Démonstration. Pour $g : \mathbb{R}^n \to \mathbb{R}$, on note $P_t g$ la fonction définie sur $\mathbb{R}^n$ par $P_t g(x) = \mathbb{E}(g(X_t^x))$ où X^x est la solution de l'EDS

$$dX_t^x = \sigma(X_t^x)dW_t + b(X_t^x)dt, \ X_0^x = x.$$

On suppose désormais que g est une fonction régulière. En remplaçant y par x dans (3.21), on obtient que la différence $P_t g(x) - Q_t g(x)$ vaut

$$\mathbb{E}\left[g\left(x + \sum_{1\le\|\alpha\|\le m} V_{j_1}\dots V_{j_k} I(x) W_t^\alpha + R_{m,I}^{W,x}(t)\right) - g\left(x + \sum_{1\le\|\alpha\|\le m} t^{\|\alpha\|/2} V_{j_1}\dots V_{j_k} I(x)\zeta^\alpha\right)\right].$$

En supposant pour simplifier que $n = 1$ et en effectuant un développement de Taylor standard à l'ordre m, on obtient

$$\begin{aligned} P_t g(x) - Q_t g(x) = \sum_{l=1}^m \frac{g^{(l)}(x)}{l!} \mathbb{E}\Bigg[\Bigg(&\sum_{1\le\|\alpha\|\le m} V_{j_1}\dots V_{j_k} I(x) W_t^\alpha + R_{m,I}^W(t)\Bigg)^l \\ &- \Bigg(\sum_{1\le\|\alpha\|\le m} t^{\|\alpha\|/2} V_{j_1}\dots V_{j_k} I(x)\zeta^\alpha\Bigg)^l\Bigg] + \mathcal{O}(t^{\frac{m+1}{2}}) \end{aligned}$$

Le fait que le reste est un $\mathcal{O}(t^{\frac{m+1}{2}})$ se déduit de $\mathbb{E}(|X_t^x - x|^{m+1}) \le Ct^{\frac{m+1}{2}}$ (voir Proposition 3.1.6) et de

$$\mathbb{E}\left(\left|\sum_{1\le\|\alpha\|\le m} t^{\frac{\|\alpha\|}{2}} V_{j_1}\dots V_{j_k} I(x)\zeta^\alpha\right|^{m+1}\right) = t^{\frac{m+1}{2}} \mathbb{E}\left(\left|\sum_{1\le\|\alpha\|\le m} t^{\frac{\|\alpha\|-1}{2}} V_{j_1}\dots V_{j_k} I(x)\zeta^\alpha\right|^{m+1}\right).$$

D'après la proposition 3.5.1, $R_{m,I}^{W,x}(t)$ se comporte en $\mathcal{O}(t^{\frac{m+1}{2}})$. En développant les puissances l-ièmes à l'intérieur de l'espérance et en utilisant (3.22), on obtient pour $j \le m$, que tous les termes en $\mathcal{O}(t^{j/2})$ disparaissent. On conclut que lorsque g est suffisamment régulière,

$$\exists C > 0, \forall t > 0, \ \|P_t g - Q_t g\|_\infty \le Ct^{(m+1)/2}.$$

Lorsque f est régulière, les fonctions $P_{t_{N-k}} f$, $k \in \{1, \dots, N\}$ le sont aussi. On en déduit

$$\exists C > 0, \forall N \in \mathbb{N}^*, \forall k \in \{1, \dots, N\}, \ \|P_{\frac{T}{N}} P_{t_{N-k}} f - Q_{\frac{T}{N}} P_{t_{N-k}} f\|_\infty \le C\left(\frac{T}{N}\right)^{(m+1)/2}.$$

Comme pour $h : \mathbb{R}^n \to \mathbb{R}$ bornée, pour tout $t \geq 0$, $\|Q_t h\|_\infty \leq \|h\|_\infty$, cela implique

$$\exists C > 0, \forall N \in \mathbb{N}^*, \forall k \in \{1, \dots, N\},\ \|Q^{k-1}_{\frac{T}{N}}(P_{\frac{T}{N}} P_{t_{N-k}} f - Q_{\frac{T}{N}} P_{t_{N-k}} f)\|_\infty \leq C \left(\frac{T}{N}\right)^{(m+1)/2}.$$

On conclut en reportant cette estimation dans la décomposition suivante de l'erreur

$$\mathbb{E}(f(X_T)) - Q^N_{\frac{T}{N}} f(y) = P_T f(y) - Q^N_{\frac{T}{N}} f(y) = \sum_{k=1}^{N} Q^{k-1}_{\frac{T}{N}} \left(P_{\frac{T}{N}} P_{t_{N-k}} f - Q_{\frac{T}{N}} P_{t_{N-k}} f\right)(y).$$

□

Pour terminer, nous allons présenter rapidement l'approximation de $\mathbb{E}(f(X_T))$ qui repose sur la notion de cubature proposée par Lyons et Victoir [51] et que l'on peut relier aux familles qui préservent les moments jusqu'à l'ordre m.

Définition 3.5.7. *Soit $m \in \mathbb{N}^*$ et $t > 0$. Des trajectoires continues $\omega_{t,1}, \dots, \omega_{t,L}$ à variation bornée de $[0,t]$ dans $\mathbb{R}^d$ et des poids positifs $\lambda_1, \dots, \lambda_L$ de somme 1 fournissent une formule de cubature de degré m au temps t si*

$$\forall \alpha = (j_1, \dots, j_k) \in \mathcal{A} \text{ t.q. } \|\alpha\| \leq m,\ \mathbb{E}(W^\alpha_t) = \sum_{l=1}^{L} \lambda_l \omega^\alpha_{t,l} \tag{3.23}$$

où $\omega^\alpha_{t,l} = \int_{0 \leq s_1 \leq \dots \leq s_k \leq t} d\omega^{j_1}_{t,l}(s_1) \dots d\omega^{j_k}_{t,l}(s_k)$ avec $\omega^j_{t,l}(s)$ désignant la j-ième coordonnée de $\omega_{t,l}(s)$ lorsque $1 \leq j \leq d$ et $\omega^0_{t,l}(s) = s$.

D'après [51], il existe une formule de cubature de degré m au temps 1 t.q. L est plus petit que $\mathrm{Card}\{\alpha \in \mathcal{A} : \|\alpha\| \leq m\}$. En outre, on en déduit facilement une formule de cubature de degré m au temps t en posant $\omega_{t,l}(s) = (t\omega^0_{1,l}(s/t), \sqrt{t}\omega^1_{1,l}(s/t), \dots, \sqrt{t}\omega^d_{1,l}(s/t))$.

Pour $l \in \{1, \dots, L\}$ soit $(\xi_{t,l}(s,y))_{s \leq t}$ la solution de l'équation différentielle ordinaire :

$$\xi_{t,l}(0,y) = y \text{ et } \forall s \in [0,t],\ d\xi_{t,l}(s,y) = \sum_{j=0}^{d} \sigma_j(\xi_{t,l}(s,y)) d\omega^j_{t,l}(s).$$

Lyons et Victoir proposent d'approcher $\mathbb{E}(g(X^x_t))$ par $Q_t g(x) = \sum_{l=1}^{L} \lambda_l g(\xi_{t,l}(t,x))$. Le théorème 3.5.5 reste vrai pour cette nouvelle définition de $Q_t g$. La preuve repose sur une décomposition analogue de l'erreur mais le contrôle de $P_t g(x) - Q_t g(x)$ est plus facile. En effet, le développement de Taylor (3.21) reste valable lorsque l'on remplace (X, W) par $(\xi_{t,l}, w_{t,l})$:

$$g(\xi_{t,l}(s,x)) = g(x) + \sum_{\alpha : \|\alpha\| \leq m} V_{j_1} \dots V_{j_k} g(x) \omega^\alpha_{t,l} + R^{\omega_{t,l},x}_{m,g}(t).$$

En multipliant cette égalité par λ_l, en sommant sur l, et en soustrayant l'espérance du développement (3.21) pour la fonction g, on obtient

$$Q_t g(x) - P_t g(x) = \sum_{\alpha : \|\alpha\| \leq m} V_{j_1} \dots V_{j_k} g(x) \left(\sum_{l=1}^{L} \lambda_l \omega^\alpha_{t,l} - \mathbb{E}(W^\alpha_t)\right) + \sum_{l=1}^{L} \lambda_l R^{\omega_{t,l},x}_{m,g}(t) - \mathbb{E}(R^{W,x}_{m,g}(t)).$$

La première somme est nulle d'après (3.23). Les termes de restes sont en $\mathcal{O}(t^{(m+1)/2})$ d'après la proposition 3.5.1 et le scaling qui permet d'obtenir la formule de cubature au temps t à partir de la formule de cubature au temps 1.

3.6 Pont brownien pour les options barrière ou lookback

En dimension $n = d = 1$, on condidère l'EDS

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt,\ X_0 = x_0.$$

On suppose que les coefficients σ et b satisfont (Lip) et que σ ne s'annule pas. Alors σ est de signe constant et quitte à remplacer W_t par $-W_t$ on suppose que $\boxed{\forall x \in \mathbb{R},\ \sigma(x) > 0}$. On souhaite calculer $\mathbb{E}\left(\varphi(X_T, M_T)\right)$ où pour $t \geq 0$, M_t désigne soit $\max_{s\in[0,t]} X_s$ soit $\min_{s\in[0,t]} X_s$. Pour fixer les idées, nous supposerons désormais que $M_t = \max_{s\in[0,t]} X_s$ mais le cas du minimum se traite par symétrie.

Exemple. *Lorsque $(X_t)_t$ modélise le cours d'un actif financier,*

— *le choix $\varphi(x, y) = y - x$ correspond à une option lookback,*

— *le choix $\varphi(x, y) = f(x)1_{\{y<L\}}$ correspond à une option barrière up and out.*

Une approche possible consiste à discrétiser l'EDS suivant le schéma d'Euler

$$\begin{cases} \bar{X}_0 = y \text{ et pour } k \in \{0, \ldots, N-1\},\ \forall t \in [t_k, t_{k+1}], \\ \bar{X}_t = \bar{X}_{t_k} + \sigma(\bar{X}_{t_k})(W_t - W_{t_k}) + b(\bar{X}_{t_k})(t - t_k) \end{cases}$$

et à poser $\bar{M}_T = \max_{0\leq k\leq N} \bar{X}_{t_k}$. Alors on peut approcher $\mathbb{E}\left(\varphi(X_T, M_T)\right)$ par $\mathbb{E}\left(\varphi(\bar{X}_T, \bar{M}_T)\right)$ (la vitesse de convergence forte de $\bar{M}_T$ vers M_T fait l'objet du problème du paragraphe 3.8.2). Malheureusement le maximum discrétisé converge lentement vers sa limite, ce qui conduit à un biais important. Pour remédier à cette difficulté, nous allons voir qu'il est possible de simuler le couple $(\bar{X}_T, \max_{t\in[0,T]} \bar{X}_t)$ où le maximum est maintenant pris sur l'ensemble de la trajectoire du schéma d'Euler en temps continu et non plus seulement sur les valeurs de ce schéma aux instants de discrétisation. Cela améliore grandement la performance.

Le résultat suivant montre que l'on peut en fait se ramener à la simulation du couple $(W_{t_1}, \max_{t\in[0,t_1]} W_t)$. Puis nous verrons comment simuler ce couple.

Proposition 3.6.1. *Conditionnellement à $(\bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T) = (x_0, x_1, \ldots, x_N)$, les variables aléatoires $\max_{t\in[t_k,t_{k+1}]} \bar{X}_t,\ k \in \{0, \ldots, N-1\}$ sont indépendantes et ont respectivement même loi que $x_k + \sigma(x_k)\max_{t\in[0,t_1]} W_t$ conditionnellement à $W_{t_1} = \frac{x_{k+1}-x_k}{\sigma(x_k)}$.*

Remarque 3.6.2. *On constate que le coefficient de dérive b n'intervient pas dans la loi conditionnelle de $\max_{t\in[t_k,t_{k+1}]} \bar{X}_t$ sachant $(\bar{X}_{t_k}, \bar{X}_{t_{k+1}}) = (x_k, x_{k+1})$.*

Démonstration : L'indépendance conditionnelle découle de l'indépendance des accroissements browniens.

Intéressons nous à la loi conditionnelle de $\max_{t\in[t_k,t_{k+1}]} \bar{X}_t$ sachant $(\bar{X}_{t_k}, \bar{X}_{t_{k+1}}) = (x_k, x_{k+1})$. On a

$$\max_{t\in[t_k,t_{k+1}]} \bar{X}_t = \bar{X}_{t_k} + \sigma(\bar{X}_{t_k}) \max_{t\in[t_k,t_{k+1}]} \left(W_t - W_{t_k} + \frac{b}{\sigma}(\bar{X}_{t_k})(t - t_k) \right)$$

$$W_{t_{k+1}} - W_{t_k} + \frac{b}{\sigma}(\bar{X}_{t_k})(t_{k+1} - t_k) = \frac{\bar{X}_{t_{k+1}} - \bar{X}_{t_k}}{\sigma(\bar{X}_{t_k})}.$$

Il suffit donc de montrer que pour $\alpha = \frac{b}{\sigma}(x_k)$, la loi conditionnelle de $\max_{t\in[t_k,t_{k+1}]}(W_t - W_{t_k} + \alpha(t-t_k))$ sachant $W_{t_{k+1}} - W_{t_k} + \alpha(t_{k+1}-t_k) = \frac{x_{k+1}-x_k}{\sigma(x_k)}$ est égale à la loi conditionnelle de $\max_{t\in[0,t_1]} W_t$ sachant $W_{t_1} = \frac{x_{k+1}-x_k}{\sigma(x_k)}$.
La stationnarité des accroissements browniens implique que pour $g : \mathbb{R}^2 \to \mathbb{R}$ bornée,

$$\begin{aligned}&\mathbb{E}\left[g\left(W_{t_{k+1}} - W_{t_k} + \alpha(t_{k+1}-t_k), \max_{t\in[t_k,t_{k+1}]}(W_t - W_{t_k} + \alpha(t-t_k))\right)\right]\\ &= \mathbb{E}\left[g\left(W_{t_1} + \alpha t_1, \max_{t\in[0,t_1]}(W_t + \alpha t)\right)\right].\end{aligned}$$

D'après le théorème de Girsanov, sous $\mathbb{Q}$ de densité $\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{-\alpha W_{t_1} - \frac{\alpha^2}{2}t_1}$ par rapport à $\mathbb{P}$, $(B_t = W_t + \alpha t)_{t\le t_1}$ est un mouvement brownien, ce qui implique que

$$\begin{aligned}\mathbb{E}\left[g\left(W_{t_1} + \alpha t_1, \max_{t\in[0,t_1]}(W_t + \alpha t)\right)\right] &= \mathbb{E}^{\mathbb{Q}}\left[g\left(W_{t_1} + \alpha t_1, \max_{t\in[0,t_1]}(W_t + \alpha t)\right)\frac{d\mathbb{P}}{d\mathbb{Q}}\right]\\ &= \mathbb{E}^{\mathbb{Q}}\left[g\left(B_{t_1}, \max_{t\in[0,t_1]} B_t\right)e^{\alpha W_{t_1} + \frac{\alpha^2}{2}t_1}\right]\\ &= \mathbb{E}^{\mathbb{Q}}\left[g\left(B_{t_1}, \max_{t\in[0,t_1]} B_t\right)e^{\alpha B_{t_1} - \frac{\alpha^2}{2}t_1}\right]\\ &= \mathbb{E}\left[g\left(W_{t_1}, \max_{t\in[0,t_1]} W_t\right)e^{\alpha W_{t_1} - \frac{\alpha^2}{2}t_1}\right].\end{aligned}$$

Donc si on note $p(x,y)$ la densité du couple $(W_{t_1}, \max_{t\in[0,t_1]} W_t)$ dont l'existence est assurée par le lemme 3.6.3 ci-dessous, on a pour toute fonction $g : \mathbb{R}^2 \to \mathbb{R}$ bornée,

$$\begin{aligned}&\mathbb{E}\left[g\left(W_{t_{k+1}} - W_{t_k} + \alpha(t_{k+1}-t_k), \max_{t\in[t_k,t_{k+1}]}(W_t - W_{t_k} + \alpha(t-t_k))\right)\right]\\ &= \mathbb{E}\left[g\left(W_{t_1}, \max_{t\in[0,t_1]} W_t\right)e^{\alpha W_{t_1} - \frac{\alpha^2}{2}t_1}\right] = \int_{\mathbb{R}^2} g(x,y)p(x,y)e^{\alpha x - \frac{\alpha^2}{2}t_1}dxdy.\end{aligned}$$

On en déduit que le couple $\left(W_{t_{k+1}} - W_{t_k} + \alpha(t_{k+1}-t_k), \max_{t\in[t_k,t_{k+1}]}(W_t - W_{t_k} + \alpha(t-t_k))\right)$ admet pour densité $p(x,y)e^{\alpha x - \frac{\alpha^2}{2}t_1}$. Donc la densité conditionnelle de $\max_{t\in[t_k,t_{k+1}]}(W_t - W_{t_k} + \alpha(t-t_k))$ sachant $W_{t_{k+1}} - W_{t_k} + \alpha(t_{k+1}-t_k) = x$ est égale à

$$\frac{p(x,y)e^{\alpha x - \frac{\alpha^2}{2}t_1}}{\int_{\mathbb{R}} p(x,z)e^{\alpha x - \frac{\alpha^2}{2}t_1}dz} = \frac{p(x,y)}{\int_{\mathbb{R}} p(x,z)dz}$$

c'est-à-dire à la densité conditionnelle de $\max_{t\in[0,t_1]} W_t$ sachant $W_{t_1} = x$. □

Lemme 3.6.3. *Soit $(W_t)_t$ un mouvement brownien réel. Le couple $(W_{t_1}, \max_{t\in[0,t_1]} W_t)$ possède la densité $p(z,w) = 1_{\{w>z^+\}}\frac{2(2w-z)}{t_1\sqrt{2\pi t_1}}e^{-\frac{(2w-z)^2}{2t_1}}$. En outre,*

$$\mathbb{P}\left(\max_{t\in[0,t_1]} W_t \ge y \middle| W_{t_1} = x\right) = \begin{cases} 1 \text{ si } y \le \max(0,x) \\ e^{-\frac{2y(y-x)}{t_1}} \text{ sinon}\end{cases}. \tag{3.24}$$

En particulier, conditionnellement à $W_{t_1} = x$, $\max_{t\in[0,t_1]} W_t$ a même loi que $\frac{1}{2}\left(x + \sqrt{x^2 - 2t_1\ln(U)}\right)$ où U suit la loi uniforme sur $[0,1]$.

Démonstration : L'égalité (3.24) est claire si $y \leq \max(0, x)$. Lorsque $y > \max(0, x)$, on va utiliser le principe de réflexion pour le mouvement brownien. D'après la propriété de Markov forte, si $\tau_y = \inf\{t \geq 0 : W_t = y\}$, le processus $(B_t)_{t \geq 0}$ défini par

$$B_t = \begin{cases} W_t \text{ si } t \leq \tau_y \\ W_{\tau_y} - (W_t - W_{\tau_y}) = 2y - W_t \text{ si } t \geq \tau_y \end{cases},$$

est un mouvement brownien. Donc lorsque $y \geq \max(x, 0)$,

$$\begin{aligned} \mathbb{P}\left(\max_{t \in [0,t_1]} W_t \geq y, W_{t_1} \leq x\right) &= \mathbb{P}\left(\max_{t \in [0,t_1]} B_t \geq y, B_{t_1} \leq x\right) = \mathbb{P}\left(\tau_y \leq t_1, 2y - W_{t_1} \leq x\right) \\ &= \mathbb{P}\left(\max_{t \in [0,t_1]} W_t \geq y, W_{t_1} \geq 2y - x\right) = \mathbb{P}\left(W_{t_1} \geq 2y - x\right) \end{aligned}$$

où on a utilisé la définition de $(B_t)_{t \geq 0}$ pour la deuxième égalité et l'inégalité $2y - x \geq y$ pour la dernière. Si $x > y \geq 0$, alors en utilisant le résultat précédent, on obtient

$$\begin{aligned} \mathbb{P}\left(\max_{t \in [0,t_1]} W_t \geq y, W_{t_1} \leq x\right) &= \mathbb{P}\left(\max_{t \in [0,t_1]} W_t \geq y, W_{t_1} \leq y\right) + \mathbb{P}\left(y < W_{t_1} \leq x\right) \\ &= \mathbb{P}\left(W_{t_1} \geq y\right) + \mathbb{P}\left(y < W_{t_1} \leq x\right). \end{aligned}$$

Enfin si $y < 0$, $\mathbb{P}\left(\max_{t \in [0,t_1]} W_t \geq y, W_{t_1} \leq x\right) = \mathbb{P}\left(W_{t_1} \leq x\right)$. Comme W_{t_1} possède la densité $e^{-\frac{z^2}{2t_1}}/\sqrt{2\pi t_1}$, on vérifie alors facilement que

$$\forall (x,y) \in \mathbb{R}^2,\ \mathbb{P}\left(\max_{t \in [0,t_1]} W_t \geq y, W_{t_1} \leq x\right) = \int_{z=-\infty}^{x} \int_{w=y}^{+\infty} p(z,w)\,dw\,dz \qquad (3.25)$$

pour $p(z,w) = 1_{\{w > z^+\}} \frac{2(2w-z)}{t_1\sqrt{2\pi t_1}} e^{-\frac{(2w-z)^2}{2t_1}}$. En effet

$$\int_y^{+\infty} p(z,w)\,dw = \frac{1}{\sqrt{2\pi t_1}} e^{-\frac{(2(y \vee z^+) - z)^2}{2t_1}} = \begin{cases} \frac{1}{\sqrt{2\pi t_1}} e^{-\frac{(2y-z)^2}{2t_1}} \text{ si } y \geq z^+ \\ \frac{1}{\sqrt{2\pi t_1}} e^{-\frac{z^2}{2t_1}} \text{ sinon} \end{cases}.$$

Ainsi lorsque $y < 0$, $\int_{z=-\infty}^{x} \int_{w=y}^{+\infty} p(z,w)\,dw\,dz = \mathbb{P}(W_{t_1} \leq x)$. Lorsque $y \geq x^+$,

$$\int_{z=-\infty}^{x} \int_{w=y}^{+\infty} p(z,w)\,dw\,dz = \int_{-\infty}^{x} e^{-\frac{(2y-z)^2}{2t_1}} \frac{dz}{\sqrt{2\pi t_1}} = \int_{2y-x}^{+\infty} e^{-\frac{z^2}{2t_1}} \frac{dz}{\sqrt{2\pi t_1}} = \mathbb{P}(W_{t_1} \geq 2y - x).$$

Enfin, lorsque $0 \leq y < x$,

$$\begin{aligned} \int_{z=-\infty}^{x} \int_{w=y}^{+\infty} p(z,w)\,dw\,dz &= \int_{z=-\infty}^{y} \int_{w=y}^{+\infty} p(z,w)\,dw\,dz + \int_y^x \frac{1}{\sqrt{2\pi t_1}} e^{-\frac{z^2}{2t_1}} dz \\ &= \mathbb{P}(W_{t_1} \geq 2y - y) + \mathbb{P}(y < W_{t_1} \leq x). \end{aligned}$$

On déduit de (3.25) que le couple $(W_{t_1}, \max_{t \in [0,t_1]} W_t)$ possède la densité $p(z,w)$.

Donc pour $y > \max(x, 0)$,

$$\mathbb{P}\left(\max_{t\in[0,t_1]} W_t \geq y \middle| W_{t_1} = x\right) = \frac{\int_y^{+\infty} p(x,w)dw}{e^{-\frac{x^2}{2t_1}}/\sqrt{2\pi t_1}} = \frac{\partial_x \mathbb{P}\left(\max_{t\in[0,t_1]} W_t \geq y, W_{t_1} \leq x\right)}{e^{-\frac{x^2}{2t_1}}/\sqrt{2\pi t_1}}$$
$$= \frac{\partial_x \mathbb{P}\left(W_{t_1} \geq 2y - x\right)}{e^{-\frac{x^2}{2t_1}}/\sqrt{2\pi t_1}} = e^{-\frac{(2y-x)^2 - x^2}{2t_1}} = e^{-\frac{2y(y-x)}{t_1}}.$$

La dernière assertion du lemme repose sur la méthode d'inversion de la fonction de répartition. Soit $u \in]0,1[$. On cherche $y > \max(x, 0)$ tel que

$$\mathbb{P}\left(\max_{t\in[0,t_1]} W_t \leq y \middle| W_{t_1} = x\right) = u.$$

En utilisant (3.24), cette égalité se récrit

$$1 - e^{-\frac{2y(y-x)}{t_1}} = u \Leftrightarrow \ln(1-u) = -\frac{2y(y-x)}{t_1} \Leftrightarrow y^2 - xy + \frac{t_1}{2}\ln(1-u) = 0.$$

Le discriminant de l'équation du second degré est $\Delta = x^2 - 2t_1 \ln(1-u)$ et la seule racine supérieure à $\max(x, 0)$ est $y = \frac{1}{2}\left(x + \sqrt{x^2 - 2t_1\ln(1-u)}\right)$.
D'après la méthode d'inversion de la fonction de répartition, si U suit la loi uniforme sur $[0,1]$, $\frac{1}{2}\left(x + \sqrt{x^2 - 2t_1\ln(1-U)}\right)$ suit la loi conditionnelle de $\max_{t\in[0,t_1]} W_t$ sachant $W_{t_1} = x$. On conclut en remarquant que $1 - U$ a même loi que U. $\square$

Exercice 3.6.4. *Vérifier que conditionnellement à $W_{t_1} = x$, $\min_{t\in[0,t_1]} W_t$ a même loi que $\frac{1}{2}\left(x - \sqrt{x^2 - 2t_1\ln(U)}\right)$ où U suit la loi uniforme sur $[0,1]$ (On pourra appliquer le lemme 3.6.3 au mouvement brownien $(-W_t)_{t\geq 0}$).*

3.6.1 Application aux options barrière

On va approcher $\mathbb{E}\left(f(X_T)1_{\{\max_{t\in[0,T]} X_t < L\}}\right)$ par $\mathbb{E}\left(f(\bar{X}_T)1_{\{\max_{t\in[0,T]} \bar{X}_t < L\}}\right)$. Nous allons voir que l'on peut calculer l'espérance conditionnelle de la variable aléatoire dans la seconde espérance sachant les valeurs $(\bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T)$ du schéma d'Euler aux instants de discrétisation. Cela permet d'évaluer cette espérance par la méthode de Monte Carlo en simulant le schéma d'Euler. Comme $1_{\{\max_{t\in[0,T]} \bar{X}_t < L\}} = \prod_{k=0}^{N-1} 1_{\{\max_{t\in[t_k,t_{k+1}]} \bar{X}_t < L\}}$, d'après la proposition 3.6.1, on a

$$\mathbb{E}\left(f(\bar{X}_T)1_{\{\max_{t\in[0,T]} \bar{X}_t < L\}} \middle| \bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T\right) = f(\bar{X}_T)\prod_{k=0}^{N-1}\left(1 - \psi(\bar{X}_{t_k}, \bar{X}_{t_{k+1}})\right)$$

où

$$\psi(x_k, x_{k+1}) = \mathbb{P}\left(\max_{t\in[t_k,t_{k+1}]} \bar{X}_t \geq L \middle| (\bar{X}_{t_k}, \bar{X}_{t_{k+1}}) = (x_k, x_{k+1})\right)$$
$$= \mathbb{P}\left(\max_{t\in[0,t_1]} W_t \geq \frac{L - x_k}{\sigma(x_k)} \middle| W_{t_1} = \frac{x_{k+1} - x_k}{\sigma(x_k)}\right).$$

D'après (3.24),

$$\psi(x_k, x_{k+1}) = \begin{cases} 1 \text{ si } \frac{L-x_k}{\sigma(x_k)} \leq \max\left(\frac{x_{k+1}-x_k}{\sigma(x_k)}, 0\right) \text{ i.e. si } L \leq \max(x_k, x_{k+1}) \\ e^{-\frac{2(L-x_k)(L-x_{k+1})}{\sigma^2(x_k)t_1}} \quad \text{sinon .} \end{cases}$$

On conclut finalement que

$$\begin{aligned} &\mathbb{E}\left(f(\bar{X}_T)1_{\{\max_{t\in[0,T]}\bar{X}_t<L\}}\middle|\bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T\right) \\ &= f(\bar{X}_T)1_{\{\max_{0\leq k\leq N}\bar{X}_{t_k}<L\}}\prod_{k=0}^{N-1}\left[1-\exp\left(-\frac{2(L-\bar{X}_{t_k})(L-\bar{X}_{t_{k+1}})}{t_1\sigma^2(\bar{X}_{t_k})}\right)\right]. \end{aligned} \quad (3.26)$$

Pour approcher $\mathbb{E}\left(f(X_T)1_{\{\max_{t\in[0,T]}X_t<L\}}\right)$, on évalue l'espérance du terme de droite par la méthode de Monte Carlo. Il suffit pour cela de générer le schéma d'Euler aux instants de discrétisation. Par rapport à l'approche naïve évoquée au début du paragraphe qui consiste à calculer $\mathbb{E}\left(f(\bar{X}_T)1_{\{\max_{0\leq k\leq N}\bar{X}_{t_k}<L\}}\right)$ par la méthode de Monte Carlo, nous voyons que nous avons introduit un facteur multiplicatif plus petit que 1 dans l'espérance pour tenir compte du fait que le schéma d'Euler en temps continu peut dépasser la barrière alors qu'il est en dessous de cette barrière à tous les instants de discrétisation. Cela améliore nettement le biais.

Exercice 3.6.5. *Déterminer* $\mathbb{E}\left(1_{\{\min_{t\in[0,T]}\bar{X}_t>L\}}\middle|\bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T\right)$.

3.6.2 Application aux options lookback

Si on se donne $(U_k)_{0\leq k\leq N-1}$ des variables indépendantes uniformes sur $[0,1]$ et indépendantes du brownien W, alors la proposition 3.6.1 et le lemme 3.6.3, assurent que conditionnellement à $(\bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T) = (x_0, \ldots, x_N)$, le vecteur $(\max_{t\in[t_k,t_{k+1}]}\bar{X}_t)_{0\leq k\leq N-1}$ a même loi que le vecteur

$$\begin{aligned} &\left(x_k + \frac{\sigma(x_k)}{2}\left(\frac{x_{k+1}-x_k}{\sigma(x_k)} + \sqrt{\frac{(x_{k+1}-x_k)^2}{\sigma^2(x_k)} - 2t_1\ln(U_k)}\right)\right)_{0\leq k\leq N-1} \\ &= \left(\frac{1}{2}\left(x_k + x_{k+1} + \sqrt{(x_{k+1}-x_k)^2 - 2\sigma^2(x_k)t_1\ln(U_k)}\right)\right)_{0\leq k\leq N-1}. \end{aligned}$$

En particulier, cela implique que le vecteur $(\bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T, \max_{t\in[0,T]}\bar{X}_t)$ a même loi que le vecteur

$$\left(\bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T, \frac{1}{2}\max_{0\leq k\leq N-1}\left(\bar{X}_{t_k} + \bar{X}_{t_{k+1}} + \sqrt{(\bar{X}_{t_{k+1}}-\bar{X}_{t_k})^2 - 2\sigma^2(\bar{X}_{t_k})t_1\ln(U_k)}\right)\right).$$

Ainsi on sait simuler suivant la loi de $(\bar{X}_T, \max_{t\in[0,T]}\bar{X}_t)$. Cela permet d'évaluer $\mathbb{E}\left(\varphi(\bar{X}_T, \max_{t\in[0,T]}\bar{X}_t)\right)$ par la méthode de Monte-Carlo pour approcher $\mathbb{E}\left(\varphi(X_T, \max_{t\in[0,T]}X_t)\right)$

Remarque 3.6.6. *Dans le cas des options barrières, on pourrait utiliser cette approche avec* $\varphi(x, y) = f(x)1_{\{y<L\}}$. *Mais l'égalité en loi qui précède et* (3.26) *assurent que*

$$\mathbb{E}\left(f(\bar{X}_T)1_{\{\frac{1}{2}\max_{0\leq k\leq N-1}\left(\bar{X}_{t_k}+\bar{X}_{t_{k+1}}+\sqrt{(\bar{X}_{t_{k+1}}-\bar{X}_{t_k})^2-2\sigma^2(\bar{X}_{t_k})t_1\ln(U_k)}\right)<L\}}\middle| \bar{X}_0, \bar{X}_{t_1}, \ldots, \bar{X}_T\right)$$
$$= f(\bar{X}_T)1_{\{\max_{0\leq k\leq N}\bar{X}_{t_k}<L\}}\prod_{k=0}^{N-1}\left[1-\exp\left(-\frac{2(L-\bar{X}_{t_k})(L-\bar{X}_{t_{k+1}})}{t_1\sigma^2(\bar{X}_{t_k})}\right)\right].$$

Donc l'approche proposée au paragraphe précédent qui consiste à évaluer l'espérance de la variable aléatoire au second membre par la méthode de Monte Carlo est préférable puisqu'elle est plus précise (réduction de variance par conditionnement) et nécessite moins de calculs (pas besoin de générer les U_k*).*

3.6.3 Le cas des options barrière sur plusieurs sous-jacents

On suppose que les n sous-jacents évoluent suivant l'EDS

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt$$

avec $(W_t)_t$ un mouvement brownien à valeurs $\mathbb{R}^d$ et que l'on souhaite évaluer

$C = \mathbb{E}\left(f(X_T)1_{\{\forall t\in[0,T], X_t\in D\}}\right)$ où $D = \{x \in \mathbb{R}^n : e.(x-z) < 0\}$ pour $e, z \in \mathbb{R}^n$ t.q. $e \neq 0$.

On va approcher cette quantité à l'aide du schéma d'Euler en temps continu $(\bar{X}_t)_{t\in[0,T]}$ en calculant

$$C_N = \mathbb{E}\left(f(\bar{X}_T)1_{\{\forall t\in[0,T], \bar{X}_t\in D\}}\right).$$

En généralisant l'approche développée en dimension 1, on peut montrer que

$$C_N = \mathbb{E}\left(f(\bar{X}_T)\prod_{k=0}^{N-1}\left(1-\psi(\bar{X}_{t_k}, \bar{X}_{t_{k+1}})\right)\right)$$

où

$$\psi(x_k, x_{k+1}) = \mathbb{P}\left(\exists t \in [t_k, t_{k+1}] : \bar{X}_t \notin D \middle| (\bar{X}_{t_k}, \bar{X}_{t_{k+1}}) = (x_k, x_{k+1})\right)$$
$$= \mathbb{P}\left(\exists t \in [t_k, t_{k+1}] : e.\bar{X}_t \geq e.z \middle| (\bar{X}_{t_k}, \bar{X}_{t_{k+1}}) = (x_k, x_{k+1})\right).$$

Pour $t \in [t_k, t_{k+1}]$, conditionnellement à $\bar{X}_{t_k} = x_k$, on a

$$e.\bar{X}_t = e.x_k + (\sigma^*(x_k)e).(W_t - W_{t_k}) + (t-t_k)b(x_k).e$$

On distingue alors deux cas :

— Lorsque le vecteur $\sigma^*(x_k)e \in \mathbb{R}^d$ est nul, $t \in [t_k, t_{k+1}] \to e.\bar{X}_t$ est une fonction affine ce qui entraîne que $\psi(x_k, x_{k+1}) = 1_{\{x_k\notin D \text{ ou } x_{k+1}\notin D\}}$.

— Sinon, on remarque que $\frac{\sigma^*(x_k)e}{|\sigma^*(x_k)e|}.(W_t - W_{t_k})$ est un mouvement brownien unidimensionnel. Comme dans la proposition 3.6.1, on en déduit que conditionnellement à $(\bar{X}_{t_k}, \bar{X}_{t_{k+1}}) = (x_k, x_{k+1})$, $\max_{t\in[t_k,t_{k+1}]} e.\bar{X}_t$ a même loi que $e.x_k +$

$|\sigma^*(x_k)e| \max_{t\in[0,t_1]} B_t$ sachant $B_{t_1} = \frac{e.(x_{k+1}-x_k)}{|\sigma^*(x_k)e|}$ avec $(B_t)_t$ un mouvement brownien réel. Donc

$$\psi(x_k, x_{k+1}) = \mathbb{P}\left(\max_{t\in[0,t_1]} B_t \geq \frac{e.(z-x_k)}{|\sigma^*(x_k)e|}\middle| B_{t_1} = \frac{e.(x_{k+1}-x_k)}{|\sigma^*(x_k)e|}\right).$$

En utilisant (3.24), on conclut que

$$\psi(x_k, x_{k+1}) = \begin{cases} 1 \text{ si } e.x_k \geq e.z \text{ ou } e.x_{k+1} \geq e.z \\ e^{-\frac{2(e.(z-x_k))(e.(z-x_{k+1}))}{t_1|\sigma^*(x_k)e|^2}} \text{ sinon.} \end{cases}$$

On conclut que

$$C_N = \mathbb{E}\left(f(\bar{X}_T) 1_{\{\forall 0\leq k\leq N,\ \bar{X}_{t_k}\in D\}} \prod_{k=0}^{N-1}\left[1 - \exp\left(-\frac{2(e.(z-\bar{X}_{t_k}))(e.(z-\bar{X}_{t_{k+1}}))}{t_1|\sigma^*(\bar{X}_{t_k})e|^2}\right)\right]\right).$$

Cette formule généralise celle obtenue en dimension 1.

Remarque 3.6.7. *Lorsque D est un domaine régulier de $\mathbb{R}^n$ et non un demi-espace comme supposé plus haut, on peut utiliser l'approche précédente sur chaque pas de temps en remplaçant la frontière de D par un hyperplan tangent en un point bien choisi. Plus précisément pour $x_k, x_{k+1} \in D$, on pourra approcher $\mathbb{P}\left(\exists t \in [t_k, t_{k+1}] : \bar{X}_t \notin D \middle| (\bar{X}_{t_k}, \bar{X}_{t_{k+1}}) = (x_k, x_{k+1})\right)$ en remplaçant la frontière ∂D de D par l'hyperplan tangent à cette frontière en la projection de x_k sur ∂D et en utilisant la formule obtenue plus haut.*

3.7 Options asiatiques dans le modèle de Black-Scholes

En dimension $n = d = 1$, on condidère l'EDS

$$dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt,\ X_0 = x_0.$$

On s'intéresse à une option asiatique dont le payoff porte sur le couple $(X_T, \frac{1}{T}\int_0^T X_s ds) = (X_T, \frac{1}{T}I_T)$ où $I_t \stackrel{\text{def}}{=} \int_0^t X_s ds$ vérifie $dI_t = X_t dt$. Appliqué à l'EDS en dimension $n = 2$ satisfaite par le couple (X_t, I_t) le schéma d'Euler s'écrit $(\bar{X}_0, \bar{I}_0) = (x_0, 0)$ et

$$\forall k \in \{0, \ldots, N-1\} \begin{cases} \bar{X}_{t_{k+1}} = \bar{X}_{t_k} + \sigma(t_k, \bar{X}_{t_k})(W_{t_{k+1}} - W_{t_k}) + b(t_k, \bar{X}_{t_k})(t_{k+1} - t_k) \\ \bar{I}_{t_{k+1}} = \bar{I}_{t_k} + \bar{X}_{t_k}(t_{k+1} - t_k) \end{cases}.$$

Les résultats de convergence forte et faible énoncés précédemment s'appliquent pour contrôler l'erreur commise en approchant $(X_T, \frac{1}{T}\int_0^T X_s ds)$ par $(\bar{X}_T, \frac{1}{T}\bar{I}_T = \frac{1}{N}\sum_{k=0}^{N-1}\bar{X}_{t_k})$. En particulier, si les fonctions σ et b satisfont l'hypothèse (Lip), la vitesse d'approximation forte de $\frac{1}{T}\int_0^T X_s ds$ par $\frac{1}{N}\sum_{k=0}^{N-1}\bar{X}_{t_k}$ est en $\frac{C}{N^{\min(\frac{1}{2},\alpha)}}$ où α est l'indice de Hölder donnant la régularité en temps des coefficients σ et b du modèle.

Dans le cas du modèle de Black-Scholes

$$dX_t = \sigma X_t dW_t + rX_t dt,\ X_0 = y > 0,$$

il est possible de simuler sans commettre d'erreur de discrétisation le vecteur $(X_0, X_{t_1}, X_{t_2}, \ldots, X_T) = y \times (1, e^{\sigma W_{t_1} + (r - \frac{\sigma^2}{2})t_1}, e^{\sigma W_{t_2} + (r - \frac{\sigma^2}{2})t_2}, \ldots, e^{\sigma W_T + (r - \frac{\sigma^2}{2})T})$. Il faut bien sûr tirer parti de cette possibilité comme l'indique le résultat suivant :

Proposition 3.7.1.

$$\forall p \geq 1,\ \exists C_p > 0,\ \forall N \geq 1,\ \mathbb{E}\left(\left|\frac{1}{T}\int_0^T X_t dt - \frac{1}{N}\sum_{k=0}^{N-1} X_{t_k}\right|^{2p}\right) \leq \frac{C_p}{N^{2p}}.$$

Remarque 3.7.2. *En particulier pour une fonction de payoff φ lipschitzienne sur $\mathbb{R}^2$ comme $\varphi(x,y) = (x-y)^+$ (Put floating Strike) ou $\varphi(x,y) = (y-K)^+$ (Call fixed Strike) on en déduit*

$$\left|\mathbb{E}\left(e^{-rT}\varphi\left(X_T, \frac{1}{T}\int_0^T X_t dt\right)\right) - \mathbb{E}\left(e^{-rT}\varphi\left(X_T, \frac{1}{N}\sum_{k=0}^{N-1} X_{t_k}\right)\right)\right| \leq \frac{C}{N}.$$

Démonstration : Soit $k \in \{0, \ldots, N-1\}$. Si on pose $Y_t = X_t - X_{t_k} = \int_{t_k}^t (\sigma X_s dW_s + rX_s ds)$, $Z_t = t_{k+1} - t$, la formule d'intégration par parties

$$Y_{t_{k+1}} Z_{t_{k+1}} - Y_{t_k} Z_{t_k} = \int_{t_k}^{t_{k+1}} Y_t dZ_t + \int_{t_k}^{t_{k+1}} Z_t dY_t + < Y, Z >_{t_{k+1}} - < Y, Z >_{t_k}$$

s'écrit

$$0 - 0 = -\int_{t_k}^{t_{k+1}} (X_t - X_{t_k}) dt + \int_{t_k}^{t_{k+1}} (t_{k+1} - t)(\sigma X_t dW_t + rX_t dt) + 0 - 0.$$

On en déduit que

$$\begin{aligned}\frac{1}{T}\int_0^T X_t dt - \frac{1}{N}\sum_{k=0}^{N-1} X_{t_k} &= \frac{1}{T}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} (t_{k+1} - t)(\sigma X_t dW_t + rX_t dt) \\ &= \frac{\sigma}{T}\int_0^T (\bar{\tau}_t - t) X_t dW_t + \frac{r}{T}\int_0^T (\bar{\tau}_t - t) X_t dt,\end{aligned}$$

où $\bar{\tau}_t = \left\lceil \frac{tN}{T} \right\rceil \frac{T}{N}$ désigne le premier instant de discrétisation après t (pour $x \in \mathbb{R}$, on note $\lceil x \rceil$ le plus petit entier plus grand que x). En procédant comme dans la preuve de la Proposition 3.1.6, on en déduit que

$$\mathbb{E}\left(\left|\frac{1}{T}\int_0^T X_t dt - \frac{1}{N}\sum_{k=0}^{N-1} X_{t_k}\right|^{2p}\right) \leq C\int_0^T (\bar{\tau}_t - t)^{2p}\mathbb{E}(X_t^{2p}) dt \leq C\left(\frac{T}{N}\right)^{2p}\int_0^T \mathbb{E}(X_t^{2p}) dt.$$

On conclut en remarquant que $\mathbb{E}(X_t^{2p}) = y^{2p} e^{2p(r - \frac{\sigma^2}{2})t}\mathbb{E}(e^{2p\sigma W_t}) = y^{2p} e^{p(2r + (2p-1)\sigma^2)t}$ assure que $\int_0^T \mathbb{E}(X_t^{2p}) dt < +\infty$. □

Approcher $m_T = \frac{1}{T}\int_0^T X_t dt$ par $\frac{1}{N}\sum_{k=0}^{N-1} X_{t_k}$ consiste à discrétiser l'intégrale par la méthode des rectangles. D'après les travaux de Lapeyre et Temam [50], il vaut beaucoup mieux discrétiser l'intégrale par la méthode des trapèzes et approcher la moyenne du cours de l'actif par $m_T^N = \frac{1}{N}\left(\frac{X_0+X_T}{2} + \sum_{j=1}^{N-1} X_{t_j}\right)$. L'erreur forte commise est toujours en $\frac{C}{N}$ mais la constante au numérateur est bien plus petite. En fait, d'après [50],

$$\mathbb{E}\left((m_T - \mathbb{E}\left(m_T|(W_{t_1}, W_{t_2}, \ldots, W_T)\right))^2\right) = \mathcal{O}\left(\frac{1}{N^2}\right)$$
$$\mathbb{E}\left(\left(\mathbb{E}\left(m_T|(W_{t_1}, W_{t_2}, \ldots, W_T)\right) - m_T^N\right)^2\right) = \mathcal{O}\left(\frac{1}{N^3}\right),$$

Comme l'espérance conditionnelle est la projection orthogonale au sens L^2 sur les variables aléatoires mesurables par rapport à la tribu conditionnante, cela montre que m_T^N est très proche de la meilleure approximation (au sens quadratique) $\mathbb{E}\left(m_T|(W_{t_1}, W_{t_2}, \ldots, W_T)\right)$ de m_T ne faisant intervenir que $(W_{t_1}, W_{t_2}, \ldots, W_T)$.

