Processus de Markov

et inégalités fonctionnelles

Cours de Master 2 - 2005/2006

Florent Malrieu

Table des matières

Chapitre 1

Le semi-groupe d'Ornstein-Uhlenbeck

Ce chapitre d'introduction présente l'exemple fondamental du processus d'Ornstein-Uhlenbeck. Pour la partie modélisation, on lira avec intérêt le livre de Nelson [Nel67] qui se trouve être un des initiateurs de la théorie des inégalités de Sobolev logarithmiques. Pour le reste de l'étude, on pourra consulter [ABC+00].

1.1 Un peu de modélisation

On voudrait, comme l'a fait avant nous Robert Brown, modéliser la dynamique d'une particule de pollen dans l'eau. Pour cela, on applique le principe fondamental de la dynamique qui assure que le produit de l'accélération d'un corps par sa masse est égal à la somme de forces qui s'exercent sur ce corps. La particule de pollen est soumise à :

1. son poids,
2. des forces dues aux chocs avec les molécules d'eau,
3. une force de frottement qui s'oppose et est proportionnelle à sa vitesse.

Les molécules d'eau sont de taille tout à fait négligeables par rapport à la particule de pollen qui elle est visible à l'œil nu. Durant une seconde (ou même un laps de temps plus petit), un très grand nombre de molécules d'eau sont venues heurter la particule de pollen. On modélise la variation de vitesse due aux chocs pendant un intervalle de temps t par une v.a. gaussienne de variance t. De plus, lorsque la particule s'est déplacée à l'échelle visible elle interagit avec des molécules d'eau totalement différentes de l'instant précédent. Il semble donc raisonnable de supposer que les accroissements de la vitesse dus aux chocs sont indépendants. On les modélisent donc par un mouvement brownien.

On peut alors modéliser la vitesse de la particule de pollen par l'équation différentielle stochastique suivante :

$$mdV_t = \sigma dB_t - \lambda V_t dt - m\overrightarrow{g}dt,$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien standard sur $\mathbb{R}^3$, m est la masse de la particule de pollen, σ et λ sont des constantes positives et $\overrightarrow{g}$ est le vecteur $(0,0,g)$ où g est la constante de gravitation.

Puisque ■la vitesse est la dérivée de la position■, la position de la particule est quant à elle solution de l'équation :

$$dX_t = V_t dt.$$

Remarquons enfin que les coefficients σ et λ seront d'autant plus grands que le volume de la particule sera grand. Le pollen ayant une faible masse volumique les constantes σ/m et λ/m

seront grandes devant g. Ceci nous indique qu'il peut être légitime de négliger le poids (au moins dans un premier temps).

On s'intéresse donc au système d'équation suivant :

$$\begin{cases} dX_t = V_t dt, \\ dV_t = \sigma dB_t - \lambda V_t dt. \end{cases}$$

En fait, nous allons principalement nous intéresser au processus des vitesses. Celui-ci est appelé processus d'Ornstein-Uhlenbeck. Il s'agit de l'exemple phare de ce cours.

1.2 Un peu de calcul stochastique

Considérons le processus d'Ornstein-Uhlenbeck $(X_t)_{t\geq 0}$ défini comme la solution de l'équation différentielle stochastique suivante :

$$dX_t = \sqrt{2}\, dB_t - X_t\, dt,$$

où X_0 est une variable aléatoire donnée et $(B_t)_{t\geq 0}$ est un mouvement brownien standard sur $\mathbb{R}$ indépendant de X_0. On notera $(X_t^x)_{t\geq 0}$ le processus issu de $x \in \mathbb{R}$ (*i.e.* tel que $X_0 = x$ p.s.).

Soit $\mathcal{A}$ l'ensemble des fonctions de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ à dérivées à croissance lente. Une fonction f est dite à croissance lente s'il existe un polynôme P tel que $|f| \leq P$. Dans toute la section, les fonctions considérées seront supposées appartenir à l'ensemble $\mathcal{A}$.

Définition 1.2.1. *Le semi-groupe d'Ornstein-Uhlenbeck sur $\mathbb{R}$ est la famille $(N_t)_{t\geq 0}$ d'opérateurs agissant sur les fonctions f de $\mathcal{A}$ par :*

$$N_t(f)(x) \overset{déf.}{=} \int_{\mathbb{R}} f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) \gamma(dy),$$

où γ désigne la mesure gaussienne standard sur $\mathbb{R}$.

Proposition 1.2.2. *Le semi-groupe d'Ornstein-Uhlenbeck s'exprime en fonction du processus d'Ornstein-Uhlenbeck de la manière suivante :*

$$N_t f(x) = \mathbb{E} f(X_t^x).$$

Réciproquement, si X_0 est une v.a. indépendante de B de carré intégrable de loi ν alors

$$\mathbb{E}_\nu f(X_t) = \int_{\mathbb{R}} N_t f(x)\, \nu(dx).$$

Preuve. La formule d'Itô appliquée à la fonction $f(t,x) = e^t x$ assure que

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \left[\partial_s f(s, X_s) + \partial_x f(s, X_s) + \partial_x^2 f(s, X_s)\right] ds + \int_0^t \partial_x f(s, X_s)\, dB_s \\ &= x + \int_0^t (e^s X_s - e^s X_s)\, ds + \int_0^t e^s\, dB_s. \end{aligned}$$

Donc le processus d'Ornstein-Uhlenbeck issu de x peut encore se représenter par la formule suivante :

$$X_t = xe^{-t} + \sqrt{2} \int_0^t e^{-(t-s)}\, dB_s.$$

D'après le théorème de représentation des martingales, la variable aléatoire

$$M_t = \sqrt{2}\int_0^t e^s\, dB_s$$

est de loi gaussienne centrée de variance

$$\langle M \rangle_t = \int_0^t 2e^s\, ds = e^{2t} - 1.$$

Ceci entraîne que la loi de X_t^x est la loi $\mathcal{N}(xe^{-t}, 1 - e^{-2t})$, ce qui assure bien que $N_t f(x) = \mathbb{E} f(X_t^x)$. □

Proposition 1.2.3. *La famille $(N_t)_{t\geq 0}$ vérifie les propriétés suivantes :*

(i) pour tout $t \geq 0$, $N_t \mathbf{1}(x) = 1$,

(ii) pour tous $t \geq 0$ et $f \in \mathcal{A}$, $f \geq 0$ entraîne $N_t f \geq 0$,

(iii) pour toute fonction $f \in \mathcal{A}$, $N_0 f = f$,

(iv) pour tous $s, t \geq 0$ et $f \in \mathcal{A}$, $N_t N_s f(x) = N_{t+s} f(x)$.

Preuve. Les points (i), (ii) et (iii) sont évidents. Établissons le point (iv). Soient X et Y deux variables aléatoires indépendantes de loi γ. Pour alléger les calculs nous noterons :

$$\sigma_t = \sqrt{1 - e^{-2t}}.$$

Par définition du semi-groupe N_t,

$$\begin{aligned} N_s \circ N_t(f)(x) &= \mathbb{E}\big[N_t(f)(e^{-s}x + \sigma_s X)\big] \\ &= \mathbb{E}\big[f(e^{-t}(e^{-s}x + \sigma_s X) + \sigma_t Y)\big] \\ &= \mathbb{E}\big[f(e^{-s-t}x + e^{-t}\sigma_s X + \sigma_t Y)\big]. \end{aligned}$$

La loi de $e^{-t}\sigma_s X + \sigma_t Y$ est la loi gaussienne $\mathcal{N}(0, \sigma_{s+t}^2)$ donc

$$N_s \circ N_t(f)(x) = \mathbb{E}\big[f(e^{-s-t}x + \sigma_{s+t} Z)\big] = N_{t+s}(f)(x),$$

où Z est de loi γ, ce qui est le résultat escompté. □

1.3 Générateur infinitésimal

La proposition 1.2.3 suggère que le processus d'Ornstein-Uhlenbeck est un processus de Markov et en particulier que ■conditionnellement au présent, passé et futur sont indépendants■. Une autre façon de s'en rendre compte est de reprendre le calcul de la preuve de la proposition 1.2.2 pour montrer que pour tous $0 \leq s \leq t$,

$$X_t = X_s e^{-(t-s)} + \int_s^t e^{-(t-u)} dB_u.$$

Par indépendance des accroissements du mouvement brownien,

$$\int_s^t e^{-(t-u)} dB_u$$

est indépendant de $\sigma(B_r,\ 0 \le r \le s)$ et donc de $\sigma(X_r,\ 0 \le r \le s) = \mathcal{F}_s$. On a donc

$$\mathbb{E}_x(f(X_t)|\mathcal{F}_s) = X_s e^{-(t-s)} + \mathbb{E}\left(\int_s^t e^{-(t-u)} dB_u\right) = X_s e^{-(t-s)}.$$

En d'autres termes la position à l'instant t ne dépend de l'information jusqu'au temps s que par la donnée de la position à l'instant s. La dynamique du processus d'Ornstein-Uhlenbeck, comme celle des processus de Markov plus généraux peut donc être décrite à partir de la description de l'évolution infinitésimale du processus.

Proposition 1.3.1. *Pour toute fonction f de $\mathcal{A}$,*

$$\lim_{t\to 0} \frac{1}{t}(N_t f - f)(x) = f''(x) - xf'(x).$$

L'opérateur A défini sur $\mathcal{A}$ par

$$Af(x) = f''(x) - xf'(x)$$

sera appelé générateur infinitésimal du semi-groupe d'Ornstein-Uhlenbeck.

Preuve. Soit $f \in \mathcal{A}$, la formule de Taylor à l'ordre 2 avec reste borné au voisinage de x assure que

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + O(|h|^3). \tag{1.1}$$

Remarquons dans un premier temps que pour $h = (e^{-t} - 1)x + \sigma_t y$, on a

$$\int h\,\gamma(dy) = (e^{-t} - 1)x, \quad \int h^2\,\gamma(dy) = (e^{-t} - 1)^2 x^2 + \sigma_t^2 \quad \text{et} \quad \int |h|^3\,\gamma(dy) = O(t^{3/2}).$$

L'équation (1.1) assure donc :

$$N_t f(x) = f(x) + (e^{-t} - 1)xf'(x) + \frac{1}{2}\big((e^{-t} - 1)^2 x^2 + \sigma_t^2\big) f''(x) + O(t^{3/2}).$$

On a donc bien la limite annoncée :

$$\lim_{t\to 0} \frac{1}{t}(N_t f(x) - f(x)) = f''(x) - xf'(x).$$

□

Il est possible de retrouver directement le générateur infinitésimal de $\mathbb{E}f(X_t^x)$. Soit f dans $\mathcal{A}$. La formule d'Itô assure que

$$f(X_t^x) = f(x) + \int_0^t [f''(X_s^x) - X_s^x f'(X_s^x)]\,ds + \underbrace{\sqrt{2}\int_0^t f'(X_s^x)\,dB_s}_{=M_t},$$

où M_t est une martingale (donc d'espérance nulle). En prenant l'espérance, on obtient :

$$\mathbb{E}f(X_t^x) = f(x) + \int_0^t \mathbb{E}[Af(X_s^x)]\,ds,$$

où $Af(x) = f''(x) - xf'(x)$ est le générateur infinitésimal du semi-groupe d'Ornstein-Uhlenbeck. Donc

$$\lim_{t\to 0} \frac{1}{t}(\mathbb{E}f(X_t^x) - f(x)) = \lim_{t\to 0} \frac{1}{t}\int_0^t \mathbb{E}[Af(X_s^x)]\,ds = Af(x).$$

Proposition 1.3.2. *Le semi-groupe d'Ornstein-Uhlenbeck vérifie la propriété de dérivation suivante :*

$$\partial_t N_t f = A(N_t f) = N_t(Af).$$

Exercice 1.3.3. *Démontrer la proposition 1.3.2.*

Cette propriété est très important car elle montre en particulier que, pour f_0 dans $\mathcal{A}$, la fonction f définie sur $[0,+\infty[\times\mathbb{R}$ par $f(t,x) \stackrel{\text{déf.}}{=} N_t f_0(x)$ est solution de l'équation aux dérivées partielles suivante :

$$\begin{cases} \partial_t f(t,x) = Af(t,x) = \partial^2_{xx} f(t,x) - x\partial_x f(t,x) & \text{pour } (t,x) \in]0,+\infty[\times\mathbb{R}, \\ f(0,x) = f_0(x). \end{cases}$$

1.4 Mesure invariante et symétrique

Une des questions principales auxquelles nous essaierons de répondre dans la suite est l'étude de l'évolution de la loi de X_t au cours du temps. La première étape est de chercher une loi telle que si X_0 suit cette loi alors c'est aussi le cas de X_t pour tout $t \geq 0$. On dit alors que la mesure est invariante. La réponse ici est facile : la mesure gaussienne centrée réduite γ est invariante pour le semi-groupe d'Ornstein-Uhlenbeck, c'est-à-dire que pour toute fonction $f \in \mathcal{A}$,

$$\int Af(x)\,\gamma(dx) = 0,$$

ou de manière équivalente, pour tout $t \geq 0$,

$$\int N_t f(x)\,\gamma(dx) = \int f(x)\,\gamma(dx).$$

D'après la définition du semi-groupe d'Ornstein-Uhlenbeck, pour tout $x \in \mathbb{R}$, $N_t f(x)$ converge vers $\int f\,d\gamma$. La mesure invariante apparaît donc comme ■le bon candidat■ pour le comportement du processus en temps grand. Dans la suite nous montrerons comment étudier cette convergence et donner notamment des vitesses explicites de convergence à l'équilibre.

En fait le lien entre la mesure γ et le semi-groupe (N_t) est plus fort comme le montre le résultat suivant.

Proposition 1.4.1. *Le semi-groupe d'Ornstein-Uhlenbeck est symétrique pour la mesure γ et A est auto-adjoint dans $L^2(\gamma)$ c'est-à-dire que pour toutes fonctions $f, g \in \mathcal{A}$,*

$$\int N_t f(x) g(x)\,\gamma(dx) = \int f(x) N_t g(x)\,\gamma(dx),$$

et

$$\int Af(x) g(x)\,\gamma(dx) = \int f(x) Ag(x)\,\gamma(dx) = -\int \nabla f(x) \nabla g(x)\,\gamma(dx).$$

Preuve. Par définition du semi-groupe d'Ornstein-Uhlenbeck,

$$\int_{\mathbb{R}} f(x) N_t g(x)\,\gamma(dx) = \mathbb{E}\big[f(X) g\big(e^{-t}X + \sigma_t Y\big)\big], \tag{1.2}$$

où X et Y sont deux v.a. indépendantes de loi γ. Le couple $(X, e^{-t}X + \sigma_t Y)$ est un vecteur gaussien centré de matrice de covariance

$$\begin{pmatrix} 1 & e^{-t} \\ e^{-t} & 1 \end{pmatrix}$$

donc sa loi est symétrique et l'on peut donc échanger les rôles de f et g dans (1.2). □

En choisissant $g = \mathbf{1}$, on se rend compte immédiatement que la propriété de symétrie implique celle d'invariance, c'est-à-dire que

$$\int N_t f \, d\gamma = \int f \, d\gamma.$$

1.5 Propriété de contractivité

Dans ce paragraphe, $\|\cdot\|_p$ désigne la norme de l'espace $L^p(\gamma)$ pour $\in [1, \infty]$. De la définition du semi-groupe d'Ornstein-Uhlenbeck, il découle que

$$\|N_t f\|_\infty \leq \|f\|_\infty,$$

c'est-à-dire que N_t est une contraction de L^∞. De plus, pour $p \geq 1$, l'inégalité de Jensen et la propriété d'invariance de γ assurent que

$$\|N_t f\|_p^p = \int (N_t f)^p \, d\gamma \leq \int N_t(|f|^p) \, d\gamma = \int |f|^p \, d\gamma = \|f\|_p^p.$$

Ainsi, nous avons établi le résultat suivant.

Lemme 1.5.1. *L'opérateur N_t est une contraction dans tous les espaces $L^p(\gamma)$ pour $p \in [1, \infty]$.*

Cette propriété n'est pas propre au processus d'Ornstein-Uhlenbeck. Les processus de Markov dans leur grande majorité la partage. Par contre, nous montrerons que N_t a des propriétés régularisantes : si $f \in L^p(\gamma)$ alors $N_t f \in L^q(\gamma)$ avec $q > p$. Cette propriété dite d'hypercontractivité a été une des grandes motivations de l'étude des processus de Markov.

1.6 Polynômes de Hermite

Les polynômes forment un sous-espace dense de $L^2(\gamma)$. En effet, soit $f \in L^2(\gamma)$ orthogonale à tous les polynômes. La transformée de Laplace de la mesure $\nu = f\gamma$ est finie sur $\mathbb{R}$ tout entier : pour $\lambda \in \mathbb{R}$

$$\int e^{\lambda x} \, \nu(dx) \leq \int f^2 \, d\gamma \int e^{2\lambda x} \, \gamma(dx) < \infty.$$

La transformée de Laplace de ν est donc en particulier analytique au voisinage de 0. Puisque f est orthogonale à tous les polynômes, toutes les dérivées de cette transformée sont nulles en 0 et par suite elle est donc également nulle. Ceci implique la nullité de f et la densité des polynômes.