Pour aller plus loin, il faut rajouter de l'information dans le schéma. Pour $t \in [t_k, t_{k+1}[$, X_t est proche de sa discrétisation par schéma d'Euler : $X_{t_k}(1+\sigma(W_t - W_{t_k}) + r(t-t_k))$. Ainsi $\int_{t_k}^{t_{k+1}} X_t dt$ est proche de $X_{t_k}\left(\frac{T}{N} + \sigma\int_{t_k}^{t_{k+1}}(W_t - W_{t_k})dt + \frac{rT^2}{2N^2}\right)$ ce qui conduit à approcher m_T par

$$\hat{m}_T^N = \frac{1}{N}\sum_{k=0}^{N-1} X_{t_k}\left(1 + \frac{\sigma N}{T}\int_{t_k}^{t_{k+1}}(W_t - W_{t_k})dt + \frac{rT}{2N}\right).$$

D'après [50], l'erreur forte de ce schéma est en $\frac{C}{N^{3/2}}$.

Proposition 3.7.3.

$$\forall p \geq 1,\ \exists C_p > 0,\ \forall N \geq 1,\ \mathbb{E}\left(\left|m_T - \hat{m}_T^N\right|^{2p}\right) \leq \frac{C_p}{N^{3p}}.$$

En plus des accroissements $\Delta W_{k+1} = W_{t_{k+1}} - W_{t_k}$ pour $k \in \{0, \ldots, N-1\}$, ce schéma fait intervenir les variables $\int_{t_k}^{t_{k+1}}(W_t - W_{t_k})dt$. Si on pose $Y_{k+1} = \int_{t_k}^{t_{k+1}}(W_t - W_{t_k})dt - \frac{T}{2N}\Delta W_{k+1}$ il est facile de simuler le vecteur $(Y_1, \Delta W_1, Y_2, \Delta W_2, \ldots, Y_N, \Delta W_N)$ d'après l'exercice qui suit.

Exercice 3.7.4. *Montrer que les variables* $(Y_{k+1}, \Delta W_{k+1})_{k\in\{0,\ldots,N-1\}}$ *sont i.i.d. suivant la loi gaussienne* $\mathcal{N}_2\left(0, \begin{pmatrix} \frac{T^3}{12N^3} & 0 \\ 0 & \frac{T}{N} \end{pmatrix}\right)$.

3.8 Problèmes

3.8.1 Vitesse forte du schéma d'Euler dans le cas d'un coefficient de diffusion constant

Soit $T > 0$, $x \in \mathbb{R}$, $b : \mathbb{R} \to \mathbb{R}$ une fonction C_b^2 (deux fois continuement dérivable avec des dérivées première et seconde bornées) et sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, $(W_t)_{t\in[0,T]}$ un mouvement brownien réel standard. On s'intéresse à l'équation différentielle stochastique

$$\begin{cases} X_0 = x, \\ dX_t = dW_t + b(X_t)dt. \end{cases} \tag{3.27}$$

On se donne $N \in \mathbb{N}^*$ et pour $0 \leq k \leq N$, on pose $t_k = k\Delta t$ où $\Delta t = T/N$. Le schéma d'Euler à N pas de temps associé à (3.27) est défini de façon récurrente par

$$\begin{cases} \bar{X}_0 = x \\ \forall 0 \leq k \leq N-1,\ \forall t \in [t_k, t_{k+1}], \bar{X}_t = \bar{X}_{t_k} + (W_t - W_{t_k}) + b(\bar{X}_{t_k})(t - t_k). \end{cases}$$

Les 2 questions peuvent être traitées de façon indépendante.

1. **Cas général :**

   (a) Quel résultat du cours assure que $\sup_{t\in[0,T]} \mathbb{E}\left(|X_t - \bar{X}_t|\right) \leq C/N$, où C ne dépend pas de N ? On se propose d'améliorer cette majoration en prouvant que

   $$\mathbb{E}\left(\sup_{t\in[0,T]} |X_t - \bar{X}_t|\right) \leq \frac{C}{N}. \tag{3.28}$$

   (b) Pour $s \in [0,T]$, on note $\tau_s = [s/\Delta t] \times \Delta t$ le dernier instant de discrétisation avant s (pour $y \in \mathbb{R}$, $[y]$ désigne la partie entière de y). Vérifier que

   $$\forall t \in [0,T],\ |X_t - \bar{X}_t| \leq \sup|b'| \int_0^t |X_{\tau_s} - \bar{X}_{\tau_s}|ds + \left|\int_0^t b(X_s) - b(X_{\tau_s})ds\right|.$$

   (c) Vérifier que pour $1 \leq k \leq N$,

   $$\begin{aligned}\int_{t_{k-1}}^{t_k} b(X_s) - b(X_{t_{k-1}})ds = & \int_{t_{k-1}}^{t_k} (t_k - r)\left(b(X_r)b'(X_r) + \frac{1}{2}b''(X_r)\right)dr \\ & + \int_{t_{k-1}}^{t_k} (t_k - r)\ b'(X_r)dW_r.\end{aligned}$$

   (On pourra commencer par calculer $b(X_s) - b(X_{t_{k-1}})$).

   (d) En déduire que

   $$\begin{aligned}\int_0^t b(X_s) - b(X_{\tau_s})ds = & \int_0^t (\min(t, \tau_r + \Delta t) - r)\left(b(X_r)b'(X_r) + \frac{1}{2}b''(X_r)\right)dr \\ & + \int_0^t (\min(t, \tau_r + \Delta t) - r)\ b'(X_r)dW_r,\end{aligned}$$

puis que

$$\mathbb{E}\left(\sup_{u\in[0,t]}\left|\int_0^u b(X_s)-b(X_{\tau_s})ds\right|\right)$$
$$\leq \Delta t\left(\int_0^t \mathbb{E}\left|b(X_r)b'(X_r)+\frac{1}{2}b''(X_r)\right|dr+2\sqrt{\int_0^t \mathbb{E}\left((b'(X_r))^2\right)dr}\right).$$

(e) On pose $z(t)=\mathbb{E}\left(\sup_{u\in[0,t]}|X_u-\bar{X}_u|\right)$. Montrer que

$$\forall t\in[0,T],\ z(t)\leq C\left(\Delta t+\int_0^t z(s)ds\right),$$

où la constante C ne dépend pas de N et conclure que l'inégalité (3.28) est satisfaite.

2. **Cas d'un coefficient de dérive linéaire :** $\forall y\in\mathbb{R}$, $b(y)=cy$ où $c\in\mathbb{R}^*$.
   (a) Montrer que la solution de l'équation différentielle stochastique (3.27) est donnée par

$$\forall t\in[0,T],\ X_t=xe^{ct}+\int_0^t e^{c(t-s)}dW_s.$$

   (b) Que peut-on dire du vecteur $(X_T,W_{t_1}-W_0,W_{t_2}-W_{t_1},\ldots,W_{t_N}-W_{t_{N-1}})$?
   Pour $1\leq k\leq N$, calculer $\mathrm{Cov}\,(X_T,W_{t_k}-W_{t_{k-1}})$.
   (c) On note $\mathcal{F}_N$ la tribu engendrée par $(W_{t_1},W_{t_2},\ldots,W_{t_N})$. Montrer que

$$\mathbb{E}(X_T|\mathcal{F}_N)=xe^{cT}+\frac{1}{c\Delta t}\sum_{k=1}^N e^{c(N-k)\Delta t}(e^{c\Delta t}-1)(W_{t_k}-W_{t_{k-1}}).$$

   (d) Vérifier que

$$\mathbb{E}\left((X_T-\mathbb{E}(X_T|\mathcal{F}_N))^2\right)=\frac{1}{2c}(e^{2cT}-1)\left(1-\frac{2}{c\Delta t}\ \frac{e^{c\Delta t}-1}{e^{c\Delta t}+1}\right).$$

   (e) Vérifier que pour $\alpha\to 0$,

$$\frac{e^\alpha-1}{\alpha(e^\alpha+1)}=\frac{1}{2}-\frac{\alpha^2}{24}+o(\alpha^2).$$

   (f) En déduire l'équivalent pour N tendant vers $+\infty$ de l'erreur quadratique minimale $\mathbb{E}((X_T-\tilde{X}_T)^2)$ commise en générant $\tilde{X}_T$ à l'aide d'un schéma "simple" qui ne fait intervenir que les accroissements $W_{t_k}-W_{t_{k-1}},\ 1\leq k\leq N$. Conclure que l'ordre de convergence du schéma d'Euler est l'ordre optimal parmi les schémas "simples" (la constante multiplicative n'étant pas nécessairement optimale).
   (g) Que peut-on dire du vecteur $(X_{t_1},X_{t_2},\ldots,X_{t_n})$?
   Soit $1\leq k\leq N$. Vérifier que

$$X_{t_k}=e^{c\Delta t}X_{t_{k-1}}+\int_{t_{k-1}}^{t_k}e^{c(t_k-s)}dW_s$$

   et en déduire α t.q. la variable $Z_k=X_{t_k}-\alpha X_{t_{k-1}}$ soit indépendante de $X_{t_{k-1}}$.
   Vérifier qu'en fait Z_k est indépendante du vecteur $(X_{t_1},\ldots,X_{t_{k-1}})$.
   Quelle est la loi de Z_k ?
   (h) En déduire une méthode permettant de générer un vecteur $(Y_1,\ldots,Y_N)$ qui a même loi que $(X_{t_1},\ldots,X_{t_N})$.

3.8.2 Approximation du maximum par le schéma d'Euler

On considère l'équation différentielle stochastique unidimensionnelle

$$dX_t = \sigma(X_t)\, dW_t + b(X_t)\, dt,$$

où b et σ sont des fonctions lipschitziennes. Le processus (W_t) est un mouvement brownien standard unidimensionnel.

On se donne une condition initiale $X_0 = x$ déterministe. Enfin, on se fixe un horizon en temps T et un nombre $N \in \mathbb{N}^*$ de pas de temps. Pour $k \in \{0, \dots, N\}$ on note $t_k = \frac{kT}{N}$ le k-ème instant de discrétisation et $\bar{X}^N_{t_k}$ la valeur en t_k du schéma d'Euler de pas $\frac{T}{N}$.

Le but de ce problème est de montrer l'estimation suivante (issue de la thèse de P. Seumen Tonou) : pour tout $p \geq 1$

$$\mathbb{E}\left(\left|\max_{0\leq t\leq T} X_t - \max_{0\leq k\leq N} \bar{X}^N_{t_k}\right|^{2p}\right) \leq C\left(\frac{\ln(N)}{N}\right)^p. \tag{3.29}$$

Dans l'inégalité précédente et dans tout le reste de l'énoncé on notera C des constantes indépendantes de N qui peuvent varier de ligne en ligne. On pose

$$\begin{aligned}
\varepsilon &= \max_{0\leq t\leq T} X_t - \max_{0\leq k\leq N} \bar{X}^N_{t_k},\\
\varepsilon_1 &= \max_{0\leq t\leq T} X_t - \max_{0\leq k\leq N} X_{t_k},\\
\varepsilon_2 &= \max_{0\leq k\leq N} X_{t_k} - \max_{0\leq k\leq N} \bar{X}^N_{t_k}.
\end{aligned}$$

1. Montrer
$$\varepsilon_2 \leq \max_{0\leq k\leq N} (X_{t_k} - \bar{X}^N_{t_k}),$$
puis
$$\varepsilon_2 \geq -\max_{0\leq k\leq N} (\bar{X}^N_{t_k} - X_{t_k}).$$
Appliquer un résultat du cours pour en déduire
$$\mathbb{E}\left(|\varepsilon_2|^{2p}\right) \leq \frac{C}{N^p}.$$

2. Montrer que
$$\varepsilon_1 \leq \max_{0\leq k\leq N-1}\left(\max_{t_k\leq t\leq t_{k+1}} (X_t - X_{t_k})\right).$$
En déduire que
$$|\varepsilon_1|^{2p} \leq C \max_{0\leq k\leq N-1}\left(\max_{t_k\leq t\leq t_{k+1}} \int_{t_k}^t b(X_s)\, ds\right)^{2p} + C \max_{0\leq k\leq N-1}\left(\max_{t_k\leq t\leq t_{k+1}} \int_{t_k}^t \sigma(X_s)\, dW_s\right)^{2p}.$$

3. En utilisant l'inégalité
$$\max_{0\leq k\leq N-1} |x_k| \leq \sum_{k=0}^{N-1} |x_k| \tag{3.30}$$
montrer que
$$\mathbb{E}\left(\max_{0\leq k\leq N-1}\left(\max_{t_k\leq t\leq t_{k+1}} \int_{t_k}^t b(X_s)\, ds\right)^{2p}\right) \leq \frac{C}{N^{2p-1}}.$$

4. (a) Montrer que

$$\mathbb{E}\left(\max_{0\leq k\leq N-1}\left(\max_{t_k\leq t\leq t_{k+1}}\int_{t_k}^{t}[\sigma(X_s)\ -\sigma(X_{t_k})]dW_s\right)^{2p}\right)\leq\frac{C}{N^{2p-1}}.$$

(b) Montrer que

$$\mathbb{E}\left(\max_{0\leq k\leq N-1}\left(\max_{t_k\leq t\leq t_{k+1}}\int_{t_k}^{t}\sigma(X_{t_k})dW_s\right)^{2p}\right)$$
$$\leq C\sqrt{\mathbb{E}\left(\left(\max_{0\leq k\leq N-1}\max_{t_k\leq t\leq t_{k+1}}|W_t-W_{t_k}|\right)^{4p}\right)}.$$

(c) Vérifier que pour $y\geq 0$, $\mathbb{P}(\max_{0\leq t\leq t_1}W_t\geq y)=2\mathbb{P}(W_{t_1}\geq y)$. En déduire que $\max_{0\leq k\leq N-1}\max_{t_k\leq t\leq t_{k+1}}(W_t-W_{t_k})$ a même loi que $\sqrt{t_1}\max_{0\leq k\leq N-1}|G_k|$ où les variables aléatoires $(G_k)_{0\leq k\leq N-1}$ sont i.i.d. suivant la loi normale centrée réduite $\mathcal{N}_1(0,1)$.

(d) Soit q le plus petit entier impair qui majore $4p$. Montrer que pour $x>0$, $\mathbb{E}(|G_0|^q 1_{\{|G_0|\geq x\}})=P(x)e^{-\frac{x^2}{2}}$ où $P(.)$ est un polynôme de degré $q-1$. En déduire à l'aide de (3.30) que pour $\alpha>2$,

$$\lim_{N\to\infty}\mathbb{E}\left(\max_{0\leq k\leq N-1}\left(|G_k|^{4p}1_{\{|G_k|\geq\sqrt{\alpha\ln(N)}\}}\right)\right)=0,$$

puis que pour $N\geq 2$, $\mathbb{E}\left(\max_{0\leq k\leq N-1}|G_k|^{4p}\right)\leq C(\ln(N))^{2p}$.

(e) Conclure que

$$\mathbb{E}\left(\max_{0\leq k\leq N-1}\left(\max_{t_k\leq t\leq t_{k+1}}\int_{t_k}^{t}\sigma(X_s)\ dW_s\right)^{2p}\right)\leq C\left(\frac{\ln(N)}{N}\right)^p.$$

5. Vérifier l'estimation (3.29).
6. Comment peut-on améliorer l'approximation de $\max_{0\leq t\leq T}X_t$?

3.8.3 Schéma randomisé pour un coefficient de dérive irrégulier en temps

Soit $T>0$ un horizon de temps. Sur un espace de probabilité $(\Omega,\mathcal{A},\mathbb{P})$ muni d'une filtration $(\mathcal{F}_t)_{t\geq 0}$, on considère un $\mathcal{F}_t$-mouvement brownien W_t à valeurs $\mathbb{R}^d$. On se donne également des coefficients $\sigma:[0,T]\times\mathbb{R}^n\to\mathbb{R}^{n\times d}$ et $b:[0,T]\times\mathbb{R}^n\to\mathbb{R}^n$ mesurables et qui vérifient

$$(\mathrm{Lip})\ \exists\mathrm{K}<\infty,\ \forall\mathrm{t}\in[0,\mathrm{T}],\ \begin{cases}\forall x\in\mathbb{R}^n,\ |\sigma(t,x)|+|b(t,x)|\leq K(1+|x|)\\ \forall x,y\in\mathbb{R}^n,\ |\sigma(t,x)-\sigma(t,y)|+|b(t,x)-b(t,y)|\leq K|x-y|\end{cases}.$$

On s'intéresse à l'Équation Différentielle Stochastique

$$\begin{cases}dX_t=\sigma(t,X_t)dW_t+b(t,X_t)dt,\ t\in[0,T]\\ X_0=x_0\in\mathbb{R}^n\end{cases}$$

1. Quel résultat assure l'existence d'une unique solution à cette équation ?

Soit $N \in \mathbb{N}^*$, $(t_k = \frac{kT}{N})_{0\leq k\leq N}$ et $\tau_t = \lfloor \frac{Nt}{T} \rfloor \frac{T}{N}$ l'instant de cette grille uniforme qui précède $t \in [0,T]$.

2. Écrire le schéma d'Euler sur la grille $(t_k)_{0\leq k\leq N}$.
3. Rappeler le résultat de convergence forte de ce schéma.

On se place désormais en dimension $\boxed{n = d = 1}$, on suppose (Lip) et

$$\exists K < \infty,\ \forall x \in \mathbb{R}^n,\ \forall 0 \leq s \leq t \leq T,\ |\sigma(t,x) - \sigma(s,x)| \leq K(1+|x|)\sqrt{t-s}, \tag{3.31}$$

mais on ne fait pas d'hypothèse sur la régularité du coefficient de dérive b en la variable temporelle t. On se donne $(U_k)_{0\leq k\leq N-1}$ une suite de variables aléatoires indépendantes et identiquement distribuées sur $[0,1]$ indépendantes de $(W_t)_{t\geq 0}$. On pose $\eta_k = t_k + \frac{T}{N}U_k$ pour $k \in \{0,\ldots,N-1\}$ et on définit le schéma d'Euler avec randomisation de la variable temporelle dans la dérive en posant $\tilde{X}_0^N = x_0$ puis pour $k \in \{0,\ldots,N-1\}$,

$$\tilde{X}^N_{t_{k+1}} = \tilde{X}^N_{t_k} + \sigma(t_k, \tilde{X}^N_{t_k})(W_{t_{k+1}} - W_{t_k}) + b(\eta_k, \tilde{X}^N_{t_k})(t_{k+1} - t_k).$$

4. Quelle est la loi de la suite $(\eta_k)_{0\leq k\leq N-1}$?
5. Montrer que pour $k \in \{0,\ldots,N-1\}$,
   (a)
$$\mathbb{E}\left[|\tilde{X}^N_{t_{k+1}}|^2\right] = \mathbb{E}\left[|\tilde{X}^N_{t_k}|^2\right] + \frac{T}{N}\mathbb{E}\left[|\sigma(t_k,\tilde{X}^N_{t_k})|^2 + 2\tilde{X}^N_{t_k}.b(\eta_k,\tilde{X}^N_{t_k})\right] + \frac{T^2}{N^2}\mathbb{E}\left[|b(\eta_k,\tilde{X}^N_{t_k})|^2\right],$$
   (b) puis que
$$\mathbb{E}\left[|\tilde{X}^N_{t_{k+1}}|^2\right] \leq \left(1 + \frac{K(3+2K)T}{N} + \frac{2K^2T^2}{N^2}\right)\mathbb{E}\left[|\tilde{X}^N_{t_k}|^2\right] + \frac{K(1+2K)T}{N} + \frac{2K^2T^2}{N^2}$$
   (c) et enfin que $1 + \mathbb{E}\left[|\tilde{X}^N_{t_{k+1}}|^2\right] \leq \left(1 + \frac{K(3+2K+2KT)T}{N}\right)\left(1 + \mathbb{E}\left[|\tilde{X}^N_{t_k}|^2\right]\right)$.
   (d) Conclure que $\sup_{N\in\mathbb{N}^*} \max_{k\in\{0,\ldots,N\}} \mathbb{E}\left[|\tilde{X}^N_{t_k}|^2\right] \leq e^{K(3+2K+2KT)T}(1+|x_0|^2) - 1$.

Pour $k \in \{0,\ldots,N\}$, on pose $M^N_k = \frac{T}{N}\sum_{\ell=0}^{k-1} b(\eta_\ell, \tilde{X}^N_{t_\ell}) - \int_0^{t_k} b(s, \tilde{X}^N_{\tau_s})ds$ (Convention : $M^N_0 = 0$) et $\mathcal{G}^N_k = \mathcal{F}_{t_k} \vee \sigma(\eta_\ell, 0 \leq \ell \leq k-1)$.

6. (a) Montrer que $(M^N_k)_{k\in\{0,\ldots,N\}}$ est une $\mathcal{G}^N_k$-martingale de carré intégrable.
   (b) Calculer $\langle M^N\rangle_k$ pour $k \in \{0,\ldots,N\}$.
   (c) En déduire à l'aide de l'inégalité de Doob que
$$\mathbb{E}\left[\max_{k\in\{0,\ldots,N\}} |M^N_k|^2\right] \leq 4\left(\frac{T}{N}\int_0^T \mathbb{E}[|b(s,\tilde{X}^N_{\tau_s})|^2]ds - \sum_{k=0}^{N-1}\mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}} b(s,\tilde{X}^N_{\tau_s})ds\right|^2\right]\right).$$
   (d) Conclure que $\sup_{N\in\mathbb{N}^*} N\mathbb{E}\left[\max_{k\in\{0,\ldots,N\}} |M^N_k|^2\right] < \infty$.

Pour $t \in [0,T]$, on pose $\hat{X}^N_t = x_0 + \int_0^t \sigma(\tau_s, \tilde{X}^N_{\tau_s})dW_s + \int_0^t b(s, \tilde{X}^N_{\tau_s})ds$.

7. (a) Que peut-on dire de $C := \sup_{N\in\mathbb{N}^*} N\mathbb{E}\left[\max_{k\in\{0,\ldots,N\}} |\hat{X}^N_{t_k} - \tilde{X}^N_{t_k}|^2\right]$?
   (b) Montrer que pour tout $N \in \mathbb{N}^*$ et tout $t \in [0,T]$,
$$\mathbb{E}\left[\sup_{u\in[0,t]} |X_u - \hat{X}^N_u|^2\right] \leq 2\int_0^t 4\mathbb{E}[|\sigma(s,X_s) - \sigma(\tau_s,\tilde{X}^N_{\tau_s})|^2] + t\mathbb{E}[|b(s,X_s) - b(s,\tilde{X}^N_{\tau_s})|^2]ds$$

(c) Vérifier que

$$\mathbb{E}[|\sigma(s,X_s)-\sigma(\tau_s,\tilde{X}^N_{\tau_s})|^2] \leq 4K^2\bigg(\mathbb{E}[|X_s-X_{\tau_s}|^2]+(s-\tau_s)\mathbb{E}[(1+|X_{\tau_s}|)^2] + \mathbb{E}[|X_{\tau_s}-\hat{X}^N_{\tau_s}|^2]+\frac{C}{N}\bigg).$$

(d) Comment majorer $\mathbb{E}[|b(s,X_s)-b(s,\tilde{X}^N_{\tau_s})|^2]$?

(e) En déduire que $\sup_{N\in\mathbb{N}^*} N\mathbb{E}\left[\sup_{t\in[0,T]}|X_t-\hat{X}^N_t|^2\right]<\infty$.

(f) Quel est l'ordre de convergence forte du schéma $(\tilde{X}^N_{t_k})_{k\in\{0,\ldots,N\}}$? Comment se dégrade-t-il si on suppose seulement σ höldérienne d'ordre $\alpha\in]0,1/2[$ en la variable temporelle t (i.e. $\sqrt{t-s}$ est remplacé par $(t-s)^\alpha$ au membre de droite de (3.31)) ?

(g) Pourquoi la randomisation du coefficient de diffusion ne conduit-elle pas au même ordre de convergence forte sous l'hypothèse (Lip) ?

Toutes les majorations précédentes et donc l'ordre fort du schéma $(\tilde{X}^N_{t_k})_{k\in\{0,\ldots,N\}}$ restent valables en dimension n et d quelconques lorsque $|.|$ désigne la norme euclidienne sur $\mathbb{R}^n$ et $\mathbb{R}^d$ et pour $\varsigma\in\mathbb{R}^{n\times d}$, $|\varsigma|=\left(\sum_{i=1}^n\sum_{j=1}^d\varsigma_{ij}^2\right)^{1/2}$.

3.8.4 Discrétisation d'une équation différentielle stochastique à coefficients localement lipschitziens

Pour $(W_t)_{t\in[0,T]}$ un mouvement brownien à valeurs $\mathbb{R}^d$ et $\sigma:\mathbb{R}^n\to\mathbb{R}^{n\times d}$ et $b:\mathbb{R}^n\to\mathbb{R}^n$, on s'intéresse à l'EDS

$$dX_t=\sigma(X_t)dW_t+b(X_t)dt,\ t\leq T,\ X_0=y. \tag{3.32}$$

On suppose que les fonctions σ et b sont localement lipschitziennes au sens où pour tout $m\in\mathbb{N}^*$ il existe une constante $C_m<+\infty$ telle que

$$\forall x,y\in\mathbb{R}^n \text{ t.q. } |x|\leq m \text{ et } |y|\leq m,\ |\sigma(x)-\sigma(y)|+|b(x)-b(y)|\leq C_m|x-y|.$$

On note

$$\sigma_m(x)=\sigma(P_mx)\ \text{ et }\ b_m(x)=b(P_mx)\ \text{ où }\ P_mx=\frac{|x|\wedge m}{|x|}x \tag{3.33}$$

désigne la projection orthogonale de x sur la boule fermée de rayon m centrée à l'origine. On pose également $t_k=k\Delta t$ avec $\Delta t=\frac{T}{N}$ où $N\in\mathbb{N}^*$ est un nombre de pas de temps.

Absence de convergence faible et dans L^p dans un cas particulier

On s'intéresse à l'EDS en dimension $n=d=1$

$$X_t=W_t-\int_0^t X_s^3ds. \tag{3.34}$$

On admet pour l'instant qu'elle possède une unique solution $(X_t)_{t\in[0,T]}$ et que cette solution vérifie $\mathbb{E}\left(\sup_{t\in[0,T]}|X_t|^4\right)<+\infty$.
On note $(\bar{X}^N_t)_{t\in[0,T]}$ le schéma d'Euler en temps continu associé et on pose $A_N=\{|W_{t_1}|\geq \frac{3N}{T},\ \sup_{1\leq k\leq N-1}|W_{t_{k+1}}-W_{t_k}|\leq 1\}$.

1. Vérifier que $\mathbb{P}(A_N) \geq \mathbb{P}(|W_{t_1}| \geq \frac{3N}{T})\mathbb{P}(\sup_{t\in[0,T]} |W_t| \leq \frac{1}{2})$. En remarquant que pour $x > 0$, $\frac{e^{-x^2/2}}{x} = \int_x^\infty (1 + \frac{1}{y^2})e^{-y^2/2}dy$, vérifier que $\int_x^{+\infty} e^{-y^2/2}dy \geq \frac{xe^{-x^2/2}}{1+x^2}$.
   Conclure que $\mathbb{P}(A_N) \geq \mathbb{P}(\sup_{t\in[0,T]} |W_t| \leq \frac{1}{2}) \times \frac{6(NT)^{\frac{3}{2}}}{T^3+9N^3} \times \frac{e^{-\frac{9N^3}{2T^3}}}{\sqrt{2\pi}}$.
2. Vérifier que sur l'événement A_N, si pour $k \geq 1$, $|\bar{X}^N_{t_k}| \geq 1$, alors $|\bar{X}^N_{t_{k+1}}| \geq |\bar{X}^N_{t_k}|^2 \left(\frac{T}{N}|\bar{X}^N_{t_k}| - 2\right)$. En déduire que pour $N \geq \frac{T}{3}$, sur l'événement A_N, $\forall k \in \{1,\ldots,N\}$, $|\bar{X}^N_{t_k}| \geq \left(\frac{3N}{T}\right)^{2^{k-1}}$.
3. Conclure que $\lim_{N\to\infty} \mathbb{E}(|\bar{X}^N_T|) = +\infty$. En déduire $\lim_{N\to\infty} \mathbb{E}(|\bar{X}^N_T - X_T|)$ puis $\lim_{N\to\infty} \mathbb{E}(\sup_{t\in[0,T]} |\bar{X}^N_t - X_t|^p)$ pour $p \geq 1$.
4. En remarquant que la loi de $\bar{X}^N_T$ est symétrique, montrer que pour $K > 0$, $\mathbb{E}\left((\bar{X}^N_T - K)^+\right) \geq \frac{1}{2}\mathbb{E}(|\bar{X}^N_T|) - K$ et en déduire que $\mathbb{E}\left((\bar{X}^N_T - K)^+\right)$ ne converge pas vers $\mathbb{E}\left((X_T - K)^+\right)$ lorsque $N \to \infty$.

Convergence presque sûre du schéma d'Euler

On suppose maintenant qu'il existe une solution à l'EDS (3.32) (voir le paragraphe suivant pour une condition suffisante). On pose $\nu_0 = 0$ et pour $m \in \mathbb{N}^*$, on note $\nu_m = \inf\{t \in [0,T] : |X_t| > m\}$ avec la convention $\inf \emptyset = T$.

5. Écrire le schéma d'Euler en temps continu $(\bar{X}^N_t)_{t\in[0,T]}$ associé à (3.32).
6. Soit $m \in \mathbb{N}^*$ et $x, y \in \mathbb{R}^n$. Vérifier que $(x - P_m x, P_m y - P_m x) \leq 0$ (*dans le cas où* $|x| > m$, *on pourra vérifier que* $(x - P_m x, P_m x) = m|x - P_m x|$).
   En déduire que $|P_m y - P_m x|^2 \leq (y - x, P_m y - P_m x)$ et conclure que les fonctions σ_m et b_m sont globalement lipschitziennes de constante de Lipschitz C_m.
7. En déduire l'existence d'une unique solution $(X^m_t)_{t\in[0,T]}$ à l'EDS
   $$dX^m_t = \sigma_m(X^m_t)dW_t + b_m(X^m_t)dt,\ X^m_0 = y. \tag{3.35}$$
   Décrire l'événement $\{\nu_m = T\}$ à l'aide de la variable aléatoire $\sup_{t\in[0,T]} |X_t|$ et vérifier que sur cet événement $(X^m_t)_{t\in[0,T]}$, $(X_t)_{t\in[0,T]}$ et $(X^{m+1}_t)_{t\in[0,T]}$ coincident. Que vaut $\mathbb{P}(\bigcup_{m\in\mathbb{N}^*}\{\nu_m = T\})$? Y-a-t-il unicité pour (3.32) ?

On note $(\bar{X}^{m,N}_t)_{t\in[0,T]}$ le schéma d'Euler correspondant à (3.35) et $\nu^N_m = \min\{t_k : |\bar{X}^{N,m}_{t_k}| > m\}$.

8. Pour quelles valeurs de γ, la suite $N^\gamma \sup_{t\leq T} |X^m_t - \bar{X}^{m,N}_t|$ converge-t-elle presque sûrement lorsque $N \to \infty$ à m fixé ?
   En déduire que si $\nu_m(\omega) = T$, alors il existe $\mathcal{N}(\omega) < +\infty$ tel que pour $N \geq \mathcal{N}(\omega)$, $\nu^N_{m+1}(\omega) = T$ et $\sup_{t\leq T} |X_t(\omega) - \bar{X}^N_t(\omega)| = \sup_{t\leq T} |X^{m+1}_t(\omega) - \bar{X}^{m+1,N}_t(\omega)|$.
9. Conclure que pour tout $\gamma < \frac{1}{2}$, $N^\gamma \sup_{t\leq T} |X_t - \bar{X}^N_t|$ converge p.s. vers 0 lorsque $N \to \infty$.

Existence pour l'EDS (3.32) **lorsque la dérive est rentrante**

On suppose maintenant que σ est **globalement lipschitzienne** et que la fonction de dérive b est localement lipschitzienne et rentrante au sens où il existe $\beta \in [0, +\infty[$ tel que
$$\forall x \in \mathbb{R}^n,\ (x, b(x)) \leq \beta(|x| + |x|^2). \tag{3.36}$$

avec (x, y) qui désigne le produit scalaire de x et y dans $\mathbb{R}^n$. On note $(Y_t^m)_{t\in[0,T]}$ l'unique solution de l'EDS

$$dY_t^m = \sigma(Y_t^m)dW_t + b_m(Y_t^m)dt,\ Y_0^m = y.$$

10. Vérifier que toute fonction de dérive globalement lipschitzienne est rentrante. Donner un exemple de fonction de dérive b localement lipschitzienne rentrante mais pas globalement lipschitzienne.
11. Pour $m \in \mathbb{N}^*$, vérifier que la fonction b_m définie dans (3.70) satisfait (3.36).
12. Calculer $|Y_u^m|^2$ à l'aide de la formule d'Itô et en déduire que

$$\mathbb{E}\left(\sup_{u\leq t}|Y_u^m|^4\right) \leq 4\Bigg\{|y|^4 + 4\beta^2\mathbb{E}\left(\left(\int_0^t |Y_s^m| + |Y_s^m|^2 ds\right)^2\right) + \mathbb{E}\left(\left(\int_0^t |\sigma(Y_s^m)|^2 ds\right)^2\right) + 16\mathbb{E}\left(\int_0^t |\sigma^*(Y_s^m)Y_s^m|^2 ds\right)\Bigg\}$$

où $|\sigma(x)|^2 = \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq d}} |\sigma_{ij}(x)|^2$.

13. En déduire que $\sup_m \mathbb{E}\left(\sup_{t\in[0,T]}|Y_t^m|^4\right) < +\infty$ puis $\mathbb{P}\left(\bigcup_{m\in\mathbb{N}^*}\{\tau_m = T\}\right) = 1$ où $\tau_m = \inf\{t \in [0, T] : |Y_t^m| > m\}$ (convention $\inf \emptyset = T$)
Conclure à l'existence d'une unique solution $(X_t)_{t\in[0,T]}$ pour (3.32) et que cette solution vérifie $\mathbb{E}\left(\sup_{t\in[0,T]}|X_t|^4\right) < +\infty$ [1].

C'est l'absence de contrôle des moments du schéma d'Euler qui peut empêcher sa convergence dans L^p et sa convergence faible alors que l'on a convergence presque sûre à la vitesse $N^{-\gamma}$ pour tout $\gamma < 1/2$. Dans [36], Hutzenthaler, Jentzen et Kloeden ont proposé le schéma explicite suivant :

$$\tilde{X}_{t_{k+1}}^N = \tilde{X}_{t_k}^N + \sigma(\tilde{X}_{t_k}^N)(W_{t_{k+1}} - W_{t_k}) + \frac{b(\tilde{X}_{t_k}^N)\Delta t}{1 + |b(\tilde{X}_{t_k}^N)|\Delta t}.$$

Lorsque la fonction σ est globalement lipschitzienne, la matrice jacobienne de b est à croissance polynomiale et b vérifie la condition

$$\exists\alpha \in (0, +\infty),\ \forall x, y \in \mathbb{R}^n,\ (x - y, b(x) - b(y)) \leq \alpha|x - y|^2$$

qui implique (3.36) pour le choix $y = 0$, alors ils montrent que ce schéma converge à la vitesse forte $\frac{1}{\sqrt{N}}$ dans tous les espaces L^p. Ce résultat de convergence forte reste valable pour le schéma d'Euler semi-implicite

$$\hat{X}_{t_{k+1}}^N - b(\hat{X}_{t_{k+1}}^N)\Delta t = \hat{X}_{t_k}^N + \sigma(\hat{X}_{t_k}^N)(W_{t_{k+1}} - W_{t_k})$$

sous les mêmes hypothèses.

3.8.5 Convergence en loi de l'erreur renormalisée du schéma d'Euler dans le modèle de Black-Scholes

On s'intéresse à la discrétisation du modèle de Black-Scholes $X_t = e^{\sigma W_t + (\mu - \frac{\sigma^2}{2})t}$ où $(W_t)_{t\in[0,T]}$ est un mouvement brownien réel. On se donne un horizon T et un nombre $N \in \mathbb{N}^*$ de pas de discrétisation. Pour $k \in \{0, \ldots, N\}$, on pose également $t_k = \frac{kT}{N}$. Pour une suite réelle $(x_N)_{N\in\mathbb{N}^*}$ et $\alpha \in \mathbb{R}$, on note $x_N = \mathcal{O}(N^\alpha)$ lorsque la suite $(N^{-\alpha}x_N)_{N\in\mathbb{N}^*}$ est bornée.

1. Ce résultat s'applique en particulier à l'EDS (3.34).

1. Rappeler l'équation différentielle stochastique satisfaite par $(X_t)_{t\geq 0}$.
2. Écrire le schéma d'Euler $(X^N_{t_k})_{k\in\{0,\ldots,N\}}$ associé à cette équation. Donner également le schéma de Milstein $(\tilde{X}^N_{t_k})_{k\in\{0,\ldots,N\}}$.
3. Calculer $\mathbb{E}(X^N_T)$ et vérifier que $\mathbb{E}(X^N_T) = \mathbb{E}(X_T) + \mathcal{O}(\frac{1}{N})$. Commenter ce résultat.
4. On pose
$$A^N_k = \frac{X^N_{t_k}}{X_{t_{k+1}}}\left(e^{\sigma(W_{t_{k+1}}-W_{t_k})+(\mu-\frac{\sigma^2}{2})t_1} - 1 - \sigma(W_{t_{k+1}}-W_{t_k}) - \mu t_1\right).$$
Exprimer A^N_k à l'aide de $(\frac{X^N_{t_k}}{X_{t_k}}, \frac{X^N_{t_{k+1}}}{X_{t_{k+1}}})$ et en déduire que $X_T - X^N_T = X_T \sum_{k=0}^{N-1} A^N_k$.
5. On pose
$$R^N_k = e^{\sigma(W_{t_{k+1}}-W_{t_k})+(\mu-\frac{\sigma^2}{2})t_1} - 1 - \sigma(W_{t_{k+1}}-W_{t_k}) - (\mu - \frac{\sigma^2}{2})t_1 - \frac{\sigma^2}{2}(W_{t_{k+1}}-W_{t_k})^2$$
$$V^N_k = \left(\frac{X_{t_k}}{X_{t_{k+1}}} - 1\right)\left(e^{\sigma(W_{t_{k+1}}-W_{t_k})+(\mu-\frac{\sigma^2}{2})t_1} - 1 - \sigma(W_{t_{k+1}}-W_{t_k}) - \mu t_1\right)$$
$$Z^N_k = \frac{X_T}{X_{t_{k+1}}}\left(e^{\sigma(W_{t_{k+1}}-W_{t_k})+(\mu-\frac{\sigma^2}{2})t_1} - 1 - \sigma(W_{t_{k+1}}-W_{t_k}) - \mu t_1\right)(X^N_{t_k} - X_{t_k}).$$
Vérifier que
$$R^N_k + V^N_k + \frac{Z^N_k}{X_T} = A^N_k - \frac{\sigma^2}{2}((W_{t_{k+1}}-W_{t_k})^2 - t_1).$$
En déduire que
$$\sqrt{N}(X_T - X^N_T) = X_T\left(\frac{\sigma^2}{2}S_N + \sqrt{N}\sum_{k=0}^{N-1}(R^N_k + V^N_k)\right) + \sqrt{N}\sum_{k=0}^{N-1} Z^N_k$$
où $S_N = \sqrt{N}\sum_{k=0}^{N-1}[(W_{t_{k+1}}-W_{t_k})^2 - t_1]$.

On se donne $(G_k)_{k\geq 0}$ suite de variables gaussiennes centrées de variance T indépendantes de $(W_t)_{t\geq 0}$.

6. (a) Montrer que (W_T, S_N) et $(\mathcal{W}_N, \Sigma_N) = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}(G_k, G_k^2 - T)$ ont même loi.
   (b) Montrer que lorsque $N \to \infty$, $(\mathcal{W}_N, \Sigma_N)$ converge en loi vers un couple $(\mathcal{W}, \Sigma)$ dont on précisera la loi.
   (c) Conclure que $\frac{\sigma^2}{2}X_T S_N$ converge en loi vers $\sigma^2\sqrt{\frac{T}{2}}X_T G_0$.

D'après les questions 9 à 11, $\mathbb{E}\left|\sqrt{N}\sum_{k=0}^{N-1}(X_T R^N_k + X_T V^N_k + Z^N_k)\right| = \mathcal{O}(\frac{1}{\sqrt{N}})$.