On peut donc trouver une base hilbertienne de $L^2(\gamma)$ formée de polynômes. Les polynômes de Hermite, que nous noterons $(H_n)_n$, en sont un exemple. Ils forment en effet une famille de polynômes orthogonaux dans $L^2(\gamma)$, c'est-à-dire que si $m \neq n$,

$$\int H_m(x) H_n(x) \, \gamma(dx) = 0.$$

On peut les définir à partir de leur série génératrice :

$$G(s,x) = \exp(sx - s^2/2) = \sum_{n=0}^{\infty} \frac{s^n}{n!} H_n(x),$$

c'est-à-dire que $H_n(x) = \partial_s^n G(s,x)_{|s=0}$.

Exercice 1.6.1. *Donner les quatre premiers polynômes de Hermite en utilisant leur orthogonalité puis grâce à la formule ci-dessus.*

Ces polynômes jouent un rôle très important pour le semi-groupe d'Ornstein-Uhlenbeck. Soient $s, t \in \mathbb{R}$ fixés. Appliquons N_t à la fonction $x \mapsto G(s,x)$:

$$N_t(G(s,\cdot))(x) = \exp(-s^2/2)\mathbb{E}\big[\exp(se^{-t}x + s\sigma_t Y)\big] = \exp(se^{-t}x - s^2/2)\mathbb{E}[\exp(s\sigma_t Y)].$$

La transformée de Laplace de γ vaut $\mathbb{E}(\exp(\lambda Y)) = \exp(\lambda^2/2)$ donc

$$N_t(G(s,\cdot))(x) = \exp(se^{-t}x - s^2 e^{-2t}/2) = G(se^{-t}, x).$$

Par définition des polynômes de Hermite, on obtient

$$\begin{aligned} N_t(H_n)(x) &= N_t(\partial_s^n G(s,\cdot)_{|s=0})(x) = \partial_s^n N_t(G(s,\cdot))(x)_{|s=0} \\ &= \partial_s^n G(se^{-t}, x)_{|s=0} = e^{-nt}\partial_s^n G(se^{-t}, x)_{|s=0} = e^{-nt} H_n(x). \end{aligned}$$

Le polynôme de Hermite H_n est donc vecteur propre de N_t, de valeur propre e^{-nt}.

En utilisant la symétrie de N_t par rapport à γ, on obtient, pour tous entiers m et n et tout $t > 0$,

$$e^{-mt} \int H_m H_n \, d\gamma = \int N_t(H_m) H_n \, d\gamma = \int H_m N_t(H_n) \, d\gamma = e^{-nt} \int H_m H_n \, d\gamma.$$

En particulier,

$$m \neq n \quad \Longrightarrow \quad \int H_m H_n \, d\gamma = 0,$$

c'est-à-dire que H_m et H_n sont orthogonaux dans $L^2(\gamma)$.

La formule de Plancherel fournit la valeur de norme de H_n :

$$\sum_{n=0}^{\infty} \frac{s^{2n}}{n!^2} \|H_n\|_2^2 = \int G(s,x)^2 \, \gamma(dx) = \exp(-s^2) \int e^{-2sx} \, \gamma(dx) = \exp(s^2) = \sum_{n=0}^{\infty} \frac{s^{2n}}{n!}.$$

En identifiant les deux séries, il vient $\|H_n\|_2^2 = n!$. Les polynômes $H_n/\sqrt{n!}$ forment une base hilbertienne de $L^2(\gamma)$.

Remarque 1.6.2. *Soit f une fonction de $L^2(\gamma)$. Il existe $(a_n)_{n\geq 0}$ telle que $\sum_n n! a_n^2 < \infty$ et f décompose sous la forme $\sum_n a_n H_n$. On en déduit la décomposition de la fonction $N_t f$: $N_t f = \sum_n e^{-nt} a_n H_n$.*

Exercice 1.6.3. *Montrer que l'ensemble des valeurs propres de A est exactement $-\mathbb{N}$ (l'ensemble des entiers négatifs ou nuls). Quelle sont les multiplicités de ces valeurs propres et leurs sous-espaces propres ?*

Retenons de ce calcul les leçons suivantes :

– 0 est valeur propre de N_t et la fonction **1** est vecteur propre associée,
– les valeurs propres de A sont négatives ou nulles puisque si λ est valeur propre associée à f, alors
$$\lambda \int f^2 \, d\gamma = \int f A f \, d\gamma = - \int |\nabla f|^2 \, d\gamma \leq 0,$$
d'après la proposition 1.4.1.
– le sous espace propre de 0 ne contient que les constantes et 0 est une valeur propre isolée. Il y a donc un ■trou dans le spectre■ entre 0 et la première valeur propre non nulle.

Nous verrons dans la suite que les deux premières propriétés sont tout à fait générales tandis que la dernière joue un rôle capital dans le comportement en temps long du semi-groupe.

Chapitre 2

Semi-groupes de Markov

Le processus d'Ornstein-Uhlenbeck est un représentant pédagogique de la classe des processus de Markov à temps continu. Ces processus possède principalement la propriété d'indépendance du futur et du passé conditionnellement au présent comme c'est le cas pour les chaînes de Markov à temps discret.

Dans toute le suite, les processus que nous considérerons seront indexé par le temps continu et seront *a priori* à valeurs dans $\mathbb{R}^d$. Ceci pose de nombreux problèmes, notamment de mesurabilité, par rapport à la situation des chaînes de Markov finies à temps discret.

Les premières sections de ce chapitre concernent les définitions et premières propriétés des processus de Markov. On pourra se reporter à [RY91] pour les preuves manquantes. La section 3 donne un aperçu de l'étude des chaînes de Markov à temps continu. Seul le cas de l'espace d'état fini est abordé. L'ouvrage de Norris [Nor98] mène une étude plus complète et fournit des exemples. La dernière section traite des processus de diffusion, ou plus exactement d'exemples de processus de diffusion. Nous étudierons principalement des processus de Kolmogorov et des diffusions à valeurs dans un intervalle de $\mathbb{R}$. On pourra utilement se référer à [Roy99] pour les premiers et [KS] pour les seconds.

2.1 Processus de Markov

Par définition même des processus de Markov, la dynamique d'un tel processus après un instant t connaissant sa position à l'instant t est parfaitement déterminée. Pour décrire un processus donné, il suffit donc de décrire la loi du processus issu d'un point donné à un instant donné et en particulier il faut pourvoir décrire la loi au temps t sachant que le processus était en x au temps s. Pour cela, nous introduisons la notions de noyau de transition (ou probabilité de transition) qui est tout simplement un une famille de mesures de probabilité dépendant mesurablement d'un paramètre.

Définition 2.1.1. *Soit $(E, \mathcal{E})$ un espace mesuré. Un noyau N sur E est une application de $E \times \mathcal{E}$ dans $\mathbb{R}_+ \cup \{+\infty\}$ tel que*

(i) pour tout $x \in E$, l'application $A \mapsto N(x, A)$ est mesurable et positive sur $\mathcal{E}$,

(ii) pour tout $A \in \mathcal{E}$, l'application $x \mapsto N(x, A)$ est $\mathcal{E}$-mesurable.

Un noyau π est appelé probabilité de transition (ou noyau de transition) si $\pi(x, E) = 1$ pour tout $x \in \mathbb{E}$.

Si $f \in \mathcal{E}_+$, c'est-à-dire que f est une application de $(E, \mathcal{E})$ dans $(\mathbb{R}_+, \mathcal{B})$ mesurable, alors on définit la fonction Nf sur E par

$$Nf(x) = \int_E N(x, dy) f(y).$$

La fonction Nf appartient encore à $\mathcal{E}_+$. De plus, si M et N sont deux noyaux alors MN définit par

$$MN(x, A) = \int_E M(x, dy) N(y, A)$$

est encore un noyau.

La probabilité de transition π induit un mouvement aléatoire dans E qui est décrit de la façon suivante. Au temps initial, la ■particule■ est en x, la position x_1 au temps 1 sera choisie selon la mesure de probabilité $\pi(x, \cdot)$, la position x_2 au temps 2 selon la mesure de probabilité $\pi(x_1, \cdot)$, etc. Le processus ainsi obtenu est appelé chaîne de Markov (le temps est discret) homogène (la transition ne dépend pas du temps). Un processus de Markov est l'analogue en temps continu.

Supposons que le processus X soit tel que pour tous $s < t$, il existe une probabilité de transition $P_{s,t}$ telle que

$$\mathbb{P}(X_t \in A | \sigma(X_u, u \leq s)) = P_{s,t}(X_s, A) \quad \text{p.s.}$$

Alors, pour toute fonction $f \in \mathcal{E}_+$, $\mathbb{E}(f(X_t)|\sigma(X_u, u \leq s)) = P_{s,t}f(X_s)$. Soient $s < t < v$, alors

$$\begin{aligned} P_{s,v}(X_s, A) &= \mathbb{P}(X_v \in A | \sigma(X_u, u \leq s)) \\ &= \mathbb{E}\big(\mathbb{E}\big(\mathbf{1}_{\{X_v \in A\}} | \sigma(X_u, u \leq t)\big) | \sigma(X_u, u \leq s)\big) \\ &= \mathbb{E}(P_{t,v}(X_t, A) | \sigma(X_u, u \leq s)) \\ &= \int P_{s,t}(X_s, dy) P_{t,v}(y, A). \end{aligned}$$

Ceci nous amène à la définition suivante de fonction de transition.

Définition 2.1.2. *Une fonction de transition (en abrégé f.t.) sur* $(E, \mathcal{E})$ *est une famille* $(P_{s,t})_{0 \leq s < t}$ *de transitions de probabilité sur* $(E, \mathcal{E})$ *telle que pour tous* $s < t < v$*, on ait*

$$\int_E P_{s,t}(x, dy) P_{t,v}(y, A) = P_{s,v}(x, A) \tag{2.1}$$

pour tous $x \in E$ *et* $A \in \mathcal{E}$*.*

La relation (2.1) est appelé équation de Chapman-Kolmogorov. La f.t. est dite homogène si $P_{s,t}$ ne dépend de s et t que par la différence $t - s$. Dans ce cas, on note P_t pour $P_{0,t}$ et l'équation de Chapman-Kolmogorov s'écrit

$$P_{t+s}(x, A) = \int P_s(x, dy) P_t(y, A)$$

pour tous $s, t \geq 0$. On dit que $(P_t)_{t \geq 0}$ est un semi-groupe.

Définition 2.1.3. *Soit* $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ *un espace de probabilité filtré. Un processus adapté* X *est un processus de Markov pour la filtration* $(\mathcal{F}_t)$*, de fonction de transition* $P_{s,t}$ *si, pour toute fonction* $f \in \mathcal{E}_+$ *et* $s < t$*,*

$$\mathbb{E}(f(X_t)|\mathcal{F}_s) = P_{s,t}f(X_s) \quad \mathbb{P} \text{ p.s.}$$

Le processus est dit homogène si la f.t. est homogène et la formule précédente s'écrit alors

$$\mathbb{E}(f(X_t)|\mathcal{F}_s) = P_{t-s}f(X_s) \quad \mathbb{P} \text{ p.s.}$$

Exemple 2.1.4. *Le mouvement brownien sur $\mathbb{R}^d$ est un processus de Markov homogène de f.t.*

$$P_t f(x) = \int f(y) \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|x-y|^2}{2t}\right) dy.$$

La première remarque d'importance qu'il convient de faire est la suivante. La définition du processus de Markov prescrit la loi au temps t connaissant la position au temps $s < t$. Il convient de se demander si, pour toute fonction de transition, il existe un processus de Markov associé et si oui, s'il est unique. En d'autres termes, en quoi la description de la dynamique du temps s au temps t caractérise-t-elle toute la dynamique du processus. On remarque tout d'abord que les lois cylindriques sont entièrement déterminées par la propriété de conditionnement successif.

Proposition 2.1.5. *Un processus X est un processus de Markov par rapport à sa filtration $(\mathcal{F}_t) = (\sigma(X_u, u \le t))$ de f.t. $P_{s,t}$ et de loi initiale ν si et seulement si, pour tous $0 = t_0 < t_1 < \cdots < t_k$ et $f_0, \ldots, f_k \in \mathcal{E}_+$,*

$$\mathbb{E}\left[\prod_{i=0}^{k} f_i(X_{t_i})\right] = \int_E \nu(dx_0) f_0(x_0) \int_E P_{0,t_1}(x_0, dx_1) f_1(x_1) \ldots \int_E P_{t_{k-1},t_k}(x_{k-1}, dx_k) f_k(x_k). \quad (2.2)$$

Preuve. Supposons dans un premier temps que X est un processus de Markov. On peut alors écrire

$$\mathbb{E}\left[\prod_{i=0}^{k} f_i(X_{t_i})\right] = \mathbb{E}\left[\prod_{i=0}^{k-1} f_i(X_{t_i}) \mathbb{E}\big[f_k(X_{t_k}) | \mathcal{F}_{t_{k-1}}\big]\right] = \mathbb{E}\left[\prod_{i=0}^{k-1} f_i(X_{t_i}) P_{t_{k-1},t_k} f_k(X_{t_{k-1}})\right].$$

Cette expression ne fait plus intervenir que les temps $t_0, \ldots, t_{k-1}$ et la fonction f_{k-1} est remplacée par $f_{k-1} P_{t_{k-1},t_k} f_k$. On peut alors procéder par récurrence pour obtenir (2.2).

Réciproquement, pour prouver que X est un processus de Markov, il faut montrer que

$$\mathbb{E}(f(X_t)|\mathcal{F}_s) = P_{s,t} f(X_s) \quad \mathbb{P} \text{ p.s.}$$

Pour cela, en vertu du théorème des classes monotones, il suffit de montrer que pour tous $t_1 < \cdots < t_k \le s < t$ et $f_1, \ldots, f_k, g \in \mathcal{E}_+$,

$$\mathbb{E}\left[\prod_{i=0}^{k} f_i(X_{t_i}) g(X_t)\right] = \mathbb{E}\left[\prod_{i=0}^{k} f_i(X_{t_i}) P_{s,t} g(X_s)\right],$$

ce qui est clairement assuré par (2.2). □

Remarque 2.1.6. *On pourrait réécrire formellement la relation (2.2) comme :*

$$\mathbb{P}(X_{t_0} \in dx_0, X_{t_1} \in dx_1, \ldots, X_{t_k} \in dx_k) = \nu(dx_0) P_{0,t_1}(x_0, dx_1) \ldots P_{t_{k-1},t_k}(x_{k-1}, dx_k)$$

qui est l'analogue de la formule classique donnant la loi de $(X_0, \ldots, X_k)$ pour une chaîne de Markov à espace d'états dénombrable de matrice de transition P et de loi initiale ν :

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_k = i_k) = \nu(i_0) P_{i_0,i_1} \ldots P_{i_{k-1},i_k}.$$

Il est à présent possible de répondre à la question de l'existence et l'unicité du processus de Markov associé à une f.t. donnée : la donnée de la f.t. fournit les lois finies-dimmensionnelles, le théorème d'extension de Kolmogorov assure alors le résultat. On va construire le processus sur l'espace $\Omega = E^{\mathbb{R}_+}$, $\mathcal{F}_\infty = \mathcal{E}^{\mathbb{R}_+}$ et $\mathcal{F}_t = \sigma(X_u, u \leq t)$ où X est le processus coordonnée *i.e.* $X_t(\omega) = \omega(t)$.

Théorème 2.1.7. *Soit $(E, \mathcal{E})$ un espace polonais muni de sa tribu borélienne. Soit $P_{s,t}$ une f.t. sur $(E, \mathcal{E})$. Pour toute mesure de probabilité ν sur $(E, \mathcal{E})$, il existe une unique mesure de probabilité $\mathbb{P}_\nu$ sur $(\Omega, \mathcal{F}_\infty)$ telle que X soit un processus de Markov pour la filtration $(\mathcal{F}_t)$ de f.t. $P_{s,t}$ et de mesure initiale ν.*

Dans toute la suite du chapitre, on se restreint aux processus homogènes en temps.

Remarquons immédiatement que l'on a alors

$$\mathbb{P}_\nu(X_0 \in A_0, X_{t_1} \in A_1 \dots X_{t_k} \in A_k) \\ = \int_{A_0} \nu(dx_0) \int_{A_1} P_{t_1}(x_0, dx_1) \int_{A_2} P_{t_2 - t_1}(x_1, dx_2) \dots \int_{A_k} P_{t_k - t_{k-1}}(x_{k-1}, dx_k).$$

En particulier, pour tout $x \in E$, la mesure de probabilité $\mathbb{P}_{\delta_x}$, que nous noterons plus simplement $\mathbb{P}_x$, est la loi du processus X issu de x. Si Z est une v.a. $\mathcal{F}_\infty$-mesurable et positive (ou bornée), nous noterons son espérance par rapport à $\mathbb{P}_x$ (resp. $\mathbb{P}_\nu$) $\mathbb{E}_x(Z)$ (resp. $\mathbb{E}_\nu(Z)$). En particulier, si $Z = \mathbf{1}_{\{X_t \in A\}}$, on obtient

$$\mathbb{E}_x(Z) = \mathbb{P}_x(X_t \in A) = P_t(x, A),$$

ce qui se lit ■ la probabilité que le processus issu de x soit en A au temps t est donnée par $P_t(x, A)$■. Ceci prouve en particulier que la fonction $x \mapsto \mathbb{P}_x(X_t \in A)$ est mesurable. Il est de plus possible d'exprimer $\mathbb{E}_\nu(Z)$ en fonction de $x \mapsto \mathbb{E}_x(Z)$: l'espérance de Z sous la loi de mesure initiale ν est la moyenne, pondérée par ν, des espérance de Z sous la loi de mesure initiale δ_x.