7. En admettant provisoirement ce résultat, montrer que $\sqrt{N}(X_T - X^N_T)$ converge en loi vers $\sigma^2\sqrt{\frac{T}{2}}X_T G_0$.

Sous des hypothèses de régularité sur $\eta, b : \mathbb{R} \to \mathbb{R}$, Kurtz et Protter [46] ont montré que si $(X_t)_{t\in[0,T]}$ est solution de l'équation différentielle stochastique $dX_t = \eta(X_t)dW_t + b(X_t)dt$ et $(X^N_t)_{t\in[0,T]}$ désigne son schéma d'Euler en temps continu, alors le processus $(\sqrt{N}(X_t - X^N_t))_{t\in[0,T]}$ converge en loi vers $(Y_t)_{t\in[0,T]}$ solution de
$$Y_t = \int_0^t Y_s(\eta'(X_s)dW_s + b'(X_s)ds) + \sqrt{\frac{T}{2}}\int_0^t \eta\eta'(X_s)dB_s$$
où $(B_t)_{t\in[0,T]}$ est un mouvement brownien indépendant de $(W_t)_{t\in[0,T]}$.

8. Dans le cas particulier du modèle de Black-Scholes, calculer $d\frac{1}{X_t}$ puis $d\frac{Y_t}{X_t}$. En déduire Y_T et vérifier que la convergence établie à la question 7 est bien une conséquence du résultat général de Kurtz et Protter.
9. (a) Calculer $\mathbb{E}(R_0^N)$ et en déduire que $\mathbb{E}(R_k^N) = \mathcal{O}\left(\frac{1}{N^2}\right)$.
   Montrer que $R_0^N = \sigma\int_0^{t_1}(X_s - 1 - \sigma W_s)dW_s + \mu\int_0^{t_1}(X_s - 1)ds$. Calculer $\mathbb{E}((X_s - 1 - \sigma W_s)^2)$, vérifier que $\mathbb{E}((X_s - 1 - \sigma W_s)^2) = \mathcal{O}(s^2)$ pour $s \to 0^+$ et en déduire que $\mathbb{E}((R_0^N)^2) = \mathcal{O}\left(\frac{1}{N^3}\right)$.
   (b) Montrer que $\mathbb{E}\left(\left(\sum_{k=0}^{N-1} R_k^N\right)^2\right) = N(N-1)\mathbb{E}((R_0^N))^2 + N\mathbb{E}((R_0^N)^2)$.
   (c) Conclure que $\mathbb{E}\left|X_T\sqrt{N}\sum_{k=0}^{N-1} R_k^N\right| = \mathcal{O}(\frac{1}{\sqrt{N}})$.

On pose $D_k^N = e^{\sigma(W_{t_{k+1}} - W_{t_k}) + (\mu - \frac{\sigma^2}{2})t_1} - 1 - \sigma(W_{t_{k+1}} - W_{t_k}) - \mu t_1$.

10. (a) Calculer $\mathbb{E}(V_0^N)$ et vérifier que $\mathbb{E}(V_0^N) = \mathcal{O}(\frac{1}{N^2})$.
   (b) Rappeler l'équation différentielle stochastique satisfaite par $\frac{1}{X_t}$ et vérifier que $\mathbb{E}\left(\left(\frac{1}{X_{t_1}} - 1\right)^4\right) = \mathcal{O}(\frac{1}{N^2})$.
   En remarquant que $D_0^N = \int_0^{t_1}(X_s - 1)(\sigma dW_s + \mu ds)$, vérifier que $\mathbb{E}((D_0^N)^4) = \mathcal{O}(\frac{1}{N^4})$. En déduire que $\mathbb{E}((V_0^N)^2) = \mathcal{O}(\frac{1}{N^3})$.
   (c) Conclure que $\mathbb{E}\left|X_T\sqrt{N}\sum_{k=0}^{N-1} V_k^N\right| = \mathcal{O}(\frac{1}{\sqrt{N}})$.
11. (a) Vérifier que $\mathbb{E}((D_k^N)^2) = \mathcal{O}(\frac{1}{N^2})$.
   (b) À l'aide de la vitesse forte, en déduire $\mathbb{E}((Z_k^N)^2) = \mathcal{O}(\frac{1}{N^3})$.
   (c) Pour $0 \leq l < k \leq N-1$, vérifier que
   $$\mathbb{E}(Z_k^N Z_l^N) = \mathbb{E}(X_{T-t_{k+1}}^2)\mathbb{E}\left(\frac{X_{t_{k+1}}}{X_{t_k}}D_k^N\right)\mathbb{E}\left((X_{t_k}^N - X_{t_k})\frac{X_{t_k}}{X_{t_{l+1}}}D_l^N(X_{t_l}^N - X_{t_l})\right).$$
   Vérifier que $\mathbb{E}\left(\frac{X_{t_{k+1}}}{X_{t_k}}D_k^N\right) = e^{\mu t_1}\left(e^{(\mu+\sigma^2)t_1} - 1 - (\mu+\sigma^2)t_1\right)$. Remarquer que
   $$\mathbb{E}\left|(X_{t_k}^N - X_{t_k})\frac{X_{t_k}}{X_{t_{l+1}}}D_l^N(X_{t_l}^N - X_{t_l})\right| \leq \sqrt{\mathbb{E}\left(\sup_{t\leq T}(X_t^N - X_t)^4\right)\mathbb{E}(X_{t_k - t_{l+1}}^2)\mathbb{E}((D_l^N)^2)}$$
   et en déduire que $\mathbb{E}(Z_k^N Z_l^N) = \mathcal{O}(\frac{1}{N^4})$.
   (d) Conclure que $\mathbb{E}((\sqrt{N}\sum_{k=0}^{N-1} Z_k^N)^2) = \mathcal{O}(\frac{1}{N})$ et que
   $$\mathbb{E}\left|\sqrt{N}\sum_{k=0}^{N-1}(X_T R_k^N + X_T V_k^N + Z_k^N)\right| = \mathcal{O}(\frac{1}{\sqrt{N}}).$$

3.8.6 Vitesse forte du schéma de Milstein

On s'intéresse à la discrétisation de l'EDS uni-dimensionnelle
$$\begin{cases} dX_t = \sigma(X_t)dW_t + b(X_t)dt \\ X_0 = y \end{cases} \tag{3.37}$$
où $(W_t)_{t\geq 0}$ est un mouvement brownien réel de filtration naturelle $(\mathcal{F}_t)_{t\geq 0}$, $\sigma, b : \mathbb{R} \to \mathbb{R}$ sont des fonctions C^2 avec des dérivées premières et secondes bornées et $y \in \mathbb{R}$. On se

donne un horizon T et un nombre $N \in \mathbb{N}^*$ de pas de discrétisation. Pour $k \in \{0, \ldots, N\}$, on pose $t_k = \frac{kT}{N}$. Pour $t \in [0,T]$, on note respectivement $\tau_t = \lfloor \frac{Nt}{T} \rfloor \frac{T}{N}$ et $\bar{\tau}_t = \lceil \frac{Nt}{T} \rceil \frac{T}{N}$ l'instant de discrétisation juste avant t et l'instant de discrétisation juste après t.

1. Que peut-on dire de $\sup_{t\in[0,T]} \mathbb{E}(X_t^4)$? Et de $\sup_{0\leq s<t\leq T} \frac{\mathbb{E}[(X_t-X_s)^2]}{t-s}$?
2. Écrire le schéma de Milstein $(\tilde{X}_{t_k}^N)_{k\in\{0,\ldots,N\}}$ associé à cette équation. Donner également son extension en temps continu $(\tilde{X}_t^N)_{t\in[0,T]}$.
3. Vérifier que
$$\forall u \in [0,T],\ \tilde{X}_u^N = y + \int_0^u \left(\sigma(\tilde{X}_{\tau_s}^N) + \sigma\sigma'(\tilde{X}_{\tau_s}^N)(W_s - W_{\tau_s})\right) dW_s + \int_0^u b(\tilde{X}_{\tau_s}^N) ds.$$
4. En déduire que pour $t \in [0,T]$,
$$\begin{aligned}\mathbb{E}\left[\sup_{u\in[0,t]} (X_u - \tilde{X}_u^N)^2\right] \leq 3\bigg(4\int_0^t \mathbb{E}\left[\left(\sigma(X_s) - \sigma(\tilde{X}_{\tau_s}^N) - \sigma\sigma'(\tilde{X}_{\tau_s}^N)(W_s - W_{\tau_s})\right)^2\right] ds \\ + \mathbb{E}\left[\sup_{u\in[0,T]} \left(\int_0^u (b(X_s) - b(X_{\tau_s}))ds\right)^2\right] + t\int_0^t \mathbb{E}[(b(X_{\tau_s}) - b(\tilde{X}_{\tau_s}^N))^2] ds\bigg).\end{aligned} \tag{3.38}$$

Dans toute la suite $C \in]0, +\infty[$ désigne une constante qui ne dépend pas de N et peut changer d'une ligne à l'autre.

5. Majoration du second terme du second membre de (3.38) :
   (a) Calculer $db(X_r)$ et en déduire que pour $u \in [0,T]$
   $$\int_0^{\tau_u} (b(X_s) - b(X_{\tau_s}))ds = \int_0^{\tau_u} (\bar{\tau}_r - r)\left(\sigma b'(X_r)dW_r + [bb' + \frac{\sigma^2 b''}{2}](X_r)dr\right)$$
   puis que $\mathbb{E}\left[\sup_{u\in[0,T]} \left(\int_0^{\tau_u} (b(X_s) - b(X_{\tau_s}))ds\right)^2\right] \leq \frac{C}{N^2}$.
   (b) Vérifier que
   $$\begin{aligned}\sup_{u\in[0,T]} \left(\int_{\tau_u}^u |b(X_s) - b(X_{\tau_s})|ds\right)^2 &\leq \frac{T}{N} \max_{0\leq k\leq N-1} \int_{t_k}^{t_{k+1}} (b(X_s) - b(X_{\tau_s}))^2 ds \\ &\leq \frac{T}{N} \int_0^T (b(X_s) - b(X_{\tau_s}))^2 ds,\end{aligned}$$
   et conclure que $\mathbb{E}\left[\sup_{u\in[0,T]} \left(\int_0^u (b(X_s) - b(X_{\tau_s}))ds\right)^2\right] \leq \frac{C}{N^2}$.
6. Majoration du premier terme du second membre de (3.38) **en supposant $\sigma\sigma'$ lipschitzienne** :
   (a) Calculer $d\sigma(X_r)$ par la formule d'Itô.
   (b) En déduire que pour $s \in [0,T]$,
   $$\begin{aligned}\sigma(X_s) - \sigma(X_{\tau_s}) - \sigma\sigma'(X_{\tau_s})(W_s - W_{\tau_s}) = &\int_{\tau_s}^s (\sigma\sigma'(X_r) - \sigma\sigma'(X_{\tau_s}))dW_r \\ &+ \int_{\tau_s}^s [b\sigma' + \frac{\sigma^2\sigma''}{2}](X_r)dr,\end{aligned}$$
   puis que $\forall s \in [0,T],\ \mathbb{E}\left[(\sigma(X_s) - \sigma(X_{\tau_s}) - \sigma\sigma'(X_{\tau_s})(W_s - W_{\tau_s}))^2\right] \leq \frac{C}{N^2}$.

(c) Conclure que pour tout $s \in [0,T]$,

$$\mathbb{E}\left[\left(\sigma(X_s) - \sigma(\tilde{X}^N_{\tau_s}) - \sigma\sigma'(\tilde{X}^N_{\tau_s})(W_s - W_{\tau_s})\right)^2\right] \leq C\left(\mathbb{E}[(X_{\tau_s} - \tilde{X}^N_{\tau_s})^2] + \frac{1}{N^2}\right).$$

7. En supposant $\sigma\sigma'$ lipschitzienne, vérifier que

$$\forall t \in [0,T],\ \mathbb{E}\left[\sup_{u\in[0,t]}(X_u - \tilde{X}^N_u)^2\right] \leq C\left(\int_0^t \mathbb{E}\left[\sup_{u\in[0,s]}(X_u - \tilde{X}^N_u)^2\right]ds + \frac{1}{N^2}\right)$$

et conclure que $\mathbb{E}\left[\sup_{u\in[0,T]}(X_u - \tilde{X}^N_u)^2\right] \leq \frac{C}{N^2}$ si $\int_0^T \mathbb{E}\left[\sup_{u\in[0,t]}(X_u - \tilde{X}^N_u)^2\right]dt < +\infty$. Comment se passer de cette hypothèse d'intégrabilité ?

8. On ne suppose plus que la fonction $\sigma\sigma'$ est lipschitzienne.
   (a) Vérifier que pour $r \in [0,T]$, $\mathbb{E}[(\sigma\sigma'(X_r) - \sigma\sigma'(X_{\tau_r}))^2]$ est majoré par

   $$2\left(\|\sigma'\|^2_\infty \mathbb{E}[(\sigma(X_r) - \sigma(X_{\tau_r}))^2] + \|\sigma''\|^2_\infty \mathbb{E}[\sigma^2(X_{\tau_r})\mathbb{E}[(X_r - X_{\tau_r})^2|\mathcal{F}_{\tau_r}]]\right),$$

   puis par $\frac{C}{N}$.
   (b) Pour $s \in [0,T]$, majorer $\mathbb{E}[(\sigma\sigma'(X_{\tau_s}) - \sigma\sigma'(\tilde{X}^N_{\tau_s}))^2(W_s - W_{\tau_s})^2]$ par

   $$\frac{2T}{N}\left(\|\sigma'\|^2_\infty \mathbb{E}[(\sigma(X_{\tau_s}) - \sigma(\tilde{X}^N_{\tau_s}))^2] + 2\|\sigma'\|_\infty \mathbb{E}^{1/2}[\sigma^4(X_{\tau_s})]\mathbb{E}^{1/2}[(\sigma'(X_{\tau_s}) - \sigma'(\tilde{X}^N_{\tau_s}))^2]\right)$$

   puis par $C\left(\mathbb{E}[(X_{\tau_s} - \tilde{X}^N_{\tau_s})^2] + \frac{1}{N^2}\right)$ (*on pourra utiliser que* $\forall a,b \in \mathbb{R}$, $ab \leq \frac{a^2+b^2}{2}$).
   (c) Conclure que le contrôle de la vitesse forte quadratique obtenu à la question 7 reste vrai.

9. Expliquer comment généraliser à toute constante $p \geq 1$ l'analyse de l'erreur forte $\mathbb{E}[\sup_{u\in[0,T]}|X_u - \tilde{X}^N_u|^{2p}]$ effectuée ci-dessus dans le cas particulier $p = 1$.

3.8.7 Méthode de Monte Carlo multipas et schéma de Giles et Szpruch

On s'intéresse au calcul de $\mathbb{E}[f(X_T)]$ où $f : \mathbb{R}^n \to \mathbb{R}$ est une fonction C^4 **lipschitzienne** avec des dérivées à croissance polynomiale et $(X_t)_{t\in[0,T]}$ est la solution de l'EDS

$$\begin{cases} dX_t = \sigma(X_t)dW_t + b(X_t)dt \\ X_0 = x_0 \end{cases}$$

où $(W_t)_{t\geq 0}$ est un mouvement brownien à valeurs $\mathbb{R}^d$, $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$ et $b : \mathbb{R}^n \to \mathbb{R}^n$ sont des fonctions C^4 à dérivées bornées et $x_0 \in \mathbb{R}^n$. On note X^N_T l'approximation de X_T obtenue par un schéma de discrétisation comportant N pas de discrétisation et **dont le temps de calcul est proportionnel à** N. On suppose que les ordres de convergence forte et de convergence faible pour la fonction f de ce schéma sont respectivement égaux à $\alpha/2$ et β avec $\alpha, \beta \geq 1$:

$$\exists C < +\infty,\ \forall N \in \mathbb{N}^*,\ \mathbb{E}[|X_T - X^N_T|^2] \leq \frac{C}{N^\alpha} \text{ et } |\mathbb{E}[f(X_T) - f(X^N_T)]| \leq \frac{C}{N^\beta}.$$

Méthode de Monte Carlo multipas

1. Soit Y un estimateur de $\mathbb{E}[f(X_T)]$ de carré intégrable. Montrer la décomposition biais/variance
$$\mathbb{E}[(\mathbb{E}(f(X_T)) - Y)^2] = (\mathbb{E}[f(X_T) - Y])^2 + \mathrm{Var}\,(Y)$$
de l'erreur quadratique.
2. On suppose que $Y = \frac{1}{M}\sum_{i=1}^{M} f(X_T^{i,N})$ est la moyenne empirique de M copies indépendantes de $f(X_T^N)$.
   (a) Combien de pas N faut-il choisir pour que $(\mathbb{E}[f(X_T) - Y])^2$ soit égal à ε^2 où ε est un niveau de précision donné petit ?
   (b) Vérifier que $\lim_{N\to\infty} \mathbb{E}[(f(X_T) - f(X_T^N))^2] = 0$.
   En déduire que $\lim_{N\to\infty} \mathrm{Var}\,(f(X_T^N)) = \mathrm{Var}\,(f(X_T))$. Combien de copies M faut-il choisir pour que $\mathrm{Var}\,(Y)$ soit égale à ε^2 ?
   (c) Conclure que le temps de calcul nécessaire pour atteindre la précision ε (i.e. une erreur quadratique égale à $2\varepsilon^2$) est proportionnel à $\varepsilon^{-(2+1/\beta)}$?
3. On s'intéresse maintenant à l'estimateur multipas proposé par Giles [28] $Y = \sum_{l=0}^{L} Y_{l,M_l}$ où les variables Y_{l,M_l} sont indépendantes et
   — Y_{0,M_0} est la moyenne empirique de M_0 copies indépendantes de $f(X_T^1)$,
   — pour $l \in \{1, \dots, L\}$, Y_{l,M_l} est la moyenne empirique de M_l copies indépendantes de $f(X_T^{2^l}) - f(X_T^{2^{l-1}})$.

   (a) Vérifier que $\mathbb{E}[Y] = \mathbb{E}[f(X_T^{2^L})]$ et en déduire que
$$\exists C > 0, \forall \varepsilon \in]0,1[,\ L \geq -\frac{\log_2(\varepsilon/C)}{\beta} \Rightarrow |\mathbb{E}[f(X_T) - Y]| \leq \varepsilon.$$
   (b) En remarquant que pour $l \in \mathbb{N}^*$,
$$(f(X_T^{2^l}) - f(X_T^{2^{l-1}}))^2 \leq 2\left((f(X_T^{2^l}) - f(X_T))^2 + (f(X_T) - f(X_T^{2^{l-1}}))^2\right),$$
   vérifier que $\sup_{l\in\mathbb{N}^*} 2^{l\alpha}\mathbb{E}[(f(X_T^{2^l}) - f(X_T^{2^{l-1}}))^2] < +\infty$. En déduire
$$\exists \tilde{C} < +\infty \text{ t.q. } \forall L \in \mathbb{N}^*,\ \forall (M_0, \dots, M_L) \in {\mathbb{N}^*}^L,\ \mathrm{Var}\,(Y) \leq \tilde{C}\sum_{l=0}^{L} \frac{1}{M_l 2^{l\alpha}}.$$

   On pose $p_l = \frac{\tilde{C}}{\varepsilon^2 M_l 2^{l\alpha}}$ pour $l \in \{0, \dots, L\}$.
   (c) Vérifier que $\sum_{l=0}^{L} p_l \leq 1 \Rightarrow \mathrm{Var}\,(Y) \leq \varepsilon^2$.
   Justifier que le temps de calcul de Y est proportionnel à $\tau = \sum_{l=0}^{L} M_l 2^l$.

   Pour approcher la minimisation du temps de calcul sous la contrainte $\mathrm{Var}\,(Y) \leq \varepsilon^2$, nous allons minimiser τ sous la contrainte $\sum_{l=0}^{L} p_l \leq 1$.
   (d) Vérifier que $\tau = \frac{\varepsilon^2}{\tilde{C}}\left(\sum_{l=0}^{L}(M_l 2^{l(1+\alpha)/2})^2 p_l\right)$ et en déduire à l'aide de l'inégalité de Cauchy-Schwarz que
$$\tau \geq \frac{\tilde{C}}{\varepsilon^2}\left(\sum_{l=0}^{L} 2^{l(1-\alpha)/2}\right)^2.$$
   Vérifier que $\left(M_l = \frac{\tilde{C}\sum_{k=0}^{L} 2^{k(1-\alpha)/2}}{\varepsilon^2 2^{l(1+\alpha)/2}}\right)_{0\leq l\leq L}$ atteint ce minorant et satisfait la contrainte.

(e) En déduire que le temps de calcul pour atteindre la précision ε est d'ordre $\frac{(\log_2(\varepsilon))^2}{\varepsilon^2}$ si $\alpha = 1$ et $\frac{1}{\varepsilon^2}$ si $\alpha > 1$.

4. Comparer les temps de calcul de l'estimateur monopas et de l'estimateur multipas pour le schéma d'Euler.

Pour $j \in \{1, \dots, d\}$ et $x \in \mathbb{R}^n$, on note $\sigma_j(x) \in \mathbb{R}^n$ la j-ème colonne de la matrice $\sigma(x)$ et $\partial\sigma_j \in \mathbb{R}^{n\times n}$ la matrice $\left(\frac{\partial \sigma_{ij}}{\partial x_l}(x)\right)_{il}$.

5. Comparer les temps de calcul des deux estimateurs pour le schéma de Milstein (pour lequel $\beta = 1$). Sous quelle condition peut-on effectivement implémenter le schéma de Milstein ?

Schéma de Giles et Szpruch et transpositions d'accroissements browniens
Pour $j, m \in \{1, \dots, d\}$, on définit $\theta_{jm}, \eta_{jm} : \mathbb{R}^n \to \mathbb{R}^n$ par $\theta_{jm}(x) = \frac{\partial\sigma_j\sigma_m + \partial\sigma_m\sigma_j}{4}(x)$ et $\eta_{jm}(x) = \frac{\partial\sigma_j\sigma_m - \partial\sigma_m\sigma_j}{4}(x)$. On suppose que les fonctions θ_{jm} sont **lipschitziennes**. On pose

$$\hat{X}_t = x_0 + \sigma(x_0)W_t + b(x_0)t + \sum_{j,m=1}^{d} \theta_{jm}(x_0)W_t^j W_t^m$$

$$\tilde{X}_t = \hat{X}_{\frac{t}{2}} + \sigma(\hat{X}_{\frac{t}{2}})(W_t - W_{\frac{t}{2}}) + b(\hat{X}_{\frac{t}{2}})\frac{t}{2} + \sum_{j,m=1}^{d} \theta_{jm}(\hat{X}_{\frac{t}{2}})(W_t^j - W_{\frac{t}{2}}^j)(W_t^m - W_{\frac{t}{2}}^m)$$

6. Vérifier que
$$\exists C < +\infty,\ \forall t \in [0,T],\ \mathbb{E}[|\hat{X}_t - x_0 - \sigma(x_0)W_t|^2] \leq Ct^2 \text{ et } \mathbb{E}[|\hat{X}_t - x_0|^2] \leq Ct.$$

7. Montrer que
$$\mathbb{E}\left[\left|(\theta_{jm}(\hat{X}_{\frac{t}{2}}) - \theta_{jm}(x_0))(W_t^j - W_{\frac{t}{2}}^j)(W_t^m - W_{\frac{t}{2}}^m)\right|^2\right] = \frac{t^2(1 + 2\times 1_{\{j=m\}})}{4} \times \mathbb{E}\left[\left|\theta_{jm}(\hat{X}_{\frac{t}{2}}) - \theta_{jm}(x_0)\right|^2\right].$$
En déduire l'existence d'une constante $C < +\infty$ telle que pour tout $t \in [0,T]$,
$$\mathbb{E}\left[\left|b(\hat{X}_{\frac{t}{2}})\frac{t}{2} + \sum_{j,m=1}^{d} \theta_{jm}(\hat{X}_{\frac{t}{2}})(W_t^j - W_{\frac{t}{2}}^j)(W_t^m - W_{\frac{t}{2}}^m) - b(x_0)\frac{t}{2} - \sum_{j,m=1}^{d} \theta_{jm}(x_0)(W_t^j - W_{\frac{t}{2}}^j)(W_t^m - W_{\frac{t}{2}}^m)\right|^2\right] \leq Ct^3.$$

8. Pour $\varphi : \mathbb{R}^n \to \mathbb{R}$ fonction C^1 telle que φ et $\nabla\varphi$ sont Lipschitziennes, en remarquant l'existence d'une variable aléatoire α_t à valeurs dans $[0,1]$ t.q.
$$\varphi(x_0+\sigma(x_0)W_t)-\varphi(x_0)-\nabla\varphi(x_0).\sigma(x_0)W_t = \left(\nabla\varphi(x_0 + \sigma(x_0)\alpha_t W_t) - \nabla\varphi(x_0)\right).\sigma(x_0)W_t,$$
montrer
$$\exists C < +\infty,\ \forall t \in [0,T],\ \mathbb{E}[(\varphi(\hat{X}_t) - \varphi(x_0) - \nabla\varphi(x_0).\sigma(x_0)W_t)^2] \leq Ct^2.$$
En déduire l'existence d'une constante $C < +\infty$ telle que pour tout $t \in [0,T]$,
$$\mathbb{E}\left[\left|(\sigma(\hat{X}_{\frac{t}{2}}) - \sigma(x_0))(W_t - W_{\frac{t}{2}}) - \sum_{j,m=1}^{d} \partial\sigma_j\sigma_m(x_0)W_{\frac{t}{2}}^m(W_t^j - W_{\frac{t}{2}}^j)\right|^2\right] \leq Ct^3.$$

9. En utilisant les propriétés de symétrie de θ et η, vérifier que

$$\begin{aligned}\sum_{j,m=1}^{d} \partial\sigma_j\sigma_m(x_0)W^m_{\frac{t}{2}}(W^j_t - W^j_{\frac{t}{2}}) = &\sum_{j,m=1}^{d} \theta_{jm}(x_0)\left[W^m_{\frac{t}{2}}(W^j_t - W^j_{\frac{t}{2}}) + W^j_{\frac{t}{2}}(W^m_t - W^m_{\frac{t}{2}})\right] \\ &+ \sum_{j,m=1}^{d} \eta_{jm}(x_0)\left[W^m_{\frac{t}{2}}(W^j_t - W^j_{\frac{t}{2}}) - W^j_{\frac{t}{2}}(W^m_t - W^m_{\frac{t}{2}})\right]\end{aligned}$$

10. Que vaut $W^j_{\frac{t}{2}}W^m_{\frac{t}{2}} + W^m_{\frac{t}{2}}(W^j_t - W^j_{\frac{t}{2}}) + W^j_{\frac{t}{2}}(W^m_t - W^m_{\frac{t}{2}}) + (W^j_t - W^j_{\frac{t}{2}})(W^m_t - W^m_{\frac{t}{2}})$?
11. En déduire l'existence d'une constante $C < +\infty$ telle que $\forall t \in [0, T]$,

$$\mathbb{E}\left[\left|\tilde{X}_t - \hat{X}_t - \sum_{j,m=1}^{d} \eta_{jm}(x_0)[W^m_{\frac{t}{2}}(W^j_t - W^j_{\frac{t}{2}}) - W^j_{\frac{t}{2}}(W^m_t - W^m_{\frac{t}{2}})]\right|^2\right] \leq Ct^3.$$

On pose

$$\bar{X}_{\frac{t}{2}} = x_0 + \sigma(x_0)(W_t - W_{\frac{t}{2}}) + b(x_0)\frac{t}{2} + \sum_{j,m=1}^{d} \theta_{jm}(x_0)(W^j_t - W^j_{\frac{t}{2}})(W^m_t - W^m_{\frac{t}{2}})$$

$$\bar{X}_t = \bar{X}_{\frac{t}{2}} + \sigma(\bar{X}_{\frac{t}{2}})W_{\frac{t}{2}} + b(\bar{X}_{\frac{t}{2}})\frac{t}{2} + \sum_{j,m=1}^{d} \theta_{jm}(\bar{X}_{\frac{t}{2}})W^j_{\frac{t}{2}}W^m_{\frac{t}{2}}$$

12. Vérifier que $\bar{X}_t$ a même loi que $\tilde{X}_t$ et montrer sans calcul l'existence d'une constante $C < +\infty$ telle que $\forall t \in [0, T]$,

$$\mathbb{E}\left[\left|\bar{X}_t - \hat{X}_t - \sum_{j,m=1}^{d} \eta_{jm}(x_0)[(W^m_t - W^m_{\frac{t}{2}})W^j_{\frac{t}{2}} - (W^j_t - W^j_{\frac{t}{2}})W^m_{\frac{t}{2}}]\right|^2\right] \leq Ct^3.$$

13. Conclure que

$$\exists C < +\infty,\ \forall t \in [0, T],\ \mathbb{E}\left[\left|\hat{X}_t - \frac{\tilde{X}_t + \bar{X}_t}{2}\right|^2\right] \leq Ct^3.$$

On note maintenant $(\hat{X}^N)$ le schéma défini par $\hat{X}^N_0 = x_0$ et la relation de récurrence : $\forall k \in \{0, \ldots, N-1\}$,

$$\hat{X}^N_{t_{k+1}} = \hat{X}^N_{t_k} + \sigma(\hat{X}^N_{t_k})(W_{t_{k+1}} - W_{t_k}) + b(\hat{X}^N_{t_{k+1}})\frac{T}{N} + \sum_{j,m=1}^{d} \theta_{jm}(\hat{X}^N_{t_k})(W^m_{t_{k+1}} - W^m_{t_k})(W^j_{t_{k+1}} - W^j_{t_k}).$$

Soit $(\bar{X}^{2N})$ le schéma défini comme $(\hat{X}^{2N})$ mais en transposant chaque paire d'accroissements browniens successifs i.e. en remplaçant

$$(W_{\frac{T}{2N}}, W_{\frac{2T}{2N}} - W_{\frac{T}{2N}}, W_{\frac{3T}{2N}} - W_{\frac{2T}{2N}}, W_{\frac{4T}{2N}} - W_{\frac{3T}{2N}}, \ldots, W_{\frac{(2N-1)T}{2N}} - W_{\frac{(2N-2)T}{2N}}, W_{\frac{2NT}{2N}} - W_{\frac{(2N-1)T}{2N}})$$

par $(W_{\frac{2T}{2N}} - W_{\frac{T}{2N}}, W_{\frac{T}{2N}}, W_{\frac{4T}{2N}} - W_{\frac{3T}{2N}}, W_{\frac{3T}{2N}} - W_{\frac{2T}{2N}}, \ldots, W_{\frac{2NT}{2N}} - W_{\frac{(2N-1)T}{2N}}, W_{\frac{(2N-1)T}{2N}} - W_{\frac{(2N-2)T}{2N}})$.

14. Pourquoi peut-on anticiper le résultat prouvé par Giles et Szpruch [29] [2]

$$\exists C < +\infty,\ \forall N \in \mathbb{N}^*,\ \mathbb{E}\left[\left|\hat{X}_T^N - \frac{\hat{X}_T^{2N} + \bar{X}_T^{2N}}{2}\right|^2\right] \leq \frac{C}{N^2}$$

$$\mathbb{E}\left[\left|\hat{X}_T^N - \hat{X}_T^{2N}\right|^2 + \left|\hat{X}_T^N - \bar{X}_T^{2N}\right|^2\right] \leq \frac{C}{N}?$$

15. On suppose que les dérivées d'ordre 2 de f sont bornées. Montrer que

$$\exists C < +\infty,\ \forall x, y \in \mathbb{R}^n,\ \left|\frac{f(x)+f(y)}{2} - f\left(\frac{x+y}{2}\right)\right| \leq C|x-y|^2. \quad (3.39)$$

En déduire que $\mathbb{E}\left[\left(f(\hat{X}_T^N) - \frac{f(\hat{X}_T^{2N})+f(\bar{X}_T^{2N})}{2}\right)^2\right] \leq \frac{C}{N^2}$.

On admet que $\beta = 1$ pour le schéma $\hat{X}^N$. Quel est l'ordre du temps de calcul nécessaire pour atteindre la précision ε à l'estimateur multipas $Y = \sum_{l=0}^L Y_{l,M_l}$ de $\mathbb{E}[f(X_T)]$ où les variables Y_{l,M_l} sont indépendantes et

— Y_{0,M_0} est la moyenne empirique de M_0 copies indépendantes de $f(\hat{X}_T^1)$,

— pour $l \in \{1, \ldots, L\}$, Y_{l,M_l} est la moyenne empirique de M_l copies indépendantes de $\frac{f(\hat{X}_T^{2^l})+f(\bar{X}_T^{2^l})}{2} - f(\hat{X}_T^{2^{l-1}})$?

3.8.8 Discrétisation d'un modèle à volatilité stochastique

On considère le modèle à volatilité stochastique

$$\begin{cases} dS_t = rS_t dt + f(Y_t)S_t(\rho dW_t + \sqrt{1-\rho^2}dB_t); \quad S_0 = s_0 > 0 \\ dY_t = b(Y_t)dt + \sigma(Y_t)dW_t; \quad Y_0 = y_0 \end{cases}, \quad (3.40)$$

où $(B_t)_{t\geq 0}$ et $(W_t)_{t\geq 0}$ sont deux mouvements browniens indépendants, $r \in \mathbb{R}$, $\rho \in [-1,1]$ et $f, b, \sigma : \mathbb{R} \to \mathbb{R}$ sont des fonctions régulières (c'est-à-dire bornées et dérivables autant de fois qu'on le souhaite avec des dérivées bornées) avec σ qui ne s'annule pas.
On pose $X_t = \ln(S_t)$, $t_k = k\Delta t$ avec $\Delta t = \frac{T}{N}$ où $N \in \mathbb{N}^*$ est un nombre de pas de temps et $T > 0$ une maturité.

1. Que représente le coefficient ρ ?
2. Écrire l'équation différentielle stochastique satisfaite par le couple $(X_t, Y_t)_{t\geq 0}$.
3. Comment s'écrit le schéma de Milstein $(\bar{Y}_{t_k}^N)_{0\leq k\leq N}$ de pas Δt pour le processus $(Y_t)_{t\in[0,T]}$? Rappeler sans preuve le résultat de convergence forte du cours pour ce schéma.
4. Déterminer la condition de commutativité qui permet d'implémenter le schéma de Milstein pour le couple $(X_t, Y_t)_{t\in[0,T]}$. Que peut-on dire de la volatilité de $(S_t)_{t\geq 0}$ lorsque la fonction σ est nulle ? Et lorsque f' est nulle ? Écrire le schéma de Milstein de pas Δt pour le couple $(X_t, Y_t)_{t\in[0,T]}$ lorsque $|\rho| = 1$.

2. Notons que l'estimation (3.39) appliquée à $f = \sigma$ permet de contrôler l'erreur liée au terme d'ordre principal des schémas i.e. le terme linéaire en l'accroissement brownien.

5. On note $\boxed{F(y) = \int_{y_0}^{y} \frac{f}{\sigma}(z)dz}$, fonction que l'on suppose régulière dans la suite. Vérifier que
$$dX_t = \rho dF(Y_t) + \sqrt{1-\rho^2} f(Y_t)dB_t + h(Y_t)dt,$$
pour une fonction h que l'on précisera.

On définit par récurrence $(\bar{X}_{t_k})_{0\leq k\leq N}$ en posant $\bar{X}_0 = \ln(s_0)$ et $\forall k \in \{0,\dots,N-1\}$,
$$\bar{X}_{t_{k+1}} = \bar{X}_{t_k} + \rho(F(Y_{t_{k+1}}) - F(Y_{t_k})) + \int_{t_k}^{t_{k+1}} h(Y_s)ds + \sqrt{\frac{1-\rho^2}{\Delta t}\int_{t_k}^{t_{k+1}} f^2(Y_s)ds}(B_{t_{k+1}} - B_{t_k}).$$

6. Vérifier que les vecteurs aléatoires $(\sum_{j=0}^{k-1}\sqrt{\frac{1}{\Delta t}\int_{t_j}^{t_{j+1}} f^2(Y_s)ds}(B_{t_{j+1}} - B_{t_j}))_{1\leq k\leq N}$ et $(\int_0^{t_k} f(Y_s)dB_s)_{1\leq k\leq N}$ ont même loi conditionnelle sachant $(W_t)_{t\in[0,T]}$. En déduire que les vecteurs $(\bar{X}_{t_k})_{0\leq k\leq N}$ et $(X_{t_k})_{0\leq k\leq N}$ ont même loi.

Le schéma proposé et étudié dans [41] est le suivant : $\bar{X}_0^N = \ln(s_0)$ et pour $k \in \{0,\dots,N-1\}$,
$$\bar{X}_{t_{k+1}}^N = \bar{X}_{t_k}^N + \rho\left(F(\bar{Y}_{t_{k+1}}^N) - F(\bar{Y}_{t_k}^N)\right) + h(\bar{Y}_{t_k}^N)\Delta t + \sqrt{1-\rho^2}\eta_k^N \times (B_{t_{k+1}} - B_{t_k})$$
$$\text{où } \eta_k^N = \sqrt{\left(\psi(\bar{Y}_{t_k}^N) + \frac{\sigma\psi'(\bar{Y}_{t_k}^N)}{\Delta t}\int_{t_k}^{t_{k+1}}(W_s - W_{t_k})ds\right) \vee \underline{\psi}}$$
avec $\psi(y) = f^2(y)$ et $\underline{\psi} = \inf_{y\in\mathbb{R}} \psi(y) = \inf_{y\in\mathbb{R}} f^2(y) \geq 0$. L'objectif final de cet énoncé est de montrer que si $\underline{\psi} > 0$ alors
$$\exists C > 0,\ \forall N \in \mathbb{N}^*,\ \mathbb{E}\left(\max_{0\leq k\leq N} |\bar{X}_{t_k} - \bar{X}_{t_k}^N|^2\right) \leq \frac{C}{N^2} \tag{3.41}$$

Dans le reste de l'énoncé on notera C des constantes indépendantes de N qui peuvent varier de ligne en ligne.

7. Comment peut-on simuler le vecteur $(W_{t_{k+1}} - W_{t_k}, \int_{t_k}^{t_{k+1}}(W_s - W_{t_k})ds)_{0\leq k\leq N-1}$? En quoi, pour le pricing d'options exotiques, le schéma $(\bar{X}_{t_k}^N)_{0\leq k\leq N}$ est-il aussi performant que le schéma de Milstein ?
8. Vérifier que
$$\max_{0\leq k\leq N} |\bar{X}_{t_k} - \bar{X}_{t_k}^N|^2 \leq C\left(T_1^N + \max_{1\leq k\leq N}(T_{2,k}^N)^2 + \max_{1\leq k\leq N}(T_{3,k}^N)^2\right)$$
$$\text{où } T_1^N = \max_{0\leq k\leq N}(F(Y_{t_k}) - F(\bar{Y}_{t_k}^N))^2 + \max_{1\leq k\leq N}\left(\Delta t\sum_{j=0}^{k-1}(h(Y_{t_j}) - h(\bar{Y}_{t_j}^N))\right)^2,$$
$$T_{2,k}^N = \int_0^{t_k} h(Y_s)ds - \Delta t\sum_{j=0}^{k-1} h(Y_{t_j}),$$
$$T_{3,k}^N = \sum_{j=0}^{k-1}\left(\eta_j^N - \sqrt{\frac{1}{\Delta t}\int_{t_j}^{t_{j+1}}\psi(Y_s)ds}\right)(B_{t_{j+1}} - B_{t_j}).$$
9. Vérifier que $\max_{1\leq k\leq N}\left(\frac{1}{N}\sum_{j=0}^{k-1}(h(Y_{t_j}) - h(\bar{Y}_{t_j}^N))\right)^2 \leq \max_{1\leq k\leq N-1}(h(Y_{t_k}) - h(\bar{Y}_{t_k}^N))^2$ et en déduire que $\mathbb{E}(T_1^N) \leq \frac{C}{N^2}$?

10. Vérifier que $T_{2,k}^N = \int_0^{t_k} (\bar{\tau}_s - s) \left((bh' + \frac{\sigma^2 h''}{2})(Y_s)ds + \sigma h'(Y_s)dW_s \right)$ où, pour $s \in [0,T]$, $\bar{\tau}_s = \lceil \frac{s}{\Delta t} \rceil \Delta t$ désigne l'instant de discrétisation juste après s. En déduire que $\mathbb{E}(\max_{1 \leq k \leq N} (T_{2,k}^N)^2) \leq \frac{C}{N^2}$.
11. (a) Que peut-on dire du processus $(T_{3,k}^N)_{1 \leq k \leq N}$? En utilisant l'inégalité de Doob et le caractère lipschitzien de la racine carrée sur $[\bar{\psi}, +\infty[$, en déduire que

$$\mathbb{E}\left(\max_{1 \leq k \leq N} (T_{3,k}^N)^2 \right) \leq \frac{C}{N} \mathbb{E}\left(\sum_{j=0}^{N-1} ((\bar{T}_3^j)^2 + (\tilde{T}_3^j)^2) \right)$$

$$\text{où } \bar{T}_3^j = \psi(\bar{Y}_{t_j}^N) - \psi(Y_{t_j}) + \frac{\sigma\psi'(\bar{Y}_{t_j}^N) - \sigma\psi'(Y_{t_j})}{\Delta t} \int_{t_j}^{t_{j+1}} (W_s - W_{t_j})ds$$

$$\text{et } \tilde{T}_3^j = \psi(Y_{t_j}) + \frac{\sigma\psi'(Y_{t_j})}{\Delta t} \int_{t_j}^{t_{j+1}} (W_s - W_{t_j})ds - \frac{1}{\Delta t} \int_{t_j}^{t_{j+1}} \psi(Y_s)ds.$$

(b) Calculer $\mathbb{E}\left(\left(\frac{1}{\Delta t} \int_{t_j}^{t_{j+1}} (W_s - W_{t_j})ds \right)^2 \middle| (W_u)_{u \leq t_j} \right)$ et en déduire que $\mathbb{E}((\bar{T}_3^j)^2) \leq \frac{C}{N^2}$.

(c) Vérifier que $\tilde{T}_3^j = \frac{1}{\Delta t} \int_{t_j}^{t_{j+1}} (s - \bar{\tau}_s) \left((b\psi' + \frac{\sigma^2 \psi''}{2})(Y_s)ds + (\sigma\psi'(Y_s) - \sigma\psi'(Y_{t_j}))dW_s \right)$ et en déduire que $\mathbb{E}((\tilde{T}_3^j)^2) \leq \frac{C}{N^2}$.

12. Conclure que (3.41) est vraie.

3.8.9 Simulation du modèle de Cox-Ingersoll-Ross I

On s'intéresse à l'équation différentielle stochastique

$$\begin{cases} dX_t = a(b - X_t)dt + \sigma\sqrt{X_t}\, dW_t \\ X_0 = x_0 \end{cases} \tag{3.42}$$

où $a, b, \sigma, x_0 > 0$ et $(W_t)_{t \geq 0}$ est un mouvement brownien réel.
Pour une étude approfondie de schémas de discrétisation pour ce modèle, nous renvoyons à [1].
Soit $T > 0$ et $N \in \mathbb{N}^*$. On pose $\Delta t = T/N$ et pour $0 \leq k \leq N$, on note $t_k = kT/N = k\Delta t$.