Proposition 2.1.8. *Si Z est $\mathcal{F}_\infty$-mesurable et positive (ou bornée) alors l'application $x \mapsto \mathbb{E}_x(Z)$ est $\mathcal{E}$-mesurable et*

$$\mathbb{E}_\nu(Z) = \int_E \nu(dx) \mathbb{E}_x(Z).$$

Preuve. Argument de classes monotones. □

2.2 Processus de Feller

Dans ce chapitre on se restreint aux f.t. et processus homogènes. Soit E un espace localement compact séparable et $\mathcal{C}_0(E)$ l'espace des fonctions continues sur E qui tendent vers 0 à l'infini.

Définition 2.2.1. *Un semi-groupe de Feller sur $\mathcal{C}_0(E)$ est une famille $(T_t)_{t \geq 0}$ d'opérateurs positifs sur $\mathcal{C}_0(E)$ telle que*

(i) $T_0 = \mathrm{Id}$ et pour tout t, $\|T_t\| \leq 1$,

(ii) pour $s, t \geq 0$, $T_{t+s} = T_t \circ T_s$,

(iii) pour toute fonction $f \in \mathcal{C}_0(E)$, $\lim_{t \downarrow 0} \|T_t f - f\| = 0$.

Il existe un lien très fort avec les f.t.

Proposition 2.2.2. *Pour chaque semi-groupe de Feller sur* E*, il existe une unique fonction de transition* $(P_t)_{t \geq 0}$ *sur* $(E, \mathcal{E})$ *telle que*

$$T_t f(x) = P_t f(x)$$

pour toute fonction $f \in \mathcal{C}_0(E)$ *et tout* $x \in E$.

Preuve. Pour tout $x \in E$ l'application $f \mapsto T_t f(x)$ est une forme linéaire positive sur $\mathcal{C}_0(E)$. Le théorème de Riesz assure l'existence d'une mesure $P_t(x, \cdot)$ sur $\mathcal{E}$ telle que

$$T_t f(x) = \int_E P_t(x, dy) f(y)$$

pour toute fonction $f \in \mathcal{C}_0(E)$. L'application $x \mapsto \int P_t(x, dy) f(y)$ appartient à $\mathcal{C}_0(E)$ donc $x \mapsto P_t(x, A)$ est borélienne pour tout $A \in \mathcal{E}$. Les opérateurs P_t sont bien des probabilités de transition. Elles forment une fonction de transition en vertu de la propriété de semi-groupe (ii) couplée à un argument de classe monotone qui permet de passer d'une propriété vraie pour les fonctions de $\mathcal{C}_0(E)$ aux ensembles $A \in \mathcal{E}$. □

Définition 2.2.3. *Une f.t. associée à un semi-groupe de Feller est appelée fonction de transition de Feller.*

Pour montrer qu'une f.t. est de Feller il est pratique d'utiliser la caractérisation suivante.

Proposition 2.2.4. *Une f.t. est de Feller si et seulement si*

(i) pour tout t, $P_t \mathcal{C}_0 \subset \mathcal{C}_0$,

(ii) pour toute fonction $f \in \mathcal{C}_0$*, tout* $x \in E$, $\lim_{t \downarrow 0} P_t f(x) = f(x)$.

Théorème 2.2.5. *Tout processus de Feller* X *admet une modification càdlàg : c'est-à-dire que les trajectoires de* X *sont continues à droites avec limites à gauche.*

2.3 Générateur infinitésimal

Définition 2.3.1. *Soit* X *un processus de Feller. On dit qu'une fonction* $f \in \mathcal{C}_0$ *appartient au domaine* $D(L)$ *du générateur infinitésimal de* X *si la limite*

$$Lf = \lim_{t \downarrow 0} \frac{1}{t}(P_t f - f)$$

existe dans $\mathcal{C}_0$*. L'opérateur* $L : D(L) \to \mathcal{C}_0$ *ainsi défini est appelé générateur infinitésimal du processus* X *ou du semi-groupe* (P_t).

La définition des processus de Markov assure que pour toute fonction f borélienne et $t, h > 0$,

$$\mathbb{E}(f(X_{t+h}) - f(X_t) | \mathcal{F}_t) = P_h f(X_t) - f(X_t).$$

Donc, si de plus $f \in D(L)$, alors

$$\mathbb{E}(f(X_{t+h}) - f(X_t) | \mathcal{F}_t) = hLf(X_t) + o(h).$$

Ainsi L décrit l'évolution infinitésimale du processus sur un petit intervalle de temps.

Étudions à présent quelques propriétés remarquables du générateur infinitésimal.

Proposition 2.3.2. *Si $f \in D(L)$, alors,*

(i) $P_t f \in D(L)$ pour tout t ;

(ii) la fonction $t \mapsto P_t f$ est fortement dérivable dans $\mathcal{C}_0$ et

$$\text{pour tout } t \geq 0, \quad \frac{d}{dt} P_t f = L P_t f = P_t L f;$$

(iii) pour tout $t > 0$, $P_t f - f = \int_0^t L P_s f \, ds$.

Preuve. Soit $t > 0$. La propriété de semi-groupe assure que

$$\lim_{s \to 0} \frac{1}{s}(P_s P_t f - P_t f) = \lim_{s \to 0} P_t \left[\frac{1}{s}(P_s f - f) \right] = P_t L f,$$

ce qui prouve i) et la relation de commutation $L P_t f = P_t L f$. De plus, $t \mapsto P_t f$ admet une dérivée à droite égale à $P_t L f$.

Considérons alors la fonction

$$t \mapsto \int_0^t P_s L f \, ds.$$

Elle est dérivable et sa dérivée vaut $P_t L f$. Comme la différence de deux fonctions qui ont la même dérivée à droite est nécessairement une constante, il existe g telle que

$$P_t f = \int_0^t P_s L f \, ds + g,$$

ce qui achève la preuve du point (ii) et en faisant $t = 0$, on obtient que $g = f$, ce qui prouve (iii). $\square$

Proposition 2.3.3. *L'espace $D(L)$ est dense dans $\mathcal{C}_0$ et L est un opérateur fermé.*

Preuve. Soit

$$A_h = \frac{1}{h}(P_h f - f) \quad \text{et} \quad B_s f = \frac{1}{s} \int_0^s P_u f \, du.$$

Les opérateurs A_h et B_s sont bornés sur $\mathcal{C}_0$ et de plus,

$$A_h B_s = B_s A_h = A_s B_h = B_h A_s.$$

Pour tout $s > 0$ et $f \in \mathcal{C}_0$,

$$\lim_{h \to 0} A_h B_s f = \lim_{h \to 0} A_s (B_h f) = A_s f,$$

donc $B_s f \in D(L)$ et comme $B_s f \to f$ quand $s \to 0$, $D(L)$ est dense dans $\mathcal{C}_0$.

Soit à présent une suite $(f_n)_n$ d'éléments de $D(L)$ qui converge vers f et telle que $(A f_n)_n$ converge vers g. Alors

$$B_s g = \lim_n B_s A f_n = \lim_n B_s \Big(\lim_h A_h f_n \Big) = \lim_n \lim_h A_s (B_h f_n) = \lim A_s f_n = A_s f.$$

Ainsi la fonction f est dans le domaine de L et $Lf = g$, ce qui prouve que L est un opérateur fermé. $\square$

Proposition 2.3.4. *Si $f \in D(L)$ alors le processus*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Lf(X_s)\,ds$$

est une $(\mathcal{F}_t, \mathbb{P}_\nu)$-martingale pour toute mesure de probabilité initiale ν. Si, en particulier, $Lf = 0$, alors $f(X_t)$ est une martingale.

Preuve. Puisque f et Lf sont bornées, M_t^f est intégrable pour tout t. De plus,

$$\mathbb{E}_\nu\Big(M_t^f|\mathcal{F}_s\Big) = M_s^f + \mathbb{E}_\nu\Big[f(X_t) - f(X_s) - \int_s^t Lf(X_u)\,du\Big|\mathcal{F}_s\Big].$$

D'après la propriété de Markov, l'espérance conditionnelle dans le membre de droite vaut

$$\mathbb{E}_{X_s}\Big[f(X_{t-s}) - f(X_0) - \int_0^{t-s} Lf(X_u)\,du\Big] = P_{t-s}f(X_s) - f(X_s) - \int_0^{t-s} P_uLf(X_s)\,ds,$$

qui vaut 0 d'après la proposition 2.3.2. □

Il est possible d'énoncer une réciproque.

Proposition 2.3.5. *Si $f \in \mathcal{C}_0$ et s'il existe une fonction $g \in \mathcal{C}_0$ telle que*

$$f(X_t) - f(X_0) - \int_0^t g(X_s)\,ds$$

est une $(\mathcal{F}_t, \mathbb{P}_x)$-martingale pour tout x, alors $f \in D(L)$ et $Lf = g$.

Preuve. Pour tout x,

$$P_tf(x) - f(x) - \int_0^t P_sg(x)ds = 0,$$

donc

$$\left\|\frac{1}{t}(P_tf - f) - g\right\| = \left\|\frac{1}{t}\int_0^t (P_sg - g)\,ds\right\| \le \frac{1}{t}\int_0^t \|P_sg - g\|\,ds,$$

qui tend vers 0 quand t tend vers 0. □

2.4 Mesures invariante et symétrique

Définition 2.4.1. *Soit $\mathcal{A}$ une algèbre de fonctions incluse dans tous les espaces $L^p(\mu)$ (pour $1 < p < \infty$) et dense dans chacun d'eux. On dira que $\mathcal{A}$ est une algèbre standard si $\mathcal{A}$ est stable par tous les opérateurs P_t, par L et par composition avec les fonctions de classe $\mathcal{C}^\infty$ qui sont nulles en 0. Lorsque la mesure μ est une probabilité, nous demanderons de plus que cette algèbre contienne les constantes, et soit stable par composition avec toutes les fonctions de classe $\mathcal{C}^\infty$ (nulles ou non en 0).*

Dans toute la suite nous ferons l'hypothèse de l'existence d'une algèbre standard $\mathcal{A}$.

Remarquons que cette hypothèse est facilement vérifiée dans des cas simples. Par exemple, pour un semi-groupe de Markov sur un ensemble fini, on pourra prendre pour $\mathcal{A}$ l'ensemble de toutes les fonctions. Cependant il est difficile de spécifier un tel ensemble dans un cas général. Il faut procéder au cas par cas. Dans la suite, nous préciserons le cadre au fur et à mesure de l'étude de processus particuliers.

Définition 2.4.2. *Une mesure μ sur E est dite invariante (ou stationnaire) pour le semi-groupe $(P_t)_t$ si, pour tout $t \geq 0$ et toute fonction $f \in \mathcal{A}$,*

$$\int P_t f \, d\mu = \int f \, d\mu,$$

ou de manière équivalente, si pour toute fonction f de $\mathcal{A}$,

$$\int Lf \, d\mu = 0.$$

Elle est dite symétrique (ou réversible) si, pour tout $t \geq 0$ et toutes fonctions f et g de $\mathcal{A}$,

$$\int g P_t f \, d\mu = \int f P_t g \, d\mu,$$

ou de manière équivalente, si pour toute fonction f de $\mathcal{A}$,

$$\int g L f \, d\mu = \int f L g \, d\mu,$$

c'est-à-dire que les opérateurs P_t (ou de manière équivalente l'opérateur L) sont auto-adjoints dans $L^2(\mu)$.

Exercice 2.4.3. *Montrer que les définitions d'invariance (resp de symétrie) sont bien équivalentes. Montrer que le caractère symétrique d'une mesure entraîne son invariance.*

Définition 2.4.4. *On appelle opérateur* carré du champ *la forme bilinéaire symétrique suivante : pour f et g dans $\mathcal{A}$,*

$$\mathbf{\Gamma}(f,g) \overset{\text{déf.}}{=} \frac{1}{2}(L(fg) - fLg - gLf).$$

Nous noterons $\mathbf{\Gamma}(f)$ pour $\mathbf{\Gamma}(f,f)$.

Proposition 2.4.5. *Soit A le générateur infinitésimal de mesure invariante μ. Pour toute fonction $f \in \mathcal{A}$,*

1. $\mathbf{\Gamma}(f) \geq 0$,
2. $\displaystyle\int \mathbf{\Gamma}(f) \, d\mu = -\int f L f \, d\mu.$

Preuve. Comme P_t préserve la positivité, pour toute fonction $f \in \mathcal{A}$, pour tout $t \geq 0$ et tout $a \in \mathbb{R}$,

$$0 \leq P_t((f+a)^2) = P_t(f^2) - 2aP_t f + a^2.$$

Le discriminant du trinôme du membre de droite est donc négatif, ce qui implique que $P_t(f^2) \geq (P_t f)^2$. Enfin, l'opérateur carré du champ peut aussi s'écrire de la manière suivante :

$$\mathbf{\Gamma}(f) = \lim_{t \to 0} \frac{1}{2t}\big(P_t(f^2) - (P_t f)^2\big) \geq 0.$$

La propriété (2) découle directement de la propriété d'invariance de μ et de la définition de $\mathbf{\Gamma}$. □

Définition 2.4.6. *On appelle énergie la forme quadratique associée à L, définie par :*

$$D(\mathcal{E}) = D(L) \quad \text{et, pour } f \in D(\mathcal{E}), \quad \mathcal{E}_\mu(f,f) \overset{\text{déf.}}{=} -\int f L f \, d\mu = \int \mathbf{\Gamma}(f) \, d\mu.$$

Chapitre 3

Chaînes de Markov à temps continu

Dans tout le chapitre E désigne un ensemble fini de cardinal d et $(X_t)_{t\geq 0}$ un processus de Markov défini sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans E. Il est aisé de voir que X est nécessairement un processus de Feller : puisque E est fini, la convergence uniforme et la convergence simple sont équivalentes. On peut donc supposer que les trajectoires de X sont càdlàg (continues à droite avec limites à gauche). Plus précisément, pour tout $\omega \in \Omega$,

(i) $t \mapsto X_t(\omega)$ est continu à droite,

(ii) $\exists 0 < T_1(\omega) < \ldots < T_n(\omega) < \ldots$ tels que $(T_n)_n$ n'ait pas de point d'accumulation et sur chaque intervalle $[T_i, T_{i+1}[$, X soit constant.

Soit H l'espace vectoriel des fonctions de E dans $\mathbb{R}$ (sa dimension est égale à d). À partir du processus X on définit, pour tout $t \geq 0$, l'opérateur P_t de H dans H par

$$P_t f(x) = \mathbb{E}(f(X_t)|X_0 = x).$$

Remarquons que P_t s'identifie à une matrice de taille $d \times d$ que l'on notera $(P_t(x,y))_{x,y\in E}$. La matrice P_t est markovienne, c'est-à-dire que

$$\begin{cases} P_t(x,y) \geq 0 & \text{pour tous } x, y \in E, \\ \displaystyle\sum_{y\in E} P_t(x,y) = 1 & \text{et pour tout } x \in E. \end{cases}$$

On notera I la matrice identité. On munira l'ensemble des matrices de taille $d \times d$ de la norme

$$\|M\| \stackrel{\text{déf.}}{=} \sup_{x\in E} \sum_{y\in E} |M(x,y)|.$$

Cette norme est multiplicative (la norme du produit MN est majorée par le produit des normes de M et N) et P_t est de norme 1 pour tout $t \geq 0$.

La première section de ce chapitre rassemble les propriétés des lois exponentielles dont nous aurons besoin dans la suite.

3.1 Lois exponentielles

Définition 3.1.1. *Une variable aléatoire T à valeurs dans $\mathbb{R}_+$ suit la loi exponentielle de paramètre $\lambda \geq 0$, notée $\mathcal{E}(\lambda)$ si sa fonction de répartition est donnée par $F(t) = (1 - e^{-\lambda t})\mathbf{1}_{t\geq 0}$ (si $\lambda = 0$, T vaut $+\infty$ presque sûrement).*

Sa loi admet $t \mapsto \lambda e^{-\lambda t}\mathbf{1}_{\{t>0\}}$ pour densité par rapport à la mesure de Lebesgue. De plus,

$$\mathbb{E}(T) = \frac{1}{\lambda}, \quad \mathbb{V}(T) = \frac{1}{\lambda^2} \quad \textit{et, pour } t < \lambda \quad \mathbb{E}\big(e^{tT}\big) = \frac{\lambda}{\lambda - t}.$$

Proposition 3.1.2 (Absence de mémoire). *Une variable aléatoire T à valeurs dans $\mathbb{R}_+ \cup \{+\infty\}$ suit une loi exponentielle si et seulement si elle vérifie la propriété d'absence de mémoire :*

$$\forall s, t \geq 0, \quad \mathbb{P}(T > t + s | T > s) = \mathbb{P}(T > t).$$

Proposition 3.1.3. *Soit $(S_n)_{n\geq 1}$ une suite de v.a. indépendantes de lois exponentielles de paramètres respectifs $(\lambda_n)_{n\geq 1}$ strictement positifs.*

$$\textit{Si} \quad \sum_{n\geq 1} \frac{1}{\lambda_n} < \infty \quad \textit{alors} \quad \mathbb{P}\left(\sum_{n\geq 1} S_n < \infty\right) = 1$$

$$\textit{et si} \quad \sum_{n\geq 1} \frac{1}{\lambda_n} = \infty \quad \textit{alors} \quad \mathbb{P}\left(\sum_{n\geq 1} S_n = \infty\right) = 1.$$

Preuve. Si $\sum_{n\geq 1} 1/\lambda_n$ est fini alors, par convergence monotone, la variable aléatoire $\sum_{n\geq 1} S_n$ est intégrable donc elle est finie p.s.