1. (a) Quel résultat assure l'existence d'une unique solution à l'équation

$$\begin{cases} dY_t = -\frac{a}{2} Y_t dt + \frac{\sigma}{2}\, dB_t \\ Y_0 = \sqrt{x_0} \end{cases}$$

où $(B_t)_{t \geq 0}$ est un mouvement brownien standard. Que peut-on dire du processus $W_t = \int_0^t \left(1_{\{Y_s \geq 0\}} - 1_{\{Y_s < 0\}} \right) dB_s$? Vérifier que $X_t = (Y_t)^2$ est solution de l'équation (3.42) pour $b = \frac{\sigma^2}{4a}$.

(b) On se donne maintenant $(W_t^1, \ldots, W_t^d)_{t \geq 0}$ un mouvement brownien à valeurs $\mathbb{R}^d$ et pour $1 \leq i \leq d$ on note Y_t^i la solution de

$$\begin{cases} dY_t^i = -\frac{a}{2} Y_t^i dt + \frac{\sigma}{2}\, dW_t^i \\ Y_0^i = 1_{\{i=1\}} \sqrt{x_0}. \end{cases}$$

On pose $X_t = \sum_{i=1}^d (Y_t^i)^2$. Que peut-on dire du processus

$$W_t = \int_0^t 1_{\{X_s>0\}} \frac{1}{\sqrt{X_s}} \sum_{i=1}^d Y_s^i dW_s^i + \int_0^t 1_{\{X_s=0\}} dW_s^1 ?$$

En déduire que X_t est solution de l'équation (3.42) pour $b = \frac{d\sigma^2}{4a}$.

On admet désormais que **l'équation** (3.42) **possède une unique solution** $(X_t)_{t\geq 0}$ **telle que** $\mathbb{P}(\forall t \geq 0,\ X_t \geq 0) = 1$.

2. Vérifier que la variable $\bar{X}_{t_1}^N$ obtenue par le schéma d'Euler prend des valeurs strictement négatives avec probabilité strictement positive. Quel problème cela pose-t-il ?
3. On pose $\alpha \stackrel{\text{déf}}{=} \frac{2ab}{\sigma^2}$ et
$$\forall x > 0,\ s(x) = \int_1^x y^{-\alpha} e^{\alpha y/b} dy.$$
Pour $k, n \in \mathbb{N}^*$ tels que $1/n \leq x_0 \leq k$, on pose
$$\begin{aligned} \tau_n^k &= \inf\{t \geq 0,\ X_t \notin]1/n, k[\} \\ \tau^k &= \inf\{t \geq 0,\ X_t \geq k\} \\ \tau_n &= \inf\{t \geq 0,\ X_t \leq 1/n\}. \end{aligned}$$
On admet que $\mathbb{P}(\tau_n^k < +\infty) = 1$.
   (a) Vérifier que $\forall x > 0,\ s''(x) = -\alpha \frac{(b-x)}{bx} s'(x)$.
   (b) Calculer $ds(X_t)$. En déduire que $\mathbb{E}\left(s\left(X_{\tau_n^k}\right)\right) = s(x_0)$.
   (c) On suppose que $\alpha > 1$. Donner $\lim_{n\to+\infty} s(1/n)$.
   En déduire $\lim_{n\to+\infty} \mathbb{P}\left(X_{\tau_n^k} = 1/n\right)$ puis $\lim_{n\to+\infty} \mathbb{P}(\tau_n > \tau^k)$. Conclure que $\mathbb{P}(\forall t \leq \tau^k,\ X_t > 0) = 1$ puis que $\mathbb{P}(\forall t \geq 0,\ X_t > 0) = 1$.
4. **On se place désormais dans le cas** $\alpha > 1$.
   (a) Calculer $< \sqrt{X}, W >_t$.
   (b) En remarquant que pour $0 \leq k \leq N-1$, $\sqrt{X_{t_{k+1}}}\left(W_{t_{k+1}} - W_{t_k}\right)$ est égal à
   $$\left(\sqrt{X_{t_{k+1}}} - \sqrt{X_{t_k}}\right)\left(W_{t_{k+1}} - W_{t_k}\right) + \sqrt{X_{t_k}}\left(W_{t_{k+1}} - W_{t_k}\right),$$
   donner la limite en probabilité de $\sum_{k=0}^{N-1} \sqrt{X_{t_{k+1}}}\left(W_{t_{k+1}} - W_{t_k}\right)$ lorsque $N \to +\infty$.
   (c) Conclure que
   $$x_0 + \sum_{k=0}^{N-1} \left[\left(a(b - X_{t_{k+1}}) - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{X_{t_{k+1}}}\left(W_{t_{k+1}} - W_{t_k}\right)\right]$$
   converge vers X_T lorsque $N \to +\infty$.
   Ce résultat suggère l'utilisation du schéma implicite suivant
   $$\begin{cases} \tilde{X}_0 = x_0 \text{ et pour } 0 \leq k \leq N-1, \\ \tilde{X}_{t_{k+1}} = \tilde{X}_{t_k} + \left(a(b - \tilde{X}_{t_{k+1}}) - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\tilde{X}_{t_{k+1}}}\left(W_{t_{k+1}} - W_{t_k}\right) \end{cases} \tag{3.43}$$
   pour l'équation (3.42). En raison du caractère implicite, il faut vérifier que l'équation de récurrence qui précède a bien une solution.

(d) Vérifier que pour $x > 0$ et $w \in \mathbb{R}$, l'équation

$$(1 + a\Delta t)y^2 - wy - ((\alpha - 1)\sigma^2 \Delta t/2 + x) = 0$$

portant sur la variable y admet une unique racine strictement positive $f(x, w)$ que l'exprimera.

(e) En déduire l'existence d'une suite de variables strictement positives $(\tilde{X}_{t_k})_{0 \leq k \leq N}$ vérifiant (3.43).

(f) Pour insister sur la dépendance en la condition initiale x_0, on note $X_T^{x_0}$ la solution de l'équation (3.42) en T et $\tilde{X}_T^{x_0}$ son approximation obtenue par le schéma implicite (3.43). On peut démontrer que $x_0 \to X_T^{x_0}$ est croissante. Vérifier que $x_0 \to \tilde{X}_T^{x_0}$ satisfait la même propriété (*Indication :* on pourra commencer par s'intéresser à la monotonie de la fonction $f(x, w)$ en sa première variable).

5. Pour $\lambda \geq 0$, on note $\varphi(t, \lambda) = \mathbb{E}\left(e^{-\lambda X_t}\right)$ la transformée de Laplace de X_t pour le paramètre λ. En calculant $de^{-\lambda X_t}$, vérifier que φ est solution de l'équation

$$\begin{cases} \partial_t \varphi(t, \lambda) + \lambda ab\varphi(t, \lambda) + \left(\lambda a + \frac{\sigma^2 \lambda^2}{2}\right) \partial_\lambda \varphi(t, \lambda) = 0, \ (t, \lambda) \in \mathbb{R}_+ \times \mathbb{R}_+ \\ \varphi(0, \lambda) = e^{-\lambda x_0}, \ \lambda \in \mathbb{R}_+ \\ \varphi(t, 0) = 1, \ t \in \mathbb{R}_+. \end{cases}$$

L'obtention d'une telle équation aux dérivées partielles suggère que l'on peut déterminer la loi de X_t. C'est effectivement le cas et il est possible de simuler suivant cette loi, ce qui constitue une alternative au schéma implicite (3.43).

3.8.10 Simulation du modèle de Cox-Ingersoll-Ross II

On s'intéresse à l'équation différentielle stochastique

$$\begin{cases} dX_t = adt + \sigma\sqrt{X_t}\, dW_t \\ X_0 = x_0 \end{cases} \tag{3.44}$$

où $a, \sigma, x_0 > 0$ et $(W_t)_{t \geq 0}$ est un mouvement brownien réel.

1. On pose $Y_t = \sqrt{x_0} + \frac{\sigma}{2} B_t$ où $(B_t)_{t \geq 0}$ est un mouvement brownien standard. Que peut-on dire du processus $W_t = \int_0^t \left(1_{\{Y_s \geq 0\}} - 1_{\{Y_s < 0\}}\right) dB_s$? Vérifier que $X_t = (Y_t)^2$ est solution de l'équation (3.44) pour $a = \frac{\sigma^2}{4}$.

   On admet désormais que pour tout $a > 0$, l'équation (3.44) admet une solution $(X_t)_{t \geq 0}$ telle que $\mathbb{P}(\forall t \geq 0, \ X_t \geq 0) = 1$.

2. On souhaite maintenant montrer l'unicité pour (3.44). On suppose donc que $(X_t)_{t \geq 0}$ et $(\tilde{X}_t)_{t \geq 0}$ sont deux solutions telles que $\mathbb{P}(\forall t \geq 0, \ \min(X_t, \tilde{X}_t) \geq 0) = 1$.

   (a) Soit N un entier supérieur ou égal à 2. Calculer $\int_{1/(N^{1/4} \ln N)}^{1/\ln N} \frac{du}{u}$.
   En déduire sans l'expliciter l'existence d'une suite de fonctions $\rho_N : \mathbb{R}_+ \to \mathbb{R}_+$ continues avec ρ_N nulle en dehors de $[1/(N^{1/4} \ln N), 1/\ln N]$, vérifiant

$$\int_0^{1/\ln(N)} \rho_N(u) du = 1 \ \text{ et } \ \forall u \in \mathbb{R}_+^*, \ \rho_N(u) \leq \frac{8}{u \ln N}.$$

(b) Pour $N \geq 2$ et $x \in \mathbb{R}$, on pose $\varphi_N(x) = \int_0^{|x|} \int_0^y \rho_N(u) du dy$. Vérifier que φ_N est une fonction C^2 paire telle que

$$\forall x \in \mathbb{R},\ |\varphi_N''(x)| \leq \min\left(\frac{8}{|x| \ln N}, 8N^{1/4}\right). \tag{3.45}$$

Verifier que $\forall x \in \mathbb{R}_+$, $0 \leq \varphi_N'(x) \leq 1$. Que vaut $\varphi_N'(x)$ pour $x \geq 1/\ln N$? En déduire que

$$\forall x \in \mathbb{R},\ |x| - \frac{1}{\ln N} \leq \varphi_N(x) \leq |x|. \tag{3.46}$$

Ces fonctions régulières qui approchent la fonction valeur absolue s'appellent fonctions de Yamada.

(c) En remarquant que pour $x, \tilde{x} \geq 0$, $(\sqrt{x} - \sqrt{\tilde{x}})^2 \leq |x - \tilde{x}|$, vérifier que pour $t \geq 0$,

$$\varphi_N(X_t - \tilde{X}_t) \leq \sigma \int_0^t \varphi_N'(X_s - \tilde{X}_s)\left(\sqrt{X_s} - \sqrt{\tilde{X}_s}\right) dW_s + \frac{4\sigma^2 t}{\ln N}.$$

(d) En introduisant les temps d'arrêt, $\eta_M = \inf\left\{s \geq 0 : \left|\sqrt{X_s} - \sqrt{\tilde{X}_s}\right| \geq M\right\}$, montrer que $\mathbb{E}\left(\varphi_N(X_t - \tilde{X}_t)\right) \leq \frac{4\sigma^2 t}{\ln N}$.

(e) Conclure que $\mathbb{E}|X_t - \tilde{X}_t| = 0$.

3. On souhaite désormais discrétiser en temps l'équation (3.44) dans le cas où $\boxed{a \geq \frac{\sigma^2}{4}}$.

On se donne $T > 0$, un entier N supérieur ou égal à 2 et pour $k \in \{0, \ldots, N\}$, on pose $t_k = kT/N$.

(a) Vérifier que la variable $\bar{X}_{t_1}^N$ obtenue par le schéma d'Euler prend des valeurs strictement négatives avec probabilité strictement positive. Quel problème cela pose-t-il ?

(b) On note $(X_{t_k}^N)_{0 \leq k \leq N}$ la discrétisation de (3.44) à l'aide du schéma de Milshtein. En remarquant que $X_{t_{k+1}}^N - \left(a - \frac{\sigma^2}{4}\right) \frac{T}{N}$ est une variable aléatoire positive, vérifier que ce schéma est bien défini.

(c) Vérifier que si on note $\tau_s = \left[\frac{Ns}{T}\right] \frac{T}{N}$ le dernier instant de discrétisation avant $s \in [0, T]$, alors

$$X_t^N = x_0 + at + \int_0^t \left(\sigma \sqrt{X_{\tau_s}^N} + \frac{\sigma^2}{2}(W_s - W_{\tau_s})\right) dW_s$$

coïncide avec le schéma de Milshtein aux instant t_k.

(d) Soit $t \in [0, T]$. Calculer $\mathbb{E}(X_t^N)$ et vérifier que

$$\mathbb{E}\left((X_t^N - X_{\tau_t}^N)^2\right) = \left(a^2 + \frac{\sigma^4}{8}\right)(t - \tau_t)^2 + \sigma^2(x_0 + a\tau_t)(t - \tau_t).$$

En déduire qu'il existe une constante $C > 0$ telle que

$$\forall N \in \mathbb{N}^*,\ \forall t \in [0, T],\ \mathbb{E}|X_t^N - X_{\tau_t}^N| \leq \frac{C}{\sqrt{N}}.$$

(e) Vérifier que pour $t \leq T$,

$$\begin{aligned}\varphi_N(X_t - X_t^N) \leq & \int_0^t \varphi_N'(X_s - X_s^N)\left(\sigma\sqrt{X_s} - \sigma\sqrt{X_{\tau_s}^N} - \frac{\sigma^2}{2}(W_s - W_{\tau_s})\right) dW_s \\ & + \int_0^t \varphi_N''(X_s - X_s^N)\left(\sigma^2|X_s - X_{\tau_s}^N| + \frac{\sigma^4}{4}(W_s - W_{\tau_s})^2\right) ds.\end{aligned}$$

(f) En utilisant notamment (3.45) et (3.46), en déduire que

$$\mathbb{E}|X_t - X_t^N| \leq \frac{1 + 8\sigma^2 t}{\ln N} + 8\sigma^2 N^{1/4} \int_0^t \mathbb{E}\left(|X_s^N - X_{\tau_s}^N| + \frac{\sigma^2}{4}(W_s - W_{\tau_s})^2\right) ds.$$

(g) Conclure à l'existence d'une constante $C > 0$ telle que

$$\forall N \in \mathbb{N}^*,\ \forall t \in [0, T],\ \mathbb{E}|X_t - X_t^N| \leq \frac{C}{\ln N}.$$

En fait, on peut par cette approche montrer que $(X_t^N, t \in [0, T])_{N \in \mathbb{N}^*}$ est une suite de Cauchy dans un espace complet bien choisi et en déduire l'existence pour (3.44).

4. On se donne maintenant $b \in \mathbb{R}^*$ et on note

$$\frac{1}{b^-} = \begin{cases} -\frac{1}{b} \text{ si } b < 0 \\ +\infty \text{ si } b \geq 0 \end{cases}.$$

(a) Vérifier que pour $t \geq 0$, $\frac{1}{b}(e^{bt} - 1) < \frac{1}{b^-}$ (où par convention $1/0$ est égal à $+\infty$) et montrer que $\left(B_t = \int_0^{\frac{1}{b}(e^{bt}-1)} \frac{dW_u}{\sqrt{1+bu}}\right)_{t \geq 0}$ est un mouvement brownien réel.

(b) Montrer que pour $t < 1/b^-$, $Y_t = \frac{X_t}{1+bt}$ vérifie

$$Y_t = x_0 + \int_0^t (a - bY_u)\frac{du}{1 + bu} + \sigma \int_0^t \sqrt{Y_u}\frac{dW_u}{\sqrt{1 + bu}}.$$

(c) Vérifier que pour $t \geq 0$, $\max_{0 \leq k \leq N-1} \frac{1}{b}\left(e^{b(k+1)t/N} - e^{bkt/N}\right)$ tend vers 0 lorsque $N \to +\infty$. En déduire que $\sum_{k=0}^{N-1} \sqrt{Y_{\frac{1}{b}(e^{bkt/N}-1)}} \int_{\frac{1}{b}(e^{bkt/N}-1)}^{\frac{1}{b}(e^{b(k+1)t/N}-1)} \frac{dW_u}{\sqrt{1+bu}}$ converge en probabilité vers $\int_0^{\frac{1}{b}(e^{bt}-1)} \sqrt{Y_u}\frac{dW_u}{\sqrt{1+bu}}$.

Conclure que $\left(Z_t = Y_{\frac{1}{b}(e^{bt}-1)}\right)_{t \geq 0}$ est solution de

$$\begin{cases} dZ_t = (a - bZ_t)dt + \sigma\sqrt{Z_t}\, dB_t \\ Z_0 = x_0, \end{cases}$$

équation pour laquelle on peut montrer l'unicité en procédant comme dans la question 2.

(d) En déduire une méthode pour approcher Z_S où $S > 0$.

3.8.11 Simulation exacte en loi d'une EDS en dimension 1

Ce problème présente dans un cadre simplifié la technique de simulation exacte proposée par Beskos Papaspiliopoulos et Roberts [13].

1. **Méthode du rejet**
   Soit $(W^i)_{i\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs dans un espace mesurable $(C, \mathcal{C})$ et $f : C \to [0, M]$ (où $M > 0$) mesurable telle que $\mathbb{E}(f(W^1)) = 1$.
   On se donne également $(A_i)_{i\geq 1}$ une suite d'événements tels que conditionnellement à $(W^j)_{j\geq 1}$ les A_i sont indépendants et de probabilités respectives $1 - \frac{f(W^i)}{M}$, ce qui signifie que

$$\forall I \subset \mathbb{N}^*,\ \mathbb{E}\left(\prod_{i\in I} 1_{A_i}\middle|(W^j)_{j\geq 1}\right) = \prod_{i\in I}\left(1 - \frac{f(W^i)}{M}\right). \tag{3.47}$$

   (a) Pourquoi a-t-on $M \geq 1$? On pose

$$N = \sum_{n\in\mathbb{N}^*} n 1_{A_1\cap A_2\cap\ldots\cap A_{n-1}\cap A_n^c}.$$

   Pour $n \in \mathbb{N}^*$, vérifier que

$$\mathbb{E}\left(1_{A_1\cap A_2\cap\ldots\cap A_{n-1}\cap A_n^c}\middle|(W^j)_{j\geq 1}\right) = \frac{f(W^n)}{M}\prod_{i=1}^{n-1}\left(1 - \frac{f(W^i)}{M}\right)$$

   En déduire la loi de N et vérifier que $\mathbb{P}(N = 0) = 0$.

   (b) On pose $Y = W^N$. Pour $\varphi : C \to \mathbb{R}_+$ mesurable et $n \in \mathbb{N}^*$, calculer $\mathbb{E}\left(1_{\{N=n\}}\varphi(Y)\right)$. En déduire que les variables aléatoires Y et N sont indépendantes et que

$$\mathbb{E}(\varphi(Y)) = \mathbb{E}(\varphi(W^1)f(W^1)).$$

   (c) Si on dispose d'une suite $(U_i)_{i\geq 1}$ de variables aléatoires i.i.d. suivant la loi uniforme sur $[0, 1]$ indépendante de $(W^i)_{i\geq 1}$, comment génère-t-on usuellement les événements A_i ?

2. **Transformation de l'EDS**
   On se donne sur $(\Omega, \mathcal{F}, \mathbb{P})$ un mouvement Brownien réel $(W_t, t \leq T)$. On note $\mathbb{E}$ l'espérance sous $\mathbb{P}$. On s'intéresse à l'EDS

$$X_t = x_0 + \int_0^t \sigma(X_s)dW_s + \int_0^t b(X_s)ds,\ t \leq T$$

   où $\sigma : \mathbb{R} \to \mathbb{R}_+^*$ et $b : \mathbb{R} \to \mathbb{R}$ sont des fonctions régulières et $x_0 \in \mathbb{R}$.

   (a) On pose $\eta(x) = \int_{x_0}^x \frac{dz}{\sigma(z)}$. Pourquoi la fonction η est-elle inversible ?
   Vérifier que $Y_t = \eta(X_t)$ est solution de l'EDS

$$Y_t = W_t + \int_0^t \gamma(Y_s)ds,\ t \leq T$$

   pour une fonction γ à préciser.

(b) On suppose que $\mathbb{E}\left(\exp\left(-\int_0^T \gamma(Y_s)dW_s - \frac{1}{2}\int_0^T \gamma^2(Y_s)ds\right)\right) = 1$ et on note $\mathbb{Q}$ la probabilité sur $(\Omega, \mathcal{F})$ définie par

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_0^T \gamma(Y_s)dW_s - \frac{1}{2}\int_0^T \gamma^2(Y_s)ds\right).$$

Que peut-on dire du processus $(Y_t, t \leq T)$ sous $\mathbb{Q}$?
En déduire que si $\phi : C([0,T],\mathbb{R}) \to \mathbb{R}$ est mesurable et positive,

$$\mathbb{E}\left(\phi(Y_t, t \leq T)\frac{d\mathbb{Q}}{d\mathbb{P}}\right) = \mathbb{E}\left(\phi(W_t, t \leq T)\right). \tag{3.48}$$

(c) Vérifier à l'aide de la formule d'Itô que si on pose $G(y) = \int_0^y \gamma(z)dz$,

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-G(Y_T) + \frac{1}{2}\int_0^T [\gamma^2 + \gamma'](Y_s)ds\right).$$

Soit $\varphi : C([0,T],\mathbb{R}) \to \mathbb{R}$ mesurable et positive. En choisissant bien ϕ dans (3.48), conclure que

$$\mathbb{E}\left(\varphi(Y_t, t \leq T)\right) = \mathbb{E}\left(\varphi(W_t, t \leq T) f(W_t, t \leq T)\right),$$

où pour $(z_t, t \leq T) \in C([0,T],\mathbb{R})$,

$$f(z_t, t \leq T) = \exp\left(G(z_T) - \frac{1}{2}\int_0^T [\gamma^2 + \gamma'](z_s)ds\right). \tag{3.49}$$

3. **Simulation exacte en loi**
On suppose désormais [3] que la fonction $z \in \mathbb{R} \to H(z) = \exp(G(z))$ est majorée par M_H et que la fonction $z \in \mathbb{R} \to \frac{1}{2}[\gamma^2 + \gamma'](z)$ est à valeurs dans $[m_\gamma, m_\gamma + \lambda]$ avec $M_H, \lambda > 0$ et $m_\gamma \in \mathbb{R}$.

(a) Vérifier que la fonction f définie par (3.49) est majorée par $M = M_H \exp(-m_\gamma T)$ et que la fonction $z \in \mathbb{R} \to l(z) = \frac{1}{2}[\gamma^2 + \gamma'](z) - m_\gamma$ est à valeurs dans $[0, \lambda]$.

(b) On se donne indépendamment de $(W_t, t \leq T)$ deux suites indépendantes de variables aléatoires i.i.d. $(U_i)_{i\geq 0}$ et $(\tau_i)_{i\geq 1}$ avec U_0 de loi uniforme sur $[0,1]$ et τ_1 exponentielle de paramètre λ.
On pose $S_0 = 0$ et pour $n \geq 1$, $S_n = \tau_1 + \ldots + \tau_n$. Montrer que pour $n \geq 1$ et $(z_t, t \leq T) \in C([0,T],\mathbb{R})$,

$$\begin{aligned}\mathbb{P}\Bigg(S_n \leq T < S_{n+1}, U_1 \geq \frac{l(z_{S_1})}{\lambda}, \ldots, U_n \geq \frac{l(z_{S_n})}{\lambda}\Bigg)\\ = e^{-\lambda T}\int_{D_T^n}\prod_{k=1}^n (\lambda - l(z_{t_1+\ldots+t_k}))dt_1\ldots dt_n\end{aligned}$$

3. Les hypothèses que nous faisons ici sont restrictives dans un but de simplification. Mais cette technique de simulation proposée par Beskos, Papaspiliopoulos et Roberts peut être adaptée dès que

- la fonction $\exp\left(G(z) - \frac{z^2}{2T}\right)$ est intégrable sur $\mathbb{R}$,
- la fonction $\frac{1}{2}[\gamma^2 + \gamma'](z)$ est minorée par m_γ et l'une de ses limites supérieures pour $z \to +\infty$ ou pour $z \to -\infty$ est finie.

où $D_T^n = \{(t_1, \dots, t_n) \in \mathbb{R}_+^n : \ t_1 + \dots + t_n \leq T\}$.
En effectuant un changement de variables approprié dans cette intégrale, en déduire que

$$\mathbb{P}\left(S_n \leq T < S_{n+1}, \ U_1 \geq \frac{l(z_{S_1})}{\lambda}, \dots, U_n \geq \frac{l(z_{S_n})}{\lambda}\right) = e^{-\lambda T} \frac{\left(\lambda T - \int_0^T l(z_s)ds\right)^n}{n!}.$$

Vérifier que cette formule reste vraie pour $n = 0$.

(c) En déduire que si $\nu = \sum_{n \geq 1} n 1_{\{S_n \leq T < S_{n+1}\}}$,

$$\mathbb{P}\left(U_0 \leq \frac{H(z_T)}{M_H}, \ U_1 \geq \frac{l(z_{S_1})}{\lambda}, \dots, U_\nu \geq \frac{l(z_{S_\nu})}{\lambda}\right) = \frac{f(z_t, t \leq T)}{M}.$$

(d) Soit A l'événement de contraire $A^c = \left\{U_0 \leq \frac{H(W_T)}{M_H}, U_1 \geq \frac{l(W_{S_1})}{\lambda}, \dots, U_\nu \geq \frac{l(W_{S_\nu})}{\lambda}\right\}$. Calculer $\mathbb{E}(1_A | (W_t, t \leq T))$.

(e) En utilisant la question 1, conclure comment simuler en loi Y_T et donc X_T.

3.8.12 Méthode de Monte Carlo pour les options asiatiques

Sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, on se donne $(W_t)_{t \in [0,T]}$ un mouvement brownien réel standard. On se place dans le modèle de Black-Scholes avec taux d'intérêt r sous la probabilité risque-neutre où le cours à l'instant t de l'actif risqué est $S_t = S_0 e^{\sigma W_t + (r - \frac{\sigma^2}{2})t}$. L'objectif de ce problème est de récrire le prix

$$\mathcal{P} = \mathbb{E}\left(e^{-rT} \varphi\left(\frac{1}{T} \int_0^T S_t dt\right)\right)$$

de l'option asiatique de payoff φ sous une forme préparant l'application de la méthode de simulation exacte décrite dans le problème 3.8.11.

1. On pose $\gamma = r - \frac{\sigma^2}{2}$ et $X_t = \frac{S_t}{t} \int_0^t e^{-\sigma W_u - \gamma u} du$ pour $t > 0$.
   (a) Vérifier que $X_T = \frac{1}{T} \int_0^T S_0 e^{\sigma(W_T - W_{T-t}) + \gamma t} dt$.
   (b) Que peut-on dire du processus $(W_T - W_{T-t})_{t \in [0,T]}$?
   (c) En déduire que $\mathcal{P} = \mathbb{E}(e^{-rT} \varphi(X_T))$.
   (d) Donner $\lim_{t \to 0^+} X_t$ et calculer dX_t.
   (e) En déduire que le processus $Y_t = \ln(X_t / S_0)$ est solution de l'EDS

$$dY_t = \sigma dW_t + \gamma dt + \frac{e^{-Y_t} - 1}{t} dt, \ Y_0 = 0. \tag{3.50}$$

   (f) Pour $\tilde{Y}$ une autre solution de cette équation, vérifier que $d(Y_t - \tilde{Y}_t)^2 \leq 0$ et en déduire l'unicité trajectorielle pour (3.50).
2. Pour $t > 0$ soit $Z_t = \frac{\sigma}{t} \int_0^t s dW_s + \frac{\gamma}{2} t$.
   (a) Pour $u, t \geq 0$, calculer $\mathbb{E}(Z_t)$ et Cov (Z_u, Z_t).
   (b) Pour $0 \leq t_1 < t_2 < \dots < t_n \leq T$, que peut-on dire du vecteur $(Z_{t_1}, \dots, Z_{t_n})$? Comment peut-on le simuler ?
   (c) Vérifier que $\frac{1}{t} \int_0^t s dW_s = W_t - \frac{1}{t} \int_0^t W_s ds$ et en déduire $\lim_{t \to 0^+} Z_t$.

(d) Montrer que Z_t est solution de l'EDS

$$dZ_t = \sigma dW_t + \gamma dt - \frac{Z_t}{t}dt,\ Z_0 = 0.$$

Pourquoi est-ce l'unique solution de cette équation ?

3. Dans cette question, on admettra que les intégrales que l'on est amené à considérer sont bien définies.

(a) Pour un processus $(H_t)_{t\in[0,T]}$ adapté à la filtration de W et suffisamment intégrable, donner $\frac{d\mathbb{Q}}{d\mathbb{P}}$ tel que sous la probabilité $\mathbb{Q}$, $(B_t = W_t - \int_0^t H_s ds)_{t\le T}$ est un mouvement Brownien.

(b) Préciser H_t pour que Z_t soit solution de l'EDS

$$dZ_t = \sigma dB_t + \gamma dt + \frac{e^{-Z_t} - 1}{t}dt,\ Z_0 = 0.$$

(c) On pose

$$A(t,z) = \frac{1 - z + \frac{z^2}{2} - e^{-z}}{\sigma^2 t} \quad \text{et}\ f(t,z) = \left[\frac{A}{t} + \left(\frac{z}{t} - \gamma - \frac{\sigma^2}{2}\frac{\partial A}{\partial z}\right)\frac{\partial A}{\partial z} - \frac{\sigma^2}{2}\frac{\partial^2 A}{\partial z^2}\right](t,z).$$

En admettant que d'après la loi du logarithme itéré, pour tout $\varepsilon > 0$, $Z_t = o(t^{\frac{1}{2}-\varepsilon})$ en 0^+, donner $\lim_{t\to 0^+} A(t,Z_t)$.

(d) Calculer $dA(t,Z_t)$. En déduire que $\int_0^T H_t dW_t - \frac{1}{2}\int_0^T H_t^2 dt = A(T,Z_T) + \int_0^T f(t,Z_t)dt$ puis que

$$\mathcal{P} = \mathbb{E}\left(\psi(Z_T)e^{\int_0^T f(t,Z_t)dt}\right) \quad \text{où} \quad \psi(z) = e^{-rT}\varphi(S_0 e^z)e^{A(T,z)}.$$

3.8.13 Schéma de Ninomiya Victoir [61]

Dans cette partie, par fonction régulière, on entend une fonction dérivable autant de fois que nécessaire avec des dérivées bornées.

Soit $T > 0$, $y \in \mathbb{R}$, $b : \mathbb{R} \to \mathbb{R}$ une fonction régulière et sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, $(W_t)_{t\in[0,T]}$ un mouvement brownien réel standard. On s'intéresse à l'Equation Différentielle Stochastique

$$\begin{cases} X_0 = y, \\ dX_t = dW_t + b(X_t)dt. \end{cases} \tag{3.51}$$

Pour $(t,z) \in \mathbb{R}_+ \times \mathbb{R}$, on introduit $x(t,z)$ la solution à l'instant t de l'Equation Différentielle Ordinaire issue de z à l'instant initial et obtenue en enlevant le terme brownien de l'EDS :

$$\begin{cases} x(0,z) = z, \\ \frac{d}{dt}x(t,z) = b(x(t,z)). \end{cases} \tag{3.52}$$

On se donne $N \in \mathbb{N}^*$ et pour $0 \le k \le N$, on pose $t_k = kT/N$. On s'intéresse à un schéma de discrétisation proposé récemment par Ninomiya et Victoir. Dans l'exemple simple qui nous intéresse, pour passer du temps t_k au temps t_{k+1}, ce schéma consiste à intégrer [4]

4. Si la solution de l'EDO (3.52) n'a pas d'expression analytique, on peut recourir à un schéma de discrétisation pour cette étape. Il faut choisir ce schéma avec soin pour préserver l'ordre faible du schéma qui en découle pour l'EDS (3.51).

l'EDO (3.52) sur l'intervalle $[t_k, \frac{(2k+1)T}{2N}]$ puis à ajouter l'accroissement brownien et enfin à intégrer l'EDO (3.52) sur l'intervalle de temps $[\frac{(2k+1)T}{2N}, t_{k+1}]$:

$$\begin{cases} \bar{X}_0 = y, \\ \forall k \in \{0, \ldots, N-1\},\ \bar{X}_{t_{k+1}} = x(\frac{T}{2N}, Z_{k+\frac{1}{2}}) \text{ où } Z_{k+\frac{1}{2}} = x(\frac{T}{2N}, \bar{X}_{t_k}) + W_{t_{k+1}} - W_{t_k}. \end{cases} \tag{3.53}$$

L'objectif de ce sujet est de vérifier que l'ordre faible de ce schéma est en $\frac{1}{N^2}$. Pour cela, on s'intéressera au cas d'un coefficient de dérive linéaire avant de traiter le cas général.

Les 2 questions peuvent être traitées de façon indépendante.

1. **Cas d'un coefficient de dérive linéaire :** $\forall y \in \mathbb{R}$, $b(y) = cy$ où $c \in \mathbb{R}^*$.
   (a) Quelle est la solution $x(t, z)$ de l'EDO (3.52) ? Préciser le schéma (3.53) dans ce cas particulier. Vérifier que
   $$\forall k \in \{0, \ldots, N\},\ \bar{X}_{t_k} = ye^{ct_k} + e^{\frac{cT}{2N}} \sum_{l=1}^{k} e^{c(t_k - t_l)}(W_{t_l} - W_{t_{l-1}}).$$
   En déduire que Var $(\bar{X}_T) = \frac{e^{2cT}-1}{2c} \times \frac{\frac{cT}{N}}{\sinh(\frac{cT}{N})}$.
   (b) Quelle est la loi de X_T ? En déduire que X_T a même loi que $\bar{X}_T + \sqrt{\gamma_N}\, G$ où G est une gaussienne centrée réduite indépendante de $(W_t)_{t\in[0,T]}$ et γ_N une constante que l'on précisera.
   (c) En remarquant que si $f : \mathbb{R} \to \mathbb{R}$ est une fonction C^2, alors
   $$\forall x, z \in \mathbb{R},\ f(x+z) = f(x) + zf'(x) + z^2 \int_0^1 (1-\alpha) f''(x + \alpha z) d\alpha,$$
   en déduire que si f est bornée ainsi que ses dérivées,
   $$\left|\mathbb{E}(f(X_T)) - \mathbb{E}(f(\bar{X}_T))\right| \le \frac{\sup_{z\in\mathbb{R}} |f''(z)|}{2} \gamma_N.$$
   Conclure que l'ordre faible du schéma est en $1/N^2$.
   (d) On note $\hat{X}_{t_k}$ la valeur obtenue en t_k par le schéma d'Euler avec pas de temps $\frac{T}{N}$. Calculer $\mathbb{E}(\hat{X}_T)$ et Var $(\hat{X}_T)$.
   (e) On suppose $c > 0$. Vérifier que X_T a même loi que $\hat{X}_T + \hat{\eta}_N + \sqrt{\hat{\gamma}_N}\, G$ pour des constantes $\hat{\eta}_N$ et $\hat{\gamma}_N$ à préciser. Retrouver que l'ordre faible du schéma d'Euler est en $1/N$. Comment peut-on améliorer la convergence de ce schéma ?
2. **Cas général :** Pour $f : \mathbb{R} \to \mathbb{R}$ une fonction régulière, on introduit la solution u, supposée régulière, de l'Equation aux Dérivées Partielles
   $$\begin{cases} \frac{\partial u}{\partial t}(t,x) + Lu(t,x) = 0,\ (t,x) \in [0,T] \times \mathbb{R} \\ u(T,x) = f(x),\ x \in \mathbb{R} \end{cases}, \tag{3.54}$$
   où L est le générateur infinitésimal de l'EDS (3.51) : $Lg(x) = \frac{1}{2}g''(x) + b(x)g'(x)$.
   (a) Vérifier que
   $$\mathbb{E}(f(\bar{X}_T)) - \mathbb{E}(f(X_T)) = \sum_{k=0}^{N-1} \mathcal{E}_k \text{ où } \mathcal{E}_k = \mathbb{E}\left(u(t_{k+1}, \bar{X}_{t_{k+1}}) - u(t_k, \bar{X}_{t_k})\right).$$

(b) Vérifier que $\frac{\partial^2 u}{\partial t^2}(t,x) = L(Lu)(t,x)$. En admettant provisoirement que

$$\begin{aligned}\mathbb{E}(u(t_{k+1},\bar{X}_{t_{k+1}})|\bar{X}_{t_k}) =& u(t_{k+1},\bar{X}_{t_k}) + Lu(t_{k+1},\bar{X}_{t_k})\frac{T}{N}\\ &+ \frac{1}{2}L(Lu)(t_{k+1},\bar{X}_{t_k})\frac{T^2}{N^2} + \mathcal{O}\left(\frac{1}{N^3}\right),\end{aligned} \tag{3.55}$$

en déduire que $\mathcal{E}_k = \mathcal{O}\left(\frac{1}{N^3}\right)$ et conclure que l'ordre faible du schéma est $1/N^2$.

(c) L'objectif de cette question est de vérifier que pour $g : \mathbb{R} \to \mathbb{R}$ régulière

$$\mathbb{E}(g(\bar{X}_{t_1})) = g(y) + Lg(y)t_1 + L(Lg)(y)\frac{t_1^2}{2} + \mathcal{O}\left(\frac{1}{N^3}\right), \tag{3.56}$$

propriété qui se généralise facilement en (3.55).

i. Vérifier que

$$g(x(t,z)) = g(z) + (bg')(z)t + \int_0^t \int_0^s b(bg')'(x(r,z))drds$$

et en déduire que pour t au voisinage de 0,

$$g(x(t,z)) = g(z) + (bg')(z)t + b(bg')'(z)\frac{t^2}{2} + \mathcal{O}\left(t^3\right)$$
$$\text{et } x(t,z) = z + b(z)t + bb'(z)\frac{t^2}{2} + \mathcal{O}\left(t^3\right).$$

ii. En déduire que

$$\begin{aligned}g(\bar{X}_{t_1}) =& g\left(y + b(y)\frac{t_1}{2} + bb'(y)\frac{t_1^2}{8} + W_{t_1} + \mathcal{O}\left(\frac{1}{N^3}\right)\right)\\ &+ (bg')\left(y + b(y)\frac{t_1}{2} + bb'(y)\frac{t_1^2}{8} + W_{t_1} + \mathcal{O}\left(\frac{1}{N^3}\right)\right)\frac{t_1}{2}\\ &+ b(bg')'\left(y + b(y)\frac{t_1}{2} + bb'(y)\frac{t_1^2}{8} + W_{t_1} + \mathcal{O}\left(\frac{1}{N^3}\right)\right)\frac{t_1^2}{8} + \mathcal{O}\left(\frac{1}{N^3}\right).\end{aligned}$$

iii. Vérifier que $\mathbb{E}\left((bg')\left(y + b(y)\frac{t_1}{2} + bb'(y)\frac{t_1^2}{8} + W_{t_1} + \mathcal{O}\left(\frac{1}{N^3}\right)\right)\right)$ est égal à

$$bg'(y) + (bg')'(y)b(y)\frac{t_1}{2} + \frac{1}{2}(bg')''(y)t_1 + \mathcal{O}\left(\frac{1}{N^2}\right).$$

iv. Vérifier que $\mathbb{E}\left(g\left(y + b(y)\frac{t_1}{2} + bb'(y)\frac{t_1^2}{8} + W_{t_1} + \mathcal{O}\left(\frac{1}{N^3}\right)\right)\right)$ est égal à un terme en $\mathcal{O}\left(\frac{1}{N^3}\right)$ près à la fonction

$$g + g' \times \left(b\frac{t_1}{2} + bb'\frac{t_1^2}{8}\right) + \frac{1}{2}g'' \times \left(b^2\frac{t_1^2}{4} + t_1\right) + \frac{1}{6}g^{(3)} \times \left(3b\frac{t_1^2}{2}\right) + \frac{3t_1^2}{24}g^{(4)}$$

prise au point y ($g^{(k)}$ désigne la dérivée d'ordre k de la fonction g).

v. En remarquant que

$$L(Lg) = \frac{1}{4}b(b'g' + bg'') + \frac{1}{2}b(bg')' + \frac{1}{4}b(bg')' + \frac{1}{2}(bg')'' + \frac{1}{2}bg^{(3)} + \frac{1}{4}g^{(4)},$$

conclure que (3.56) est vérifiée.

Dans le cas de l'EDS générale posée en dimension n avec $(W^1, \ldots, W^d)$ un mouvement brownien standard de dimension d :

$$dX_t = \sum_{j=1}^{d} \sigma_j(X_t)dW_t^j + b(X_t)dt$$

où b et les $\sigma_j = (\sigma_{1j}, \ldots, \sigma_{nj})^*$, $1 \leq j \leq d$ sont des fonctions régulières de $\mathbb{R}^n$ dans $\mathbb{R}^n$, le schéma de Ninomiya et Victoir nécessite de générer une suite $(U_k)_{1\leq k\leq N}$ de variables uniformes sur $[0,1]$ indépendantes et indépendantes de $(W^1, \ldots, W^d)$. Pour passer de l'instant t_k à l'instant t_{k+1}, il consiste à

1. intégrer sur la durée $\frac{T}{2N}$ l'EDO $\frac{d}{dt}x(t) = \left[b - \frac{1}{2}\sum_{j=1}^{d} \partial\sigma_j\sigma_j\right](x(t))$ où $\partial\sigma_j = \left(\frac{\partial\sigma_{ij}}{\partial\sigma_{x_l}}\right)_{1\leq i,l\leq n}$.
2. Si $U_{k+1} \leq \frac{1}{2}$, intégrer successivement pour j croissant de 1 à d l'EDO $\frac{d}{dt}x(t) = \sigma_j(x(t))$ sur la durée aléatoire $W^j_{t_{k+1}} - W^j_{t_k}$. Si $U_{k+1} > \frac{1}{2}$, effectuer la même opération mais pour j décroissant de d à 1.
3. Reprendre la première étape.

3.8.14 Schémas de Ninomiya-Victoir [61] et de Ninomiya-Ninomiya [62]

Le résultat de la question préliminaire qui suit sera utilisé à la fin du problème dans les questions 8a et 8c.

1. Soit (X, Y) un couple qui suit la loi gaussienne $\mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} a & c \\ c & b \end{pmatrix}\right)$. On suppose $a > 0$ et on pose $Z = Y - \frac{c}{a}X$.
   (a) Quelle est la loi de (X, Z) ?
   (b) Vérifier que $\mathbb{E}(X^4) = 3a^2$, $\mathbb{E}(X^2Y^2) = ab + 2c^2$, $\mathbb{E}(X^3Y) = 3ac$ et $\mathbb{E}(XY^3) = 3bc$.
   (c) Montrer que pour $l, m \in \mathbb{N}$, $\mathbb{E}(X^lY^m) = 0$ dès lors que $l + m$ est impair.

Dans toute la suite, par fonction régulière, on entend une fonction dérivable autant de fois que nécessaire avec des dérivées bornées et φ désigne une fonction régulière de $\mathbb{R}^n$ dans $\mathbb{R}$.