Si $\sum_{n\geq 1} 1/\lambda_n$ est infini alors $\prod_{n\geq 1}(1 + 1/\lambda_n)$ est infini et

$$\mathbb{E}\left(\exp\left\{-\sum_{n\geq 1} S_n\right\}\right) = \prod_{n\geq 1} \mathbb{E}(\exp\{-S_n\}) = \prod_{n\geq 1}(1 + 1/\lambda_n)^{-1} = 0$$

par convergence monotone et indépendance. □

Proposition 3.1.4. *Soit I un ensemble (au plus) dénombrable et $(T_k)_{k\in I}$ des v.a. exponentielles de paramètres respectifs $(\lambda_k)_{k\in I}$ vérifiant $\lambda = \sum_k \lambda_k < \infty$. Soit $T = \inf_k T_k$. Alors, avec probabilité 1, l'infimum est atteint en un unique K (aléatoire) élément de I. De plus, T et K sont indépendantes, T suit la loi $\mathcal{E}(\lambda)$ et pour tout $k \in I$, $\mathbb{P}(K = k) = \lambda_k/\lambda$.*

Preuve. Soit $k \in I$ et $t \geq 0$.

$$\begin{aligned}
\mathbb{P}(K = k, T \geq t) &= \mathbb{P}(T_k \geq t, T_j > T_k, j \neq k) \\
&= \int_t^\infty \lambda_k e^{-\lambda_k s} \mathbb{P}(T_j > s, j \neq k)\, ds \\
&= \int_t^\infty \lambda_k e^{-\lambda_k s} \prod_{j\neq k} e^{-\lambda_j s}\, ds \\
&= \int_t^\infty \lambda_k e^{-\lambda s}\, ds = \frac{\lambda_k}{\lambda} e^{-\lambda t}.
\end{aligned}$$

On conclut en remarquant que les ensembles $\{K = k\} \times \{T \geq t\}$ engendre la tribu produit.
□

Exercice 1.
Soit $(T_n)_{n\geq 1}$ une suite de v.a. i.i.d. de loi $\mathcal{E}(\lambda)$ et N une v.a. indépendante de loi géométrique de paramètre β. Montrer que $T = \sum_{n=1}^N T_n$ suit une loi exponentielle de paramètre $\lambda\beta$.

3.2 Caractérisation du générateur infinitésimal

Nous allons montrer à présent que le générateur infinitésimal d'un processus de Markov à temps continu sur un espace d'état fini s'identifie à une matrice dont les sommes des lignes sont nulles et les coefficients hors-diagonaux sont positifs.

Définition 3.2.1. *On notera $\mathcal{A}$ l'ensemble des matrices A telles que*

(i) $\forall x \neq y, \quad A_{xy} \geq 0,$

(ii) $\forall x \in E, \quad \sum_{y \in E} A_{x,y} = 0.$

Proposition 3.2.2. *Soit X un processus de Markov sur E. Il existe $A \in \mathcal{A}$ tel que pour tout $t \geq 0$,*

$$P_t = e^{tA} \overset{\text{déf.}}{=} \sum_{k=0}^{\infty} \frac{t^n}{n!} A^n. \tag{3.1}$$

La matrice A est le générateur infinitésimal du semi-groupe (P_t), c'est-à-dire que

$$A(x,y) = \lim_{t \to 0} \frac{P_t(x,y) - \mathbf{1}_{\{x=y\}}}{t}.$$

Preuve. L'application $t \mapsto P_t$ est continue à droite. En effet, si $t, h \geq 0$, alors

$$\|P_{t+h} - P_t\| = \|P_t \circ P_h - P_t\| \leq \|P_t\| \|P_h - I\| \leq \|P_h - I\| \xrightarrow[h \to 0]{} 0.$$

On peut montrer de même qu'elle est continue à gauche et donc continue sur $\mathbb{R}^*$. Pour montrer la relation (3.1), on montre que $t \mapsto P_t$ est dérivable en 0. Sa dérivée ne sera rien d'autre que A. Pour cela il faut introduire la résolvante $(R_\lambda)_{\lambda>0}$ du semi-groupe $(P_t)_{t\geq 0}$ et étudier ses premières propriétés. Pour $\lambda > 0$, on pose

$$R_\lambda = \int_0^\infty e^{-\lambda t} P_t \, dt, \quad \text{ou pour } x, y \in E \quad R_\lambda(x,y) = \int_0^\infty e^{-\lambda t} P_t(x,y) \, dt.$$

Pour tous $\lambda, \mu > 0$,

$$(\mu - \lambda) R_\lambda R_\mu = R_\lambda - R_\mu. \tag{3.2}$$

En particulier, R_λ et R_μ commuttent et $R_\mu = R_\lambda(I + (\lambda - \mu)R_\mu)$. Ceci entraîne que $R_\mu(H)$ est inclus dans $R_\lambda(H)$ et, par symétrie, que $R_\mu(H) = R_\lambda(H)$. Pour établir (3.2), on écrit

$$\begin{aligned}
R_\lambda R_\mu &= R_\lambda \Big(\int_0^\infty e^{-\mu t} P_t \, dt \Big) = \int_0^\infty e^{-\mu t} R_\lambda P_t \, dt = \int_0^\infty e^{-\mu t} \Big(\int_0^\infty e^{-\lambda s} P_s \, ds \Big) P_t \, dt \\
&= \int_0^\infty \int_0^\infty e^{-\mu t - \lambda s} P_{t+s} \, ds \, dt = \int_0^\infty \int_t^\infty e^{-\mu t - \lambda(v-t)} P_v \, dv \, dt \\
&= \int_0^\infty e^{-\lambda v} \Big(\int_0^v e^{(\lambda-\mu)t} \, dt \Big) P_v \, dv = \frac{1}{\mu - \lambda}(R_\lambda - R_\mu).
\end{aligned}$$

On souhaite montrer que $R_\mu(H) = H$. Il est clair que $R_\mu(H) = \mu R_\mu(H)$. Or, par définition de R_μ,

$$\mu R_\mu = \mu \int_0^\infty e^{-\mu t} P_t \, dt = \int_0^\infty e^{-t} P_{t/\mu} \, dt \xrightarrow[\mu \to \infty]{} \Big(\int_0^\infty e^{-t} dt \Big) I = I.$$

Donc, pour tout $f \in H$, $f = \lim_{\mu\to\infty} \mu R_\mu f$. Le sous-espace $R_\mu(H)$ est un sous-espace fermé de H (de dimension fini) donc $f \in R_\mu(H)$. En d'autres termes, $R_\mu(H)$ coïncide avec H. Nous avons donc montré que

$$\forall \mu > 0,\ \forall f \in H,\ \exists g \in H, \quad f = R_\mu g.$$

L'idée est à présent d'intercaler R_μ dans la différence $P_t f - f$.

$$\begin{aligned} P_t R_\mu &= \int_0^\infty e^{-\mu s} P_{t+s}\, ds = \int_t^\infty e^{-\mu(v-t)} P_v\, dv \\ &= e^{\mu t} \int_t^\infty e^{-\mu v} P_v\, dv = e^{\mu t}\Big(R_\mu - \int_0^t e^{-\mu v} P_v\, dv\Big). \end{aligned}$$

Donc

$$\frac{P_t R_\mu - R_\mu}{t} = \frac{e^{\mu t} - 1}{t} R_\mu - e^{\mu t}\frac{1}{t}\int_0^t e^{-\mu v} P_v\, dv \underset{t\to 0}{\longrightarrow} \mu R_\mu - I,$$

puisque $t \mapsto P_t$ est continue à droite. On peut donc écrire pour $f \in H$ et $\mu > 0$,

$$\frac{P_t f - f}{t} = \frac{P_t R_\mu g - R\mu g}{t} \underset{t\to 0}{\longrightarrow} \mu R_\mu g - g,$$

où g est donnée par $f = R_\mu g$. On a donc montré que $t \mapsto P_t$ est dérivable en 0 et cette limite, notée A vérifie $AR_\mu = \mu R_\mu - I$ ou encore $(\mu I - A)R_\mu = I$. Ainsi, dans le cas fini, A est défini sur tout l'ensemble H des fonctions de E dans $\mathbb{R}$. Dans le cas d'un processus de Feller général, on peut montrer que $\mathcal{D}(A)$ est dense dans l'ensemble des fonctions continues. Il est en général impossible de l'expliciter complètement.

Comme nous l'avons remarqué dans un cadre plus général,

$$\lim_{s\to 0} \frac{P_{t+s} - P_t}{s} = P_t\Big(\lim_{s\to 0}\frac{P_s - I}{s}\Big) = \lim_{s\to 0}\Big(\frac{P_s - I}{s}\Big)P_t,$$

et par conséquent P_t et A commuttent et $\frac{d}{dt}P_t = P_t A = AP_t$. Notons pour $0 \le s \le t$

$$M_t = e^{tA} \quad \text{et} \quad N_s = P_s M_{t-s}.$$

Dérivons à présent N_s :

$$\frac{d}{ds}N_s = P_s A M_{t-s} - P_s A M_{t-s} = 0,$$

donc N_s est constant. De plus $N_0 = M_t$ et $N_t = P_t$ et par suite $P_t = e^{tA}$.

Remarquons à présent que A appartient à $\mathcal{A}$. Pour cela, il suffit de revenir à la définition de A en choisissant f de la forme $f_x = \mathbf{1}_{\{x\}}$. On obtient alors si $x \neq y$

$$A(x,y) = Af_x(y) = \lim_{t\to 0}\frac{P_t(x,y)}{t} \ge 0,$$

et, de même, pour les termes diagonaux,

$$A(x,x) = Af_x(x) = \lim_{t\to 0}\frac{P_t(x,x) - 1}{t} \le 0.$$

Enfin, le fait que P_t soit markovienne assure que la somme des termes de chaque ligne de A est nulle :

$$\sum_{y\in E} A(x,y) = \sum_{y\in E} Af_x(y) = \lim_{t\to 0}\frac{\sum_{y\in E} P_t(x,y) - 1}{t} = 0,$$

ce qui achève la preuve. □

Après avoir montré que le générateur infinitésimal d'un processus de Markov sur E est nécessairement un élément de $\mathcal{A}$, il reste à remarquer qu'un élément de $\mathcal{A}$ engendre un semi-groupe de Markov.

Proposition 3.2.3. *Soit $A \in \mathcal{A}$. On lui associe les matrices $P_t := e^{tA}$ pour $t \geq 0$. Alors P_t est markovienne et $(P_t)_t$ est un semi-groupe de Feller.*

Preuve. Si λ est valeur propre de A associée à un vecteur propre v alors $e^{\lambda t}$ est valeur propre de P_t associée au vecteur propre v :

$$P_t v = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n v = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} v = e^{\lambda t} v.$$

En particulier 0 est valeur propre de A associée au vecteur propre constant donc la somme de chaque ligne de P_t est égale à 1.

Soit à présent $m = \sup_x \{-A(x,x)\}$. Tous les coefficients de la matrice $mI + A$ sont positifs. Il en est donc de même pour

$$N_t = e^{t(mI+A)} = \sum_{n=0}^{\infty} \frac{t^n}{n!} (mI + A)^n.$$

Or N_t est égale à $e^{tmI} e^{tA} = e^{tm} e^{tA}$ ou encore $P_t = e^{-tm} N_t$ ce qui assure que P_t a des coefficients positifs.

Enfin il est clair la propriété de semi-groupe est satisfaite par P_t. □

3.3 Temps d'arrêt

Définition 3.3.1. *Une v.a. T à valeurs dans $[0, +\infty]$ est un temps d'arrêt si*

$$\forall t \geq 0, \quad \{T \leq t\} \in \mathcal{F}_t.$$

Remarque 3.3.2. *Puisque*

$$\{T = t\} = \{T \leq t\} \setminus \bigcup_{n \geq 1} \underbrace{\{T \leq t - 1/n\}}_{\in \mathcal{F}_{t-1/n}},$$

l'ensemble $\{T = t\}$ appartient lui aussi à $\mathcal{F}_t$.

Proposition 3.3.3. *Si T et S sont des temps d'arrêt alors il en est de même pour $T + S$, $T \vee S$, $T \wedge S$. La v.a. constante $T = t$ est un temps d'arrêt.*

Attention : $T - \varepsilon$ n'est pas un temps d'arrêt.

Exemple 3.3.4. *Soit $a \in E$. La v.a. $T = \inf \{s,\ X_s = a\}$ est un temps d'arrêt. Pour s'en convaincre il suffit de remarquer que*

$$\{T \geq t\} = \bigcap_{s<t} \{X_s \neq a\} = \bigcap_{\substack{s<t \\ s \in \mathbb{Q}}} \{X_s \neq a\} \in \mathcal{F}_t.$$

3.4 Propriété de Markov

Nous avons déjà donné un sens à la notion de propriété de Markov. Nous en donnons ici une autre forme très utile tant en pratique qu'au niveau de l'intuition. Soit

$$\hat{\Omega} = \{\hat{\omega} \ : \ \mathbb{R}_+ \to E \text{ càd}\}.$$

Il est appelé espace canonique. Posons $\hat{X}_t(\omega) = \omega(t)$ et notons $\mathcal{F}_\infty$ la tribu

$$\mathcal{F}_\infty = \sigma(\hat{X}_s, \ s \in \mathbb{R}) = \sigma(\hat{X}_s, \ s \in \mathbb{Q}).$$

Soit X un processus càd (continu à droite) de $(\Omega, \mathcal{F}, \mathbb{P})$ dans E. Alors l'application Φ de Ω dans $\hat{\Omega}$ qui à ω associe la trajectoire $(t \mapsto X_t(\omega) = \hat{\omega}(t))$ est mesurable et l'image de Φ est la loi du processus X. En d'autres termes la loi d'un processus X càd à valeurs dans E peut être vue comme une mesure de probabilité sur $\hat{\Omega}$ l'ensemble des fonctions càd sur E.

Dans la suite Ω désignera l'espace canonique. On définit l'opérateur de translation θ_t de Ω dans lui-même par :

$$\text{pour tout } s \geq 0, \quad \theta_t(\omega)(s) = \omega(t+s).$$

La trajectoire $\theta_t(\omega)$ est donc la trajectoire ω à laquelle on a ôté le début (jusqu'à l'instant t).

Proposition 3.4.1. *Si Z est une v.a. de $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ dans $\mathbb{R}$ mesurable (c'est-à-dire une fonction mesurable par rapport à la tribu engendrée par les v.a. X_s pour tout s où X est un processus de Markov) bornée alors*

$$\mathbb{E}(\theta_t Z | \mathcal{F}_t) = \mathbb{E}_{X_t}(Z) = K(X_t), \tag{3.3}$$

où $K(x) = \mathbb{E}(Z | X_0 = x) = \mathbb{E}_x(Z)$.

Remarque 3.4.2. *La quantité $\theta_t Z$ représente le futur : on considère des événements qui se passent après t (mais on remet la pendule à 0). Donc $\mathbb{E}(\theta_t Z | \mathcal{F}_t)$ représente le futur sachant le tout le passé jusqu'au temps t. La formule (3.3) est donc bien une nouvelle formulation de la propriété de Markov.*

Preuve. Grâce au théorème des classes monotones, il suffit de vérifier la propriété pour Z de la forme

$$Z = f_1(X_{t_1}) \dots f_n(X_{t_n}),$$

avec $t_1 < \cdots < t_n$. Par définition de θ_t, $\theta_t Z = f_1(X_{t+t_1}) \dots f_n(X_{t+t_n})$. Donc

$$\begin{aligned}
\mathbb{E}(\theta_t Z | \mathcal{F}_t) &= \mathbb{E}(\mathbb{E}(\theta_t Z | \mathcal{F}_{t+t_{n-1}}) | \mathcal{F}_t) \\
&= \mathbb{E}(f_1(X_{t+t_1}) \dots f_{n-1}(X_{t+t_{n-1}}) \mathbb{E}(f_n(X_{t+t_n}) | \mathcal{F}_{t+t_{n-1}}) | \mathcal{F}_t) \\
&= \mathbb{E}(f_1(X_{t+t_1}) \dots f_{n-1}(X_{t+t_{n-1}}) P_{t_n - t_{n-1}}(f_n)(X_{t+t_{n-1}}) | \mathcal{F}_t).
\end{aligned}$$

Donc

$$\mathbb{E}(\theta_t Z | \mathcal{F}_t) = P_{t_1}(f_1 P_{t_2 - t_1}(f_2 P_{t_3 - t_2}(f_4 \dots (f_{n-1} P_{t_n - t_{n-1}} f_n) \dots)))(X_t) = \mathbb{E}_{X_t}(Z).$$

□

La propriété de Markov peut ici s'étendre à des temps d'arrêt. On l'appelle alors propriété de Markov forte et on dit que X est un processus de Markov fort. Il existe des processus de Markov qui ne sont pas forts mais ceux que nous étudierons dans le cadre de ce cours le seront toujours.

Théorème 3.4.3. *Si T est un temps d'arrêt et Z est une v.a. de $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ dans $\mathbb{R}$ mesurable bornée (ou positive) alors*
$$\mathbb{E}(\theta_T Z | \mathcal{F}_T) = \mathbb{E}_{X_T}(Z). \tag{3.4}$$

Preuve. L'idée est de montrer la proposition pour des temps d'arrêt prenant un nombre au plus dénombrable de valeurs (ceci est toujours vrai pour un processus de Markov). On passe ensuite aux processus d'arrêt généraux grâce au fait que les trajectoires de X sont càdlàg. En particulier, les processus de Feller vérifient la propriété de Markov forte. Par contre, on peut montrer que la réciproque est fausse. □

3.5 Premier instant de saut

Soit $T =\in \{s,\ X_s \neq X_0\}$ le premier instant de saut de la chaîne.

Proposition 3.5.1. *Il existe $\lambda(x) \in [0, +\infty[$ tel que, pour tout $t \geq 0$,*
$$\mathbb{P}_x(T > t) = e^{-\lambda(x)t}.$$
De plus, X_T et T sont indépendants.

Remarque 3.5.2. *La proposition assure que si la chaîne est issue de x alors elle y reste (si $\lambda(x) = 0$) ou elle quitte x après un temps exponentiel de paramètre $\lambda(x) > 0$.*

Preuve (proposition 3.5.1). On peut réécrire la v.a. $T\mathbf{1}_{\{T>t\}}$ grâce à l'opérateur de translation :
$$T\mathbf{1}_{\{T>t\}} = (t + \theta_t T)\mathbf{1}_{\{T>t\}}.$$
Discutons tout d'abord de la finitude de T : si $T(\omega)$ est infini alors la trajectoire est constante égale à x. Cela revient à dire que pour tout instant t le premier saut n'est pas survenu avant t et ne surviendra pas après t ! Ceci s'écrit : pour tout $t \geq 0$,
$$\{T = \infty\} = \theta_t\{T = \infty\} \cap \{T > t\}.$$
En prenant l'espérance il vient, en vertu de la propriété de Markov :
$$\begin{aligned}
\mathbb{P}_x(T = \infty) &= \mathbb{E}(\mathbf{1}_{\{T>t\}}\theta_t\mathbf{1}_{\{T=\infty\}}) = \mathbb{E}(\mathbf{1}_{\{T>t\}}\mathbb{E}(\theta_t\mathbf{1}_{\{T=\infty\}}|\mathcal{F}_t)) \\
&= \mathbb{E}(\mathbf{1}_{\{T>t\}}\mathbb{E}_{X_t}(\mathbf{1}_{\{T=\infty\}})) = \mathbb{E}(\mathbf{1}_{\{T>t\}}\mathbb{E}_x(\mathbf{1}_{\{T=\infty\}})) \\
&= \mathbb{P}_x(T = \infty)\mathbb{P}_x(T > t).
\end{aligned}$$
Si $\mathbb{P}_x(T = \infty) \neq 0$ alors en particulier pour tout $n \in \mathbb{N}$, $\mathbb{P}_x(T > t) = 1$. Donc
$$\mathbb{P}_x(T = \infty) = \mathbb{P}_x(\bigcap_n\{T > n\}) = 1,$$
puisque l'intersection dénombrable d'événements de probabilité 1 est encore de probabilité 1. Ainsi, l'événement $\{T = \infty\}$ est-il nécessairement de probabilité 1 ou 0. Le cas où cette probabilité vaut 1 correspond au cas $\lambda(x) = 0$ de la proposition.

Dans la suite nous supposerons que $\mathbb{P}_x(T = \infty) = 0$. Soit $\varphi(t) = \mathbb{P}_x(T > t)$. Pour $t, s \geq 0$,
$$\{T > t + s\} = \{T > t\} \cap \{\theta_t T > s\}.$$

En prenant l'espérance, il vient donc

$$\begin{aligned}\varphi(t+s) &= \mathbb{E}_x(\mathbf{1}_{\{T>t\}}\mathbb{E}(\theta_t\mathbf{1}_{\{T>s\}}|\mathcal{F}_t)) = \mathbb{E}_x(\mathbf{1}_{\{T>t\}}\mathbb{E}_{X_t}(\mathbf{1}_{\{T>s\}}))\\ &= \mathbb{E}_x(\mathbf{1}_{\{T>t\}}\mathbb{E}_x(\mathbf{1}_{\{T>s\}})) = \varphi(t)\varphi(s).\end{aligned}$$

De plus, $\varphi(t) \xrightarrow[t\to 0]{} 1$ car les trajectoires de X sont continues à droite. Il existe donc $\lambda(x) > 0$ tel que pour tout t,

$$\mathbb{P}_x(T > t) = e^{-\lambda(x)t}.$$

Pour montrer l'indépendance de X_T et T, il suffit de montrer que pour tout $t > 0$ et $A \subset E$,

$$\mathbb{P}_x(X_T \in A,\ T > t) = \mathbb{P}_x(X_T \in A)\mathbb{P}_x(T > t).$$

Or sur l'événement $\{T > t\}$,

$$T = t + \theta_t T \quad \text{et} \quad X_T = X_{t+\theta_t T} = \theta_t(X_T).$$

Donc

$$\begin{aligned}\mathbb{P}_x(X_T \in A,\ T > t) &= \mathbb{E}_x(\mathbf{1}_{\{T>t\}}\mathbb{E}(\mathbf{1}_{X_T\in A}|\mathcal{F}_t)) = \mathbb{E}_x(\mathbf{1}_{\{T>t\}}\mathbb{E}(\theta_t\mathbf{1}_{X_T\in A}|\mathcal{F}_t))\\ &= \mathbb{E}_x(\mathbf{1}_{\{T>t\}}\mathbb{E}_{X_t}(\mathbf{1}_{\{X_T\in A\}})) = \mathbb{E}_x(\mathbf{1}_{\{T>t\}}\mathbb{E}_x(\mathbf{1}_{\{X_T\in A\}}))\\ &= \mathbb{P}_x(X_T \in A)\mathbb{P}_x(T > t).\end{aligned}$$

□

Définition 3.5.3. *Le premier instant de saut est défini par*

$$T_1 = \inf\{s > 0,\ X_s \neq X_0\}.$$

Les temps de saut suivants sont définis par récurrence :

$$T_{k+1} = \inf\{s > T_k,\ X_s \neq X_{T_k}\} = T_k + \theta_{T_k}T = T + \theta_T(T_k).$$

3.6 Dynamique du processus

Nous sommes à présent en mesure de relier les quantités suivantes :

$$A(x,y) = \lim_{t\to 0}\frac{P_t(x,y) - \delta_{xy}}{t}, \quad \lambda(x) \quad \text{et} \quad p(x,y) = \mathbb{P}_x(X_T = y).$$

Proposition 3.6.1. *Si X est un processus de Markov sur E de générateur infinitésimal A alors*

$$\lambda(x) = -A(x,x) \quad \textit{et} \quad p(x,y) = -\frac{A(x,y)}{A(x,x)}.$$

Preuve. Soit $x \in E$. Nous allons tout d'abord montrer la probabilité que deux sauts au moins apparaissent avant l'instant t est négligeable devant la probabilité qu'un saut seulement apparaisse. Ceci est en particulier assuré si l'on obtient le contrôle suivant :

$$\mathbb{P}_x(T_2 < t) \leq t^2 \sup_x \lambda(x)^2.$$

Si deux sauts au moins ont eu lieu avant t, c'est qu'il a eu un premier saut avant t et un deuxième entre T_1 et t :

$$\{T_2 < t\} = \{T_1 < t\} \cap \{T_1 + \theta_{T_1} T_1 < t\}.$$

En particulier

$$\{T_2 < t\} \subset \{T_1 < t\} \cap \{\theta_{T_1} T_1 < t\}$$

En prenant l'espérance, il vient, avec la notation $k = \sup_x \lambda(x)$,

$$\begin{aligned}
\mathbb{P}_x(T_{(2)} < t) &\leq \mathbb{E}\big(\mathbf{1}_{\{T<t\}} \theta_T \mathbf{1}_{\{T<t\}}\big) \leq \mathbb{E}\big(\mathbf{1}_{\{T<t\}} \mathbb{E}\big(\theta_T \mathbf{1}_{\{T<t\}} | \mathcal{F}_T\big)\big) \\
&\leq \mathbb{E}\big(\mathbf{1}_{\{T<t\}} \mathbb{E}_{X_T}\big(\mathbf{1}_{\{T<t\}}\big)\big) \leq \mathbb{E}\big(\mathbf{1}_{\{T<t\}} \big(1 - e^{-\lambda(X_T)t}\big)\big) \\
&\leq kt\big(1 - e^{-\lambda(x)t}\big) \leq k^2 t^2,
\end{aligned}$$

puisque $1 - e^{-u} \leq u$ dès que $u \geq 0$.

Soit à présent x et y deux éléments distincts de E.

$$\begin{aligned}
P_t(x, y) &= \mathbb{P}_x(X_t = y) = \mathbb{P}_x(X_t = y, T < t) = \mathbb{P}_x(T < t, T_2 > t, X_t = y) + \mathbb{P}_x(T_2 \leq t, X_t = y) \\
&= \mathbb{P}_x(T < t, X_T = y) - \mathbb{P}_x(T < t, T_2 \leq t, X_t = y) + \mathbb{P}_x(T_2 \leq t, X_t = y).
\end{aligned}$$

On a donc

$$|\mathbb{P}_x(X_t = y) - \mathbb{P}_x(T < t, X_T = y)| \leq \mathbb{P}_x(T_{(2)} < t) \leq k^2 t^2.$$

Enfin, par indépendance de T et X_T,

$$\mathbb{P}_x(T < t, X_T = y) = \mathbb{P}_x(T < t)\mathbb{P}_x(X_T = y) = (1 - e^{-\lambda(x)t})p(x, y).$$

On a donc

$$A(x, y) = \lim_{t \to 0} \frac{P_t(x, y)}{t} \lim_{t \to 0} \frac{1 - e^{-\lambda(x)t}}{t} p(x, y) = \lambda(x)p(x).$$

On montre de même que $P_t(x, x) = e^{-\lambda(x)t} + o(t)$ et ainsi que

$$A(x, x) = -\lambda(x),$$

ce qui achève la preuve. □

Remarque 3.6.2. *Nous nous sommes contentés ici d'un exposé dans le cas où E est fini. Lorsqu'il est dénombrable, des phénomènes d'explosion peuvent apparaître lorsque les paramètres des lois des temps inter-sauts sont trop grands. Le processus se met alors à sauter de plus en plus vite et il peut ainsi apparaître un nombre infini de sauts en un temps fini (sans remettre en cause le caractère càdlàg des trajectoires). On pourra se reporter à Norris pour des exemples et l'adaptation de ce qui précède au cas dénombrable.*

3.7 Classes

Les notions de classes sont les mêmes à temps continu et à temps discret. Plus précisément, $(X_t)_{t\geq 0}$ a la même structure de classe que sa chaîne incluse $(Y_n)_{n\geq 0}$.

On dit que i mène à j, et on note $i \to j$ si

$$\mathbb{P}_i(X_t = j, \text{ pour un certain temps } t \geq 0) > 0.$$

On dit que i et j communiquent, et on note $i \leftrightarrow j$ si $i \to j$ et $j \to i$.

Théorème 3.7.1. *Pour deux états distincts i et j, les propriétés suivantes sont équivalentes :*

1. $i \to j$,
2. $i \to j$ *pour la chaîne incluse,*
3. *il existe* $n \in \mathbb{N}$ *et* $i_0, \ldots, i_n$ *dans* I *tels que* $q(i_0, i_1) \ldots q_(i_{n-1}, i_n) > 0$,
4. *pour tout* $t > 0$, $p_t(i,j) > 0$,
5. *il existe* $t > 0$, $p_t(i,j) > 0$.

Démonstration. Les implications 4. ⇒ 5. ⇒ 1. ⇒ 2. découlent des définitions. Si 2. est vraie, il existe $n \in \mathbb{N}$ et $i_0, \ldots, i_n$ dans I tels que $\pi_{i_0 i_1} \ldots \pi_{i_{n-1} i_n} > 0$, ce qui implique 3. Si $q_{ij} > 0$ alors

$$p_t(i,j) \geq \mathbb{P}_i(J_1 \leq t, Y_1 = j, J_2 < t) = (1 - e^{-q_i t}) \pi_{ij} e^{-q_j t} > 0,$$

pour tout $t > 0$. Donc si 3. est vraie, alors

$$p_t(i,j) \geq p_{t/n}(i_0, i_1) \cdots p_{t/n}(i_{n-1}, i_n) > 0$$

pour tout $t > 0$ et 4. est établie. □

Remarque 3.7.2. *Le point 4. montre que la situation est plus simple en temps continu qu'en temps discret. En temps continu, si i mène à j alors avec une probabilité strictement positive, X_t peut valoir j (sachant que $X_0 = i$). En temps, si i mène à j, il faut parfois un certain temps avant que la chaîne issue de i atteigne j et des phénomènes de périodicité apparaissent. Par exemple, pour la marche aléatoire simple issue de 0, le temps d'atteinte de 4 sera supérieur ou égal à 4 et pair.*

3.8 Temps d'atteinte et d'absorption

Soit $(X_t)_{t \geq 0}$ une chaîne de Markov de générateur Q. Le temps d'atteinte d'un sous-ensemble A de I est la v.a. D^A définie par

$$D^A(\omega) = \inf\{t \geq 0 \; ; \; X_t(\omega) \in A\}, \quad \text{avec la convention } \inf \emptyset = +\infty.$$

Si H^A est le temps d'atteinte de A pour la chaîne incluse, alors $\{D^A < \infty\} = \{H^A < \infty\}$, et sur cet ensemble, on a $D^A = J_{H^A}$.

Exemple 3.8.1. *On considère la chaîne de générateur*

$$Q = \begin{pmatrix} -2 & 1 & 1 & 0 \\ 2 & -6 & 2 & 2 \\ 3 & 3 & -9 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Quel est le temps moyen mis par X pour atteindre 4 depuis 1 ? Notons $k_i = \mathbb{E}_i(H^{\{4\}})$. Partant de 1, on passe en moyenne en 1 un temps $1/q_1 = 1/2$ puis on saute avec des probabilités égales en 2 ou 3. Donc

$$k_1 = \frac{1}{2} + \frac{1}{2} k_2 + \frac{1}{2} k_3.$$

De même

$$k_2 = \frac{1}{6} + \frac{1}{3} k_1 + \frac{1}{3} k_3, \quad k_3 = \frac{1}{9} + \frac{1}{3} k_1 + \frac{1}{3} k_2.$$

La résolution de ce système fournit la valeur $k_1 = 17/12$.

3.9 Récurrence et transience

Définition 3.9.1. *Soit $(X_t)_{t\geq 0}$ une chaîne de Markov de générateur Q.*

1. *L'état i est dit* ***récurrent*** *si $\mathbb{P}_i(\{t \geq 0,\ X_t = i\}$ est non borné$) = 1$.*
2. *L'état i est dit* ***transient*** *si $\mathbb{P}_i(\{t \geq 0,\ X_t = i\}$ est non borné$) = 0$.*

Remarque 3.9.2. *Si $(X_t)_{t\geq 0}$ peut exploser en étant issu de i, alors i ne peut être récurrent.*

Théorème 3.9.3. *On a*

1. *si i est récurrent pour la chaîne incluse, alors i est récurrent pour X ;*
2. *si i est transient pour la chaîne incluse alors i est transient pour X ;*
3. *tout état est soit récurrent, soit transient ;*
4. *récurrence et transience sont des propriétés de classes.*

Démonstration. Supposons que i soit récurrent pour Y. Si $X_0 = i$, X n'explose pas, donc J_n tend vers $+\infty$. De plus, $X_{J_n} = Y_n = i$ infiniment souvent et $\{t \geq 0,\ X_t = i\}$ est non borné avec probabilité 1.

Supposons que i soit transient pour Y. Si $X_0 = i$, alors

$$N = \sup\{n \in \mathbb{N},\ Y_n = i\} < +\infty \quad p.s.,$$

donc $\{t \geq 0,\ X_t = i\}$ est borné par J_{N+1}, qui est fini avec probabilité 1 puisque Y n'a pas atteint de point absorbant avant le temps N.