Soit $T > 0$, $y \in \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$ et $b : \mathbb{R}^n \to \mathbb{R}^n$ deux fonctions régulières. Sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, on se donne également $(W_t)_{t\in[0,T]}$ un mouvement brownien à valeurs $\mathbb{R}^d$. On s'intéresse à l'Equation Différentielle Stochastique

$$\begin{cases} X_0 = y, \\ dX_t = \sigma(X_t)dW_t + b(X_t)dt. \end{cases} \tag{3.57}$$

On note L l'opérateur différentiel du second ordre défini par

$$L\varphi(x) = \frac{1}{2}\sum_{i,l=1}^{n} a_{il}(x)\partial^2_{x_ix_l}\varphi(x) + \sum_{i=1}^{n} b_i(x)\partial_{x_i}\varphi(x) \text{ où } a_{il}(x) = \sum_{j=1}^{d} \sigma_{ij}(x)\sigma_{lj}(x).$$

Pour $j \in \{1, \ldots, d\}$, on note également $\sigma_j(x) = (\sigma_{ij}(x))_{1\leq i\leq n} \in \mathbb{R}^n$ la j-ème colonne de la matrice σ et $\partial\sigma_j(x) = (\partial_{x_i}\sigma_{lj}(x))_{1\leq l,i\leq n} \in \mathbb{R}^{n\times n}$.

2. Écrire le schéma de Milshtein pour cette EDS. À quelle condition peut-on l'implémenter ?
3. Pour $f : \mathbb{R}^n \to \mathbb{R}$ une fonction régulière, on introduit la solution u, supposée régulière, de l'Equation aux Dérivées Partielles

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) + Lu(t,x) = 0, \ (t,x) \in [0,T] \times \mathbb{R}^n \\ u(T,x) = f(x), \ x \in \mathbb{R}^n \end{cases} . \tag{3.58}$$

On note également $(\bar{X}_{t_k})_{0\leq k\leq N}$ le vecteur obtenu sur la grille temporelle $t_k = \frac{kT}{N}$ par discrétisation de l'EDS (3.57) par un schéma approprié tel que $\bar{X}_0 = y$.

(a) Vérifier que

$$\mathbb{E}(f(\bar{X}_T)) - \mathbb{E}(f(X_T)) = \sum_{k=0}^{N-1} \mathcal{E}_k \ \text{ où } \ \mathcal{E}_k = \mathbb{E}\left(u(t_{k+1}, \bar{X}_{t_{k+1}}) - u(t_k, \bar{X}_{t_k})\right).$$

(b) Vérifier que $\frac{\partial^2 u}{\partial t^2}(t,x) = L^2 u(t,x)$ où $L^2 u = L(Lu)$. En déduire que

$$u(t_{k+1}, \bar{X}_{t_k}) + \frac{T}{N} Lu(t_{k+1}, \bar{X}_{t_k}) + \frac{T^2}{2N^2} L^2 u(t_{k+1}, \bar{X}_{t_k}) = u(t_k, \bar{X}_{t_k}) + \mathcal{O}\left(\frac{1}{N^3}\right).$$

(c) Conclure que si pour tout $k \in \{0, \dots, N-1\}$,

$$\begin{aligned} \mathbb{E}(u(t_{k+1}, \bar{X}_{t_{k+1}}) | \bar{X}_{t_k}) = & u(t_{k+1}, \bar{X}_{t_k}) + \frac{T}{N} Lu(t_{k+1}, \bar{X}_{t_k}) \\ & + \frac{T^2}{2N^2} L^2 u(t_{k+1}, \bar{X}_{t_k}) + \mathcal{O}\left(\frac{1}{N^3}\right) \end{aligned} \tag{3.59}$$

alors $\mathcal{E}_k = \mathcal{O}\left(\frac{1}{N^3}\right)$ et l'ordre faible du schéma est $1/N^2$.

L'objectif du problème est de démontrer que (3.59) est vérifiée pour deux schémas proposés récemment par Ninomiya et Victoir [61] et par Ninomiya et Ninomiya [62].

4. Pour $j \in \{1, \dots, d\}$, on note V_j l'opérateur différentiel du premier ordre défini par $V_j\varphi(x) = \sigma_j(x).\nabla\varphi(x) = \sum_{i=1}^n \sigma_{ij}(x)\partial_{x_i}\varphi(x)$. Pour V et W deux opérateurs différentiels, VW désigne l'opérateur différentiel (en général différent de WV) défini par $VW\varphi(x) = V(W\varphi)(x)$. Pour $m \in \mathbb{N}^*$, V^m désigne l'opérateur $\underbrace{V \dots V}_{m \text{ fois}}$.

Enfin, on pose $V_0 = L - \frac{1}{2}\sum_{j=1}^d V_j^2$.

(a) Exprimer L^2 à l'aide des V_i, $i \in \{0, \dots, d\}$.

(b) Montrer que pour $j \in \{1, \dots, d\}$,

$$V_j^2\varphi(x) = \sum_{i,l=1}^n \sigma_{ij}(x)\sigma_{lj}(x)\partial^2_{x_i x_l}\varphi(x) + \sum_{l=1}^n \left(\sum_{i=1}^n \partial_{x_i}\sigma_{lj}(x)\sigma_{ij}(x)\right)\partial_{x_l}\varphi(x)$$

et en déduire que $V_0\varphi(x) = \sigma_0(x).\nabla\varphi(x)$ où $\sigma_0(x) = b(x) - \frac{1}{2}\sum_{j=1}^d \partial\sigma_j\sigma_j(x)$.

5. Pour $\eta : \mathbb{R}^n \to \mathbb{R}^n$ une fonction régulière on note V l'opérateur différentiel du premier ordre défini par $V\varphi(x) = \eta(x).\nabla\varphi(x)$ et $(e^{tV}(x))_{t\in\mathbb{R}}$ l'unique solution de l'Équation Différentielle ordinaire

$$\frac{d}{dt}y(t) = \eta(y(t)), \quad y(0) = x.$$

(a) Calculer $\frac{d\varphi(e^{tV}(x))}{dt}$ et en déduire que $\forall t \in \mathbb{R}$, $\varphi(e^{tV}(x)) = \varphi(x) + \int_0^t V\varphi(e^{sV}(x))ds$.

(b) En déduire que pour tout $l \in \mathbb{N}$,

$$\varphi(e^{tV}(x)) = \sum_{m=0}^{l} \frac{t^m V^m \varphi(x)}{m!} + \int_0^t \int_0^{s_1} \dots \int_0^{s_l} V^{l+1}\varphi(e^{s_{l+1}V}(x))ds_{l+1}\dots ds_1.$$

<u>Indication :</u> *écrire l'égalité de la question précédente en remplaçant φ par $V^m\varphi$ où $m \geq 1$.*

Justifier la notation $e^{tV}(x)$ et vérifier que $\varphi(e^{tV}(x)) = \sum_{m=0}^{l} \frac{t^m V^m \varphi(x)}{m!} + \mathcal{O}(|t|^{l+1})$.

(c) Soit W un autre opérateur différentiel du premier ordre associé à une fonction régulière de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Montrer que pour $s, t \in \mathbb{R}$,

$$\varphi(e^{sW}e^{tV}(x)) = \sum_{\substack{m_1,m_2\geq 0\\ m_1+m_2\leq l}} \frac{s^{m_1}t^{m_2}V^{m_2}W^{m_1}\varphi(x)}{m_1!m_2!} + \mathcal{O}((|s|\vee|t|)^{l+1}).$$

<u>Indication :</u> *écrire le développement à l'ordre l de $\varphi(e^{sW}(y))$ en une valeur bien choisie de $y \in \mathbb{R}^n$.*

6. Soit $t > 0$, $G_1, \dots, G_d$ des gaussiennes centrées réduites indépendantes et pour $j \in \{1, \dots, d\}$, $Z_j = \sqrt{t}G_j$.

(a) Pour $m = (m_0, \dots, m_{d+1}) \in \mathbb{N}^{d+2}$ on note $\|m\| = 2(m_0 + m_{d+1}) + \sum_{j=1}^d m_j$. Expliquer comment obtenir l'égalité

$$\begin{aligned}&\mathbb{E}\left[\varphi(e^{\frac{t}{2}V_0}e^{Z_1V_1}\dots e^{Z_dV_d}e^{\frac{t}{2}V_0}(x))\right]\\ &= \sum_{m:\|m\|\leq 5} \frac{t^{\frac{\|m\|}{2}}}{2^{m_0+m_{d+1}}} \frac{V_0^{m_{d+1}}V_d^{m_d}\dots V_1^{m_1}V_0^{m_0}\varphi(x)}{m_0!m_1!\dots m_d!m_{d+1}!}\mathbb{E}[G_1^{m_1}\times\dots\times G_d^{m_d}] + \mathcal{O}(t^3).\end{aligned}$$

(b) Remarquer que s'il existe $j \in \{1, \dots, d\}$ tel que m_j est impair, alors $\mathbb{E}[G_1^{m_1} \times \dots \times G_d^{m_d}] = 0$ et en déduire que la somme précédente est restreinte aux $(d+2)$-uplets m tels que $\|m\| \in \{0, 2, 4\}$. Vérifier qu'en dehors du cas $m = (0, \dots, 0)$, les seuls termes non nuls de cette somme correspondent aux cas suivants où on ne précise que les m_i non nuls :

(1) $m_0 + m_{d+1} = 1$,
(2) $m_i = 2$ pour un indice $i \in \{1, \dots, d\}$,
(3) $m_i = 4$ pour un indice $i \in \{1, \dots, d\}$,
(4) $m_i = m_j = 2$ pour un couple d'indices $i < j \in \{1, \dots, d\}$,
(5) $m_i = 2$ pour un indice $i \in \{1, \dots, d\}$ et $m_0 + m_{d+1} = 1$,
(6) $m_0 + m_{d+1} = 2$.

(c) Conclure que

$$\begin{aligned}&\mathbb{E}\left[\varphi(e^{\frac{t}{2}V_0}e^{Z_1V_1}\dots e^{Z_dV_d}e^{\frac{t}{2}V_0}(x))\right] = \varphi(x) + tL\varphi(x)\\ &+\frac{t^2}{2}\left(\frac{1}{4}\sum_{i=1}^d V_i^4 + \frac{1}{2}\sum_{1\leq i<j\leq d} V_j^2V_i^2 + \frac{1}{2}\left\{V_0\sum_{i=1}^d V_i^2 + \sum_{i=1}^d V_i^2V_0\right\} + V_0^2\right)\varphi(x) + \mathcal{O}(t^3).\end{aligned}$$

(d) En déduire que

$$\mathbb{E}\left[\frac{1}{2}\varphi(e^{\frac{t}{2}V_0}e^{Z_1V_1}\dots e^{Z_dV_d}e^{\frac{t}{2}V_0}(x))+\frac{1}{2}\varphi(e^{\frac{t}{2}V_0}e^{Z_dV_d}\dots e^{Z_1V_1}e^{\frac{t}{2}V_0}(x))\right]$$
$$=\varphi(x)+tL\varphi(x)+\frac{t^2}{2}L^2\varphi(x)+\mathcal{O}(t^3).$$

7. Le schéma de Ninomiya et Victoir [61] nécessite de générer une suite $(U_k)_{1\leq k\leq N}$ de variables uniformes sur $[0,1]$ indépendantes et indépendantes de $(W^1,\dots,W^d)$. Pour passer de $\bar{X}_{t_k}$ à $\bar{X}_{t_{k+1}}$, il consiste à

(i) Intégrer sur la durée $\frac{T}{2N}$ l'EDO $\frac{d}{dt}y(t)=\sigma_0(y(t))$,

(ii) Si $U_{k+1}\leq\frac{1}{2}$, intégrer successivement pour j croissant de 1 à d l'EDO $\frac{d}{dt}y(t)=\sigma_j(y(t))$ sur la durée aléatoire $W^j_{t_{k+1}}-W^j_{t_k}$. Si $U_{k+1}>\frac{1}{2}$, effectuer la même opération mais pour j décroissant de d à 1.

(iii) Reprendre l'étape **(i)**.

Vérifier que pour ce schéma,

$$\mathbb{E}(\varphi(\bar{X}_{t_1}))=\varphi(y)+\frac{T}{N}L\varphi(y)+\frac{T^2}{2N^2}L^2\varphi(y)+\mathcal{O}\left(\frac{1}{N^3}\right),$$

propriété qui se généralise facilement en (3.59).

8. L'inconvenient du schéma de Ninomiya et Victoir est qu'il faut intégrer $(d+2)$ EDOs à chaque pas de temps, ce qui peut s'avérer très coûteux. Ninomiya et Ninomiya [62] ont proposé un schéma qui préserve l'ordre de convergence faible en $1/N^2$ mais dans lequel il suffit d'intégrer deux EDOs à chaque pas de temps.

(a) Soit $(G_{1,j},G_{2,j})_{1\leq j\leq d}$ des couples i.i.d. suivant la loi $\mathcal{N}_2\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}a&c\\c&b\end{pmatrix}\right)$. À l'aide de la question 1, montrer que $\mathbb{E}(\prod_{j=1}^d G_{1,j}^{l_j}G_{2,j}^{m_j})=0$ dès que la somme des deux coordonnées de l'un des couples $(l_j,m_j)\in\mathbb{N}^2$ est impaire.

(b) Soient $\alpha, \beta \in \mathbb{R}$. Expliquer comment obtenir l'égalité

$$
\begin{aligned}
&\mathbb{E}\left(\varphi(e^{t\beta V_0+\sqrt{t}\sum_{j=1}^{d} G_{2,j}V_j}e^{t\alpha V_0+\sqrt{t}\sum_{j=1}^{d} G_{1,j}V_j}(x))\right)\\
&= \varphi(x) + tV_0\varphi(x)(\alpha+\beta) + t\sum_{j=1}^{d} V_j^2\varphi(x)\left(\frac{\mathbb{E}(G_{1,j}^2)}{2} + \mathbb{E}(G_{1,j}G_{2,j}) + \frac{\mathbb{E}(G_{1,j}^2)}{2}\right)\\
&+ t^2V_0^2\varphi(x)\left(\frac{\alpha^2}{2} + \alpha\beta + \frac{\beta^2}{2}\right)\\
&+ t^2\sum_{j=1}^{d}\Bigg\{V_jV_0V_j\varphi(x)\left(\alpha\frac{\mathbb{E}(G_{1,j}^2)}{6} + (\alpha+\beta)\frac{\mathbb{E}(G_{1,j}G_{2,j})}{2} + \beta\frac{\mathbb{E}(G_{2,j}^2)}{6}\right)\\
&\qquad + V_0V_j^2\varphi(x)\left(\alpha\frac{\mathbb{E}(G_{1,j}^2)}{6} + \alpha\frac{\mathbb{E}(G_{1,j}G_{2,j}) + \mathbb{E}(G_{2,j}^2)}{2} + \beta\frac{\mathbb{E}(G_{2,j}^2)}{6}\right)\\
&\qquad + V_j^2V_0\varphi(x)\left(\alpha\frac{\mathbb{E}(G_{1,j}^2)}{6} + \beta\frac{\mathbb{E}(G_{1,j}^2) + \mathbb{E}(G_{1,j}G_{2,j})}{2} + \beta\frac{\mathbb{E}(G_{2,j}^2)}{6}\right)\\
&\qquad + V_j^4\varphi(x)\left(\frac{\mathbb{E}(G_{1,j}^4)}{24} + \frac{\mathbb{E}(G_{1,j}^3G_{2,j})}{6} + \frac{\mathbb{E}(G_{1,j}^2G_{2,j}^2)}{4} + \frac{\mathbb{E}(G_{1,j}G_{2,j}^3)}{6} + \frac{\mathbb{E}(G_{2,j}^4)}{24}\right)\Bigg\}\\
&+ t^2\sum_{\substack{i,j=1\\ i\neq j}}^{d}\Bigg\{V_i^2V_j^2\varphi(x)\Bigg(\frac{\mathbb{E}(G_{1,i}^2G_{1,j}^2)}{24} + \frac{\mathbb{E}(G_{1,i}^2G_{1,j}G_{2,j})}{6} + \frac{\mathbb{E}(G_{1,i}^2G_{2,j}^2)}{4}\\
&\qquad\qquad + \frac{\mathbb{E}(G_{1,i}G_{2,i}G_{2,j}^2)}{6} + \frac{\mathbb{E}(G_{2,i}^2G_{2,j}^2)}{24}\Bigg)\\
&\qquad + V_iV_jV_iV_j\varphi(x)\Bigg(\frac{\mathbb{E}(G_{1,i}^2G_{1,j}^2)}{24} + \frac{\mathbb{E}(G_{1,i}^2G_{1,j}G_{2,j})}{6} + \frac{\mathbb{E}(G_{1,i}G_{1,j}G_{2,i}G_{2,j})}{4}\\
&\qquad\qquad + \frac{\mathbb{E}(G_{1,i}G_{2,i}G_{2,j}^2)}{6} + \frac{\mathbb{E}(G_{2,i}^2G_{2,j}^2)}{24}\Bigg)\Bigg\} + \mathcal{O}(t^3)
\end{aligned}
$$

(c) Soit $u \geq 1/2$ et $\varepsilon \in \{-1, 1\}$. Vérifier que pour le choix $a = u$, $b = 1 + u - \varepsilon\sqrt{2(2u-1)}$ et $c = -u + \varepsilon\frac{\sqrt{2(2u-1)}}{2}$, $\begin{pmatrix} a & c \\ c & b \end{pmatrix}$ est bien une matrice de covariance. En posant également $\alpha = \varepsilon\frac{\sqrt{2(2u-1)}}{2}$ et $\beta = 1 - \alpha$, vérifier que

$$
\begin{cases}
a + b + 2c = 1 = a^2 + 4ac + 2ab + 4c^2 + 4bc + b^2\\
\alpha a + \beta b + 3(\alpha+\beta)c = 0\\
\alpha a + \beta b + 3\alpha(c+b) = 3/2\\
\alpha a + \beta b + 3\beta(c+a) = 3/2\\
a^2 + 4ac + 6ab + 4bc + b^2 = 3\\
a^2 + 4ac + 6c^2 + 4bc + b^2 = 0
\end{cases}.
$$

En déduire que pour ce choix

$$
\mathbb{E}\left(\varphi(e^{t\beta V_0+\sqrt{t}\sum_{j=1}^{d} G_{2,j}V_j}e^{t\alpha V_0+\sqrt{t}\sum_{j=1}^{d} G_{1,j}V_j}(x))\right) = \varphi(x) + tL\varphi(x) + \frac{t^2}{2}L^2\varphi(x) + \mathcal{O}(t^3).
$$

(d) Quel est le schéma de Ninomiya et Ninomiya ?

3.8.15 Estimateur parametrix du prix d'une option vanille

Ce problème est consacré à l'estimateur parametrix introduit dans [10] du prix $\mathbb{E}[f(X_S)]$ d'une option vanille de maturité S déterministe, de fonction de payoff $f : \mathbb{R} \to \mathbb{R}$, C^∞ à dérivées à croissance polynomiale, portant sur le sous-jacent $(X_t)_{t\leq S}$ dont l'évolution temporelle est donnée par l'EDS

$$\begin{cases} dX_t = \sigma(X_t)dW_t + b(X_t)dt \\ X_0 = x_0 \end{cases}$$

où $(W_t)_{t\geq 0}$ est un mouvement brownien standard de dimension 1, $\sigma, b : \mathbb{R} \to \mathbb{R}$ sont des fonctions C^∞ à dérivées bornées et $\inf_{x\in\mathbb{R}} \sigma(x) > 0$. On pose $\boxed{a(x) = \sigma^2(x)}$.

Représentation Poissonnienne d'une somme d'intégrales multiples Soit $(N_t)_{t\geq 0}$ un processus de Poisson de paramètre λ, de temps de sauts $(T_k)_{k\geq 1}$, indépendant de $(W_t)_{t\geq 0}$ et pour tout $k \in \mathbb{N}$, $\psi_k : [0, S]^k \to \mathbb{R}_+$ une fonction mesurable ($\psi_0 \in \mathbb{R}_+$ est une constante).

1. On pose $T_0 = 0$. Que peut-on dire des variables aléatoires $(\tau_k = T_k - T_{k-1})_{k\geq 1}$?
2. Soit $n \in \mathbb{N}^*$. Exprimer à l'aide du vecteur $(\tau_1, \ldots, \tau_{n+1})$ la variable aléatoire $1_{\{N_S=n\}}\psi_n(T_1, \ldots, T_n)$. En déduire que
$$\mathbb{E}\left[1_{\{N_S=n\}}\psi_n(T_1, \ldots, T_n)\right] = e^{-\lambda S} \int_{0\leq t_1\leq t_2\leq \ldots\leq t_n\leq S} \lambda^n \psi_n(t_1, \ldots, t_n)dt_1 \ldots dt_n.$$
3. En déduire, avec la convention $\psi_0(T_1, \ldots, T_0) = \psi_0$, que
$$\begin{aligned} &e^{\lambda S}\mathbb{E}\left[\psi_{N_S}(T_1, \ldots, T_{N_s})\right] \\ &\quad = \psi_0 + \sum_{n\in\mathbb{N}^*} \int_{0\leq u_n\leq u_{n-1}\leq \ldots\leq u_1\leq S} \lambda^n \psi_n(S - u_1, \ldots, S - u_n)du_n \ldots du_1. \end{aligned}$$

Méthode parametrix Pour $s \geq 0$ et $x \in \mathbb{R}$, on note X^x_s la solution à l'instant s de l'EDS
$$X^x_s = x + \int_0^s \sigma(X^x_u)dW_u + \int_0^s b(X^x_u)du$$
et $\bar{X}^x_s = x + \sigma(x)W_s + b(x)s$. Pour $\varphi : \mathbb{R} \to \mathbb{R}$ régulière, on note également, $\bar{P}_s\varphi(x) = \mathbb{E}[\varphi(\bar{X}^x_s)]$ et $P_s\varphi(x)$ la solution de l'EDP

$$\begin{cases} \partial_s P_s\varphi(x) = L(P_s\varphi)(x),\ (s, x) \in \mathbb{R}_+ \times \mathbb{R} \\ P_0\varphi(x) = \varphi(x),\ x \in \mathbb{R} \end{cases} \quad \text{où } L\psi(x) = \frac{a(x)}{2}\psi''(x) + b(x)\psi'(x).$$

On admet que cette solution est C^∞ avec des dérivées bornées.

4. Soit $t > 0$ et $x \in \mathbb{R}$. Calculer $dP_{t-s}\varphi(X^x_s)$ par la formule d'Itô et en déduire que $P_t\varphi(x) = \mathbb{E}[\varphi(X^x_t)]$.
5. Vérifier que pour $s > 0$, $\bar{X}^x_s$ possède la densité $g_s(x, y) = \frac{1}{\sigma(x)\sqrt{2\pi s}} e^{-\frac{(y-x-b(x)s)^2}{2a(x)s}}$. Calculer $\partial_s \ln(g_s(x, y))$ et en déduire que $\partial_s\bar{P}_s\varphi(x) = \int_{\mathbb{R}} \varphi(y)\bar{\theta}_s(x, y)g_s(x, y)dy$ où
$$\bar{\theta}_s(x, y) = -\frac{1}{2s} + \frac{(y - x - b(x)s)^2}{2a(x)s^2} + \frac{b(x)(y - x - b(x)s)}{a(x)s}.$$

6. Calculer $\partial_y \ln(g_s(x,y))$ et vérifier que

$$\partial^2_{yy} g_s(x,y) = \left(\frac{(y-x-b(x)s)^2}{a^2(x)s^2} - \frac{1}{a(x)s}\right) g_s(x,y).$$

Expliquer, sans détailler les calculs, comment en déduire que $\bar{P}_s L\psi(x) = \int_{\mathbb{R}} \psi(y)\tilde{\theta}_s(x,y)g_s(x,y)dy$ où

$$\tilde{\theta}_s(x,y) = \frac{1}{2}a''(y) - b'(y) + \frac{(y-x-b(x)s)(b-a')(y)}{a(x)s} + \frac{a(y)}{2}\left(\frac{(y-x-b(x)s)^2}{a^2(x)s^2} - \frac{1}{a(x)s}\right).$$

7. Remarquer que pour $t \geq u \geq 0$, $\partial_u(\bar{P}_{t-u}P_u\varphi(x)) = -(\partial_s \bar{P}_s|_{s=t-u})P_u\varphi(x) + \bar{P}_{t-u}LP_u\varphi(x)$ et en déduire que

$$P_t\varphi(x) - \bar{P}_t\varphi(x) = \int_0^t Q_{t-u}P_u\varphi(x)du$$

où $Q_s\psi(x) = \int_{\mathbb{R}} \psi(y)\theta_s(x,y)g_s(x,y)dy$ pour une fonction θ_s que l'on exprimera à l'aide de $\bar{\theta}_s$ et $\tilde{\theta}_s$.
8. Montrer que $(P_t - \bar{P}_t)\varphi(x) = \int_0^t Q_{t-u}\bar{P}_u\varphi(x)du + \int_0^t Q_{t-u}(P_u - \bar{P}_u)\varphi(x)du$.
9. En déduire que pour tout $m \in \mathbb{N}^*$,

$$\mathbb{E}[f(X_S)] = \bar{P}_S f(x_0) + \sum_{n=1}^{m} \int_{0\leq u_n\leq u_{n-1}\leq \ldots \leq u_1 \leq S} Q_{S-u_1}Q_{u_1-u_2}\ldots Q_{u_{n-1}-u_n}\bar{P}_{u_n} f(x_0)du_1\ldots du_n + R_m$$

où

$$R_m = \int_{0\leq u_m\leq u_{m-1}\leq \ldots \leq u_1 \leq S} Q_{S-u_1}Q_{u_1-u_2}\ldots Q_{u_{m-1}-u_m}(P_{u_m} - \bar{P}_{u_m})f(x_0)du_1\ldots du_m.$$

10. En admettant que $\lim_{m\to\infty} R_m = 0$, en déduire à l'aide de la question 3 que

$$\mathbb{E}[f(X_S)] = e^{\lambda S}\mathbb{E}\left[\lambda^{-N_S}Q_{T_1}Q_{T_2-T_1}\ldots Q_{T_{N_S}-T_{N_S-1}}\bar{P}_{S-T_{N_S}}f(x_0)\right],$$

où, par convention, $Q_{T_1}Q_{T_2-T_1}\ldots Q_{T_{N_S}-T_{N_S-1}}\bar{P}_{S-T_{N_S}}f(x_0) = \bar{P}_{S-T_0}f(x_0) = \bar{P}_S f(x_0)$ si $N_S = 0$.

On définit par récurrence

$$\begin{cases} \bar{X}_0 = x_0 \\ \forall k \in \{0,\ldots,N_S-1\},\ \bar{X}_{T_{k+1}} = \bar{X}_{T_k} + \sigma(\bar{X}_{T_k})(W_{T_{k+1}} - W_{T_k}) + b(\bar{X}_{T_k})(T_{k+1}-T_k) \\ \bar{X}_S = \bar{X}_{T_{N_S}} + \sigma(\bar{X}_{T_{N_S}})(W_S - W_{T_{N_S}}) + b(\bar{X}_{T_{N_S}})(S - T_{N_S}) \end{cases}$$

où la seconde étape n'a pas lieu lorsque $N_S = 0$.

11. Quelle est la loi conditionnelle de $W_S - W_{T_{N_S}}$ sachant $(N_t)_{t\leq S}$ et $(W_t)_{t\leq T_{N_S}}$? En déduire que

$$\mathbb{E}\left[f(\bar{X}_S)|(N_t)_{t\leq S},(W_t)_{t\leq T_{N_S}}\right] = \bar{P}_{S-T_{N_s}}f(\bar{X}_{T_{N_s}}),$$

et sur $\{N_S \geq 1\}$,

$$\mathbb{E}\left[f(\bar{X}_S)\theta_{T_{N_S}-T_{N_S-1}}(\bar{X}_{T_{N_S-1}}\bar{X}_{T_{N_S}})|(N_t)_{t\leq S},(W_t)_{t\leq T_{N_S-1}}\right] = Q_{T_{N_S}-T_{N_S-1}}\bar{P}_{S-T_{N_s}}f(\bar{X}_{T_{N_s-1}}).$$

En déduire que sur $\{N_S \geq 1\}$,

$$\mathbb{E}\left[f(\bar{X}_S)\prod_{k=1}^{N_S}\theta_{T_k-T_{k-1}}(\bar{X}_{T_{k-1}},\bar{X}_{T_k})\middle|(N_t)_{t\leq S}\right] = Q_{T_1}Q_{T_2-T_1}\dots Q_{T_{N_S}-T_{N_S-1}}\bar{P}_{S-T_{N_S}}f(x_0).$$

12. Conclure que

$$\mathbb{E}[f(X_S)] = e^{\lambda S}\mathbb{E}\left[\lambda^{-N_S}f(\bar{X}_S)\prod_{k=1}^{N_S}\theta_{T_k-T_{k-1}}(\bar{X}_{T_{k-1}},\bar{X}_{T_k})\right]$$

avec la convention que le produit vaut 1 lorsque $N_S = 0$.

3.8.16 Schéma d'Euler avec acceptation de Metropolis-Hastings

Sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, on se donne un mouvement brownien $(W_t)_{t\geq 0}$ de dimension n et un vecteur aléatoire X_0 à valeurs dans $\mathbb{R}^n$ indépendants. Ce problème est consacré à la simulation de l'Équation Différentielle Stochastique

$$X_t = X_0 + W_t - \int_0^t \nabla V(X_s)ds$$

où $V : \mathbb{R}^n \to \mathbb{R}$ une fonction régulière telle que $\int_{\mathbb{R}^n} e^{-2V(x)}dx = 1$. On suppose en particulier que V est C^2 et que ∇V et $\nabla^2 V$ sont lipschitziennes et bornées.
Dans la seconde partie, on montrera que si X_0 possède la densité $e^{-2V(x)}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$, alors pour tout $t > 0$, X_t possède également la densité $e^{-2V(x)}$. La première partie est consacrée à l'analyse de la convergence forte du schéma d'Euler où, à chaque pas, la nouvelle position est acceptée ou rejetée suivant la règle de Metropolis-Hastings, ce qui permet d'obtenir une dynamique qui préserve la densité invariante $e^{-2V(x)}$.

Schéma d'Euler On suppose que $X_0 = x_0$ où $x_0 \in \mathbb{R}^n$ est déterministe.

1. Écrire le schéma d'Euler $(\bar{X}_{t_k})_{0\leq k\leq N}$ sur la grille $(t_k = \frac{kT}{N})_{0\leq k\leq N}$ où $T \in]0, +\infty[$ est l'horizon de simulation et $N \in \mathbb{N}^*$ le nombre de pas de temps.
2. Quelle est la vitesse forte de ce schéma pour l'EDS considérée ?
3. Quelle est la loi de $\bar{X}_{t_1}$? Donner sa densité que l'on notera $x_1 \to q(x_0, x_1)$.
4. Exprimer à l'aide de

$$\beta(x_0, x_1) = 2V(x_0) + \frac{1}{2t_1}|x_1 - x_0 + \nabla V(x_0)t_1|^2 - 2V(x_1) - \frac{1}{2t_1}|x_0 - x_1 + \nabla V(x_1)t_1|^2$$

le taux d'acceptation $\alpha(x_0, x_1)$ de l'algorithme de Metropolis-Hastings avec noyau de proposition de densité $q(x_0, x_1)$ et densité cible $e^{-2V(.)}$.

5. Montrer que

$$\exists C < \infty,\ \forall N \in \mathbb{N}^*,\ \forall x_0 \in \mathbb{R}^n, \mathbb{E}(|\bar{X}_{t_1} - x_0|^5) \leq CN^{-5/2}.$$

6. Remarquer que $x \mapsto |\nabla V(x)|^2$ est Lipschitzienne, vérifier que

$$2V(x_0) - 2V(x_1) + (x_1 - x_0)^t(\nabla V(x_0) + \nabla V(x_1))$$
$$= \frac{1}{2}\int_0^1 (1-u)(x_1-x_0)^t \left(\nabla^2 V(x_1 - \frac{u}{2}(x_1-x_0)) - \nabla^2 V(x_0 + \frac{u}{2}(x_1 - x_0))\right)(x_1-x_0)du,$$

et en déduire que

$$\exists C < \infty,\ \forall N \in \mathbb{N}^*,\ \forall x_0, x_1 \in \mathbb{R}^n,\ |\beta(x_0,x_1)| \le C(|x_1-x_0|^3 + |x_1 - x_0|/N).$$

En remarquant que $\forall x \in \mathbb{R},\ (1-e^x)^+ \le \max(-x,0) \le |x|$, conclure que

$$\exists C < \infty,\ \forall N \in \mathbb{N}^*,\ \forall x_0, x_1 \in \mathbb{R}^n,\ 1 - \alpha(x_0,x_1) \le C(|x_1-x_0|^3 + |x_1 - x_0|/N).$$

7. Si U est variable uniforme sur $[0,1]$ indépendante de $(W_t)_{t\ge 0}$ et $\hat{X}_{t_1} = x_0 1_{\{U > \alpha(x_0, \bar{X}_{t_1})\}} + \bar{X}_{t_1} 1_{\{U \le \alpha(x_0,\bar{X}_{t_1})\}}$, montrer que

$$\exists C < \infty,\ \forall N \in \mathbb{N}^*,\ \forall x_0 \in \mathbb{R}^n,\ \mathbb{E}\left(|\bar{X}_{t_1} - \hat{X}_{t_1}|^2\right) \le CN^{-5/2}.$$

8. Soit maintenant $\bar{Y}_{t_1} = y_0 + W_{t_1} - \nabla V(y_0)t_1$ où $y_0 \in \mathbb{R}^n$. On pose

$$f(x_0,y_0) = \mathbb{E}\left(|\hat{X}_{t_1} - \bar{Y}_{t_1}|^2\right).$$

En décomposant sur les événements $\{U \le \alpha(x_0,\bar{X}_{t_1})\}$ et $\{U > \alpha(x_0,\bar{X}_{t_1})\}$, montrer que

$$|\hat{X}_{t_1} - \bar{Y}_{t_1}|^2 \le (1 + 1_{\{U>\alpha(x_0,\bar{X}_{t_1})\}})|\bar{X}_{t_1} - \bar{Y}_{t_1}|^2 + 2|\hat{X}_{t_1} - \bar{X}_{t_1}|^2$$

et en déduire que

$$\exists \mathcal{C} < \infty,\ \forall N \in \mathbb{N}^*,\ \forall x_0, y_0 \in \mathbb{R}^n,\ f(x_0,y_0) \le (1 + \mathcal{C}/N)|x_0 - y_0|^2 + \mathcal{C}N^{-5/2}.$$

On introduit le schéma d'Euler avec acceptation de Metropolis-Hastings défini par $\hat{X}_0 = x_0$ et la relation de récurrence

$$\forall k \in \{0,\ldots,N-1\},\ \begin{cases} \tilde{X}_{t_{k+1}} = \hat{X}_{t_k} + W_{t_{k+1}} - W_{t_k} - \nabla V(\hat{X}_{t_k})t_1 \\ \hat{X}_{t_{k+1}} = \hat{X}_{t_k} 1_{\{U_{k+1} > \alpha(\hat{X}_{t_k}, \tilde{X}_{t_{k+1}})\}} + \tilde{X}_{t_{k+1}} 1_{\{U_{k+1} \le \alpha(\hat{X}_{t_k},\tilde{X}_{t_{k+1}})\}} \end{cases}$$

où $(U_k)_{1\le k\le N}$ est une suite indépendante de $(W_t)_{t\ge 0}$ de variables aléatoires indépendantes et uniformes sur $[0,1]$. L'intérêt de l'acceptation de Metropolis-Hastings est d'assurer que la densité $e^{-2V(x)}$ est invariante.

9. Montrer que pour $k \in \{0,\ldots,N-1\}$, $\mathbb{E}\left(|\hat{X}_{t_{k+1}} - \bar{X}_{t_{k+1}}|^2\right) = \mathbb{E}\left(f(\hat{X}_{t_k}, \bar{X}_{t_k})\right)$ et en déduire que

$$\forall k \in \{1,\ldots,N\},\ \mathbb{E}\left(|\hat{X}_{t_k} - \bar{X}_{t_k}|^2\right) \le \mathcal{C}N^{-5/2}\sum_{j=0}^{k-1}(1+\mathcal{C}/N)^j.$$

Conclure que

$$\exists C < \infty,\ \forall N \in \mathbb{N}^*,\ \sup_{k\in\{0,\ldots,N\}} \mathbb{E}(|\hat{X}_{t_k} - X_{t_k}|^2) \le CN^{-3/2}.$$

Densité invariante de l'EDS

10. Que peut-on dire de $(Z_s = X_s - X_0)_{s\in[0,t]}$ sous la probabilité $\mathbb{Q}$ de densité $\mathcal{E}_t = e^{\int_0^t \nabla V(X_s).dW_s - \frac{1}{2}\int_0^t |\nabla V(X_s)|^2 ds}$ par rapport à $\mathbb{P}$?
11. Pour $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ mesurable bornée, vérifier que
$$\mathbb{E}(\varphi(X_0, X_t)) = \mathbb{E}^{\mathbb{Q}}\left(\varphi(X_0, X_0 + Z_t)e^{-\int_0^t \nabla V(X_0+Z_s).dZ_s - \frac{1}{2}\int_0^t |\nabla V(X_0+Z_s)|^2 ds}\right)$$
et en déduire que
$$\mathbb{E}(\varphi(X_0, X_t)) = \mathbb{E}\left(\varphi(X_0, X_0 + W_t)e^{V(X_0) - V(X_0+W_t) + \frac{1}{2}\int_0^t (\Delta V - |\nabla V|^2)(X_0+W_s)ds}\right).$$
12. Soit $(B_s)_{s\in[0,t]}$ un mouvement brownien indépendant de $(W_s)_{s\geq 0}$ et de X_0. Pour $0 \leq r \leq s \leq t$, calculer $\mathrm{Cov}(B_r - \frac{r}{t}(B_t - W_t), B_s - \frac{s}{t}(B_t - W_t))$. En déduire que $(\beta_s = B_s - \frac{s}{t}(B_t - W_t))_{s\in[0,t]}$ est un mouvement brownien indépendant de X_0 tel que $\beta_t = W_t$ puis que
$$\mathbb{E}(\varphi(X_0, X_t)) = \mathbb{E}\left(\varphi(X_0, X_0 + W_t)e^{V(X_0) - V(X_0+W_t)}\psi_t(X_0, X_0 + W_t)\right)$$
où $\psi_t(x, y) = \mathbb{E}\left(e^{\frac{1}{2}\int_0^t (\Delta V - |\nabla V|^2)(\frac{t-s}{t}x + \frac{s}{t}y + B_s - \frac{s}{t}B_t)ds}\right)$.

On supose désormais que X_0 possède la densité $e^{-2V(x)}$.

13. Montrer que
$$\mathbb{E}(\varphi(X_0, X_t)) = (2\pi t)^{-n/2}\int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi(x, y)e^{-V(x) - V(y) - \frac{|x-y|^2}{2t}}\psi_t(x, y)dxdy.$$
14. En remarquant que $(B_s)_{s\in[0,t]}$ a même loi que $(B_{t-s} - B_t)_{s\in[0,t]}$, vérifier que $(B_s - \frac{s}{t}B_t)_{s\in[0,t]}$ a même loi que $(B_{t-s} - \frac{t-s}{t}B_t)_{s\in[0,t]}$. Conclure que $\psi_t(x, y) = \psi_t(y, x)$.
15. En déduire que $\mathbb{E}(\varphi(X_0, X_t)) = \mathbb{E}(\varphi(X_t, X_0))$ et conclure que X_t possède la densité $e^{-2V(x)}$.

3.8.17 Théorème de Wong-Zakai

Sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, on se donne un mouvement brownien $(W_t)_{t\geq 0}$ de dimension d. Pour un horizon $T \in]0, +\infty[$ et un nombre $N \in \mathbb{N}^*$ de pas, on considère la grille uniforme $(t_k = \frac{kT}{N})_{0\leq k\leq N}$. Pour discrétiser la solution de l'EDS
$$dX_t = \sigma(X_t)dW_t + b(X_t)dt,\ X_0 = x_0 \tag{3.60}$$
où $x_0 \in \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$ et $b : \mathbb{R}^n \to \mathbb{R}^n$ sont supposées lipschitziennes, on peut être tenté d'approcher la trajectoire brownienne par interpolation linéaire sur chaque intervalle $[t_k, t_{k+1}]$ en posant $W_t^N = W_{t_k} + \frac{W_{t_{k+1}} - W_{t_k}}{t_{k+1} - t_k}(t - t_k)$ puis de résoudre
$$dX_t^N = \sigma(X_t^N)dW_t^N + b(X_t^N)dt,\ X_0^N = x_0. \tag{3.61}$$

1. Vérifier que pour tout $k \in \{0, \ldots, N-1\}$, $(X_t^N)_{t\in[t_k,t_{k+1}]}$ satisfait une équation différentielle ordinaire et en déduire que (3.61) admet une unique solution.

On se place d'abord en dimension $n = d = 1$.