Les points 3. et 4. se déduisent des résultats analogues sur Y. □

Comme en temps discret, on peut reformuler les notions de récurrence et transience en terme de temps de premier retour. On définit le temps T_i de **premier retour** de $(X_t)_{t\geq 0}$ dans l'état i par

$$T_i(\omega) = \inf\{t \geq J_1(\omega)\ ;\ X_t(\omega) = i\}.$$

Remarque 3.9.4. *Attention à ne pas confondre les temps de premier passage et de premier retour. Dans le second cas, on s'assure que l'on a quitté le point de départ.*

Théorème 3.9.5. *On a la dichotomie suivante :*

1. *si $q_i = 0$ ou $\mathbb{P}_i(T_i < \infty) = 1$, alors i est récurrent et $\int_0^\infty p_t(i,i)\,dt = \infty$,*
2. *si $q_i > 0$ et $\mathbb{P}_i(T_i < \infty) < 1$, alors i est transient et $\int_0^\infty p_t(i,i)\,dt < \infty$.*

Démonstration. Si $q_i = 0$, alors $(X_t)_{t\geq 0}$ ne peut quitter i qui est donc récurrent. De plus, $p_t(i,i) = 1$ pour tout $t \geq 0$ et $\int_0^\infty p_t(i,i)\,dt = \infty$. Supposons que $q_i > 0$. Soit N_i le temps de premier retour en i pour la chaîne incluse $(Y_n)_{n\geq 0}$. Alors $\mathbb{P}_i(N_i < \infty) = \mathbb{P}_i(T_i < \infty)$. Comme i est récurrent pour Y ssi $\mathbb{P}_i(N_i < \infty) = 1$, i est récurrent pour X si et seulement si $\mathbb{P}_i(T_i < \infty) = 1$. De plus, en notant $\pi_{ij}^{(n)} = (\Pi^n)_{ij}$, on va montrer que

$$\int_0^\infty p_t(i,i)\,dt = \frac{1}{q_i}\sum_{n=0}^\infty \pi_{ii}^{(n)},$$

ce qui assure que i est récurrent ssi $\int_0^\infty p_t(i,i)\,dt$ est infini d'après le résultat correspondant pour la chaîne incluse. Pour établir la relation ci-dessus, on écrit, en discutant sur le nombre de sauts qui ont eu lieu avant le temps t,

$$\begin{aligned}
\int_0^\infty p_t(i,i)\,dt &= \int_0^\infty \mathbb{E}_i\big(\mathbf{1}_{\{X_t=i\}}\big)\,dt = \mathbb{E}_i\left(\int_0^\infty \mathbf{1}_{\{X_t=i\}}\,dt\right) \\
&= \mathbb{E}_i\left(\int_0^\infty \mathbf{1}_{\{X_t=i\}} \sum_{n=0}^\infty \mathbf{1}_{\{J_n\le t<J_{n+1}\}}\,dt\right) = \sum_{n=0}^\infty \mathbb{E}_i\left(\mathbf{1}_{\{Y_n=i\}} \int_0^\infty \mathbf{1}_{\{J_n\le t<J_{n+1}\}}\,dt\right) \\
&= \sum_{n=0}^\infty \mathbb{E}_i\big(S_{n+1}\mathbf{1}_{\{Y_n=i\}}\big) = \sum_{n=0}^\infty \mathbb{E}_i(S_{n+1}|Y_n=i)\mathbb{P}_i(Y_n=i) = \frac{1}{q_i}\sum_{n=0}^\infty \pi_{ii}^{(n)},
\end{aligned}$$

ce qui est le résultat annoncé. □

3.10 Mesures invariantes

On dit que λ est une mesure **invariante** pour X si

$$\lambda Q = 0.$$

Remarque 3.10.1. *Une mesure λ est « invariante » ssi, si X_0 suit la loi λ, alors X_t suit la loi λ. En effet, si X_0 suit la loi λ alors X_t suit la loi $\lambda P(t)$. Cette mesure ne dépend pas du temps si $(\lambda P(t))' = 0$, c'est-à-dire si $\lambda Q P(t) = 0$. Or $P(t)$ est une matrice inversible (penser à $P(t) = e^{tQ}$) donc la loi de X_t est indépendante du temps ssi $\lambda Q = 0$.*

On peut relier les mesures invariantes de X aux mesures invariantes de la chaîne incluse.

Théorème 3.10.2. *Soit Q un générateur, Π sa matrice de saut et λ une mesure. Les deux assertions suivantes sont équivalentes :*

1. *λ est invariante ;*
2. *$\mu\Pi = \mu$ avec $\mu_i = \lambda_i q_i$, c'est-à-dire que μ est invariante pour la chaîne incluse.*

Démonstration. On a, par définition, $q_i(\pi_{ij} - \delta_{ij}) = q_{ij}$ pour tous i, j donc

$$(\mu(\Pi - Id))_j = \sum_{i\in I} \mu_i(\pi_{ij} - \delta_{ij}) = \sum_{i\in I} \lambda_i q_{ij} = (\lambda Q)_j.$$

□

Ce lien permet de réutiliser les théorèmes d'existence et d'unicité connus pour les chaînes à temps discret.

Théorème 3.10.3. *Supposons que Q soit le générateur d'une chaîne irréductible et récurrente. Alors Q admet une mesure invariante, unique à une constante multiplicative près.*

Démonstration. Une fois exclu le cas où $I = \{i\}$, l'irréductibilité assure que pour tout $i \in I$, $q_i > 0$. Ceci entraîne que Π est irréductible et récurrente donc Π admet une mesure invariante unique à une constante multiplicative près. □

Remarque 3.10.4. *Attention, si l'espace d'état I est infini, rien ne dit que les mesures invariantes sont de masse finie : l'existence d'une mesure de probabilité invariante n'est pas assurée. Elle l'est par contre dans le cas I fini.*

Pour répondre complètement à cette question, il faut introduire la notion de récurrence positive. On dit qu'un point i est récurrent positif si $q_i = 0$ ou si le temps moyen de premier retour en i $m_i = \mathbb{E}_i(T_i)$ est fini. Un point récurrent non récurrent positif est dit récurrent nul. On obtient alors le lien entre mesure de probabilité invariante et espérance des temps de retour avec une petite subtilité par rapport au temps discret.

Théorème 3.10.5. *Soit Q un générateur irréductible. Les propriétés suivantes sont équivalentes :*

1. *tous les états sont récurrents positifs ;*
2. *il existe un état récurrent positif ;*
3. *Q est non explosif et admet une probabilité invariante λ qui charge tous les points.*

De plus, quand 3. a lieu, $m_i = 1/(\lambda_i q_i)$.

Remarque 3.10.6. *La mesure invariante affecte au point i le poids $1/(q_i m_i)$ qui s'interprète comme le temps moyen que l'on passe en i multiplié par l'inverse du temps de retour. Analogie avec le temps discret ?*

3.11 Convergence à l'équilibre et théorème ergodique

Théorème 3.11.1 (Convergence à l'équilibre). *Soit Q un générateur irréductible non-explosif de semi-groupe $(P_t)_{t\geq 0}$ et de probabilité invariante λ. Alors, pour tous états i, j,*

$$p_t(i,j) \xrightarrow[t\to\infty]{} \lambda_j = \frac{1}{q_j m_j}.$$

De plus, pour toute mesure initiale ν, $\mathbb{P}_\nu(X_t = j) \xrightarrow[t\to\infty]{} \dfrac{1}{q_j m_j}$.

Théorème 3.11.2 (Théorème ergodique). *Soit Q un générateur irréductible et ν la loi de probabilité de X_0. Alors*

$$\frac{1}{t}\int_0^t \mathbf{1}_{\{X_s=i\}}\, ds \xrightarrow[t\to\infty]{} \frac{1}{q_i m_i} \quad p.s. \tag{3.5}$$

De plus, dans le cas récurrent positif, pour toute fonction bornée $f : I \to \mathbb{R}$, on a

$$\frac{1}{t}\int_0^t f(X_s)\, ds \xrightarrow[t\to\infty]{} \overline{f} = \sum_{i\in I} \frac{1}{q_i m_i} f(i) = \sum_{i\in I} \lambda_i f(i) \quad p.s.$$

Remarque 3.11.3. *La relation (3.5) assure que le temps moyen passé en i par la chaîne jusqu'au temps t converge vers le poids que la mesure invariante affecte à i. Ce poids peut être nul si la chaîne n'est pas récurrente positive.*

3.12 Processus de naissance

Un processus de naissance est une généralisation du processus de Poisson dans laquelle le paramètre λ peut dépendre de l'état courant du processus. Les paramètres d'un processus de naissance sont les taux de naissances $(q_i)_{i\in\mathbb{N}}$ supposés à valeurs dans $\mathbb{R}_+$.

Définition 3.12.1. *Un processus continu à droite $(X_t)_{t\geq 0}$ à valeurs dans $\mathbb{N}\cup\{\infty\}$ est processus de naissance de taux $(q_i)_{i\in\mathbb{N}}$ si, conditionnellement à $X_0 = i$, ses temps d'attente $S_1, S_2, \ldots$ sont des v.a. indépendantes de lois exponentielles de paramètres respectifs $q_i, q_{i+1}, \ldots$, et sa chaîne de saut est donnée par $Y_n = i + n$.*

Remarque 3.12.2. *Pour le processus de Poisson la suite $(q_i)_{i\in\mathbb{N}}$ est constante égale à λ.*

Exercice 3.12.3 (Processus de Poisson, avec des oiseaux). *Des rouges-gorges (robins) et des merles (blackbirds), oiseaux solitaires lorsqu'ils cherchent de la nourriture, se posent parfois sur ma terrasse. Pour tout intervalle de temps petit de longueur h, un rouge-gorge se pose avec probabilité $\beta h + o(h)$ et un merle se pose avec probabilité $\rho h + o(h)$. Quelle est la probabilité que les deux premiers oiseaux qui se posent soient des rouges-gorges ? Quelle est la loi du nombre total d'oiseaux qui se sont posés avant le temps t ? Sachant que ce nombre est n, quelle est la loi du nombre de merles qui se sont posés avant le temps t ?*

Exemple 3.12.4 (Croissance d'une colonie de bactéries). *Chaque bactérie dans une colonie se divise en deux bactéries identiques après un temps exponentiel de paramètre λ et ce indépendamment des autres. Si i individus sont présents alors la première division apparaît après un temps exponentiel de paramètre λi. Il y a alors $i+1$ bactéries et, par la propriété d'absence de mémoire, les horloges des bactéries sont remises à zéros. Soit X_t la taille de la colonie au temps t et supposons que $X_0 = 1$. Soit T le temps de la première naissance. Alors*

$$\mathbb{E}(X_t) = \mathbb{E}\big(X_t \mathbf{1}_{\{T\leq t\}}\big) + \mathbb{E}\big(X_t \mathbf{1}_{\{T>t\}}\big) = \int_0^t \lambda e^{-\lambda s}\mathbb{E}(X_t|T=s)\,ds + e^{-\lambda t}.$$

Soit $\mu(t) = \mathbb{E}(X_t)$ alors $\mathbb{E}(X_t|T=s) = 2\mu(t-s)$ pour $0 \leq s \leq t$ donc μ vérifie l'équation fonctionnelle

$$\mu(t) = e^{-\lambda t} + \int_0^t \lambda e^{-\lambda s}\mu(t-s)\,ds,$$

qui peut encore s'écrire grâce au changement de variables $r = t - s$

$$e^{\lambda t}\mu(t) = 2\lambda \int_0^t e^{\lambda r}\mu(r)\,dr.$$

La fonction μ est donc solution de l'équation différentielle $\mu'(t) = \lambda\mu(t)$, avec la condition initiale $\mu(0) = 1$. Ceci assure donc que $\mathbb{E}(X_t) = e^{\lambda t}$: la taille de la population croît exponentiellement vite (mais n'explose pas en temps fini).

La principale différence entre le processus de Poisson et un processus de naissance général est la possibilité qu'a le second d'exploser. Il est très facile de caractériser les processus qui explosent grâce de la proposition 3.1.3. Remarquons qu'un processus de naissance esplose si et seulement si il tend vers $+\infty$ en temps fini.

Théorème 3.12.5. *Soit $(X_t)_{t\geq 0}$ un processus de naissance de taux $(q_i)_{i\in\mathbb{N}}$ issu de 0.*

$$\text{Si} \quad \sum_{i\geq 0}\frac{1}{q_i} < \infty \quad \text{alors} \quad \mathbb{P}(\zeta<\infty) = 1 \quad \text{et si} \quad \sum_{i\geq 0}\frac{1}{q_i} = \infty \quad \text{alors} \quad \mathbb{P}(\zeta=\infty) = 1.$$

Théorème 3.12.6. *Soit $(X_t)_{t\geq 0}$ un processus croissant, continu à droite à valeurs dans $\mathbb{N}\cup\{\infty\}$. Soit $(q_i)_{i\in\mathbb{N}}$ une suite de réels positifs. Les trois propriétés suivantes sont équivalentes :*

1. *conditionnellement à $X_0 = i$, les temps d'attente $S_1, S_2, \ldots$ sont des v.a. indépendantes de lois exponentielles de paramètres respectifs $q_i, q_{i+1}, \ldots$ et la chaîne de saut est donnée par $Y_n = i + n$ pour tout $n \in \mathbb{N}$;*
2. *pour tout $t, h \geq 0$, conditionnellement à $X_t = i$, X_{t+h} est indépendant de $(X_r)_{0\leq r\leq t}$ et, quand h décroît vers 0, uniformément en t,*

$$\mathbb{P}(X_{t+h} = i | X_t = i) = 1 - q_i h + o(h)$$
$$\mathbb{P}(X_{t+h} = i+1 | X_t = i) = q_i h + o(h) \ ;$$

Remarque 3.12.7. *Notons $p_t(i,j) = \mathbb{P}(X_t = j | X_0 = i)$. On peut montrer grâce à la formulation 2. que $P_t = (p_t(i,j))_{ij}$ est solution de*

$$\partial_t p_t(i,j) = -q_j p_t(i,j) + q_{j-1} p_t(i, j-1)$$

que l'on peut écrire sous forme compacte $P_t' = P_t Q$ où Q est le générateur

$$Q = \begin{pmatrix} -q_0 & q_0 & 0 & 0 & \ldots \\ 0 & -q_1 & q_1 & 0 & \ldots \\ 0 & 0 & -q_2 & q_2 & \ldots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Exercice 3.12.8 (Croissance d'une colonie de bactéries). *On reprend l'exemple 3.12.4 de la colonie de bactéries.*

1. *Montrer que la fonction génératrice $G(t,z) = \mathbb{E}(z^{X_t})$ vérifie l'équation fonctionnelle :*

$$G(t,z) = ze^{-\lambda t} + \int_0^t \lambda e^{-\lambda t} G(t-s, z)^2 \, ds.$$

On distinguera les deux cas : $T \geq t$ et $T < t$ où T, l'instant de première division, suit la loi exponentielle de paramètre λ.

2. *Après avoir effectué le changement de variables $u = t - s$ montrer que G est solution de l'équation différentielle*

$$\frac{\partial G}{\partial t}(t,z) = \lambda G(t,z)(G(t,z) - 1).$$

3. *En déduire que pour $q = 1 - e^{-\lambda t}$ et $n \in \mathbb{N}^*$, $\mathbb{P}(X_t = n) = q^{n-1}(1-q)$.*

3.13 Exercices

Exercice 2.
Considérons une flotte de N bus. Chaque véhicule tombe en panne indépendamment des autres avec un taux μ et et envoyé au dépôt. L'atelier ne peut en réparer qu'un à la fois et le travail est distribué selon une loi exponentielle de paramètre λ. Quelle est la mesure d'équilibre du nombre de bus en service ?

Exercice 3.
On modélise l'arrivée d'appels à un central téléphonique par un processus de Poisson de paramètre λ. Les durées des appels sont supposés indépendantes et identiquement distribuées de loi exponentielle de paramètre μ. On suppose que le central peut gérer un nombre infini d'appel.

1. Quel est le modèle utilisé ? Donner le générateur du processus.
2. Montrer qu'en temps grand la loi du nombre d'appels en cours au temps t est approximativement une loi de Poisson de paramètre λ/μ. Comment le vérifier par la simulation ?
3. Donner la longueur moyenne des périodes durant lesquelles au moins une ligne est occupée.
4. Montrer que le nombre moyen de lignes occupées au temps t, sachant que n étaient occupées à l'instant initial, est $ne^{-\mu t} + \lambda(1 - e^{-\mu t})/\mu$.
5. Montrer que, à l'équilibre, le nombre N_t d'appels qui ont pris fin dans l'intervalle de temps $[0, t]$ suit la loi de Poisson de paramètre λt. Est-ce que $(N_t)_{t\geq 0}$ est un processus de Poisson ?

Chapitre 4

Processus de diffusion

Dans la section précédente nous avons étudier des processus de Markov dont la dynamique n'a lieu que par des sauts. Nous allons ici adopter le point de vue opposé en nous intéressant aux processus de Markov possédant des trajectoires continues. L'idée intuitive attachée au terme ■diffusion■ est celle d'une dynamique composée d'un mouvement déterministe de dérive et d'un mouvement aléatoire imprimé via un mouvement brownien, les trajectoires pouvant ainsi être très irrégulières tout en restant continues. Donnons à présent une définition rigoureuse à la notion de diffusion.

4.1 Le cas générique

Dans la suite, a et b sont des fonctions de $\mathbb{R}^d$ à valeurs respectivement dans $\mathbb{R}^d \times \mathbb{R}^d$ et $\mathbb{R}^d$ telles que :

(i) les applications $x \mapsto a(x)$ et $x \mapsto b(x)$ sont mesurables (pour les tribus boréliennes) et localement bornées,

(ii) pour tout x, la matrice $a(x)$ est symétrique et positive, c'est-à-dire que pour tout $\lambda \in \mathbb{R}^d$,

$$\langle a\lambda, \lambda \rangle = \sum_{i,j=1}^{d} a_{ij} \lambda_i \lambda_j \geq 0.$$

On associe au couple (a, b) l'opérateur différentiel du second ordre

$$Lf(x) = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \partial^2_{ij} f(x) + \sum_{i=1}^{d} b(x) \partial_i f(x).$$

Définition 4.1.1. *Un processus de Markov $X = (\Omega, \mathcal{F}, \mathcal{F}_t, X_t, \mathbb{P}_x)$ à valeurs dans $\mathbb{R}^d$ est appelé processus de diffusion de générateur L si*

(i) les trajectoires de X sont continues,

(ii) pour tout $x \in \mathbb{R}^d$ et $f \in \mathcal{C}^\infty_K$,

$$\mathbb{E}_x[f(X_t)] = f(x) + \mathbb{E}_x\left[\int_0^t Lf(X_s)\, ds\right].$$

On dit dans ce cas que X admet a pour coefficient de diffusion et b pour coefficient de dérive.

En toute généralité, le problème de l'existence et l'unicité d'un processus de diffusion associé au couple (a, b) est donc relié à celui d'un problème de martingale. Pour les exemples que nous serons amenés à étudier, l'existence et l'unicité seront assurées.

Supposons donné un couple (a, b) remplissant les conditions ci-dessus et de plus supposons que les fonctions a et b sont régulières (c'est-à-dire suffisamment régulières pour que tous les calculs suivants, écrits formellement, soient justifiables). On associe à a une racine carrée σ, c'est-à-dire une fonction de $\mathbb{R}^d$ dans $\mathbb{R}^d \times \mathbb{R}^d$ telle que

$$\text{pour tout } x \in \mathbb{R}^d, \quad a(x) = \sigma(x)\sigma^*(x).$$

Considérons alors le processus X solution de l'équation différentielle stochastique suivante :

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt.$$

La formule d'Itô assure que pour toute fonction $f \in \mathcal{C}_K^\infty$,

$$f(X_t) = f(X_0) + \int_0^t Lf(X_s)ds + \sum_{i,j=1}^d \int_0^t \sigma_{ij}(X_s)\partial_i f(X_s)dB_s^j$$

et donc

$$\mathbb{E}_x[f(X_t)] = f(x) + \mathbb{E}_x\left[\int_0^t Lf(X_s)\, ds\right].$$

4.2 Exemple fondamental des processus de Kolmogorov

Soit U une fonction sur $\mathbb{R}^d$ de classe $\mathcal{C}^2$. Soit un mouvement brownien $(B_t, \mathcal{F}_t)$ issu de 0 à valeurs dans $\mathbb{R}^d$. On appelle équation de Langevin l'équation différentielle stochastique suivante :

$$dX_t = \sqrt{2}dB_t - \nabla U(X_t)dt. \tag{4.1}$$

La solution de cette équation différentielle stochastique sera appelée processus de Kolmogorov. On pourra parfois préférer à l'équation (4.1) sa version intégrale :

$$X_t = X_0 + \sqrt{2}B_t - \int_0^t \nabla U(X_s)\, ds, \tag{4.2}$$

où X_0 est une variable aléatoire $\mathcal{F}_0$-mesurable.

L'absence d'explosion des processus de Kolmogorov a lieu si U vérifie des conditions de croissance à l'infini convenables. On fera usage de deux hypothèses distinctes (qui d'ailleurs sont souvent vérifiées simultanément) :

$$U(x) \underset{|x|\to\infty}{\longrightarrow} +\infty \quad \text{et} \quad |\nabla U|^2 - \Delta U \text{ est borné inférieurement} \tag{4.3}$$

$$\exists a \in \mathbb{R},\ b \in \mathbb{R}, \quad x \cdot \nabla U(x) \geq -a|x|^2 - b. \tag{4.4}$$

Théorème 4.2.1. *On suppose l'hypothèse (4.3) ou (4.4) vérifiée. Pour toute variable aléatoire X_0 $\mathcal{F}_0$-mesurable, pour presque tout $\omega \in \Omega$, l'équation (4.2) admet une unique solution continue $t \mapsto X_t(\omega)$ définie sur $\mathbb{R}_+$.*

Preuve. Plaçons-nous tout d'abord sous l'hypothèse (4.3). Il faut montrer que le temps d'explosion T est infini. Pour cela, on applique la formule d'Itô à la fonction U, en l'arrêtant au temps

$$T_R \overset{\text{déf.}}{=} \inf\{t \;:\; U(X_t) > R\}.$$

On obtient ainsi

$$U(X_{t\wedge T_R}) - U(X_0) = \sqrt{2}\int_0^{t\wedge T_R} \nabla U(X_s)\cdot dB_s - \int_0^{t\wedge T_R} |\nabla U(X_s)|^2 ds + \int_0^{t\wedge T_R} \Delta U(X_s) ds.$$

□

Comme la première intégrale est une martingale de carré intégrable, en prenant l'espérance des deux membres, on trouve :

$$\mathbb{E}(U(X_{t\wedge T_R}) - U(X_0)) = \mathbb{E}\left[\int_0^{t\wedge T_R} \big(-|\nabla U(X_s)|^2 + \Delta U(X_s)\big) ds\right] \tag{4.5}$$

Puisque U tend vers l'infini à l'infini, U est minorée et l'encadrement

$$\inf U \leq U(X_t(\omega)) \leq R$$

est valable pour tout $t \leq T(\omega)$. Donc

$$\begin{aligned} \mathbb{E}(U(X_{t\wedge T_R})) - \mathbb{E}U(X_0) &= \mathbb{E}\big(U(X_{t\wedge T_R})\mathbf{1}_{\{T_R\leq t\}}\big) + \mathbb{E}\big(U(X_{t\wedge T_R})\mathbf{1}_{\{T_R> t\}}\big) - \mathbb{E}U(X_0) \\ &\geq R\mathbb{P}(T_R \leq t) + \inf U \mathbb{P}(T_R < t) - \mathbb{E}U(X_0). \end{aligned}$$

D'autre part, la fonction $\Delta U - |\nabla U|^2$ est majorée par hypothèse donc il existe deux constantes positives C_1 et C_2 telles que

$$R\mathbb{P}(T_R \leq t) - C_1 \leq C_2\mathbb{E}(t\wedge T_R) \leq C_2 t.$$

Comme T est la limite croissante des T_R (quand $R \to \infty$), on a :

$$\mathbb{P}(T \leq t) \leq \lim_{R\to\infty} \mathbb{P}(T_R \leq t),$$

et donc, d'après l'inégalité précédente, $\mathbb{P}(T \leq t) = 0$. Puisque t est arbitraire, c'est que presque sûrement $T = +\infty$.

Exercice 4.2.2. *Démontrer le théorème 4.2.1 sous l'hypothèse (4.4). On pourra utiliser le même type de raisonnement en appliquant la formule d'Itô à la fonction $x \mapsto |x|^2$.*

Dans toute la suite nous considérons une fonction U de classe $\mathcal{C}^2$ sur $\mathbb{R}^d$ telle que, presque sûrement, le processus de Kolmogorov associé n'explose pas. Posons pour f borélienne bornée sur $\mathbb{R}^d$,

$$N_t f(x) = \mathbb{E}(f(X_t^x)),$$

où $(X_t^x)_{t\geq 0}$ est la solution de l'équation (4.2) vérifiant $X_0^x = x$. Le résultat suivant est fondamental : il montre que les processus de Kolmogorov sont des processus de Markov sur $\mathbb{R}^d$.

Proposition 4.2.3. *La famille $(N_t)_t$ forme un semi-groupe de Markov.*

Preuve. Montrons la propriété de semi-groupe. Soit $t \geq 0$. Comme $h \mapsto B_{t+h} - B_t$ est un mouvement brownien, on peut écrire : $N_h f(y) = \mathbb{E}(f(Y_h^y))$ où le processus $(Y_h^y)_h$ est déterminé comme l'unique solution de l'équation différentielle stochastique suivante :

$$Y_h^y = y + B_{t+h} - B_t - \int_0^h \nabla U(Y_u^y)\, du.$$

D'après la définition de N_t et la propriété de Markov du mouvement brownien,

$$N_t N_h f(x) = \mathbb{E} f\left(Y_h^{X_t^x}\right).$$

Écrivons l'équation satisfaite par le processus $(Z_h)_h$ défini par $Z_h = Y_h^{X_t^x}$:

$$\begin{aligned} Z_h &= X_t^x + B_{t+h} - B_t - \int_0^h \nabla U\left(Y_s^{X_t^x}\right) ds \\ &= x + B_{t+h} - \int_0^t \nabla U(X_s^x)\, ds - \int_0^h \nabla U(Z_s)\, ds. \end{aligned}$$

D'après l'unicité trajectorielle de la solution de (4.2), $Z_h = X_{t+h}$ ce qui donne $N_t N_s f = N_{t+h} f$. □

Il est à présent naturel de chercher à déterminer le générateur infinitésimal des processus de Kolmogorov.

Proposition 4.2.4. *Le générateur infinitésimal de $(N_t)_{t\geq 0}$ est défini sur l'ensemble des fonctions de classe $\mathcal{C}^2$ à gradient borné par :*

$$Lf = \Delta f - \nabla U \cdot \nabla f.$$

Preuve. Appliquons la formule d'Itô à une fonction f de classe $\mathcal{C}^2$ sur $\mathbb{R}$:

$$f(X_t^x) - f(x) = \sqrt{2} \int_0^t \nabla f(X_s^x)\, dB_s - \int_0^t \nabla U(X_s^x) \nabla f(X_s^x)\, ds + \int_0^t \Delta f(X_s^x)\, ds.$$

En prenant l'espérance, la partie martingale s'annule et l'on a

$$N_t f(x) - f(x) = \mathbb{E}\left(\int_0^t Lf(X_s^x)\, ds \right) = \int_0^t N_s Lf(x)\, ds.$$

Lorsque s tend vers 0, $N_s Lf$ tend vers Lf donc en divisant par t et en passant à la limite, on obtient que

$$\lim_{t\to 0} \frac{N_t f(x) - f(x)}{t} = Lf(x).$$

□

L'étape suivante consiste à rechercher les mesures invariantes d'un processus de Kolmogorov. Il est clair que tous les processus de Kolmogorov n'ont pas une mesure invariante de probabilité puisque ce n'est pas le cas par exemple du mouvement brownien.

Définition 4.2.5. *Soit U une fonction sur $\mathbb{R}^d$ telle que e^{-U} soit intégrable pour la mesure de Lebesgue. La mesure de Boltzmann associée à U est la mesure de probabilité sur $\mathbb{R}^d$*

$$\mu(dx) \overset{\text{déf.}}{=} \frac{1}{Z} \exp(-U(x))\, dx, \quad \textit{avec} \quad Z = \int \exp(-U(x))\, dx.$$

Lemme 4.2.6. *La mesure de Boltzmann est réversible (et donc invariante) par rapport au semi-groupe de Komlogorov associé.*

Preuve. Soient f et g deux fonctions de classe C^2.

$$\int fLgd\mu = \int f(x)e^{-U(x)}\Delta g(x)\,dx + \int f\nabla U\nabla g\,d\mu.$$

En intégrant par partie le premier terme du second membre de l'équation ci-dessous, on obtient :

$$\int fLgd\mu = -\int \nabla f\nabla g\,d\mu.$$

Cette expression est parfaitement symétrique en f et g, ce qui achève la preuve. □

Remarque 4.2.7. *L'espace fonctionnel naturellement associé au processus de Kolmogorov est alors $L^2(\mu)$. Les opérateurs N_t peuvent s'étendre par un raisonnement de densité à $L^2(\mu)$ et la propriété de symétrie que nous venons de souligner signifie que ces opérateurs sont de plus auto-adjoints dans $L^2(\mu)$. ON trouvera une démonstration de ce résultat dans [Roy99].*

Exercice 4.2.8. *Soit (N_t) un semi-groupe de Kolmogorov de mesure invariante μ. Montrer que, pour tout $p \geq 1$ et $f \geq 0$,*

$$\int (N_t f)^p\,d\mu \leq \int f^p\,d\mu.$$

*En déduire que N_t induit une contraction de $L^p(\mu)$ (*i.e.* endomorphisme de $L^p(\mu)$ de norme inférieure à 1).*

Exercice 4.2.9. *Montrer que, pour un semi-groupe de Kolmogorov, on a $\mathbf{\Gamma}(f) = |\nabla f|^2$. Ceci explique la dénomination carré du champ (de gradient).*

4.3 Diffusions sur un intervalle

On souhaite à présent étudier des processus non plus à valeurs dans $\mathbb{R}$ tout entier mais plutôt sur un intervalle ouvert

$$I =]l, r[; \quad -\infty \leq l < r \leq +\infty.$$

Selon les coefficients de l'équation différentielle stochastique

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt, \tag{4.6}$$

le processus restera ou non dans un intervalle donné. Comme nous allons le voir le mouvement brownien ou le processus d'Ornstein-Uhlenbeck sortent presque sûrement de tout intervalle strictement inclus dans $\mathbb{R}$. Cela dit, ce n'est pas le cas de tous les processus. Considérons des EDS de la forme

$$dX_t = dB_t + \left(\frac{a}{X_t} - \lambda X_t\right)dt$$

avec $a > 0$ et $\lambda \geq 0$, ou encore

$$dX_t = \sqrt{1 - X_t^2}dB_t - (n-1)X_t dt.$$

Dans le premier cas, le processus est repoussé de 0 par le terme a/x et attiré vers 0 par le terme $-\lambda x$. Lorsque x est petit le premier terme est bien sûr prépondérant et le second ne joue aucun rôle tandis que la situation s'inverse lorsque x est grand. La question est de savoir si le terme répulsif suffit à maintenir le processus dans $]0,+\infty[$ ou si, partant de $x>0$, il est possible d'atteindre 0 en un temps fini. La réponse que nous allons pouvoir apporter est la suivante : si $a \geq 1$ alors, avec probabilité 1, le processus n'atteindra pas 0 et il l'atteindra à coup sûr dans le cas contraire.

Dans le deuxième exemple, où $I =]-1,1[$, le processus subit une force de rappel déterministe vers 0. De plus, lorsqu'il s'approche de 1 (ou -1) la partie diffusive à tendance à disparaître, ce qui permet à la force de rappel de devenir prédominante, à condition que n soit supérieur ou égal à 2.

Pour que la solution de (4.6) reste dans l'intervalle I il semble nécessaire les coefficients soient singuliers aux bords finis de I. On doit donc définir une notion de solution pour l'EDS qui prenne en compte à la fois le fait que les coefficients peuvent être singuliers et que le processus peut ne pas rester dans I. L'idée est de considérer un processus qui, tant qu'il est dans I, évolue selon la dynamique (4.6), et qui est arrêté dès qu'il touche le bord de I.

Définition 4.3.1. *Une solution faible dans l'intervalle I de (4.6) est un triplet $(X,B),(\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\}$, où*

1. *$(\Omega,\mathcal{F},\mathbb{P})$ est un espace probabilisé et $\{\mathcal{F}_t\}$ une filtration standard,*
2. *$X=(X_t,\mathcal{F}_t;0\leq t)$ est un processus continu, adapté à valeurs dans $[l,r]$ avec $X_0\in I$ p.s. et $B=(B_t,\mathcal{F}_t;0\leq t)$ est un mouvement brownien standard sur $\mathbb{R}$.*
3. *pour toutes suites strictement monotones $(l_n)_{n\in\mathbb{N}^*}$ et $(r_n)_{n\in\mathbb{N}^*}$ telles que $l<l_n<r_n<r$, $\lim l_n = l$, $\lim r_n = r$, et*
$$\textit{pour } n\geq 1,\quad S_n=\inf\{t\geq 0, X_t\notin]l_n,r_n[\},$$
on a, pour tout $n\geq 1$,
$$\mathbb{P}\left(X_{t\wedge S_n}=X_0+\int_0^t b(X_s)\mathbf{1}_{\{s\leq S_n\}}\,ds+\int_0^t \sigma(X_s)\mathbf{1}_{\{s\leq S_n\}}\,dB_s;\ \forall 0\leq t\right)=1.$$

La variable aléatoire
$$S=\inf\{t\geq 0;\quad X_t\notin]l,r[\}=\lim_{n\to\infty}S_n$$
est appelé temps de sortie de I. L'hypothèse $X_0\in I$ garantie que $\mathbb{P}(S>0)=1$.

On supposera dans toute la suite que les coefficients $\sigma: I\to\mathbb{R}$ et $b: I\to\mathbb{R}$ vérifient
$$\forall x\in I,\quad \sigma(x)>0, \tag{ND}$$
$$\forall x\in I,\ \exists\varepsilon>0\quad\text{tel que}\quad \int_{x-\varepsilon}^{x+\varepsilon}\frac{1+|b(y)|}{\sigma^2(y)}\,dy>\infty. \tag{LI}$$

Il est possible de relier la non-explosion du processus $(X_t)_{t\geq 0}$ au comportement d'une fonction des coefficients σ et b aux bords de I.

Définition 4.3.2. *On définit la fonction d'échelle (scale function) p du processus $(X_t)_{t\geq 0}$ par*
$$p(x)=\int_c^x \exp\left(-2\int_c^y\frac{b(z)\,dz}{\sigma^2(z)}\right)dy,$$

où $c \in I$.

La mesure de vitesse (speed measure) m est définie par

$$m(dx) = \mathbf{1}_{\{x\in I\}} \frac{2}{p'(x)\sigma^2(x)}\, dx = \mathbf{1}_{\{x\in I\}} \exp\left(2\int_c^x \frac{b(z)\,dz}{\sigma^2(z)}\right) dx.$$

Remarque 4.3.3. *La définition de p dépend de c mais on ne s'intéressera qu'à des comportements asymptotiques de p qui eux seront indépendants de c. Remarquons de plus que p est solution de*

$$\frac{1}{2}\sigma^2(x)p''(x) + b(x)p'(x) = Lp(x) = 0 \tag{4.7}$$

$$p(c) = 0 \quad et \quad p'(c) = 1.$$

La dépendance de m en c ne se fait qu'à une constante multiplicative près.