2. On suppose en outre que $\sigma : \mathbb{R} \to \mathbb{R}_+^*$ est C^1 et $b \equiv 0$ et on pose $\varphi(x) = \int_{x_0}^{x} \frac{dy}{\sigma(y)}$.
   (a) Vérifier que pour tout $k \in \{0, \dots, N\}$, $\varphi(X_{t_k}^N) = W_{t_k}$ et en déduire que X_T^N ne dépend pas de N. On note Y_T cette valeur.
   (b) Dans le cas $\sigma(x) \equiv x$, calculer φ et vérifier que $Y_T \neq X_T$. En déduire que le procédé d'approximation ne converge pas vers la solution de (3.60).
   (c) Calculer $(\varphi^{-1})'$ et $(\varphi^{-1})''$. En déduire que $(Y_t = \varphi^{-1}(W_t))_{t\in[0,T]}$ est solution de l'EDS $dY_t = \sigma(Y_t)dW_t + \frac{1}{2}\sigma\sigma'(Y_t)dt,\ Y_0 = x_0$.
3. On suppose que le processus d'Itô $(Y_t)_{t\in[0,T]}$ est solution de l'EDS au sens de Stratonovitch
$$dY_t = \sigma(Y_t) \circ dW_t + b(Y_t)dt,\ Y_0 = x_0 \tag{3.62}$$
   où, pour un processus d'Itô $(H_s)_{s\in[0,T]}$ et $t \in [0, T]$, on définit l'intégrale de Stratonovitch $\int_0^t H_s \circ dW_s$ comme la limite en probabilité lorsque $n \to \infty$ de $\sum_{k=0}^{N-1} \frac{H_{t_{k+1}}+H_{t_k}}{2}(W_{t_{k+1}\wedge t} - W_{t_k\wedge t})$. La fonction $\sigma : \mathbb{R} \to \mathbb{R}$ est supposée C^2 avec $\sigma\sigma'$ lipschitzienne.
   (a) Quelle est la limite en probabilité de $\sum_{k=0}^{N-1} H_{t_k}(W_{t_{k+1}\wedge t} - W_{t_k\wedge t})$? En déduire que $\int_0^t H_s \circ dW_s = \int_0^t H_s dW_s + \frac{1}{2}\langle H, W\rangle_t$.
   (b) En déduire la partie intégrale par rapport à dW_t de la décomposition de $\sigma(Y_t)$ comme processus d'Itô puis $d\langle\sigma(X), W\rangle_t$.
   (c) Conclure que $(Y_t)_{t\in[0,T]}$ est solution de l'EDS au sens d'Itô
$$dY_t = \sigma(Y_t)dW_t + \left(b + \frac{\sigma\sigma'}{2}\right)(Y_t)dt,\ Y_0 = x_0. \tag{3.63}$$
4. On suppose que σ est C^2 avec σ, σ' et σ'' bornées et que b est C^1 avec b et b' bornées. L'objectif de cette question est de montrer que sous cette hypothèse
$$\exists C < \infty,\ \forall N \in \mathbb{N}^*,\ \max_{0\le k\le N} \mathbb{E}\left[|X_{t_k}^N - Y_{t_k}|^2\right] \le \frac{C}{N}. \tag{3.64}$$
   (a) Écrire le schéma d'Euler $(Y_{t_k}^N)_{0\le k\le N}$ appliqué à (3.63) et préciser son erreur forte.
   (b) Vérifier que pour $t \in [t_k, t_{k+1}]$, $dX_t^N = f_k(X_t^N)dt$ avec $f_k(x) = \sigma(x)\frac{W_{t_{k+1}}-W_{t_k}}{t_{k+1}-t_k} + b(x)$.
   (c) En déduire que pour $t \in [t_k, t_{k+1}]$,
$$X_t^N = X_{t_k}^N + f_k(X_{t_k}^N)(t - t_k) + \int_{t_k}^{t}\int_{t_k}^{s} f_k f_k'(X_r^N)drds$$
   et que pour $r \in [t_k, t_{k+1}]$,
$$\sigma\sigma'(X_r^N) = \sigma\sigma'(X_{t_k}^N) + \int_{t_k}^{r} f_k(\sigma\sigma')'(X_u^N)du.$$
   (d) Conclure que pour $k \in \{0, \dots, N-1\}$,
$$X_{t_{k+1}}^N = X_{t_k}^N + \sigma(X_{t_k}^N)(W_{t_{k+1}} - W_{t_k}) + \left(b + \frac{\sigma\sigma'}{2}\right)(X_{t_k}^N)\frac{T}{N} + R_{k+1}^{N,1} + R_{k+1}^{N,2},$$

avec $R^{N,1}_{k+1} = \frac{\sigma\sigma'}{2}(X^N_{t_k})((W_{t_{k+1}} - W_{t_k})^2 - (t_{k+1} - t_k))$ et

$$R^{N,2}_{k+1} = \int_{t_k}^{t_{k+1}} (t_{k+1} - r)\left(\frac{W_{t_{k+1}} - W_{t_k}}{t_{k+1} - t_k}(b\sigma' + \sigma b')(X^N_r) + bb'(X^N_r)\right) dr$$
$$+ \int_{t_k}^{t_{k+1}} \frac{(t_{k+1} - u)^2(W_{t_{k+1}} - W_{t_k})^2}{2(t_{k+1} - t_k)^2}\left(\frac{W_{t_{k+1}} - W_{t_k}}{t_{k+1} - t_k}\sigma(\sigma\sigma')'(X^N_u) + b(\sigma\sigma')'(X^N_u)\right) du$$

(e) Que valent $\mathbb{E}\left[\left(W_{t_{k+1}} - W_{t_k}\right)^2\right]$, $\mathbb{E}\left[\left(W_{t_{k+1}} - W_{t_k}\right)^4\right]$ et $\mathbb{E}\left[\left(W_{t_{k+1}} - W_{t_k}\right)^6\right]$?

(f) Que vaut $\mathbb{E}[R^{N,1}_{j+1}R^{N,1}_{k+1}]$ pour $j < k$? En déduire que $\forall \ell \in \{0, \dots, N-1\}$,

$$\mathbb{E}\left[\left(\sum_{k=0}^{\ell} R^{N,1}_{k+1}\right)^2\right] \leq \|\sigma\sigma'\|^2_\infty \frac{(\ell+1)T^2}{2N^2} \leq \frac{\|\sigma\sigma'\|^2_\infty T^2}{2N}.$$

(g) Vérifier que pour $k \in \{0, \dots, N-1\}$, $\mathbb{E}\left[(R^{N,2}_{k+1})^2\right]$ est majoré par

$$\|b\sigma' + \sigma b'\|^2_\infty \frac{T^3}{N^3} + \|bb'\|^2_\infty \frac{T^4}{N^4} + \|\sigma(\sigma\sigma')'\|^2_\infty \frac{5T^3}{3N^3} + \|b(\sigma\sigma')'\|^2_\infty \frac{T^4}{3N^4}.$$

(h) En déduire l'existence d'une constante $C_R < \infty$ telle que

$$\forall N \in \mathbb{N}^*,\ \forall \ell \in \{0, \dots, N-1\},\ \mathbb{E}\left[\left(\sum_{k=0}^{\ell}(R^{N,1}_{k+1} + R^{N,2}_{k+1})\right)^2\right] \leq \frac{C_R}{N}.$$

On pose $e(t) = \sum_{\ell=0}^{N-1} \mathbb{E}\left[(X^N_{t_\ell} - Y^N_{t_\ell})^2\right] 1_{[t_\ell, t_{\ell+1}[}(t) + \mathbb{E}\left[(X^N_T - Y^N_T)^2\right] 1_{\{t=T\}}$ et $\tilde{b} = b + \frac{\sigma\sigma'}{2}$.

(i) Montrer que pour $\ell \in \{1, \dots, N\}$,

$$\mathbb{E}\left[(X^N_{t_\ell} - Y^N_{t_\ell})^2\right] \leq \frac{3T}{N}\sum_{k=0}^{\ell-1}\mathbb{E}\left[(\sigma(X^N_{t_k}) - \sigma(Y^N_{t_k}))^2 + T(\tilde{b}(X^N_{t_k}) - \tilde{b}(Y^N_{t_k}))^2\right] + \frac{3C_R}{N}.$$

(j) En déduire que $\forall t \in [0, T],\ e(t) \leq \frac{3C_R}{N} + 3(\|\sigma'\|^2_\infty + T\|\tilde{b}'\|^2_\infty)\int_0^t e(s)ds$ puis que

$$\max_{0\leq k\leq N} \mathbb{E}\left[|X^N_{t_k} - Y^N_{t_k}|^2\right] \leq \frac{3C_R}{N} e^{3(\|\sigma'\|^2_\infty + T\|\tilde{b}'\|^2_\infty)T}.$$

(k) Conclure que (3.64) est vraie.

On se place désormais en dimension quelconque et on suppose toujours σ bornée ainsi que ses dérivées jusqu'à l'ordre 2 et $b : \mathbb{R}^n \to \mathbb{R}^n$ C^1 bornée ainsi que ses dérivées d'ordre 1. Pour $j \in \{1, \dots, d\}$ et $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, on note $\sigma_j(x)$ la j-ème colonne de la matrice $\sigma(x)$ et $\partial\sigma_j(x)$ la matrice dont le terme d'indices $i, l \in \{1, \dots, n\}$ vaut $(\partial\sigma_j(x))_{il} = \frac{\partial\sigma_{ij}(x)}{\partial x_l}$.

5. Donner la forme Itô de l'EDS (3.62).
6. Vérifier que pour $j, m \in \{1, \dots, d\}$ avec $j \neq m$, et $(H_{t_k})_{0\leq k\leq N-1}$ une suite de variables aléatoires bornées par une constante C et adaptée à la filtration du mouvement brownien $(W_t)_{t\geq 0}$, pour tout $\ell \in \{0, \dots, N-1\}$, $\mathbb{E}\left[\left(\sum_{k=0}^{\ell}(H_{t_k}(W^j_{t_{k+1}} - W^j_{t_k})(W^m_{t_{k+1}} - W^m_{t_k})\right)^2\right] \leq \frac{(\ell+1)C^2T^2}{N^2}$.
7. En déduire comment montrer que (3.64) reste vraie.

3.8.18 Ordre convexe entre deux diffusions via leurs schémas d'Euler

Ce problème est consacré à l'approche développée dans [63] pour démontrer l'ordre convexe sur l'espace des trajectoires entre les solutions de deux équations différentielles stochastiques à coefficients de dérive nuls et sous des hypothèses de convexité et de comparaison de leurs coefficients de diffusion.

Soit $T > 0$ un horizon de temps. Sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ muni d'une filtration $(\mathcal{F}_t)_{t\geq 0}$, on considère un $\mathcal{F}_t$-mouvement brownien W_t à valeurs $\mathbb{R}^d$. On se donne également des coefficients $\sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ et $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ mesurables et qui vérifient

$$\text{(Lip)}\ \exists \mathrm{K} < \infty,\ \forall \mathrm{t} \in [0, \mathrm{T}],\ \begin{cases} \forall x \in \mathbb{R}^n,\ |\sigma(t,x)| + |b(t,x)| \leq K(1+|x|) \\ \forall x, y \in \mathbb{R}^n,\ |\sigma(t,x) - \sigma(t,y)| + |b(t,x) - b(t,y)| \leq K|x-y| \end{cases}.$$

On s'intéresse à l'Équation Différentielle Stochastique

$$\begin{cases} dX_t = \sigma(t, X_t) dW_t + b(t, X_t) dt,\ t \in [0,T] \\ X_0 = z \in \mathbb{R}^n \end{cases} \tag{3.65}$$

1. Quel résultat assure l'existence d'une unique solution $X = (X_t)_{t\in[0,T]}$ à (3.65) ?

Soit $N \in \mathbb{N}^*$, $(t_k = \frac{kT}{N})_{0\leq k\leq N+1}$ et $\tau_t = \lfloor \frac{Nt}{T} \rfloor \frac{T}{N}$ l'instant de cette grille uniforme qui précède $t \in [0,T]$.

2. Écrire le schéma d'Euler en temps continu $(X_t^N)_{t\in[0,T]}$ de pas T/N.

On suppose désormais qu'il existe $\alpha \in]0,1]$ et $K < \infty$ tels que

$$\forall x \in \mathbb{R}^n,\ \forall 0 \leq s \leq t \leq T,\ |\sigma(t,x) - \sigma(s,x)| + |b(t,x) - b(s,x)| \leq K(1+|x|)(t-s)^\alpha.$$

3. Rappeler le résultat de convergence forte de ce schéma.

Pour $x_{0:N} \in \mathbb{R}^{N+1}$ on note $\mathcal{L}_N[x_{0:N}]$ la fonction définie par

$$\forall t \in [0,T],\ \mathcal{L}_N[x_{0:N}](t) = \left(1 - \frac{N}{T}(t-\tau_t)\right) x_{\lfloor \frac{Nt}{T} \rfloor} + \frac{N}{T}(t-\tau_t) x_{1+\lfloor \frac{Nt}{T} \rfloor}.$$

Enfin, pour $k \in \{0, \ldots, N\}$, on note $X_{t_{0:k}} = (X_{t_0}, X_{t_1}, \ldots, X_{t_k})$ et $X^N_{t_{0:k}} = (X^N_{t_0}, X^N_{t_1}, \ldots, X^N_{t_k})$.

4. Montrer que $\sup_{t\in[0,T]} |\mathcal{L}_N[X_{t_{0:N}}](t) - X_t| \leq \sup_{s,t\in[0,T]:|t-s|\leq T/N} |X_s - X_t|$ et en déduire que $\mathbb{E}[\sup_{t\in[0,T]} |\mathcal{L}_N[X_{t_{0:N}}](t) - X_t|]$ tend vers 0 lorsque $N \to \infty$.
5. Montrer que pour $x_{0:N}, y_{0:N} \in \mathbb{R}^{N+1}$, $\sup_{t\in[0,T]} |\mathcal{L}_N[x_{0:N}](t) - \mathcal{L}_N[y_{0:N}](t)| = \max_{0\leq k\leq N} |x_k - y_k|$.
   En déduire que $\mathbb{E}[\sup_{t\in[0,T]} |\mathcal{L}_N[X^N_{t_{0:N}}](t) - X_t|]$ tend vers 0 lorsque $N \to \infty$.
6. Conclure que pour toute fonctionnelle $\Phi : \mathcal{C}([0,T], \mathbb{R}^n) \to \mathbb{R}$ lipschitzienne sur l'espace $\mathcal{C}([0,T], \mathbb{R}^n)$ des fonctions g continues de $[0,T]$ dans $\mathbb{R}^n$ muni de la norme $\sup_{t\in[0,T]} |g(t)|$, $\mathbb{E}[\Phi(\mathcal{L}_N[X^N_{t_{0:N}}])]$ tend vers $\mathbb{E}[\Phi(X)]$ lorsque $N \to \infty$.

On se place désormais en dimension $\boxed{n = d = 1}$ et on suppose que le coefficient de dérive b est nul, et que pour tout $t \in [0, T]$, $\mathbb{R} \ni x \mapsto \sigma(t, x)$ est convexe **et à valeurs dans** $\mathbb{R}_+$.

Soit $\Phi_N : \mathbb{R}^{N+1} \to \mathbb{R}$ une fonction convexe à croissance affine. Pour $k \in \{0, \ldots, N-1\}$, et $(x_{0:k}, u) \in \mathbb{R}^{k+1} \times \mathbb{R}$ avec $x_{0:k} = (x_0, x_1, \ldots, x_k)$, on définit par récurrence descendante $\Psi_k(x_{0:k}, u) = \mathbb{E}[\Phi_{k+1}(x_{0:k}, x_k + u(W_{t_{k+1}} - W_{t_k}))]$ et $\Phi_k(x_{0:k}) = \Psi_k(x_{0:k}, \sigma(t_k, x_k))$.

7. Vérifier que pour $k \in \{0, \ldots, N-1\}$, $\mathbb{E}[\Phi_{k+1}(X^N_{t_{0:k+1}})|X^N_{t_{0:k}}] = \Phi_k(X^N_{t_{0:k}})$. En déduire que $\Phi_0(z) = \mathbb{E}\left[\Phi_N(X^N_{t_{0:N}})\right]$.
8. Vérifier que pour $x_{0:k} \in \mathbb{R}^{k+1}$ et $0 \leq u \leq v$,
$$\Psi_k(x_{0:k}, v) = \mathbb{E}[\Phi_{k+1}(x_{0:k}, x_k + u(W_{t_{k+1}} - W_{t_k}) + \sqrt{v^2 - u^2}(W_{t_{k+2}} - W_{t_k+1}))]$$
et en déduire que si Φ_{k+1} est convexe alors
$$\forall x_{0:k} \in \mathbb{R}^{k+1},\ \forall 0 \leq u \leq v,\ \Psi_k(x_{0:k}, u) \leq \Psi_k(x_{0:k}, v).$$
9. Vérifier également que la convexité de Φ_{k+1} entraîne celle de Ψ_k.
10. Montrer que pour tout $k \in \{0, \ldots, N-1\}$, Φ_k est convexe. *On pourra comparer* $\Phi_k(\alpha x_{0:k} + (1-\alpha) y_{0:k})$ *à* $\alpha\Phi_k(x_{0:k}) + (1-\alpha)\Phi_k(y_{0:k})$ *avec* $\alpha \in [0, 1]$, $x_{0:k}, y_{0:k} \in \mathbb{R}^{k+1}$.

On note Y^N le schéma d'Euler de pas $\frac{T}{N}$ pour l'équation différentielle stochastique
$$dY_t = \eta(t, Y_t)dW_t,\ Y_0 = z \text{ où } \eta : [0, T] \times \mathbb{R} \to \mathbb{R}_+ \text{ est telle que}$$
$$\exists K < \infty, \forall t \in [0, T],\ \forall x \in \mathbb{R},\ |\eta(t, x)| \leq K(1 + |x|) \text{ et } \forall y \in \mathbb{R},\ |\eta(t, x) - \eta(t, y)| \leq K|x - y|,$$
$$\exists \alpha \in]0, 1],\ \forall x \in \mathbb{R},\ \forall 0 \leq s \leq t \leq T,\ |\eta(t, x) - \eta(s, x)| \leq K(1 + |x|)(t - s)^\alpha.$$

Partant de $\tilde{\Phi}_N = \Phi_N$, on définit par récurrence descendante $\tilde{\Psi}_k(x_{0:k}, u) = \mathbb{E}[\tilde{\Phi}_{k+1}(x_{0:k}, x_k + u(W_{t_{k+1}} - W_{t_k}))]$ et $\tilde{\Phi}_k(x_{0:k}) = \tilde{\Psi}_k(x_{0:k}, \eta(t_k, x_k))$ pour $k \in \{0, \ldots, N-1\}$.

11. Montrer que $\tilde{\Phi}_{k+1} \leq \Phi_{k+1}$ entraîne $\tilde{\Psi}_k \leq \Psi_k$.
12. On suppose que $\forall (t, x) \in [0, T] \times \mathbb{R},\ 0 \leq \eta(t, x) \leq \sigma(t, x)$. Montrer que pour tout $k \in \{0, \ldots, N-1\}$, $\tilde{\Phi}_k \leq \Phi_k$. En déduire que $\mathbb{E}\left[\Phi_N(Y^N_{t_{0:N}})\right] \leq \mathbb{E}\left[\Phi_N(X^N_{t_{0:N}})\right]$.
13. Soit $\Phi : \mathcal{C}([0, T], \mathbb{R}) \to \mathbb{R}$ une fonctionnelle lipschitzienne et convexe. Montrer que $x_{0:N} \mapsto \Phi(\mathcal{L}_N[x_{0:N}])$ est convexe. En déduire que $\mathbb{E}[\Phi(Y)] \leq \mathbb{E}[\Phi(X)]$. Que peut-on dire de $\mathbb{E}[\Phi(X)]$ comme fonction de la condition initiale z de l'EDS (3.65) ?
14. Comment comparer $\mathbb{E}[\Phi(Y)]$ à $\mathbb{E}[\Phi(X)]$ lorsque $\forall (t, x) \in [0, T] \times \mathbb{R},\ \eta(t, x) \geq \sigma(t, x)$?

3.8.19 Schéma d'Euler pour une dérive mesurable bornée

1. Soient $p, q : \mathbb{R}^d \to \mathbb{R}_+$ deux densités de probabilité sur $\mathbb{R}^d$. Vérifier que pour $\varphi : \mathbb{R}^d \to [-1, 1]$ mesurable,
$$\left| \int_{\mathbb{R}^d} \varphi(x)p(x)dx - \int_{\mathbb{R}^d} \varphi(x)q(x)dx \right| \leq \int_{\mathbb{R}^d} |p(x) - q(x)|dx$$
avec le majorant atteint pour le choix $\varphi(x) = 1_{\{p(x)>q(x)\}} - 1_{\{p(x)<q(x)\}}$.

Soit $T > 0$ un horizon de temps. Sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ muni d'une filtration $(\mathcal{F}_t)_{t\geq 0}$, on considère un $\mathcal{F}_t$-mouvement brownien W_t à valeurs $\mathbb{R}^d$. On se donne une fonction de dérive $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, une condition initiale déterministe $y \in \mathbb{R}^d$ et on s'intéresse à l'équation différentielle stochastique

$$X_t = y + W_t + \int_0^t b(s, X_s) ds. \tag{3.66}$$

2. Écrire le schéma d'Euler pour cette EDS. Quelle est sa vitesse de convergence forte lorsque la dérive b est une fonction lipschitzienne en espace et höldérienne d'exposant $1/2$ en temps ? Et lorsque c'est une fonction C^2 à dérivées bornées en espace ne dépendant pas de la variable temporelle ?

Pour le cas où b est une fonction mesurable bornée nous allons nous donner une suite $(U_k)_{k\in\mathbb{N}}$ de variables aléatoires i.i.d. suivant la loi uniforme sur $[0,1]$ indépendante de $(W_t)_{t\geq 0}$ et considérer le schéma d'Euler avec temps randomisé initialisé par $X_0^N = y$ et évoluant pour $k \in \{0, \ldots, N-1\}$ par

$$X^N_{\frac{(k+1)T}{N}} = X^N_{\frac{kT}{N}} + W_{\frac{(k+1)T}{N}} - W_{\frac{kT}{N}} + b\left(\frac{(k+U_k)T}{N}, X^N_{\frac{kT}{N}}\right)\frac{T}{N}.$$

Nous supposerons dans un premier temps que $y = 0$ et la fonction de dérive $b : [0,T] \to \mathbb{R}^d$ ne dépend pas de la variable d'espace et satisfait $B := \sup_{t\in[0,T]} |b(t)| < \infty$ où pour $x \in \mathbb{R}^d$, $|x|$ désigne la norme euclidienne de x. On note $I = \int_0^T b(s)ds$ et $I^N = \frac{T}{N}\sum_{k=0}^{N-1} b\left(\frac{(k+U_k)T}{N}\right)$.

3. Vérifier que $\mathbb{E}\left[I^N\right] = I$.
4. Vérifier que pour $k \in \{0, \ldots, N-1\}$, $\mathbb{E}\left[\left|\frac{T}{N} b\left(\frac{(k+U_k)T}{N}\right)\right|^2\right] \leq \frac{B^2T^2}{N^2}$ et en déduire que $\sum_{j=1}^d \mathrm{Var}\left(\frac{T}{N} b_j\left(\frac{(k+U_k)T}{N}\right)\right) \leq \frac{B^2T^2}{N^2}$ où b_j désigne la j-ème composante de la fonction b.
5. Conclure que
$$\mathbb{E}\left[\left|I^N - I\right|^2\right] \leq \frac{B^2T^2}{N}.$$
6. Soit $a > 0$. Vérifier que pour $x \in \mathbb{R}$, $|x-a| \leq |x+a| \Leftrightarrow x \geq 0$ et en déduire que
$$\int_{\mathbb{R}} |e^{-\frac{(x-a)^2}{2T}} - e^{-\frac{(x+a)^2}{2T}}| \frac{dx}{\sqrt{2\pi T}} = 2\int_{-a}^{a} e^{-\frac{x^2}{2T}} \frac{dx}{\sqrt{2\pi T}} \leq 2a\sqrt{\frac{2}{\pi T}}.$$

Pour $t > 0$ et $x \in \mathbb{R}^d$, on note $G_t(x) = (2\pi t)^{-d/2} e^{-\frac{|x|^2}{2t}}$ la densité au point x de la loi gaussienne centrée de matrice de covariance égale à t fois l'identité en dimension d.

7. Soient maintenant $y, z \in \mathbb{R}^d$ avec $y \neq z$ et $O \in \mathbb{R}^{d\times d}$ une matrice orthogonale dont la première colonne est égale à $e := \frac{y-z}{|y-z|}$. Pour $x \in \mathbb{R}^d$, on pose $w = O^{-1}(x - \frac{y+z}{2})$. Vérifier que $|x-y|^2 = |Ow - \frac{|y-z|}{2} e|^2 = (w_1 - \frac{|y-z|}{2})^2 + \sum_{i=2}^d w_i^2$ où w_i désigne la i-ème coordonnée de w puis que $|x-z|^2 = (w_1 + \frac{|y-z|}{2})^2 + \sum_{i=2}^d w_i^2$. Conclure que
$$\int_{\mathbb{R}^d} |G_T(x-y) - G_T(x-z)| dx \leq \sqrt{\frac{2}{\pi T}} |y-z|.$$
8. Vérifier que $W_T + I$ possède la densité $G_T(x - I)$ et $W_T + I^N$ possède la densité $\mathbb{E}[G_T(x - I^N)]$.

9. Vérifier que

$$\int_{\mathbb{R}^d} |\mathbb{E}[G_T(x - I^N)] - G_T(x - I)|dx \le \sqrt{\frac{2}{\pi T}}\mathbb{E}[|I^N - I|] \le B\sqrt{\frac{2T}{\pi N}}.$$

Dans le cas général où $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ est mesurable bornée, nous allons maintenant établir l'existence et l'unicité en loi pour (3.67).

10. Que peut-on dire de $(\beta_t = W_t - \int_0^t b(s, y + W_s)ds)_{t\in[0,T]}$ sous la probabilité de densité $e^{\int_0^T b(t,y+W_t).dW_t - \frac{1}{2}\int_0^T |b(t,y+W_t)|^2 dt}$ par rapport à $\mathbb{P}$? Vérifier que $(y + W_t)_{t\in[0,T]}$ est solution de l'équation $d(y + W_t) = d\beta_t + b(t, y + W_t)dt$ obtenue en remplaçant $(W_t)_{t\in[0,T]}$ par $(\beta_t)_{t\in[0,T]}$ dans (3.67).
11. Inversement, soit $(X_t)_{t\in[0,T]}$ une solution de l'EDS (3.67) sur l'intervalle de temps $[0, T]$ et $\varphi : \mathcal{C}([0, T], \mathbb{R}^d) \to \mathbb{R}$ mesurable bornée. Que peut-on dire de $(X_t - y)_{t\in[0,T]}$ sous la probabilité $\mathbb{Q}$ de densité de densité $\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{-\int_0^T b(t,X_t).dW_t - \frac{1}{2}\int_0^T |b(t,X_t)|^2 dt}$ par rapport à $\mathbb{P}$? En déduire que

$$\mathbb{E}\left[\varphi((y + W_t)_{t\in[0,T]})e^{\int_0^T b(t,y+W_t).dW_t - \frac{1}{2}\int_0^T |b(t,y+W_t)|^2 dt}\right]$$
$$= \mathbb{E}^{\mathbb{Q}}\left[\varphi((X_t)_{t\in[0,T]})e^{\int_0^T b(t,X_t).dW_t + \frac{1}{2}\int_0^T |b(t,X_t)|^2 dt}\right]$$

puis que

$$\mathbb{E}\left[\varphi((y + W_t)_{t\in[0,T]})e^{\int_0^T b(t,y+W_t).dW_t - \frac{1}{2}\int_0^T |b(t,y+W_t)|^2 dt}\right] = \mathbb{E}\left[\varphi((X_t)_{t\in[0,T]})\right],$$

ce qui établit l'unicité de la loi des solutions de (3.67).
12. Pour $\varphi : \mathbb{R}^d \to \mathbb{R}$ C^2 bornée ainsi que ses dérivées et $t \in]0, T]$, on admet que la fonction $u(s, z) = 1_{[0,t[}(s) \int_{\mathbb{R}^d} G_{t-s}(z - x)\varphi(x)dx + 1_{\{s=t\}}\varphi(z)$ est $C^{1,2}$ bornée ainsi que ses dérivées et solution de l'équation de la chaleur $\partial_s u(s, z) + \frac{1}{2}\Delta u(s, z) = 0$ sur $[0, t] \times \mathbb{R}^d$. Calculer $u(t, X_t) - u(0, y)$ par la formule d'Itô et en déduire que

$$\mathbb{E}[\varphi(X_t)] = \int_{\mathbb{R}^d} \varphi(x)\left(G_t(y - x) + \mathbb{E}\left[\int_0^t b(s, X_s).\nabla_x G_{t-s}(X_s - x)ds\right]\right)dx.$$

Conclure que X_t admet une densité par rapport à la mesure de Lebesgue solution de l'équation

$$p(t, x) = G_t(y - x) + \int_0^t \int_{\mathbb{R}^d} b(s, z).\nabla_x G_{t-s}(z - x)p(s, z)dzds.$$

En établissant une formule analogue pour la densité $p^N(\frac{kT}{N}, x)$ du schéma d'Euler au $k$ème instant de discrétisation avec $k \in \{1, \dots, N\}$, on peut montrer que $\sup_{N\ge1} \sqrt{N} \max_{1\le k\le N} \int_{\mathbb{R}^d} |p^N(\frac{kT}{N}, x) - p(\frac{kT}{N}, x)|dx < \infty$.

3.8.20 Schéma d'Euler pour des coefficients à croissance affine et localement Lipschitziens

Soit $T > 0$ un horizon de temps. Sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ muni d'une filtration $(\mathcal{F}_t)_{t\ge0}$, on considère un $\mathcal{F}_t$-mouvement brownien $W_t = \left(W_t^1, \cdots, W_t^d\right)$ à valeurs

$\mathbb{R}^d$. On se donne un coefficient de diffusion $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$, un coefficient de dérive $b : \mathbb{R}^n \to \mathbb{R}^n$, une condition initiale déterministe $x_0 \in \mathbb{R}^n$ et on s'intéresse à l'équation différentielle stochastique

$$X_t = x_0 + \int_0^t \sigma(X_s)dW_s + \int_0^t b(X_s)ds,\ t \in [0,T]. \tag{3.67}$$

Soit $N \in \mathbb{N}^*$, $(t_k = \frac{kT}{N})_{0\leq k\leq N}$ et $\tau_t = \lfloor \frac{Nt}{T} \rfloor \frac{T}{N}$ l'instant de cette grille uniforme qui précède $t \in [0,T]$.

1. (a) Écrire le schéma d'Euler en temps continu $(\bar{X}_t)_{t\in[0,T]}$ avec N pas pour cette EDS.
   (b) Quelle est sa vitesse de convergence forte lorsque les coefficients σ et b sont lipschitziens ?
   (c) Comment améliorer cette vitesse de convergence forte lorsque σ et b sont plus régulières ?

   On souhaite expliciter la dépendance de la vitesse de convergence du schéma d'Euler dans les constantes de Lipschitz respectives K_σ et K_b de σ et b. On rappelle à cet effet l'inégalité de Young : $\forall q,r > 0,\ \forall a,b \geq 0,\ a^q b^r \leq \frac{q}{q+r} a^{q+r} + \frac{r}{q+r} b^{q+r}$. On note $|x|$ la norme euclidienne de $x \in \mathbb{R}^n$.

   (d) *Dans cette question, on pourra supposer $n = d = 1$.*
   Pour $p \geq 1$ et $u \in [0,T]$, calculer $|X_u - \bar{X}_u|^{2p}$ à l'aide de la formule d'Itô.

   Avec l'inégalité de Burkholder Davis Gundy du lemme 3.1.5 du polycopié qui reste valable pour $p \geq \frac{1}{2}$, on en déduit l'existence d'une constante $C < \infty$ ne dépendant pas de (N, σ, b) telle que pour $t \in [0,T]$, si on pose $I_t = \int_0^t |X_s - \bar{X}_s|^{4p-2} |\sigma(X_s) - \sigma(\bar{X}_{\tau_s})|^2 ds$,

$$\begin{aligned}\mathbb{E}\left[\sup_{u\leq t} |X_u - \bar{X}_u|^{2p}\right] &\leq C \int_0^t \mathbb{E}[|X_s - \bar{X}_s|^{2p-1} |b(X_s) - b(\bar{X}_{\tau_s})|]ds \\ &\quad + C \int_0^t \mathbb{E}[|X_s - \bar{X}_s|^{2p-2} |\sigma(X_s) - \sigma(\bar{X}_{\tau_s})|^2]ds + C\mathbb{E}\left[\sqrt{I_t}\right],\end{aligned}$$

   estimation qui reste vraie pour des dimensions n et d générales.

   (e) Montrer que

$$\sqrt{I_t} \leq \frac{\sup_{u\leq t} |X_u - \bar{X}_u|^{2p}}{2C} + \frac{C}{2} \int_0^t |X_s - \bar{X}_s|^{2p-2} |\sigma(X_s) - \sigma(\bar{X}_{\tau_s})|^2 ds.$$

   et en déduire que

$$\begin{aligned}\mathbb{E}\left[\sup_{u\leq t} |X_u - \bar{X}_u|^{2p}\right] &\leq 2C \int_0^t \mathbb{E}[|X_s - \bar{X}_s|^{2p-1} |b(X_s) - b(\bar{X}_{\tau_s})|]ds \\ &\quad + C(2+C) \int_0^t \mathbb{E}[|X_s - \bar{X}_s|^{2p-2} |\sigma(X_s) - \sigma(\bar{X}_{\tau_s})|^2]ds.\end{aligned}$$

   (f) Montrer que

$$\begin{aligned}|X_s - \bar{X}_s|^{2p-1} |b(X_s) - b(\bar{X}_{\tau_s})| \leq K_b \bigg(&\frac{2p-1}{2p} |X_s - \bar{X}_s|^{2p} + \frac{1}{2p} |X_s - X_{\tau_s}|^{2p} \\ &+ \sup_{u\leq s} |X_u - \bar{X}_u|^{2p} \bigg).\end{aligned}$$

(g) En traitant de manière analogue $|X_s - \bar{X}_s|^{2p-2}|\sigma(X_s) - \sigma(\bar{X}_{\tau_s})|^2$, déduire l'existence d'une constante $C < \infty$ ne dépendant pas de (N, σ, b) telle que

$$\mathbb{E}\left[\sup_{u \le T} |X_u - \bar{X}_u|^{2p}\right] \le C(K_b + K_\sigma^2)e^{C(K_b+K_\sigma^2)T} \int_0^T \mathbb{E}[|X_s - X_{\tau_s}|^{2p}]ds.$$

On suppose désormais que les coefficients σ et b sont à croissance au plus affine et localement lipschitziens

$$\exists K < \infty,\ \forall x \in \mathbb{R}^n,\ |\sigma(x)| + |b(x)| \le K(1 + |x|) \tag{3.68}$$

$$\begin{aligned}&\forall m \in \mathbb{N}^*,\ \exists K_{\sigma,m}, K_{b,m} < \infty, \forall x, y \in \mathbb{R}^n \text{ tels que } |x| \vee |y| \le m,\\ &\qquad |\sigma(y) - \sigma(x)| \le K_{\sigma,m}|y - x| \text{ et } |b(y) - b(x)| \le K_{b,m}|y - x|.\end{aligned} \tag{3.69}$$

Pour $m \in \mathbb{N}^*$, on note

$$\sigma_m(x) = \sigma(P_m x) \text{ et } b_m(x) = b(P_m x) \text{ où } P_m x = 1_{\{|x|>0\}} \frac{|x| \wedge m}{|x|} x \tag{3.70}$$

désigne la projection orthogonale de x sur la boule fermée de rayon m centrée à l'origine.

2. À quelle condition sur la suite $(K_{\sigma,m})_{m \in \mathbb{N}^*}$, le coefficient σ est-il globalement lipschitzien ?
3. Soit $m \in \mathbb{N}^*$.
   (a) Pour $x \in \mathbb{R}^n$, en remarquant que $\langle P_m(x), x - P_m(x)\rangle = m|x - P_m(x)|$, vérifier que pour $z \in \mathbb{R}^n$ tel que $|z| \le m$, $\langle P_m(x) - z, x - P_m(x)\rangle \ge 0$. En déduire que

   $$\forall x, y \in \mathbb{R}^n,\ |x - y|^2 \ge |P_m(x) - P_m(y)|^2 + |x - P_m(x) + P_m(y) - y|^2$$

   et conclure que σ_m et b_m sont lipschitziennes de constantes respectives $K_{\sigma,m}$ et $K_{b,m}$ et vérifient $\forall x \in \mathbb{R}^n,\ |\sigma_m(x)| + |b_m(x)| \le K(1 + |x|)$.

   (b) Que peut-on en déduire pour l'équation différentielle stochastique

   $$X_t^{(m)} = x_0 + \int_0^t \sigma_m(X_s^{(m)})dW_s + \int_0^t b_m(X_s^{(m)})ds,\ t \in [0, T]?$$

   Les preuves du lemme 3.1.4 et de la proposition 3.1.6 du polycopié n'utilisent que la croissance affine des coefficients et pas leur caractère lipschitzien si bien que pour $p \ge 1$, $\sup_{m \in \mathbb{N}^*, t \in [0,T]} \mathbb{E}\left[|X_t^{(m)}|^{2p}\right] < \infty$ et

   $$\exists C < \infty,\ \forall 0 \le s \le t \le T,\ \sup_{m \in \mathbb{N}^*} \mathbb{E}\left[|X_t^{(m)} - X_s^{(m)}|^{2p}\right] \le C(t - s)^p.$$

   (c) Vérifier l'existence d'une constante $C < \infty$ ne dépendant pas de m telle que pour $t \in [0, T]$,

   $$\begin{aligned}\mathbb{E}\left[\sup_{u \in [0,t]} |X_u^{(m)}|^{2p}\right] \le 3^{2p-1}\Bigg(&|x_0|^{2p} + T^{2p-1} \int_0^t \mathbb{E}[|b_m(X_s^{(m)})|^{2p}]ds\\ &+ CT^{p-1} \int_0^t \mathbb{E}[|\sigma_m(X_s^{(m)})|^{2p}]ds\Bigg).\end{aligned}$$

   En déduire que $\sup_{m \in \mathbb{N}^*} \mathbb{E}\left[\sup_{u \in [0,T]} |X_u^{(m)}|^{2p}\right] < +\infty$.

En notant $(\bar{X}_t^{(m)})_{t\in[0,T]}$ le schéma d'Euler en temps continu avec N pas associé, on peut montrer de la même façon que $\sup_{m,N\in\mathbb{N}^*}\mathbb{E}\left[\sup_{t\in[0,T]}(|X_t^{(m)}|\vee|\bar{X}_t^{(m)}|)^{2p}\right]<+\infty$.

(d) Soit $\tau_{N,m}=\inf\{t\in[0,T]:|X_t^{(m)}|\vee|\bar{X}_t^{(m)}|\geq m\}$ (convention $\inf\emptyset=+\infty$). À l'aide de l'inégalité de Markov, montrer que $\forall q\in]0,+\infty[$, $\lim_{m\to\infty}m^q\sup_{N\in\mathbb{N}^*}\mathbb{P}(\tau_{N,m}\leq T)=0$.

(e) Vérifier que lorsque $\tau_{N,m}>t$, $(\bar{X}_s^{(m)})_{s\in[0,t]}$ et $(\bar{X}_s)_{s\in[0,t]}$ coincident.

On peut montrer qu'à N fixé dans $\mathbb{N}^*$, $(X_t=\sum_{m\in\mathbb{N}^*}X_t^{(m)}1_{\{\tau_{N,m-1}\leq t<\tau_{N,m}\}})_{t\in[0,T]}$ où $\tau_{N,0}=0$ est solution de (3.67) et même que c'est l'unique solution de cette équation.

(f) Vérifier que pour $t\in[0,T]$ et $p\geq 1$,

$$|X_t-\bar{X}_t|^{2p}\leq\sum_{m\in\mathbb{N}^*}\sup_{u\in[0,T]}|X_u^{(m)}-\bar{X}_u^{(m)}|^{2p}1_{\{\tau_{N,m-1}\leq t<\tau_{N,m}\}}.$$

(g) En déduire que

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X_t-\bar{X}_t|^{2p}\right]\leq\sum_{m\in\mathbb{N}^*}\mathbb{E}^{1/2}\left[\sup_{u\in[0,T]}|X_u^{(m)}-\bar{X}_u^{(m)}|^{4p}\right]\sqrt{\mathbb{P}(\tau_{N,m-1}\leq T)}.$$

(h) Conclure que lorsque pour tout $m\geq 2$, $K_{\sigma,m}\leq K\sqrt{\ln m}$ et $K_{b,m}\leq K\ln m$, alors

$$\exists C<\infty,\ \forall N\in\mathbb{N}^*,\ \mathbb{E}\left[\sup_{u\leq T}|X_u-\bar{X}_u|^{2p}\right]\leq\frac{C}{N^p},$$

résultat démontré par Mao et Yuan dans [52].

3.8.21 Schéma d'Euler semi-discrétisé

Soit $T\in]0,+\infty[$ un horizon de temps. Sur un espace de probabilité $(\Omega,\mathcal{A},\mathbb{P})$ muni d'une filtration $(\mathcal{F}_t)_{t\geq 0}$, on considère un $\mathcal{F}_t$-mouvement brownien $W_t=\left(W_t^1,\cdots,W_t^d\right)$ à valeurs $\mathbb{R}^d$. On se donne un coefficient de diffusion $\sigma:[0,T]\times\mathbb{R}^n\to\mathbb{R}^{n\times d}$ et un coefficient de dérive $b:[0,T]\times\mathbb{R}^n\to\mathbb{R}^n$. On suppose que

$$\text{(Lip) }\exists \mathrm{K}>0,\ \forall \mathrm{t}\in[0,\mathrm{T}],\ \begin{cases}\forall x\in\mathbb{R}^n,\ |\sigma(t,x)|+|b(t,x)|\leq K(1+|x|)\\ \forall x,y\in\mathbb{R}^n,\ |\sigma(t,x)-\sigma(t,y)|+|b(t,x)-b(t,y)|\leq K|x-y|\end{cases}.$$

Pour une condition initiale déterministe $x_0\in\mathbb{R}^n$, on s'intéresse à l'équation différentielle stochastique

$$X_t=x_0+\int_0^t\sigma(s,X_s)dW_s+\int_0^t b(s,X_s)ds,\ t\in[0,T].$$

Soit $N\in\mathbb{N}^*$, $(t_k=\frac{kT}{N})_{0\leq k\leq N}$ et $\tau_t=\lfloor\frac{Nt}{T}\rfloor\frac{T}{N}$ l'instant de cette grille uniforme qui précède $t\in[0,T]$.

(a) Écrire le schéma d'Euler en temps continu $(\bar{X}_t)_{t\in[0,T]}$ avec N pas pour cette EDS.

(b) Quelle est sa vitesse de convergence forte lorsque les coefficients σ et b vérifient également

$$\exists \alpha, K > 0,\ \forall x \in \mathbb{R}^n,\ \forall (s,t) \in [0,T],\ |\sigma(t,x)-\sigma(s,x)|+|b(t,x)-b(s,x)| \leq K(1+|x|)(t-s)^\alpha.$$

On suppose maintenant que $\boxed{d = n}$ et que l'on sait calculer les intégrales des coefficients b et $\sigma\sigma^*$ par rapport à leur variable temporelle.

(c) Montrer que pour $x \in \mathbb{R}^n$ et $0 \leq s < t \leq T$, il existe une matrice $\theta(s,t,x) \in \mathbb{R}^{n\times n}$ telle que $\theta\theta^*(s,t,x) = \frac{1}{t-s}\int_s^t \sigma\sigma^*(r,x)dr$.

On considère le schéma défini par : $\tilde{X}_0 = x_0$ et

$$\tilde{X}_{t_{k+1}} = \tilde{X}_{t_k} + \theta(t_k, t_{k+1}, \tilde{X}_{t_k})(W_{t_{k+1}} - W_{t_k}) + \int_{t_k}^{t_{k+1}} b(s, \tilde{X}_{t_k})ds,\ k \in \{0, \cdots, N-1\}.$$

(d) Pour $k \in \{0, \cdots, N-1\}$, donner la loi conditionnelle de $\tilde{X}_{t_{k+1}}$ sachant $(\tilde{X}_{t_0}, \tilde{X}_{t_1}, \cdots, \tilde{X}_{t_k})$.

(e) En déduire l'égalité des lois des vecteurs $(\tilde{X}_{t_0}, \tilde{X}_{t_1}, \cdots, \tilde{X}_{t_N})$ et $(\hat{X}_{t_0}, \hat{X}_{t_1}, \cdots, \hat{X}_{t_N})$ où

$$\hat{X}_t = x_0 + \int_0^t \sigma(s, \hat{X}_{\tau_s})dW_s + \int_0^t b(s, \hat{X}_{\tau_s})ds,\ t \in [0,T].$$

(f) Soit $p \geq 1$. Montrer que

$$\begin{aligned} 2^{1-2p}\mathbb{E}\left[\sup_{u\in[0,t]} |\hat{X}_u - X_u|^{2p}\right] \leq & CT^{p-1}\int_0^t \mathbb{E}\left[|\sigma(s,\hat{X}_{\tau_s}) - \sigma(s,X_s)|^{2p}\right] ds \\ & + T^{2p-1}\int_0^t \mathbb{E}\left[|b(s,\hat{X}_{\tau_s}) - b(s,X_s)|^{2p}\right] ds \end{aligned}$$

en précisant l'origine de la constante $C < +\infty$.

(g) Justifier que

$$|\sigma(s,\hat{X}_{\tau_s}) - \sigma(s,X_s)|^{2p} + |b(s,\hat{X}_{\tau_s}) - b(s,X_s)|^{2p} \leq 2^{2p}K^{2p}\left(\sup_{u\in[0,s]} |\hat{X}_u - X_u|^{2p} + |X_{\tau_s} - X_s|^{2p}\right)$$

(h) Conclure que

$$\exists C_{2p} < \infty,\ \forall N \in \mathbb{N}^*,\ \mathbb{E}\left[\sup_{u\in[0,T]} |\hat{X}_u - X_u|^{2p}\right] \leq \frac{C_{2p}}{N^p}$$

et en déduire le comportement pour $N \to \infty$ de $N^\gamma \sup_{u\in[0,T]} |\hat{X}_u - X_u|$ lorsque $\gamma < \frac{1}{2}$.

(i) Lorsque $f : \mathbb{R}^n \to \mathbb{R}$ est lipschitzienne, montrer l'existence d'une constante $C_f < +\infty$ telle que $\forall N \in \mathbb{N}^*,\ |\mathbb{E}[f(\tilde{X}_T)] - \mathbb{E}[f(X_T)]| \leq \frac{C_f}{\sqrt{N}}$.

Chapitre 4

Simulation de processus comportant des sauts

La volonté de mieux rendre compte des phénomènes observés sur les marchés a motivé l'introduction de sauts dans les modèles d'évolution d'actifs financiers. Sauf exception, cela conduit à un marché incomplet. Les modèles considérés sont de deux types :

— équations différentielles stochastiques avec sauts, qui comportent, en plus des termes de diffusion et de dérive usuels, un terme de sauts dirigé le plus souvent par un processus de Poisson.

— exponentielles de processus de Lévy : le sous-jacent évolue comme l'exponentielle d'un Processus à Accroissements Indépendants et Stationnaires. Le cas particulier du modèle de Black-Scholes se retrouve lorsque les trajectoires de ce PAIS sont continues.

4.1 Diffusions avec sauts

Nous nous intéresserons tout d'abord à la simulation du processus le plus simple qui comporte des sauts : le processus de Poisson. Puis nous verrons comment simuler le modèle de Merton qui est construit à partir d'un mouvement brownien et d'un processus de Poisson indépendants. Enfin, nous traiterons le cas des équations différentielles stochastiques avec sauts.

4.1.1 Le processus de Poisson

Définition 4.1.1. *On appelle processus de Poisson de paramètre $\lambda > 0$, un processus $(N_t)_{t\geq 0}$ à valeurs dans $\mathbb{N}$ et à trajectoires continues à droite t.q. pour tous temps $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$, les variables aléatoires $N_{t_1}, N_{t_2} - N_{t_1}, ..., N_{t_n} - N_{t_{n-1}}$ sont indépendantes et distribuées suivant les lois de Poisson de paramètres respectifs $\lambda t_1, \lambda(t_2 - t_1), ..., \lambda(t_n - t_{n-1})$.*

Remarque 4.1.2. — *La variable aléatoire N_0 étant distribuée suivant la loi de Poisson de paramètre 0, elle est presque sûrement nulle.*

— *Pour $0 \leq s \leq t$, comme $N_t - N_s$ suit la loi de Poisson de paramètre $\lambda(t-s)$, $\mathbb{P}(N_t \geq N_s) = 1$. On en déduit que $\mathbb{P}(\forall s \leq t \in \mathbb{Q}_+, N_t \geq N_s) = 1$ puis avec la* *continuité à droite des trajectoires que* $\mathbb{P}(\forall s \leq t \in \mathbb{R}_+, N_t \geq N_s) = 1$. *Ainsi les trajectoires* $t \mapsto N_t$ *sont croissantes.*

Le processus de Poisson est un PAIS (également appelé processus de Lévy) à trajectoires discontinues. La proposition suivante décrit la structure de ses sauts.

Proposition 4.1.3. *Soit* $(\tau_i)_{i\geq 1}$ *une suite de variables aléatoires indépendantes et identiquement distribuées suivant la loi exponentielle de paramètre* λ *et pour* $n \in \mathbb{N}^*$, $\bar{T}_n = \sum_{i=1}^n \tau_i$. *Alors le processus de comptage* $(\bar{N}_t = \sum_{n\geq 1} 1_{\{\bar{T}_n \leq t\}})_{t\geq 0}$ *est un processus de Poisson de paramètre* λ.
Inversement, si $(N_t)_{t\geq 0}$ *est un processus de Poisson de paramètre* λ *et si* $T_i = \inf\{t \geq 0 : N_t \geq i\}$ *pour* $i \geq 1$, *alors les variables aléatoires* $(T_i - T_{i-1})_{i\geq 1}$ *(où par convention* $T_0 = 0$*) sont i.i.d. suivant la loi exponentielle de paramètre* λ *et pour* $i \geq 1$, $N_{T_i} = i$.

Démonstration : Pour la première assertion, nous allons nous contenter de raisonner avec deux temps $0 \leq t_1 \leq t_2$. Le cas d'un nombre fini de temps s'obtient par une généralisation aisée. Pour $k, l \in \mathbb{N}$ et $f : \mathbb{R}^{k+l} \to \mathbb{R}$ bornée, en commençant par intégrer en s_{k+l+1} pour la quatrième égalité puis en effectuant le changement de variables de jacobien 1 $r_1 = s_1, r_2 = s_1 + s_2, \ldots, r_{k+l} = s_1 + \ldots + s_{k+l}$ pour la cinquième, on obtient

$$\begin{aligned}
&\mathbb{E}(f(\bar{T}_1, \ldots, \bar{T}_{k+l})1_{\{\bar{N}_{t_1}=k, \bar{N}_{t_2}-\bar{N}_{t_1}=l\}}) = \mathbb{E}(f(\bar{T}_1, \ldots, \bar{T}_{k+l})1_{\{\bar{N}_{t_1}=k, \bar{N}_{t_2}=k+l\}}) \\
&= \mathbb{E}\left(f(\bar{T}_1, \ldots, \bar{T}_{k+l})1_{\{\bar{T}_k \leq t_1 < \bar{T}_{k+1}, \bar{T}_{k+l} \leq t_2 < \bar{T}_{k+l+1}\}}\right) \\
&= \int_{\substack{0\leq s_1+\ldots+s_k \leq t_1 \leq s_1+\ldots+s_{k+1} \\ s_1+\ldots+s_{k+l}\leq t_2}} f(s_1, s_1+s_2, \ldots, s_1+\ldots+s_{k+l}) \\
&\qquad\qquad \left(\int_{t_2-(s_1+\ldots+s_{k+l})}^{+\infty} \lambda e^{-\lambda s_{k+l+1}} ds_{k+l+1}\right) \lambda^{k+l} e^{-\lambda(s_1+\ldots+s_{k+l})} \prod_{j=1}^{k+l} 1_{\{s_j\geq 0\}} ds_j \\
&= e^{-\lambda t_2}\lambda^{k+l} \int_{\substack{0\leq s_1+\ldots+s_k \leq t_1 \leq s_1+\ldots+s_{k+1} \\ s_1+\ldots+s_{k+l}\leq t_2}} f(s_1, s_1+s_2, \ldots, s_1+\ldots+s_{k+l}) \prod_{j=1}^{k+l} 1_{\{s_j\geq 0\}} ds_j \\
&= e^{-\lambda t_2}\lambda^{k+l} \int f(r_1, \ldots, r_{k+l}) 1_{\{0\leq r_1 \leq r_2 \leq \ldots \leq r_k \leq t_1\}} 1_{\{t_1 < r_{k+1} \leq r_{k+2} \leq \ldots \leq r_{k+l} \leq t_2\}} dr_1 \ldots dr_{k+l}.
\end{aligned} \tag{4.1}$$

Pour le choix $f \equiv 1$, on conclut que

$$\begin{aligned}
\mathbb{P}(\bar{N}_{t_1} = k, \bar{N}_{t_2} - \bar{N}_{t_1} = l) &= e^{-\lambda t_2}\lambda^{k+l}\left(\int_{0\leq r_1\leq\ldots\leq r_k\leq t_1} \prod_{j=1}^{k} dr_j\right)\left(\int_{t_1\leq r_{k+1}\leq\ldots\leq r_{k+l}\leq t_2} \prod_{j=k+1}^{k+l} dr_j\right) \\
&= e^{-\lambda t_2}\lambda^{k+l}\frac{t_1^k}{k!} \times \frac{(t_2-t_1)^l}{l!} = e^{-\lambda t_1}\frac{(\lambda t_1)^k}{k!} \times e^{-\lambda(t_2-t_1)}\frac{(\lambda(t_2-t_1))^l}{l!}.
\end{aligned}$$

Pour la réciproque, rappelons que d'après la remarque 4.1.2, $\mathbb{P}(t \mapsto N_t \text{ croissante}) = 1$. Introduisons la suite $(S_n)_{n\geq 1}$ des temps de sauts du processus $(N_t)_{t\geq 0}$ définie par récurrence par

$$\forall n \in \mathbb{N}^*,\ S_n = \inf\{t > S_{n-1} : N_t > N_{t^-}\} \text{ où } S_0 = 0 \text{ et } N_{t^-} = \lim_{s\to t^-} N_s.$$

Pour $t \geq 0$, on a $N_t = \sum_{n\geq 1} \nu_n 1_{\{S_n \leq t\}}$ où $\nu_n : \Omega \to \mathbb{N}^*$ désigne la taille a priori aléatoire du saut de l'instant S_n.

Comme $\mathbb{P}(S_1 > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}$, S_1 suit la loi exponentielle de paramètre λ. La propriété de Markov forte du PAIS $(N_t)_{t\geq 0}$ au temps d'arrêt S_1 assure que $(N_{S_1+t}-N_{S_1})_{t\geq 0}$ est un processus de Poisson de paramètre λ indépendant de S_1. Donc, d'après ce qui précède, l'instant de premier saut $S_2 - S_1$ de ce processus est une variable aléatoire exponentielle de paramètre λ indépendante de S_1. Plus généralement les variables $(S_n - S_{n-1})_{n\geq 1}$ sont i.i.d. suivant cette loi.
Pour $t \geq 0$, comme $N_t = \sum_{n\geq 1} \nu_n 1_{\{S_n \leq t\}} \geq \sum_{n\geq 1} 1_{\{S_n \leq t\}}$ où les variables aléatoires N_t et $\sum_{n\geq 1} 1_{\{S_n\leq t\}}$ suivent toutes les deux la loi de Poisson de paramètre λt la première par définition et la seconde d'après la première assertion, $N_t = \sum_{n\geq 1} 1_{\{S_n\leq t\}}$. □

On en déduit les deux possibilités suivantes pour simuler le processus de Poisson à partir d'une suite $(U_i)_{i\geq 1}$ de variables aléatoires i.i.d. suivant la loi uniforme sur $[0,1]$:

-génération des temps de sauts : $(-\frac{1}{\lambda}\ln(U_i))_{i\geq 1}$ est une suite de variables i.i.d. suivant la loi exponentielle de paramètre λ et $(\sum_{n\geq 1} 1_{\{-\frac{1}{\lambda}\sum_{i=1}^n \ln(U_i)\leq t\}})_{t\geq 0}$ est un processus de Poisson de paramètre λ. Notons que

$$\sum_{n\geq 1} 1_{\{-\frac{1}{\lambda}\sum_{i=1}^n \ln(U_i)\leq t\}} = \min\{n \in \mathbb{N} : \prod_{i=1}^{n+1} U_i < e^{-\lambda t}\}.$$

-simulation à des instants déterministes : pour générer le processus à des temps fixes $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$, on utilise la définition qui assure que les variables aléatoires $N_{t_1}, N_{t_2} - N_{t_1}, ..., N_{t_n} - N_{t_{n-1}}$ sont indépendantes et distribuées suivant les lois de Poisson de paramètres respectifs $\lambda t_1, \lambda(t_2 - t_1), ..., \lambda(t_n - t_{n-1})$.
Notons que d'après l'algorithme précédent, pour $\theta > 0$, $\min\{n \in \mathbb{N} : \prod_{i=1}^{n+1} U_i < e^{-\theta}\}$ suit la loi de Poisson de paramètre θ.

On peut généraliser la définition du processus de Poisson en relachant l'hypothèse de stationnarité des accroissements. Le paramètre réel est remplacé par une fonction positive sur $\mathbb{R}_+$.

Définition 4.1.4. *Soit $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ une fonction mesurable telle que pour tout $t > 0$, $\Lambda(t) = \int_0^t \lambda(s)ds < +\infty$. On appelle processus de Poisson d'intensité $\lambda(t)$ un processus $(N_t)_{t\geq 0}$ à valeurs dans $\mathbb{N}$ et à trajectoires continues à droite t.q. pour tous temps $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$, les variables aléatoires $N_{t_1}, N_{t_2} - N_{t_1}, ..., N_{t_n} - N_{t_{n-1}}$ sont indépendantes et distribuées suivant les lois de Poisson de paramètres respectifs $\Lambda(t_1), \Lambda(t_2) - \Lambda(t_1), ..., \Lambda(t_n) - \Lambda(t_{n-1})$.*

La simulation à des instants déterministes n'est pas plus compliquée que dans le cas homogène où l'intensité est constante. En revanche, la simulation des temps de sauts repose sur la proposition suivante :

Proposition 4.1.5. *Soit $(U_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d. suivant la loi uniforme sur $[0,1]$. Si la suite $(T_n)_{n\geq 1}$ est définie par*

$$\forall n \geq 1,\ T_n = \inf\left\{t \geq 0 : \Lambda(t) = -\sum_{i=1}^n \ln(U_i)\right\} \quad \textit{(convention } \inf\emptyset = +\infty\textit{)},$$

alors $(N_t = \sum_{n\geq 1} 1_{\{T_n\leq t\}})_{t\geq 0}$ est un processus de Poisson d'intensité $\lambda(t)$.

Si l'intensité $\lambda(t)$ est constante égale à λ, alors $\Lambda(t) = \lambda t$ et $T_n = -\frac{1}{\lambda}\sum_{i=1}^n \ln(U_i)$, si bien que l'on retrouve la première assertion de la proposition 4.1.3.

Démonstration : Par continuité de Λ, pour $n \geq 1$, si $T_n < +\infty$ alors $\Lambda(T_n) = -\sum_{i=1}^n \ln(U_i)$. Par croissance de Λ, on en déduit que pour $t \geq 0$, $T_n \leq t$ implique $-\sum_{i=1}^n \ln(U_i) \leq \Lambda(t)$. Inversement, si $-\sum_{i=1}^n \ln(U_i) \leq \Lambda(t)$, la définition de T_n et la continuité de Λ entraînent que $T_n \leq t$. Ainsi $\{T_n \leq t\} = \{-\sum_{i=1}^n \ln(U_i) \leq \Lambda(t)\}$.
D'après la première assertion de la proposition 4.1.3, le processus $(N_s^1 = \sum_{n\geq 1} 1_{\{-\sum_{i=1}^n \ln(U_i)\leq s\}})_{s\geq 0}$ est un processus de Poisson de paramètre 1. Comme pour tout $t \geq 0$, $N_t = \sum_{n\geq 1} 1_{\{-\sum_{i=1}^n \ln(U_i)\leq \Lambda(t)\}} = N^1_{\Lambda(t)}$, on conclut facilement. $\square$

Dans le cas inhomogène où $\lambda(t)$ est une fonction majorée par $\bar{\lambda} < +\infty$, on peut également utiliser la technique de simulation **de sauts fictifs** qui repose sur le résultat suivant.

Proposition 4.1.6. *Soit* $\lambda : \mathbb{R}_+ \to [0, \bar{\lambda}]$ *et* $(\bar{T}_j)_{j\geq 1}$ *la suite des temps de sauts successifs d'un processus de Poisson* $(\bar{N}_t)_{t\geq 0}$ *de paramètre* $\bar{\lambda}$ *indépendant d'une suite* $(U_j)_{j\geq 1}$ *de variables i.i.d. suivant la loi uniforme sur* $[0,1]$. *Si on définit par récurrence*

$$T_0 = 0 \text{ et pour } n \geq 1,\ T_n = \min\left\{\bar{T}_j > T_{n-1} : U_j \leq \frac{\lambda(\bar{T}_j)}{\bar{\lambda}}\right\},$$

alors $(N_t = \sum_{n\geq 1} 1_{\{T_n\leq t\}})_{t\geq 0}$ *est un processus de Poisson d'intensité* $\lambda(t)$.

On rejette chacun des temps de sauts du processus de Poisson dont le paramètre $\bar{\lambda}$ majore l'intensité $\lambda(t)$ avec une probabilité égale au rapport $\lambda(.)/\bar{\lambda}$ calculé au temps de saut. Les sauts rejetés sont les sauts fictifs dont la méthode tire son nom.

Démonstration : On a $N_t = \sum_{j=1}^{\bar{N}_t} 1_{\{U_j\leq \frac{\lambda(\bar{T}_j)}{\bar{\lambda}}\}}$. Pour $0 < t_1 < t_2$ et $n, m \in \mathbb{N}$, en utilisant l'indépendance de la suite $(U_j)_{j\geq 1}$ et de $\bar{N}$ pour la troisième égalité, puis (4.1) pour la quatrième, on obtient

$$\mathbb{P}(N_{t_1} = n, N_{t_2} - N_{t_1} = m) = \sum_{k\geq n, l\geq m} \mathbb{P}(N_{t_1} = n, \bar{N}_{t_1} = k, N_{t_2} - N_{t_1} = m, \bar{N}_{t_2} - \bar{N}_{t_1} = l)$$

$$= \sum_{k\geq n,l\geq m} \mathbb{P}\left(\sum_{j=1}^k 1_{\{U_j\leq\frac{\lambda(\bar{T}_j)}{\bar{\lambda}}\}} = n, \bar{N}_{t_1} = k, \sum_{j=k+1}^{k+l} 1_{\{U_j\leq\frac{\lambda(\bar{T}_j)}{\bar{\lambda}}\}} = m, \bar{N}_{t_2} - \bar{N}_{t_1} = l\right)$$

$$= \sum_{k\geq n,l\geq m} \int_{[0,1]^{k+l}} \mathbb{P}\left(\sum_{j=1}^k 1_{\{u_j\leq\frac{\lambda(\bar{T}_j)}{\bar{\lambda}}\}} = n, \bar{N}_{t_1} = k, \sum_{j=k+1}^{k+l} 1_{\{u_j\leq\frac{\lambda(\bar{T}_j)}{\bar{\lambda}}\}} = m, \bar{N}_{t_2} - \bar{N}_{t_1} = l\right) du_1 \ldots du_{k+l}$$

$$= e^{-\bar{\lambda}t_2} \sum_{k\geq n,l\geq m} \bar{\lambda}^{k+l} \int_{\mathbb{R}^k} \mathbb{P}\left(\sum_{j=1}^k 1_{\{U_j\leq\frac{\lambda(r_j)}{\bar{\lambda}}\}} = n\right) 1_{\{0\leq r_1\leq\ldots\leq r_k\leq t_1\}} dr_1 \ldots dr_k$$
$$\int_{\mathbb{R}^l} \mathbb{P}\left(\sum_{j=k+1}^{k+l} 1_{\{U_j\leq\frac{\lambda(r_j)}{\bar{\lambda}}\}} = n\right) 1_{\{t_1<r_{k+1}\leq\ldots\leq r_{k+l}\leq t_2\}} dr_{k+1}\ldots dr_{k+l}. \quad (4.2)$$

Comme $(r_1, \ldots, r_k) \mapsto \mathbb{P}\left(\sum_{j=1}^k 1_{\{U_j\leq\frac{\lambda(r_j)}{\bar{\lambda}}\}} = n\right)$ est une fonction symétrique, en notant

$\mathcal{S}_k$ l'ensemble des permutations de $\{1, 2, \dots, k\}$, on obtient

$$\int_{\mathbb{R}^k} \mathbb{P}\left(\sum_{j=1}^k 1_{\{U_j \le \frac{\lambda(r_j)}{\bar\lambda}\}} = n\right) 1_{\{0 \le r_1 \le \dots \le r_k \le t_1\}} dr_1 \dots dr_k$$

$$= \frac{1}{k!} \sum_{\sigma \in \mathcal{S}_k} \int_{\mathbb{R}^k} \mathbb{P}\left(\sum_{j=1}^k 1_{\{U_j \le \frac{\lambda(r_j)}{\bar\lambda}\}} = n\right) 1_{\{0 \le r_{\sigma(1)} \le \dots \le r_{\sigma(k)} \le t_1\}} dr_1 \dots dr_k$$

$$= \frac{1}{k!} \int_{[0,t_1]^k} \mathbb{P}\left(\sum_{j=1}^k 1_{\{U_j \le \frac{\lambda(r_j)}{\bar\lambda}\}} = n\right) dr_1 \dots dr_k = \frac{t_1^k}{k!} \mathbb{P}\left(\sum_{j=1}^k 1_{\{U_j \le \frac{\lambda(R_j)}{\bar\lambda}\}} = n\right)$$

où les variables aléatoires $(R_j)_{1 \le j \le k}$ sont i.i.d. suivant la loi uniforme sur $[0, t_1]$ et indépendantes des U_j. La variable aléatoire $\sum_{j=1}^k 1_{\{U_j \le \frac{\lambda(R_j)}{\bar\lambda}\}}$ suit la loi binomiale de paramètres k et

$$\mathbb{P}\left(U_1 \le \frac{\lambda(R_1)}{\bar\lambda}\right) = \frac{1}{t_1} \int_0^{t_1} \int_0^{\frac{\lambda(r)}{\bar\lambda}} du dr = \frac{1}{t_1} \int_0^{t_1} \frac{\lambda(r)}{\bar\lambda} dr = \frac{\Lambda(t_1)}{\bar\lambda t_1}.$$

On en déduit que la première intégrale au membre de droite de (4.2) est égale à

$$\frac{1}{\bar\lambda^k k!} \binom{k}{n} \Lambda(t_1)^n (\bar\lambda t_1 - \Lambda(t_1))^{k-n} = \frac{1}{\bar\lambda^k n!(k-n)!} \Lambda(t_1)^n (\bar\lambda t_1 - \Lambda(t_1))^{k-n}.$$

La seconde intégrale se calcule de manière analogue et on conclut que

$$\mathbb{P}(N_{t_1} = n, N_{t_2} - N_{t_1} = m)$$
$$= e^{-\bar\lambda t_2} \frac{\Lambda(t_1)^n (\Lambda(t_2) - \Lambda(t_1))^m}{n! m!} \sum_{k \ge n} \frac{(\bar\lambda t_1 - \Lambda(t_1))^{k-n}}{(k-n)!} \sum_{l \ge m} \frac{(\bar\lambda(t_2 - t_1) - \Lambda(t_2) + \Lambda(t_1))^{l-m}}{(l-m)!}$$
$$= e^{-\Lambda(t_1)} \frac{\Lambda(t_1)^n}{n!} \times e^{-(\Lambda(t_2) - \Lambda(t_1))} \frac{(\Lambda(t_2) - \Lambda(t_1))^m}{m!}.$$

□

4.1.2 Le modèle de Merton

Le modèle de Merton s'écrit

$$X_t^y = y e^{\sigma W_t + \mu t} \prod_{j=1}^{N_t} Y_j \tag{4.3}$$

où $(W_t)_{t \ge 0}$ est un mouvement brownien, $(N_t)_{t \ge 0}$ un processus de Poisson de paramètre $\lambda > 0$ et $(Y_j)_{j \ge 1}$ une suite de variables aléatoires réelles i.i.d. telles que $\mathbb{P}(Y_1 > 0) = 1$. On suppose également $(W_t)_{t \ge 0}$, $(N_t)_{t \ge 0}$ et $(Y_j)_{j \ge 1}$ indépendants.

Exemple 4.1.7. *Pour la loi commune des Y_j,*

— *le choix $\mathbb{P}(Y_j = a) = p = 1 - \mathbb{P}(Y_j = b)$ où $0 < a < 1 < b$ permet de modéliser à la fois des sauts à la baisse $\{Y_j = a\}$ et des sauts à la hausse $\{Y_j = b\}$,*

— *le choix d'une loi lognormale assure que la loi conditionnelle de X_t^y sachant $N_t = n$ est lognormale.*

Pour simuler le modèle de Merton, on peut utiliser

— la simulation des temps de sauts : on génère les temps de sauts successifs $(T_n)_{n\geq 1}$ du processus de Poisson et les variables aléatoires $(Y_n)_{n\geq 1}$ et, de façon indépendante, une suite $(G_n)_{n\geq 1}$ de gaussiennes centrées réduites indépendantes ; alors on simule le processus aux temps de sauts en posant

$$X_0^y = y, \; T_0 = 0 \text{ et pour } n \geq 1, X_{T_n}^y = X_{T_{n-1}}^y Y_n e^{\sigma G_n \sqrt{T_n - T_{n-1}} + \mu(T_n - T_{n-1})}.$$

— la simulation à des instants déterministes : d'après (4.3), pour simuler X_t^y, il suffit de générer (N_t, W_t) puis N_t variables $Y_1, \ldots, Y_{N_t}$ indépendantes suivant la loi qui donne les amplitudes de sauts.

Pour $t \geq 0$, on pose $X_{t^-}^y = \lim_{s \to t^-} X_s^y$. Entre deux temps de sauts T_{n-1} et T_n successifs, le processus de Merton (4.3) évolue suivant l'EDS de Black-Scholes $dX_t^y = \sigma X_t^y dW_t + (\mu + \frac{\sigma^2}{2}) X_t^y dt$ tandis qu'en T_n, $X_{T_n}^y = X_{T_n^-}^y Y_n$ égalité qui se récrit $X_{T_n}^y - X_{T_n^-}^y = X_{T_n^-}^y (Y_n - 1)$. On en déduit que ce processus est solution de l'équation différentielle stochastique avec sauts :

$$dX_t^y = \sigma X_t^y dW_t + (\mu + \sigma^2/2) X_t^y dt + X_{t^-}^y (Y_{N_t} - 1) dN_t.$$

Nous allons maintenant nous intéresser à la simulation d'EDS avec sauts plus générales et dont la solution n'a pas d'expression explicite.

4.1.3 Équation différentielles stochastiques avec sauts

Pour $\sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$, $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, $c : [0,T] \times \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^n$ et $y \in \mathbb{R}^n$, on s'intéresse à l'EDS avec sauts en dimension n

$$dX_t = \sigma(t, X_t) dW_t + b(t, X_t) dt + c(t, X_{t^-}, Y_{N_t}) dN_t, \; X_0 = y$$

où $(W_t)_{t\geq 0}$ mouvement brownien à valeurs $\mathbb{R}^d$, $(N_t)_{t\geq 0}$ processus de Poisson de paramètre λ et $(Y_i)_{i\geq 1}$ suite de vecteurs i.i.d. à valeurs $\mathbb{R}^q$ sont indépendants. En conditionnant par la partie sauts $((N_t)_{t\geq 0}, (Y_i)_{i\geq 1})$, on vérifie facilement que si σ et b satisfont l'hypothèse (Lip) du théorème 3.1.2, alors cette EDS admet une unique solution.

Remarque 4.1.8. *Ce cadre englobe le cas d'une intensité de sauts $\gamma : [0,T] \times \mathbb{R}^n \to [0, \lambda]$ dépendant de l'état courant mais bornée en posant $Y_i = (Z_i, U_i)$ et $c(t,x,z,u) = 1_{\{u \leq \frac{\gamma(t,x)}{\lambda}\}} f(t,x,z)$ avec $f : [0,T] \times \mathbb{R}^n \times \mathbb{R}^{q-1} \to \mathbb{R}^n$, Z_i à valeurs dans $\mathbb{R}^{q-1}$ et U_i uniforme sur $[0,1]$.*

Pour discrétiser ce processus jusqu'à l'horizon T, on peut utiliser une grille croissante d'instants qui combine les temps déterministes $t_k = \frac{kT}{K}$, $k \in \{0, \ldots, K\}$ et les temps de sauts $T_1, \ldots, T_{N_T}$ du processus de Poisson avant T. On initialise le schéma en posant $\bar{X}_0 = y$. Puis, pour passer d'un instant de discrétisation s_l à l'instant suivant s_{l+1},

— on commence par utiliser un schéma standard pour la partie EDS sans sauts, ce qui donne dans le cas du schéma d'Euler,

$$\bar{X}_{s_{l+1}^-} = \bar{X}_{s_l} + \sigma(s_l, \bar{X}_{s_l})(W_{s_{l+1}} - W_{s_l}) + b(s_l, \bar{X}_{s_l})(s_{l+1} - s_l).$$

— si $s_{l+1} \in \{T_1, \dots, T_{N_T}\}$ i.e. s_{l+1} est un temps de saut, on fait sauter le processus discrétisé suivant
$$\bar{X}_{s_{l+1}} = \bar{X}_{s_{l+1}^-} + c(s_{l+1}, \bar{X}_{s_{l+1}^-}, Y_{N_{s_{l+1}}}).$$
Sinon, on pose $\bar{X}_{s_{l+1}} = \bar{X}_{s_{l+1}^-}$.

Sous de bonnes hypothèses de régularité sur les coefficients de l'EDS, on peut vérifier que l'ordre de convergence faible de ce schéma est égal à celui du schéma utilisé pour la partie EDS sans sauts.

4.2 Processus de Lévy

Le modèle de Black-Scholes s'écrit

$$\forall t \geq 0, \; X_t^y = ye^{Z_t} \tag{4.4}$$

où $Z_t = \sigma W_t + (r - \frac{\sigma^2}{2})t$ est un Processus à Accroissements Indépendants et Stationnaires continu. On peut montrer que tous les PAIS continus à valeurs réelles sont de la forme $\sigma W_t + \mu t$ où $(W_t)_{t\geq 0}$ est un mouvement brownien standard. Une généralisation naturelle de la modélisation consiste alors à prendre pour $(Z_t)_{t\geq 0}$ un PAIS (ou processus de Lévy) non nécessairement continu dans (4.4).

Définition 4.2.1. *On appelle PAIS ou également processus de Lévy un processus $(Z_t)_{t\geq 0}$ à trajectoires continues à droite avec des limites à gauche tel que*

1. $Z_0 = 0$,
2. *pour tous temps $0 \leq t_1 \leq \ldots \leq t_n$, les accroissements $Z_{t_1}, Z_{t_2} - Z_{t_1}, \ldots, Z_{t_n} - Z_{t_{n-1}}$ sont indépendants,*
3. *pour $0 \leq s \leq t$, l'accroissement $Z_t - Z_s$ a même loi que Z_{t-s}.*

Exemple 4.2.2. *Le mouvement brownien et le processus de Poisson de paramètre λ sont des processus de Lévy.*

Si $(Z_t)_{t\geq 0}$ est un processus de Lévy, alors comme pour $n \in \mathbb{N}^*$,

$$Z_1 = \sum_{k=1}^{n} (Z_{k/n} - Z_{(k-1)/n}) \text{ où les variables } (Z_{k/n} - Z_{(k-1)/n})_{1\leq k\leq n} \text{ sont i.i.d.,} \tag{4.5}$$

la variable aléatoire Z_1 est indéfiniment divisible au sens de la définition suivante.

Définition 4.2.3. *Une variable aléatoire réelle X et sa loi sont dites indéfiniment divisibles si pour tout $n \in \mathbb{N}^*$, il existe des variables i.i.d. $(X_k^n)_{1\leq k\leq n}$ telle que $X \stackrel{\mathcal{L}}{=} \sum_{k=1}^n X_k^n$.*

En fait, il existe une bijection entre lois indéfiniment divisibles et processus de Lévy :

Théorème 4.2.4. *Pour toute loi indéfiniment divisible, il existe un processus de Lévy $(Z_t)_{t\geq 0}$ unique en loi t.q. la variable Z_1 est distribuée suivant cette loi.*

Le résultat suivant, que nous ne démontrerons pas, donne la structure de la fonction caractéristique des variables aléatoires indéfiniment divisibles qui porte le nom de formule de Lévy-Kinchine.

Proposition 4.2.5. *Une variable aléatoire réelle X est indéfiniment divisible si et seulement si il existe un unique triplet (σ, μ, ν) avec $\sigma \geq 0, \mu \in \mathbb{R}$ et ν mesure sur la droite réelle vérifiant $\nu(\{0\}) = 0$ et $\int_{\mathbb{R}}(h^2 \wedge 1)\nu(dh) < +\infty$ appelée mesure de Lévy, tel que*

$$\forall u \in \mathbb{R},\ \mathbb{E}\left(e^{iuX}\right) = e^{\psi(u)} \text{ où } \psi(u) = -\frac{\sigma^2 u^2}{2} + iu\mu + \int_{\mathbb{R}} (e^{iuh} - 1 - iuh1_{\{|h|\leq 1\}})\nu(dh).$$

Remarque 4.2.6. *Notons que $e^{-\frac{\sigma^2 u^2}{2} + iu\mu} = \mathbb{E}(e^{iuG})$ où $G \sim \mathcal{N}_1(\mu, \sigma^2)$.*

Ainsi, pour tout processus de Lévy $(Z_t)_{t\geq 0}$, il existe un triplet (σ, μ, ν) tel que $\mathbb{E}\left(e^{iuZ_1}\right) = e^{\psi(u)}$ avec ψ définie dans la proposition. Avec (4.5), on en déduit facilement que pour $n \in \mathbb{N}^*$, $\mathbb{E}\left(e^{iuZ_{1/n}}\right) = e^{\frac{\psi(u)}{n}}$ puis que pour tout $t \in \mathbb{Q}_+$, $\mathbb{E}\left(e^{iuZ_t}\right) = e^{t\psi(u)}$. Comme la continuité à droite des trajectoires de $(Z_t)_{t\geq 0}$ entraîne par convergence dominée celle de $t \to \mathbb{E}\left(e^{iuZ_t}\right)$, on conclut que $\forall t \geq 0,\ \forall u \in \mathbb{R},\ \mathbb{E}\left(e^{iuZ_t}\right) = e^{t\psi(u)}$, ce qui entraîne la propriété d'unicité en loi énoncée dans le théorème 4.2.4.
On en déduit que $(Z_t)_{t\geq 0}$ admet la décomposition $\forall t \geq 0,\ Z_t = \sigma W_t + \mu t + \bar{Z}_t$ où $(\bar{Z}_t)_{t\geq 0}$ est un processus de Lévy de sauts purs indépendant de $(W_t)_{t\geq 0}$ tel que $\forall t \geq 0$, $\forall u \in \mathbb{R}$, $\mathbb{E}\left(e^{iu\bar{Z}_t}\right) = \exp\left(t \int_{\mathbb{R}}(e^{iuh} - 1 - iuh1_{\{|h|\leq 1\}})\nu(dh)\right)$. Lorsque la mesure de Lévy ν est non nulle et bornée au sens où $\lambda = \nu(\mathbb{R}) \in]0, +\infty[$, alors $m = \frac{\nu}{\lambda}$ est une mesure de probabilité sur $\mathbb{R}$. Le processus $(\bar{Z}_t)_{t\geq 0}$ est un processus de Poisson composé avec dérive au sens où il existe un processus de Poisson $(N_t)_{t\geq 0}$ de paramètre λ ainsi qu'une suite indépendante $(\xi_i)_{i\geq 1}$ de variables i.i.d. suivant m tels que

$$\forall t \geq 0,\ \bar{Z}_t = \sum_{i=1}^{N_t} \xi_i - t \int_{[-1,1]} h\nu(dh).$$

Cela découle du théorème 4.2.4 et du résultat pour $t = 1$ du calcul suivant :

$$\begin{aligned}\mathbb{E}\left(e^{iu\left(\sum_{i=1}^{N_t} \xi_i - t\int_{[-1,1]} h\nu(dh)\right)}\right) &= e^{-t(\lambda + iu\int_{[-1,1]} h\nu(dh))} \sum_{n\in\mathbb{N}} \frac{(\lambda t)^n}{n!} \left(\int_{\mathbb{R}} e^{iuh} m(dh)\right)^n \\ &= e^{t\int_{\mathbb{R}}(e^{iuh} - 1 - iuh1_{\{|h|\leq 1\}})\nu(dh)}.\end{aligned}$$

Dans ce cas (4.4) s'écrit

$$X_t^y = ye^{\sigma W_t + \bar{\mu} t} \prod_{j=1}^{N_t} Y_j \text{ avec } Y_j = e^{\xi_j} \text{ et } \bar{\mu} = \mu - \int_{[-1,1]} h\nu(dh)$$

et on retrouve le modèle de Merton (4.3).
Pour sortir de ce modèle, il faut donc considérer le cas $\nu(\mathbb{R}) = +\infty$ où $(\bar{Z}_t)_{t\geq 0}$ possède une infinité de sauts sur tout intervalle de temps de longueur non nulle.
Pour une étude détaillée des processus de Lévy, nous revoyons aux livres de Bertoin [11], Sato [69] et Applebaum [5].

Exercice 4.2.7. *Montrer que le triplet associé au processus de Poisson de paramètre λ est $\sigma = 0$, $\mu = \lambda$ et $\nu = \lambda\delta_1$.*

4.2.1 Processus gamma et variance-gamma

Rappelons tout d'abord la définition et une technique de simulation de la loi gamma.

Définition 4.2.8. *Pour $a, \theta > 0$, on appelle loi gamma de paramètre (a, θ) et on note $\Gamma(a, \theta)$ la loi de densité $\frac{1}{\Gamma(a)} x^{a-1} \theta^a e^{-\theta x} 1_{\{x>0\}}$ sur la droite réelle où $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$ désigne la fonction gamma d'Euler.*

Propriétés :

1. La loi $\Gamma(1, \theta)$ est la loi exponentielle de paramètre θ.
2. Si $X \sim \Gamma(a, \theta)$ et $Y \sim \Gamma(b, \theta)$ sont indépendantes et $\lambda > 0$, $\lambda X \sim \Gamma(a, \frac{\theta}{\lambda})$ et $X + Y \sim \Gamma(a + b, \theta)$.
3. Si $X \sim \Gamma(a, \theta)$ alors

$$\forall u \in \mathbb{R},\ \forall v > -\theta,\ \mathbb{E}(e^{(iu-v)X}) = \left(\frac{\theta}{\theta + v - iu}\right)^a. \tag{4.6}$$

Les propriétés 1 et 2 sont à la base de la technique de simulation suivante.
Simulation d'une variable de loi $\Gamma(a, \theta)$ où $a, \theta > 0$: Soit $\lfloor a \rfloor$ la partie entière de a, $\alpha = a - \lfloor a \rfloor$ sa partie fractionnaire et $U_1, \ldots, U_{\lfloor a \rfloor}$ des variables uniformes sur $[0, 1]$ indépendantes.
D'après la propriété 2, si $Y \sim \Gamma(\alpha, 1)$ (convention $Y = 0$ si $\alpha = 0$) est indépendante de $(U_1, \ldots, U_{\lfloor a \rfloor})$, la variable aléatoire $\frac{1}{\theta}\left(Y - \ln\left(\prod_{j=1}^{\lfloor a \rfloor} U_j\right)\right)$ où $\prod_{j=1}^{0} = 1$ suit la loi $\Gamma(a, \theta)$.
Reste donc, pour $\alpha \in]0, 1[$, à savoir simuler $Y \sim \Gamma(\alpha, 1)$ de densité $\frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y} 1_{\{y>0\}} \leq \frac{1}{\Gamma(\alpha)}\left[y^{\alpha-1} 1_{\{0<y\leq 1\}} + e^{-y} 1_{\{y>1\}}\right]$. On peut pour cela effectuer du rejet sous la densité $p(y) = \frac{e}{\alpha+e}\left[\alpha y^{\alpha-1} 1_{\{0<y\leq 1\}}\right] + \frac{\alpha}{\alpha+e}\left[e^{1-y} 1_{\{y>1\}}\right]$ de fonction de répartition

$$F(y) = \frac{e}{\alpha + e}\left[y^\alpha 1_{\{0<y\leq 1\}} + (1 + \alpha(e^{-1} - e^{-y})) 1_{\{y>1\}}\right].$$

Comme $F^{-1}(u) = \left(\frac{(\alpha+e)u}{e}\right)^{1/\alpha} 1_{\{u \leq \frac{e}{\alpha+e}\}} - \ln\left(\frac{(\alpha+e)(1-u)}{\alpha e}\right) 1_{\{u > \frac{e}{\alpha+e}\}}$, on peut facilement simuler suivant la densité p. L'espérance du nombre géométrique de propositions suivant la densité p nécessaires pour générer une variable de loi $\Gamma(\alpha, 1)$ est $\frac{\alpha+e}{\alpha e \Gamma(\alpha)}$. Son maximum pour α parcourant l'intervalle $[0, 1]$ est d'environ 1.39.

Pour $a, \theta > 0$, d'après la propriété 2, la loi $\Gamma(a, \theta)$ est indéfiniment divisible. L'exercice suivant assure que son triplet caractéristique est $\left(a\frac{1-e^{-\theta}}{\theta}, 0, 1_{\{x>0\}} \frac{ae^{-\theta x}}{x} dx\right)$.

Exercice 4.2.9. *On admet que le triplet (μ, σ^2, ν) de $X \sim \Gamma(a, \theta)$ est caractérisé par l'égalité*

$$\forall \lambda \in]-\infty, \theta[,\ \ln\left(\mathbb{E}[e^{\lambda X}]\right) = \lambda\mu + \frac{\sigma^2\lambda^2}{2} + \int_{\mathbb{R}} (e^{\lambda x} - 1 - \lambda x 1_{\{|x|\leq 1\}}) \nu(dx),$$

qui s'obtient en remplaçant le paramètre imaginaire pur iu de la proposition 4.2.5 par le paramètre réel λ. On pose $f(\lambda) = \int_0^{+\infty} (e^{\lambda x} - 1) e^{-\theta x} \frac{dx}{x}$ pour $\lambda \in]-\infty, \theta[$.

1. *Que vaut $f(0)$? Calculer $f'(\lambda)$ pour $\lambda \in]-\infty, \theta[$ et en déduire que $f(\lambda) = \ln\left(\frac{\theta}{\theta-\lambda}\right)$.*
2. *Calculer $\int_0^{+\infty} x 1_{\{|x|\leq 1\}} e^{-\theta x} \frac{dx}{x}$.*
3. *En déduire que $(\mu, \sigma^2, \nu) = \left(a\frac{1-e^{-\theta}}{\theta}, 0, 1_{\{x>0\}} \frac{ae^{-\theta x}}{x} dx\right)$.*

Donc d'après le théorème 4.2.4, il existe un processus de Lévy $(U_t)_{t\geq 0}$ appelé processus gamma de paramètre (a, θ) tel que $U_1 \sim \Gamma(a, \theta)$ et $\mathbb{E}\left(e^{iuU_t}\right) = \left(\mathbb{E}\left(e^{iuU_1}\right)\right)^t$, ce qui assure que $U_t \sim \Gamma(at, \theta)$ d'après (4.6). Comme les variables aléatoires gamma sont positives, ce processus est croissant. Cette propriété de monotonie n'étant pas satisfaisante pour modéliser le cours d'un actif financier, Madan, Carr et Chang [54] proposent plutôt d'utiliser dans (4.4) $Z_t = U_t - D_t$ où $(U_t)_{t\geq 0}$ et $(D_t)_{t\geq 0}$ sont des processus gamma indépendants de paramètres respectifs (a, θ) et (a, λ). Le lemme suivant explique pourquoi le processus $(Z_t)_{t\geq 0}$ porte le nom de processus variance gamma.

Lemme 4.2.10. *Le processus $(Z_t)_{t\geq 0}$ a même loi que $(B_{Y_t})_{t\geq 0}$ où $B_t = \sqrt{\frac{2a}{\theta\lambda}}\,W_t + \frac{(\lambda-\theta)a}{\lambda\theta}t$ avec $(W_t)_{t\geq 0}$ mouvement brownien standard indépendant de $(Y_t)_{t\geq 0}$ processus gamma de paramètre (a, a).*

Démonstration : Posons $\sigma = \sqrt{\frac{2a}{\theta\lambda}}$, $\mu = \frac{(\lambda-\theta)a}{\lambda\theta}$ et notons p_t la densité de Y_t pour $t > 0$. Pour $0 < t_1 < t_2$ et $u, v \in \mathbb{R}$, en utilisant l'indépendance de Y et B puis le caractère indépendant et stationnaire des accroissements de ces processus et la positivité de ceux de Y, il vient

$$\begin{aligned}
\mathbb{E}\left(e^{i(uB_{Y_{t_1}}+v(B_{Y_{t_2}}-B_{Y_{t_1}}))}\right) &= \int_0^{+\infty}\int_0^{+\infty}\mathbb{E}\left(e^{i(uB_\tau+v(B_{\tau+r}-B_\tau))}\right)p_{t_1}(\tau)p_{t_2-t_1}(r)d\tau dr\\
&= \int_0^{+\infty} e^{(iu\mu-\frac{\sigma^2u^2}{2})\tau}p_{t_1}(\tau)d\tau\int_0^{+\infty} e^{(iv\mu-\frac{\sigma^2v^2}{2})r}p_{t_2-t_1}(r)dr\\
&= \mathbb{E}\left(e^{iuB_{Y_{t_1}}}\right)\mathbb{E}\left(e^{ivB_{Y_{t_2-t_1}}}\right).
\end{aligned}$$

Donc $(B_{Y_t})_{t\geq 0}$ est un PAIS et il suffit de vérifier que pour $t > 0$, Z_t et B_{Y_t} ont même loi pour conclure. Pour cela, on utilise la fonction caractéristique. Pour $u \in \mathbb{R}$, d'après le calcul qui précède puis (4.6) et la définition de (σ, μ),

$$\mathbb{E}(e^{iuB_{Y_t}}) = \mathbb{E}\left(e^{(iu\mu-\frac{\sigma^2u^2}{2})Y_t}\right) = \left(\frac{a}{a+\frac{\sigma^2u^2}{2}-iu\mu}\right)^{at} = \left(\frac{\theta\lambda}{\theta\lambda+i(\theta-\lambda)u+u^2}\right)^{at}.$$

D'autre part, d'après la définition de Z, l'indépendance de U et D puis (4.6),

$$\begin{aligned}
\mathbb{E}(e^{iuZ_t}) = \mathbb{E}(e^{iu(U_t-D_t)}) = \mathbb{E}(e^{iuU_t})\mathbb{E}(e^{-iuD_t}) &= \left(\frac{\theta}{\theta-iu}\right)^{at}\left(\frac{\lambda}{\lambda+iu}\right)^{at}\\
&= \left(\frac{\theta\lambda}{\theta\lambda+i(\theta-\lambda)u+u^2}\right)^{at}.
\end{aligned}$$

□

Au passage, nous avons vu que $Z_{t_2} - Z_{t_1} \overset{\mathcal{L}}{=} Z_{t_2-t_1} \overset{\mathcal{L}}{=} \sigma\sqrt{Y_{t_2-t_1}}G + \mu Y_{t_2-t_1}$ où G est une gaussienne centrée réduite indépendante de $(Y_t)_{t\geq 0}$. On peut utiliser cette remarque ou bien la définition $Z_t = U_t - D_t$ du processus variance gamma pour pouvoir simuler ses accroissements.

4.2.2 Processus Normal Inverse Gaussian

Définition 4.2.11. *Soit $\gamma \geq 0$ et $\delta > 0$. On appelle loi Inverse Gaussian de paramètre (γ, δ) et on note $IG(\gamma, \delta)$ la loi de densité*

$$\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}}x^{-3/2}e^{-\frac{1}{2}(\frac{\delta^2}{x}+\gamma^2x)}1_{\{x>0\}} = \frac{\delta}{\sqrt{2\pi}}x^{-3/2}e^{-\frac{(\delta-\gamma x)^2}{2x}}1_{\{x>0\}}$$

sur la droite réelle.

On a l'interprétation probabiliste suivante pour cette loi.

Lemme 4.2.12. *Soit $(W_t)_{t\geq 0}$ un mouvement brownien standard. Alors $\tau_\delta = \inf\{t \geq 0 : W_t + \gamma t \geq \delta\}$ suit la loi $IG(\gamma, \delta)$.*

Démonstration : On pose $B_t = W_t + \gamma t$. Soit $\alpha > 0$. Pour $n \in \mathbb{N}^*$, d'après le théorème de Girsanov, sous $\tilde{\mathbb{P}}$ de densité $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{-\gamma W_n - \frac{\gamma^2}{2}n}$, $(B_t)_{t\leq n}$ est un mouvement brownien. En outre, $(e^{\gamma B_t - \frac{\gamma^2}{2}t})_{t\leq n}$ et $(e^{(\alpha+\gamma)B_t - \frac{(\alpha+\gamma)^2}{2}t})_{t\leq n}$ sont des $\tilde{\mathbb{P}}$ martingales exponentielles. Comme $\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}} = e^{\gamma B_n - \frac{\gamma^2}{2}n}$, on a

$$\mathbb{E}\left(e^{\alpha B_{\tau_\delta\wedge n} - \frac{\alpha^2+2\alpha\gamma}{2}\tau_\delta\wedge n}\right) = \tilde{\mathbb{E}}\left(e^{\alpha B_{\tau_\delta\wedge n} - \frac{\alpha^2+2\alpha\gamma}{2}\tau_\delta\wedge n} e^{\gamma B_n - \frac{\gamma^2}{2}n}\right) = \tilde{\mathbb{E}}\left(e^{(\alpha+\gamma)B_{\tau_\delta\wedge n} - \frac{(\alpha+\gamma)^2}{2}\tau_\delta\wedge n}\right) = 1.$$

Lorsque n tend vers l'infini, la suite de variables aléatoires $\left(e^{\alpha B_{\tau_\delta\wedge n} - \frac{\alpha^2+2\alpha\gamma}{2}\tau_\delta\wedge n}\right)_{n\geq 1}$ à valeurs dans $[0, e^{\alpha\delta}]$ converge presque sûrement vers $e^{\alpha\delta - \frac{\alpha^2+2\alpha\gamma}{2}\tau_\delta}$. Le théorème de convergence dominée assure alors que

$$\mathbb{E}\left(e^{-\frac{\alpha^2+2\alpha\gamma}{2}\tau_\delta}\right) = e^{-\alpha\delta}.$$

Notons qu'en prenant la limite pour $\alpha \to 0^+$, on obtient par convergence monotone que $\mathbb{P}(\tau_\delta < +\infty) = 1$.
Pour $\lambda \geq 0$, on vérifie facilement que $-\gamma + \sqrt{\gamma^2 + 2\lambda}$ est l'unique racine positive de l'équation du second degré $\lambda = \frac{\alpha^2+2\alpha\gamma}{2}$ en α. Ainsi la transformée de Laplace de τ_δ est

$$\mathbb{E}\left(e^{-\lambda\tau_\delta}\right) = e^{(\gamma - \sqrt{\gamma^2+2\lambda})\delta}.$$

On conclut en remarquant que

$$\int_0^{+\infty} e^{-\lambda x} IG(\gamma,\delta)(x)dx = e^{\delta\gamma}e^{-\delta\sqrt{\gamma^2+2\lambda}} \int_0^{+\infty} IG(\sqrt{\gamma^2+2\lambda}, \delta)(x)dx = e^{(\gamma-\sqrt{\gamma^2+2\lambda})\delta}. \tag{4.7}$$

□

D'après la propriété de Markov forte pour le mouvement brownien avec dérive B_t, pour $\delta, \delta' > 0$, la variable aléatoire $\tau_{\delta+\delta'} - \tau_\delta$ est indépendante de τ_δ et a même loi que $\tau_{\delta'}$. Donc si $X \sim IG(\gamma, \delta)$ et $Y \sim IG(\gamma, \delta')$ sont indépendantes, alors $(X, Y) \stackrel{\mathcal{L}}{=} (\tau_\delta, \tau_{\delta+\delta'} - \tau_\delta)$ ce qui assure que $X + Y \stackrel{\mathcal{L}}{=} \tau_{\delta+\delta'} \sim IG(\gamma, \delta + \delta')$. Cette propriété qui peut également se lire sur la transformée de Laplace (4.7), assure que la loi $IG(\gamma, \delta)$ est indéfiniment divisible. D'après le théorème 4.2.4, il existe un processus de Lévy $(Y_t)_{t\geq 0}$ appelé processus Inverse Gaussian de paramètre (γ, δ) tel que $Y_1 \sim IG(\gamma, \delta)$ et que $\mathbb{E}(e^{-\lambda Y_t}) = \mathbb{E}(e^{-\lambda Y_1})^t$, ce qui avec (4.7) assure que $Y_t \sim IG(\gamma, \delta t)$. Ce processus étant croissant, Barndorff-Nielsen [9] propose plutôt d'utiliser dans (4.4) $Z_t = B_{Y_t}$ où $B_t = \alpha + W_t + \mu t$ avec $(W_t)_{t\geq 0}$ mouvement brownien indépendant de $(Y_t)_{t\geq 0}$ et $\alpha, \mu \in \mathbb{R}$.
En reprenant la preuve du lemme 4.2.10, on vérifie que le processus $(Z_t)_{t\geq 0}$ est alors un

processus de Lévy tel que $Z_t \stackrel{\mathcal{L}}{=} \alpha + \sqrt{Y_t}G + \mu Y_t$ où G est une gaussienne centrée réduite indépendante de $(Y_t)_{t\geq 0}$. Il suffit de savoir simuler suivant les lois IG pour pouvoir générer les accroissements de ce processus. C'est l'objet du lemme suivant.

Lemme 4.2.13. *Soient $\delta > 0$, $X \sim \chi^2(1) = \Gamma(\frac{1}{2}, \frac{1}{2})$ de densité $\frac{1}{\sqrt{2\pi x}}e^{-x/2}1_{\{x>0\}}$ et U uniformément répartie sur $[0,1]$ indépendantes.*
Alors $\frac{\delta^2}{X} \sim IG(0,\delta)$ et lorque $\gamma > 0$,

$$\textit{si } Y_\pm = \frac{2\gamma\delta + X \pm \sqrt{4\gamma\delta X + X^2}}{\gamma^2}, \textit{ alors } Y = 1_{\{U\leq \frac{\delta}{\delta+\gamma Y_-}\}}Y_- + 1_{\{U> \frac{\delta}{\delta+\gamma Y_-}\}}Y_+ \sim IG(\gamma,\delta).$$

Notons que si $G^2 \sim \mathcal{N}_1(0,1)$, $G^2 \sim \chi^2(1)$, ce qui, avec la méthode polaire de simulation des gaussiennes centrées réduites, permet de simuler facilement suivant la loi $\chi^2(1)$.

Démonstration : Dans le cas $\gamma > 0$, la clé du résultat consiste à remarquer que si $Y \sim IG(\gamma,\delta)$, $Z = \frac{(\gamma Y-\delta)^2}{Y} \sim \chi^2(1)$.
La fonction $f(y) = \frac{(\gamma y-\delta)^2}{y}$ de dérivée $f'(y) = \frac{(\gamma y+\delta)(\gamma y-\delta)}{y^2}$ décroit de $+\infty$ à 0 sur $]0, \frac{\delta}{\gamma}[$ et croit de 0 à $+\infty$ sur $]\frac{\delta}{\gamma}, +\infty[$. Pour $z > 0$, l'équation $z = f(y)$ se récrit $y^2 - \frac{(2\gamma\delta+z)}{\gamma^2}y + \frac{\delta^2}{\gamma^2} = 0$. Le discriminant est $\Delta = \frac{4\delta\gamma z+z^2}{\gamma^4}$ et elle admet deux racines $y_\pm = \frac{2\gamma\delta+z\pm\sqrt{4\delta\gamma z+z^2}}{2\gamma^2}$ avec $y_+ > \frac{\delta}{\gamma}$ et $y_- < \frac{\delta}{\gamma}$. Enfin $f'(y_-) = -\frac{(\gamma y_-+\delta)|\gamma y_- - \delta|}{y_-^2} = -(\gamma y_- + \delta)y_-^{-3/2}\sqrt{z}$ et $f'(y_+) = (\gamma y_+ + \delta)y_+^{-3/2}\sqrt{z}$. Donc pour $\varphi : \mathbb{R} \to \mathbb{R}$ bornée,

$$\begin{aligned}\mathbb{E}(\varphi(Z)) &= \int_0^{+\infty} \varphi(z)\left(-\frac{IG(\gamma,\delta)}{f'}(y_-) + \frac{IG(\gamma,\delta)}{f'}(y_+)\right)dz \\ &= \int_0^{+\infty} \varphi(z)\left(\frac{1}{\gamma y_- + \delta} + \frac{1}{\gamma y_+ + \delta}\right)\frac{\delta}{\sqrt{2\pi z}}e^{-z/2}dz.\end{aligned}$$

Comme

$$\frac{1}{\gamma y_- + \delta} + \frac{1}{\gamma y_+ + \delta} = \frac{\gamma(y_- + y_+) + 2\delta}{\gamma^2 y_- y_+ + \delta\gamma(y_- + y_+) + \delta^2} = \frac{\frac{2\gamma\delta+z}{\gamma} + 2\delta}{\delta^2 + \delta\frac{2\gamma\delta+z}{\gamma} + \delta^2} = \frac{1}{\delta},$$

on conclut que $\mathbb{E}(\varphi(Z)) = \int_0^{+\infty}\varphi(z)e^{-z/2}\frac{dz}{\sqrt{2\pi z}}$. En reprenant les calculs en marche arrière, on voit qu'il suffit de prendre la racine Y_- avec probabilité $\frac{\delta}{\delta+Y_-} = \frac{(\delta+Y_-)^{-1}}{(\delta+Y_-)^{-1}+(\delta+Y_+)^{-1}}$ et la racine Y_+ avec probabilité complémentaire pour obtenir une variable aléatoire de loi $IG(\gamma,\delta)$ à partir de $X \sim \chi^2(1)$.

Lorsque $\gamma \to 0^+$,

$$\begin{aligned}\sqrt{4\gamma\delta X + X^2} = X\sqrt{1 + \frac{4\gamma\delta}{X}} &= X\left(1 + \frac{1}{2}\left(\frac{4\gamma\delta}{X}\right) - \frac{1}{16}\left(\frac{4\gamma\delta}{X}\right)^2\right) + o(\gamma^2) \\ &= X + 2\gamma\delta - \frac{\gamma^2\delta^2}{X} + o(\gamma^2),\end{aligned}$$

ce qui assure que Y_- tend p.s. vers $\frac{\delta^2}{X} > 0$. Donc $1_{\{U\leq\frac{\delta}{\delta+\gamma Y_-}\}}$ tend p.s. vers 1 et Y vers $\frac{\delta^2}{X}$. Par convergence dominée, on en déduit avec (4.7) que

$$\forall\lambda \geq 0,\ \mathbb{E}\left(e^{-\frac{\lambda\delta^2}{X}}\right) = \lim_{\gamma\to 0^+} e^{(\gamma - \sqrt{\gamma^2+2\lambda})\delta} = e^{-\delta\sqrt{2\lambda}},$$

ce qui assure que $\frac{\delta^2}{X} \sim IG(0,\delta)$. □

Remarque 4.2.14. — *Comme $W_{\tau_\delta} + \gamma\tau_\delta = \delta$, on a $\frac{(\gamma\tau_\delta - \delta)^2}{\tau_\delta} = \frac{W^2_{\tau_\delta}}{\tau_\delta}$, égalité qui peut donner un peu d'intuition sur la clé de la démonstration (pour tout $t > 0$ déterministe $\frac{W_t^2}{t} \sim \chi^2(1)$).*

— *Pour $\varphi : \mathbb{R} \to \mathbb{R}$ bornée, en effectuant le changement de variables $y = \frac{\delta^2}{x}$, on obtient*

$$\mathbb{E}\left(\varphi\left(\frac{\delta^2}{X}\right)\right) = \int_0^{+\infty} \varphi\left(\frac{\delta^2}{x}\right) e^{-\frac{x}{2}} \frac{dx}{\sqrt{2\pi x}} = \int_0^{+\infty} \varphi(y)\delta y^{-3/2} e^{-\frac{\delta^2}{2y}} \frac{dy}{\sqrt{2\pi}}. \tag{4.8}$$

On retrouve bien que $\frac{\delta^2}{X} \sim IG(0, \delta)$.

4.2.3 Processus stables symétriques

Définition 4.2.15. *Une variable aléatoire X et sa loi sont dites stables si pour Y indépendante de X et de même loi*

$$\forall a, b > 0,\ \exists c > 0,\ \exists d \in \mathbb{R},\ aX + bY \stackrel{\mathcal{L}}{=} cX + d. \tag{4.9}$$

Exemple 4.2.16. *1. Si $X \sim \mathcal{N}_1(\mu, \sigma^2)$, $aX + bY \sim \mathcal{N}_1((a+b)\mu, (a^2+b^2)\sigma^2)$. Donc $aX + bY \stackrel{\mathcal{L}}{=} \sqrt{a^2+b^2}X + ((a+b) - \sqrt{a^2+b^2})\mu$. Ainsi $c = \sqrt{a^2+b^2}$ et $d = ((a+b) - \sqrt{a^2+b^2})\mu$.*

2. Si X suit la loi de Cauchy de paramètre $\sigma > 0$ notée $\mathcal{C}(\sigma)$ de densité $\frac{\sigma}{\pi(x^2+\sigma^2)}$ et de fonction caractéristique $\mathbb{E}(e^{iuX}) = e^{-\sigma|u|}$, alors $\mathbb{E}(e^{iu(aX+bY)}) = e^{-\sigma(a+b)|u|} = \mathbb{E}(e^{iu(a+b)X})$. Ainsi $c = a + b$ et $d = 0$.

3. Si $X \sim IG(0, \delta)$, $X \stackrel{\mathcal{L}}{=} \inf\{t \geq 0 : W_t \geq \delta\} = \delta^2 \inf\{t \geq 0 : \frac{1}{\delta} W_{\delta^2 t} \geq 1\} \sim \delta^2 IG(0, 1)$, résultat que l'on déduit aussi du lemme 4.2.13. Donc $aX \sim IG(0, \sqrt{a}\delta)$ et $bY \sim IG(0, \sqrt{b}\delta)$ puis $aX + bY \sim IG(0, (\sqrt{a} + \sqrt{b})\delta)$. Ainsi $aX + bY \stackrel{\mathcal{L}}{=} (\sqrt{a} + \sqrt{b})^2 X$ et $c = (\sqrt{a} + \sqrt{b})^2$, $d = 0$.

Définition 4.2.17. *Une variable aléatoire X et sa loi sont dites symétriques si $X \stackrel{\mathcal{L}}{=} -X$.*

Le résultat suivant, que nous ne démontrerons pas, caractérise les lois stables symétriques au travers de leur fonction caractéristique.

Théorème 4.2.18. *Soit X une variable aléatoire stable. Il existe $\alpha \in]0, 2]$ tel que pour tous $a, b > 0$, la constante c dans (4.9) est donnée par $c^\alpha = a^\alpha + b^\alpha$. La constante α s'appelle indice de stabilité de X. Enfin, si X est stable symétrique d'indice α,*

$$\exists \sigma \geq 0,\ \forall u \in \mathbb{R},\ \mathbb{E}(e^{iuX}) = e^{-\sigma^\alpha |u|^\alpha}$$

et on note alors $X \sim S\alpha S(\sigma)$.

Remarque 4.2.19. — *Les trois exemples précédents correspondent respectivement à des lois stables d'indice 2, 1 et $\frac{1}{2}$.*

— *Si $X \sim S\alpha S(1)$ et $\sigma > 0$, alors $\mathbb{E}(e^{i(u\sigma)X}) = e^{-|u\sigma|^\alpha}$, ce qui assure que $\sigma X \sim S\alpha S(\sigma)$.*

— *$S2S(\sigma) = \mathcal{N}_1(0, 2\sigma^2)$ et $S1S(\sigma) = \mathcal{C}(\sigma)$.*

Exemple 4.2.20. *Si G_1 et G_2 sont deux gaussiennes centrées réduites indépendantes, $\frac{\sigma}{4}\left(\frac{1}{G_1^2} - \frac{1}{G_2^2}\right) \sim S\frac{1}{2}S(\sigma)$. En effet, cette variable aléatoire est symétrique et comme, d'après* *le lemme 4.2.13,* $\frac{1}{G_1^2} \sim IG(0,1)$*, il est facile de vérifier qu'elle est stable d'indice* $\frac{1}{2}$*. Enfin, comme pour* $\lambda \geq 0$, $\mathbb{E}\left(e^{-\lambda \frac{1}{G_1^2}}\right) = e^{-\sqrt{2\lambda}}$*, par prolongement analytique,* $\mathbb{E}\left(e^{iu\frac{1}{G_1^2}}\right) = e^{-(1-\mathrm{sign}(u)i)\sqrt{|u|}}$ *et* $\mathbb{E}\left(e^{iu\left(\frac{1}{G_1^2}-\frac{1}{G_2^2}\right)}\right) = e^{-2\sqrt{|u|}}$.

Soit $\alpha \in]0,2]$ et $\sigma, \sigma' > 0$. Si $X \sim S\alpha S(\sigma)$ et $Y \sim S\alpha S(\sigma')$ sont indépendantes, alors $\mathbb{E}(e^{iu(X+Y)}) = e^{-\sigma^\alpha |u|^\alpha} e^{-\sigma'^\alpha |u|^\alpha} = e^{-(\sigma^\alpha + \sigma'^\alpha)|u|^\alpha}$ et $X + Y \sim S\alpha S((\sigma^\alpha + \sigma'^\alpha)^{1/\alpha})$. Cela assure que la loi $S\alpha S(\sigma)$ est indéfiniment divisible. Avec le théorème 4.2.4, on en déduit l'existence d'un processus de Lévy $(Z_t)_{t\geq 0}$ appelé processus stable symétrique d'indice α tel que $Z_1 \sim S\alpha S(\sigma)$ et que $\mathbb{E}(e^{iuZ_t}) = \mathbb{E}(e^{iuZ_1})^t$. D'après le théorème 4.2.18, cette dernière égalité entraîne que $Z_t \sim S\alpha S(\sigma t^{1/\alpha})$.

Exemple 4.2.21.

— *Lorsque* $\alpha = 2$ *et* $\sigma = \frac{1}{\sqrt{2}}$*, ce processus est le mouvement brownien.*

— *Lorsque* $\alpha \in]0,2[$*, par le changement de variables* $x = uh$

$$\int_{\mathbb{R}} (e^{iuh} - 1 - iuh1_{\{|h|\leq 1\}})\frac{dh}{|h|^{\alpha+1}} = |u|^\alpha \int_{\mathbb{R}} (e^{ix} - 1 - ix1_{\{|x|\leq 1\}})\frac{dx}{|x|^{\alpha+1}}$$

où par parité de la fonction $\frac{1}{|x|^{\alpha+1}}$*, la constante* $\int_{\mathbb{R}}(e^{ix}-1-ix1_{\{|x|\leq 1\}})\frac{dx}{|x|^{\alpha+1}}$ *est réelle. On en déduit que le triplet associé à* Z_1 *par la formule de Lévy Kinchine donnée dans la proposition 4.2.5 est* $(0, 0, \frac{cdx}{|x|^{\alpha+1}})$ *avec c une constante de normalisation. Le processus est discontinu et t.q.*

$$\forall \beta, t > 0,\ \mathbb{E}(|Z_t|^\beta) < +\infty \Leftrightarrow \int_1^{+\infty} |x|^\beta \frac{cdx}{|x|^{\alpha+1}} < +\infty \Leftrightarrow \beta < \alpha.$$

Pour simuler ses accroissements, il suffit d'utiliser $Z_{t+s} - Z_t \stackrel{\mathcal{L}}{=} \sigma s^{1/\alpha} X$ où la variable aléatoire $X \sim S\alpha S(1)$ peut être générée grâce au résultat suivant.

Proposition 4.2.22. *Soit* Θ *variable uniforme sur* $[-\frac{\pi}{2}, \frac{\pi}{2}]$ *et* Y *variable exponentielle de paramètre* 1 *indépendantes. Alors*

$$X = \frac{\sin(\alpha\Theta)}{(\cos(\Theta))^{1/\alpha}} \left(\frac{\cos((1-\alpha)\Theta)}{Y}\right)^{(1-\alpha)/\alpha} \sim S\alpha S(1).$$

Plutôt que de démontrer ce résultat, nous allons voir que pour des valeurs particulières de α, il permet de retrouver des techniques de simulation bien connues :

Cas $\alpha = 1$: on retrouve que $X = \tan(\Theta) \sim \mathcal{C}(1)$.

Cas $\alpha = 2$: $X = \frac{\sin(2\Theta)}{\sqrt{\cos(\Theta)}}\left(\frac{\cos(-\Theta)}{Y}\right)^{-1/2} = \sqrt{2}\sqrt{2Y}\sin(\Theta) \stackrel{\mathcal{L}}{=} \sqrt{2}\sqrt{2Y}\sin(2\Theta)$. Comme $2Y$ est une variable exponentielle de paramètre $1/2$ indépendante de 2Θ qui suit la loi uniforme sur $[-\pi, \pi]$, d'après la méthode polaire, $\sqrt{2Y}\sin(2\Theta) \sim \mathcal{N}_1(0,1)$. On retrouve ainsi que $X \sim \mathcal{N}_1(0,2)$.

Cas $\alpha = \frac{1}{2}$: $X = \frac{\sin(\Theta/2)}{\cos^2(\Theta)} \times \frac{\cos(\Theta/2)}{Y} = \frac{\sin(\Theta)}{2Y\cos^2(\Theta)}$. Toujours d'après la méthode polaire, $(G_1, G_2) = \sqrt{2Y}(\sin(2\Theta), \cos(2\Theta)) \sim \mathcal{N}_2(0, I_2)$.

$$\frac{1}{4}\left(\frac{1}{G_1^2} - \frac{1}{G_2^2}\right) = \frac{1}{8Y}\left(\frac{1}{\sin^2(2\Theta)} - \frac{1}{\cos^2(2\Theta)}\right) = \frac{\cos^2(2\Theta) - \sin^2(2\Theta)}{8Y\sin^2(2\Theta)\cos^2(2\Theta)} = \frac{\cos(4\Theta)}{2Y\sin^2(4\Theta)}.$$

Soit $\varphi : \mathbb{R} \to \mathbb{R}$ bornée. Comme la fonction $\gamma \to \frac{\cos(\gamma)}{\sin^2(\gamma)}$ est périodique de période 2π et paire, on obtient en faisant la changement de variables $\theta = \frac{\pi}{2} - \gamma$,

$$\mathbb{E}\left(\varphi\left(\frac{\cos(4\Theta)}{\sin^2(4\Theta)}\right)\right) = \frac{1}{4\pi}\int_{-2\pi}^{2\pi} \varphi\left(\frac{\cos(\gamma)}{\sin^2(\gamma)}\right) d\gamma = \frac{1}{\pi}\int_0^{\pi} \varphi\left(\frac{\cos(\gamma)}{\sin^2(\gamma)}\right) d\gamma$$
$$= \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \varphi\left(\frac{\cos(\frac{\pi}{2}-\theta)}{\sin^2(\frac{\pi}{2}-\theta)}\right) d\theta = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \varphi\left(\frac{\sin(\theta)}{\cos^2(\theta)}\right) d\theta,$$

ce qui assure que $\frac{\cos(4\Theta)}{\sin^2(4\Theta)} \stackrel{\mathcal{L}}{=} \frac{\sin(\Theta)}{\cos^2(\Theta)}$.

4.2.4 EDS dirigées par des processus de Lévy

Pour discrétiser l'EDS $dX_t = \sigma(t, X_t^-)dZ_t$ dirigée par le processus de Lévy $(Z_t)_{t\geq 0}$ on peut utiliser le schéma d'Euler qui s'écrit $\bar{X}_{t_{k+1}} = \sigma(t_k, \bar{X}_{t_k})(Z_{t_{k+1}} - Z_{t_k})$.
Pour simplifier on suppose ici que la dimension d du processus de Lévy et la dimension n de la solution X sont toutes deux égales à 1. L'ordre de convergence faible du schéma reste en $\frac{1}{N}$ comme dans le cas sans saut dès lors que le processus $(Z_t)_{t\geq 0}$ ne comporte pas trop de grands sauts. Pour quantifier cela, on demande que la mesure de Lévy ν associée à Z_1 par la formule de Lévy Kinchine donnée dans la proposition 4.2.5 possède suffisamment de moments finis [66].

Lorsqu'il n'est pas possible de simuler les accroissements $Z_{t_{k+1}} - Z_{t_k}$, on peut simuler les grands sauts qui sont en valeur absolue plus grands qu'un seuil ε et approcher la contribution des petits sauts par un mouvement brownien avec dérive avec même moments d'ordre 1 et 2 [68] [38]. Enfin, [44] est consacré à la construction de schémas d'ordre de convergence faible élevé pour des EDSs dirigées par des processus de Lévy de sauts purs dans l'esprit de ceux introduits dans le paragraphe 3.5 dans le cas brownien.

Bibliographie

[1] Alfonsi, A. : On the discretization schemes for the CIR (and Bessel squared) processes, Monte Carlo Methods and Applications 11(4) :355-384, 2005.

[2] Alfonsi, A. : High order discretization schemes for the CIR process : Application to affine term structure and Heston models, Mathematics of Computation 79 :209-237, 2010.

[3] Alfonsi, A., Jourdain, B. and Kohatsu-Higa, A. : Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme, Ann. Appl. Probab. 24(3) :1049-1080, 2014

[4] Andrieu, C., Moulines, E. and Priouret, P. : Stability of stochastic approximation under verifiable conditions. SIAM J. Control Optim., 44 :283–312, 2005.

[5] Applebaum, D. : Lévy processes and stochastic calculus. Cambridge Studies in Advanced Mathematics, 93. Cambridge University Press, Cambridge, 2004.

[6] Bally, V. and Talay, D. : The law of the Euler scheme for stochastic differential equations (I) : convergence rate of the distribution function. Probability Theory and Related Fields, 104 :43-60, 1995.

[7] Bally, V. and Talay, D. : The law of the Euler scheme for stochastic differential equations (II) : convergence rate of the density. Monte Carlo Methods and Applications, 2 :93-128, 1996.

[8] A. Barducci, G. Bussi, and M. Parrinello. Well-tempered metadynamics : A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.*, 100 :020603, 2008.

[9] Barndorff-Nielsen, O. : Processes of normal inverse Gaussian type. Finance Stoch. 2(1) :41-68, 1998.

[10] Bally, V., Kohatsu-Higa, A. : A probabilistic interpretation of the parametrix method. Ann. Appl. Probab. 25(6) :3095-3138, 2015.

[11] Bertoin, J. : Lévy processes. Cambridge Tracts in Mathematics, 121. Cambridge University Press, Cambridge, 1996.

[12] Bencheikh, O. and Jourdain, B. : Convergence in total variation of the Euler-Maruyama scheme applied to diffusion processes with measurable drift coefficient and additive noise, Preprint arXiv :2005.09354.

[13] Beskos, A., Papaspiliopoulos, O. and Roberts, G. : Retrospective exact simulation of diffusion sample paths with applications. Bernoulli 12(6) :1077-1098, 2006.

[14] Beskos, A., Papaspiliopoulos, O., Roberts, G.and Fearnhead, P. : Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes. J. R. Stat. Soc. Ser. B Stat. Methodol. 68(3)333-382, 2006.

[15] Beskos, A. and Roberts, G. : Exact simulation of diffusions. Ann. Appl. Probab. 15(4) :2422-2444, 2005.

[16] Breiman, L. : Probability. Corrected reprint of the 1968 original. Classics in Applied Mathematics, 7. Society for Industrial and Applied Mathematics (SIAM), Philadelphia.

[17] G. Bussi, A. Laio, and M. Parrinello. Equilibrium free energies from nonequilibrium metadynamics. *Phys. Rev. Lett.*, 96 :090601, 2006.

[18] Carr, P., Geman, H., Madan, D.B. and Yor, M. : From local volatility to local Lévy models. Quant. Finance 4(5) :581–588, 2004.

[19] Y. Crespo, F. Marinelli, F. Pietrucci, and A. Laio. Metadynamics convergence law in a multidimensional system. *Phys. Rev. E*, 81(5) :055701, 2010.

[20] J.F. Dama, G.M. Hocky, R. Sun and G.A. Voth. Exploring valleys without climbing every peak : More efficient and forgiving metabasin metadynamics via robust on-the-fly bias domain restriction. *J. Chem. Theory Comput.*, 11(12) :5638-5650, 2015

[21] J.F. Dama, M. Parrinello, and G.A. Voth. Well-tempered metadynamics converges asymptotically. *Phys. Rev. Lett.*, 112 :240602(1–6), 2014.

[22] Del Moral, P. : Feynman-Kac formulae. Genealogical and interacting particle systems with applications. Probability and its Applications (New York). Springer-Verlag, 2004.

[23] G. Fort, B. Jourdain, E. Kuhn, T. Lelièvre, and G. Stoltz. Efficiency of the Wang-Landau algorithm : A simple test case. *Appl. Math. Res. Express*, 2014(2) :275–311, 2014.

[24] G. Fort, B. Jourdain, E. Kuhn, T. Lelièvre, and G. Stoltz. Convergence of the Wang-Landau algorithm. *Math. Comput.*, 84(295) :2297–2327, 2015.

[25] G. Fort, B. Jourdain, T. Lelièvre, and G. Stoltz. Self-Healing Umbrella Sampling : Convergence and efficiency. *Stat Comput.*, 27(1), 147-168, 2017.

[26] G. Fort, E. Moulines, and P. Priouret. Convergence of adaptive and interacting Markov chain Monte Carlo algorithms. *Ann. Statist.*, 39(6) :3262–3289, 2012.

[27] Gaines, J. and Lyons, T. : Random generation of stochastic area integrals. SIAM J. Appl. Math. 54(4) :1132-1146, August 1994.

[28] Giles, M. : Multi-level Monte Carlo path simulation, Operations Research, 56(3) :607-617, 2008.

[29] Giles, M. and Szpruch, L. : Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation, Ann. Appl. Probab. 24(4) :1585-1620, 2014.

[30] Gobet, E. : Revisiting the Greeks for European and American options. Proceedings of the "International Symposium on Stochastic Processes and Mathematical Finance" at Ritsumeikan University, Kusatsu, Japan, March 2003. 18 pages. 2004.

[31] Gobet, E. Weak approximation of killed diffusion using Euler schemes. Stochastic Processes and their Applications, 87 :167-197, 2000.

[32] Gobet, E. Euler Schemes and Half-Space Approximation for the Simulation of Diffusion in a Domain, ESAIM P&S, 5 :261-297, 2001.

[33] Gobet, E. Menozzi, S. Exact approximation rate of killed hypoelliptic diffusions using the discrete Euler scheme, Stochastic Processes and their Applications, 112(2) :201-223, 2004.

[34] Guyon, J. : Euler scheme and tempered distributions. Stochastic Processes and their Applications, 116(6) :877-904, 2006.

[35] Hairer, M., Mattingly, J. C., Yet another look at Harris' ergodic theorem for Markov chains. Seminar on Stochastic Analysis, Random Fields and Applications VI, 109-117, Progr. Probab., 63, Birkhäuser/Springer Basel AG, Basel, 2011.

[36] Hutzenthaler, M., Jentzen, A. and Kloeden, P. E., Strong convergence of an explicit numerical method for SDEs with non-globally Lipschitz continuous coefficients, Annals of Applied Probability 22(4) :1611-1641, 2012.

[37] Jacob, P.E. and Ryder, R.J. . The Wang-Landau algorithm reaches the flat histogram criterion in finite time. *Ann. Appl. Probab.*, 24(1) :34–53, 2014.

[38] Jacod, J., Kurtz, T.G., Méléard, S., Protter, P. : The approximate Euler method for Lévy driven stochastic differential, Ann. Inst. H. Poincaré Probab. Statist. 41(3) :523-558, 2005.

[39] Jarner, S.F., Hansen, E. : Geometric ergodicity of Metropolis algorithms, Stochastic Processes and their Applications, 85 :341-361, 2000.

[40] Jourdain, B. and Sbai, M. : Exact retrospective Monte Carlo computation of arithmetic average Asian options, Monte Carlo methods and Applications 13(2) :135-171, 2007.

[41] Jourdain, B. and Sbai, M. : High order discretization schemes for stochastic volatility models, Journal of Computational Finance 17(2), 2013 .

[42] Kloeden, P. E. and Platen, E. : Numerical Solution of Stochastic Differential Equations. Springer Berlin Heidelberg New York, 1992.

[43] Kemna, A. G. Z. and Vorst, A. C. F. : A Pricing Method for Options Based on Average Asset Values. Journal of Banking and Finance, 14 :113-129, March 1990.

[44] Kohatsu-Higa, A. and Tankov, P. : Jump adapted discretization schemes for Lévy driven SDEs, Stochastic Process. Appl., 120(11) :2258-2285, 2010.

[45] Kunita, H. : Stochastic flows and stochastic differential equations. Cambridge University Press, 1990.

[46] Kurtz, T. and Protter, P. : Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. Stochastic analysis, 331–346, Academic Press, Boston, MA, 1991.

[47] Kusuoka, S. : Approximation of expectation of diffusion process and mathematical finance. Taniguchi Conference on Mathematics Nara '98, 147-165, Adv. Stud. Pure Math., 31, Math. Soc. Japan, Tokyo, 2001.

[48] Kusuoka, S. : Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. Advances in mathematical economics. Vol. 6, 69-83, 2004.

[49] A. Laio and M. Parrinello. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A*, 99 :12562–12566, 2002.

[50] Lapeyre, B. and Temam, E. : Competitive Monte Carlo methods for the pricing of Asian Options. Journal of Computational Finance, 5(1) :39-59, Fall 2001

[51] Lyons, T. and Victoir, N. : Cubature on Wiener space. Stochastic analysis with applications to mathematical finance. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 460 :169-198, 2004.

[52] Mao, X and Yuan, C. : A note on the rate of convergence of the Euler-Maruyama method for stochastic differential equations. Stochastic Analysis and Applications, 26(2) :325-333, 2008.

[53] Milstein, G. N. : Numerical integration of stochastic differential equations. Translated and revised from the 1988 Russian original. Mathematics and its Applications, 313. Kluwer Academic Publishers Group, Dordrecht, 1995

[54] Madan, D., Carr, P. and Chang, E. : The Variance Gamma Process and Option Pricing, European Finance Review 2 :79-105, 1998.

[55] S. Marsili, A. Barducci, R. Chelli, P. Procacci, and V. Schettino. Self-healing Umbrella Sampling : A non-equilibrium approach for quantitative free energy calculations. *J. Phys. Chem. B*, 110(29) :14011–14013, 2006.

[56] M. McGovern and J. de Pablo. A boundary correction algorithm for metadynamics in multiple dimensions. *J. Chem. Phys.*, 139(8) :084102, 2013.

[57] Newton, N. : Variance reduction for simulated diffusions. SIAM J. Appl. Math. 54(6) :1780-1805, 1994.

[58] Newton, N. : Asymptotically efficient Runge-Kutta methods for a class of Itô and Stratonovich equations. SIAM J. Appl. Math. 51(2) :542-567, 1991.

[59] Ninomiya, S. : A new simulation scheme of diffusion processes : application of the Kusuoka approximation to finance problems. Math. Comput. Simulation 62 :479-486, 2003.

[60] Ninomiya, S. : A partial sampling method applied to the Kusuoka approximation. Monte Carlo Methods Appl. 9(1) :27-38, 2003.

[61] Ninomiya, S. and Victoir, N. : Weak approximation of stochastic differential equations and application to derivative pricing, Applied Mathematical Finance, 15(2) :107-121, 2008.

[62] Ninomiya, M. and Ninomiya, S. : A new higher-order weak approximation scheme for stochastic differential equations and the Runge-Kutta method. Finance and Stochastics 13 :415-443, 2009.

[63] Pagès, G. Convex order for path-dependent derivatives : a dynamic programming approach. Séminaire de Probabilités XLVIII, 33–96, Lecture Notes in Math., 2168, Springer, Cham, 2016.

[64] Pagès, G. Multi-step Richardson-Romberg extrapolation : remarks on variance control and complexity. Monte Carlo Methods Appl. 13(1) :37-70, 2007.

[65] Peskun, P.H. : Optimal Monte-Carlo sampling using Markov chains. Biometrika 60(3) :607-612, 1973.

[66] Protter, P. and Talay, D. : The Euler scheme for Lévy driven stochastic differential equations. Ann. Probab. 25(1) :393-423, 1997.

[67] Roberts, G.O., Tweedie, R.L. : Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. Biometrika 83(1) :95-110, 1996.

[68] Rubenthaler, S. : Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process. Stochastic Process. Appl. 103(2) :311-349, 2003.

[69] Sato, K. : Lévy processes and infinitely divisible distributions. Cambridge Studies in Advanced Mathematics, 68. Cambridge University Press, Cambridge, 1999.

[70] Talay, D. and Tubaro, L. : Expansion of the global error for numerical schemes solving stochastic differential equations. Stochastic Anal. Appl. 8(4) :483-509, 1990.

[71] Tanaka, H. and Kohatsu-Higa, A. : An operator approach for Markov chain weak approximations with an application to infinite activity Lévy driven SDEs. Ann. Appl. Probab 19(3) :1026-1062, 2009.

[72] Temam, E. : Couverture approchée d'options exotiques. Pricing des options asiatiques. PhD thesis, Université Paris 6, 2001.

[73] Tierney, L., A note on Metropolis-Hastings kernels for general state spaces, Ann. Appl. Probab. 8(1) :1-9, 1998

[74] F. Wang and D.P. Landau. Determining the density of states for classical statistical models : A random walk algorithm to produce a flat histogram. *Phys. Rev. E*, 64 :056101, 2001.

[75] F. Wang and D.P. Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.*, 86(10) :2050–2053, 2001.