Remarquons qu'un calcul très simple (une intégration par partie) montre m est la mesure invariante (qui est d'ailleurs réversible) du processus X. Avec ces notations, la fonction

$$M_{a,b}(x) = -\int_a^x (p(x)-p(y))m(dy) + \frac{p(x)-p(a)}{p(b)-p(a)} \int_a^b (p(b)-p(y))m(dy)$$

est solution de l'équation

$$b(x)M'(x) + \frac{1}{2}\sigma^2(x)M''(x) = 1 \quad \text{si } a < x < b,$$

$$M(a) = M(b) = 0.$$

La proposition suivante permet de relier les propriétés d'explosion dans un intervalle $[a,b]$ strictement inclus dans I aux fonctions $M_{a,b}$ et p.

Proposition 4.3.4. *Si $l < a < x < b < r$, $X_0 = x$ et $T_{a,b} = \inf t \geq 0, X_t \notin]a,b[$ alors*

$$\mathbb{E}(T_{a,b}) = M_{a,b}(x). \tag{4.8}$$

De plus,

$$\mathbb{P}\big(X_{T_{a,b}} = a\big) = \frac{p(b)-p(x)}{p(b)-p(a)}, \quad et \quad \mathbb{P}\big(X_{T_{a,b}} = b\big) = \frac{p(x)-p(a)}{p(b)-p(a)}. \tag{4.9}$$

Remarque 4.3.5. *Ce résultat généralise le cas bien connu du mouvement brownien pour lequel la probabilité de sortir en a est égale à $(b-x)/(b-a)$. Toute l'astuce de la preuve est de remarquer que $p(X_t)$ est une martingale puisque $Lp = 0$.*

Preuve. Soit (X,B) une solution faible de (4.6) sur I avec $X_0 = x \in]a,b[$. Posons, pour $n \in \mathbb{N}^*$,

$$\tau_n = \inf\left\{t \geq;\ \int_0^t \sigma^2(X_s)\, ds \geq n\right\}.$$

La formule d'Itô appliquée à $M_{a,b}(X_t)$ assure que

$$M_{a,b}(X_{t\wedge\tau_n\wedge T_{a,b}}) = M_{a,b}(x) - (t\wedge\tau_n\wedge T_{a,b}) + \int_0^{t\wedge\tau_n\wedge T_{a,b}} M'_{a,b}(X_s)\sigma(X_s)dB_s.$$

En prenant l'espérance et en faisant tendre $n \to \infty$, on obtient

$$\mathbb{E}(t\wedge T_{a,b}) = M_{a,b}(x) - \mathbb{E}M_{a,b}(X_{t\wedge T_{a,b}}) \leq M_{a,b}(x) < \infty.$$

Il reste à faire tendre t vers ∞ pour obtenir que $\mathbb{E}(T_{a,b})$ est fini. Ceci nous permet de conclure que

$$\lim_{t\to\infty} \mathbb{E}M_{a,b}(X_{t\wedge T_{a,b}}) = \mathbb{E}M_{a,b}(X_{T_{a,b}}) = 0,$$

puisque $M_{a,b}(a) = M_{a,b}(b) = 0$. Ceci fournit (4.8).

Pour achever la preuve, il suffit d'appliquer de la même manière la formule d'Itô à $p(X_t)$ pour obtenir, grâce à (4.7),

$$p(x) = \mathbb{E}p(X_{T_{a,b}}) = p(a)\mathbb{P}\big(X_{T_{a,b}} = a\big) + p(b)\mathbb{P}\big(X_{T_{a,b}} = b\big).$$

Combinée avec $\mathbb{P}\big(X_{T_{a,b}} = a\big) + \mathbb{P}\big(X_{T_{a,b}} = b\big) = 1$ la relation précédente entraîne (4.9). □

La proposition suivante relie le comportement de la fonction d'échelle aux bords de l'intervalle à celui de la diffusion sous-jacente.

Proposition 4.3.6. *Supposons que (ND) et (LI) sont vérifiées et soit X une solution faible de (4.6) avec condition initiale déterministe $X_0 = x \in I$.*

1. *Si $p(l+) = -\infty$ et $p(r-) = +\infty$, alors*

$$\mathbb{P}(S = \infty) = \mathbb{P}\left(\sup_{0\le t} X_t = r\right) = \mathbb{P}\left(\inf_{0\le t} X_t = l\right) = 1.$$

En particulier, le processus n'explose pas et il est récurrent : pour tout $y \in I$,

$$\mathbb{P}(\exists t > 0, \quad X_t = y) = 1.$$

2. *Si $p(l+) > -\infty$ et $p(r-) = +\infty$, alors*

$$\mathbb{P}\left(\lim_{t\uparrow S} X_t = l\right) = \mathbb{P}\left(\sup_{0\le t<S} X_t < r\right) = 1.$$

3. *Si $p(l+) = -\infty$ et $p(r-) < +\infty$, alors*

$$\mathbb{P}\left(\lim_{t\uparrow S} X_t = r\right) = \mathbb{P}\left(\sup_{0\le t<S} X_t > l\right) = 1.$$

4. *Si $p(l+) > -\infty$ et $p(r-) < +\infty$, alors*

$$\mathbb{P}\left(\lim_{t\uparrow S} X_t = l\right) = 1 - \mathbb{P}\left(\lim_{t\uparrow S} X_t = r\right) = \frac{p(r-) - p(x)}{p(r-) - p(l+)}.$$

Preuve. D'après (4.9), pour $l < a < x < b < r$, on a

$$\mathbb{P}\left(\inf_{0\le t<S} X_t \le a\right) \ge \mathbb{P}(X_{T_{a,b}} = a) = \frac{1 - p(x)/p(b)}{1 - p(a)/p(b)},$$

puisque si X a atteint a avant b, il a atteint a avant de sortir de I donc avant S. En faisant tendre b vers r, on obtient que, pour tout $a < x$, $\inf_{0\le t<S} X_t$ est inférieur à a p.s. Il ne reste plus qu'à écrire

$$\mathbb{P}\left(\inf_{0\le t<S} X_t = l\right) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} \left\{\inf_{0\le t<S} X_t \le l + 1/n\right\}\right) = 1.$$

On raisonne de même pour le sup de X. Supposons à présent que $\mathbb{P}(S < \infty) > 0$. Alors l'événement

$$\left\{\lim_{t\to S} X_t \text{ existe et est égal à } l \text{ ou } r\right\}$$

a une probabilité positive ce qui contredit le fait que les événements

$$\left\{\inf_{0\le t<S} X_t = l\right\} \quad \text{et} \quad \left\{\sup_{0\le t<S} X_t = r\right\}$$

sont de probabilité 1. □

Remarque 4.3.7. *Le résultat précédent ne répond que partiellement à la question de la non-explosion : dans le cas 1. uniquement, on est assuré que S est infini. Par exemple, la diffusion associée à $\sigma(x) = \sigma > 0$ et $b(x) =sgn(x)$ entre dans le cas 4. et S est presque sûrement infini.*

Exercice 4.3.8. *Dans quels cas entrent les diffusions suivantes :*

1. *le mouvement brownien $X_t = \mu t + \sigma W_t$ sur $I =]-\infty, +\infty[$ (discuter selon le signe de μ),*
2. *le processus de Bessel de dimension d sur $I =]0, +\infty[$ avec $\sigma(x) = 1$ et $b(x) = (d-1)/2x$ (discuter selon la valeur de d, et justifier).*
3. *le carré de Bessel de dimension d sur $I =]0, +\infty[$ avec $\sigma(x) = 1$ et $b(x) = (d-1)/2x - \lambda x$ où $\lambda > 0$ (discuter selon la valeur de d, et justifier).*

Chapitre 5

Inégalités fonctionnelles

L'objet de ce chapitre est d'établir des inégalités fonctionnelles dite de Poincaré ou de Sobolev logarithmiques. Ces inégalités peuvent être vues comme des propriétés satisfaites par des mesures de probabilité. Elles permettent également d'étudier le comportement en temps long de processus stochastiques et d'équations aux dérivées partielles.

Dans tout le chapitre, $(E, \mathcal{F}, \mu)$ désigne un espace probabilisé. On s'intéressera principalement au cas où $E = \mathbb{R}^d$ muni de la tribu des boréliens.

On désigne par $L^2(\mu)$ l'ensemble des fonctions mesurables de E dans $\mathbb{R}$ de carré μ-intégrable.

5.1 Inégalité de Poincaré

On définit la variance d'une fonction μ-intégrable f par

$$\mathrm{Var}_\mu(f) = \mathbb{E}_\mu\big((f - \mathbb{E}_\mu(f))^2\big) = \mathbb{E}_\mu(f^2) - (\mathbb{E}_\mu f)^2.$$

La variance est toujours positive, nulle si et seulement si f est μ-presque sûrement constante et infinie si et seulement si f n'est pas de carré intégrable. Elle est homogène d'ordre 2 et est invariante par translation.

La variance possède une formule variationnelle :

$$\mathrm{Var}_\mu(f) = \inf_{a \in \mathbb{R}} \mathbb{E}_\mu\big((f - a)^2\big).$$

La quantité $\mathrm{Var}_\mu(f)$ est donc la distance de f à l'ensemble des constantes au sens des moindres carrés, et la borne inférieure est atteinte par la moyenne de f. La quantité $\mathbb{E}_\mu(f)$ apparaît alors comme la projection orthogonale de f dans $L^2(\mu)$ sur le sous-espace des fonctions constantes.

L'énergie d'une fonction f de $E = \mathbb{R}^d$ dans $\mathbb{R}$ est définie par

$$\mathcal{E}_\mu(f) = \mathbb{E}_\mu\big(|\nabla f|^2\big).$$

Définition 5.1.1. *Une mesure de probabilité μ sur $\mathbb{R}^d$ vérifie une inégalité de Poincaré de constante c si pour toute fonction régulière de $\mathbb{R}^d$ dans $\mathbb{R}$,*

$$\mathrm{Var}_\mu(f) \leq c\mathcal{E}_\mu(f). \tag{5.1}$$

Remarquons immédiatement que si μ est une mesure sur $\mathbb{R}$ qui admet une densité g_μ par rapport à la mesure de Lebesgue alors nécessairement g_μ doit être strictement positive sur tout $\mathbb{R}$. En effet, dans le cas contraire, il existe des fonctions non constantes dont le gradient est nul.

En fait, l'inégalité de Poincaré doit être vue comme une propriété satisfaite par le générateur infinitésimal A d'un processus de Markov sur E de mesure invariante μ. Elle s'écrit alors en toute généralité : pour toute fonction f régulière de E dans $\mathbb{R}$,

$$\mathrm{Var}_\mu(f) \leq -c \int f A f \, d\mu. \tag{5.2}$$

Par définition de l'opérateur carré du champ $\mathbf{\Gamma}$ et la propriété d'invariance,

$$\int f A f \, d\mu = - \int \mathbf{\Gamma}(f) \, d\mu.$$

Si μ est une mesure à densité régulière (strictement positive) sur $\mathbb{R}$, l'inégalité (5.1) peut donc être vue comme un cas particulier de l'inégalité (5.2) en choisissant

$$Af = f'' + \frac{g_\mu'}{g_\mu} f' = Af = f'' + \log(g_\mu)' f'.$$

L'opérateur $\mathbf{\Gamma}(f)$ vaut alors $|\nabla f|^2$ est l'on retrouve donc l'inégalité (5.1).

Remarquons que l'inégalité de Poincaré s'écrit encore comme un contrôle de la norme de la projection de f sur l'espace orthogonal aux constantes :

$$f \perp 1 \Longrightarrow \|f\|_2^2 \leq c\mathcal{E}_\mu(f).$$

Nous allons voir que l'inégalité de Poincaré est directement liée à une propriété fondamentale du spectre de A.

Théorème 5.1.2. *Soit $(P_t)_t$ un semi-groupe de Markov de générateur infinitésimal A et de mesure invariante μ. Les deux assertions suivantes sont équivalentes :*

1. *Il existe une constante $\lambda > 0$ telle que, pour toute fonction $f \in \mathcal{A}$,*
$$\|P_t f - \mathbb{E}_\mu f\|_2^2 \leq e^{-2\lambda t} \mathrm{Var}_\mu(f); \tag{5.3}$$
2. *La mesure μ vérifie une inégalité de Poincaré de constante $c > 0$.*

De plus, les constantes optimales sont liées par la relation $c = 1/\lambda$.

Remarque 5.1.3. *L'inégalité (5.3) porte le nom d'■inégalité de trou spectral■ pour la raison suivante. Le générateur infinitésimal A d'un semi-groupe de Markov de mesure invariante μ n'admet que des valeurs propres négatives et 0 est valeur propre associée aux fonctions constantes. La meilleure constante λ (*i.e *la plus grande) dans l'inégalité de (5.3) est l'opposé de la première valeur propre non nulle. Cela se voit grâce à l'inégalité de Poincaré. Soient $0 = \lambda_0 \geq -\lambda_1 \geq -\lambda_2 \geq \cdots$ les valeurs propres de A et $(f_n)_n$ une base orthonormée de $L^2(\mu)$ formée de vecteurs propres tels que $Af_n = -\lambda_n f_n$. Soit alors $f \in \mathcal{A}$ orthogonale au sous-espace des fonctions constantes. Elle s'écrit*

$$f = \sum_{n=1}^{\infty} a_n f_n \quad \textit{avec} \quad a_n = \int f f_n \, d\mu \quad \textit{et} \quad \sum_{n=1}^{\infty} a_n^2 \leq 1.$$

L'inégalité de Poincaré est équivalente à

$$\sum_{n=1}^{\infty} a_n^2 = \|f\|_2^2 \leq c \int f A f \, d\mu = -c \sum_{n=1}^{\infty} \lambda_n a_n^2 \leq -c\lambda_1 \sum_{n=1}^{\infty} a_n^2,$$

et assure que $\lambda_1 = 1/c$ (l'égalité étant assurée en choisissant $f = f_1$).

Corollaire 5.1.4. *La mesure γ vérifie une inégalité de Poincaré de constante optimale 1.*

Preuve. Ce résultat découle de l'analyse spectrale du semi-groupe d'Ornstein-Uhlenbeck. □

Examinons à présent le cas de la loi de Laplace ou double exponentielle sur $\mathbb{R}$. Elle admet pour densité la fonction $(1/2)e^{-|x|}$.

Proposition 5.1.5. *La mesure exponentielle (double) sur $\mathbb{R}$, notée μ, vérifie une inégalité de Poincaré de constante optimale 4 : pour toute fonction f régulière sur $\mathbb{R}$,*

$$\mathrm{Var}_\mu(f) \leq 4 \int f'^2 d\mu.$$

Preuve. Remarquons tout d'abord que, si ϕ est une fonction régulière sur $\mathbb{R}$, une intégration par parties assure que

$$\int \phi \, d\mu = \phi(0) + \int \mathrm{sgn}(x)\phi'(x) \, \mu(dx).$$

Soit $g(x) = f(x) - f(0)$. Alors, d'après la relation ci-dessus et l'inégalité de Cauchy-Schwarz,

$$\int g^2 \, d\mu = 2 \int \mathrm{sgn}(x) g'(x) g(x) \, \mu(dx) \leq 2 \left(\int g'^2 \, d\mu \right)^{1/2} \left(\int g^2 \, d\mu \right)^{1/2}.$$

Il ne reste plus qu'à remarquer que $\mathrm{Var}_\mu(f) \leq \int g^2 \, d\mu$, en vertu de la formulation variationnelle de la variance, et que f et g ont même dérivée. La mesure μ vérifie donc une inégalité de Poincaré de constante optimale inférieure ou égale à 4.

Pour montrer que 4 est bien la constante optimale, il faut remarquer que

$$\int |x|^p e^{-|x|} \, dx = p \int |x|^{p-1} e^{-|x|} \, dx.$$

On a alors

$$\mathrm{Var}_\mu(x^{2n+1}) = (4n+2)(4n+1) \int x^{4n} e^{-|x|} \, dx,$$

et de même,

$$\int \left[(x^{2n+1})' \right]^2 e^{-|x|} \, dx = (2n+1)^2 \int x^{4n} e^{-|x|} \, dx.$$

Pour que la relation

$$\mathrm{Var}_\mu(x^{2n+1}) \leq c \int \left[(x^{2n+1})' \right]^2 e^{-|x|} \, dx$$

soit vraie, il faut donc que

$$c \geq \frac{(4n+2)(4n+1)}{(2n+1)^2} = 4 \frac{(1+2/n)(1+1/(4n))}{(1+1/(2n))^2} \xrightarrow{n \to \infty} 4.$$

Si ceci est vrai pour tout n, c'est que, nécessairement, la constante c est supérieure ou égale à 4. □

Il est possible de donner un critère nécessaire et suffisant pour qu'une mesure de probabilité sur $\mathbb{R}$ vérifie une inégalité de Poincaré. Rappelons ici qu'une *médiane* m de la mesure μ sur $\mathbb{R}$ est un réel qui vérifie :

$$\mu(]-\infty, m]) \geq \frac{1}{2} \text{ et } \mu([m, +\infty[) \geq \frac{1}{2}.$$

Pour une mesure absolument continue par rapport à la mesure de Lebesgue de densité strictement positive g_μ, la médiane est unique et est caractérisée par

$$\mu(]-\infty, m]) = 1/2.$$

Pour α réel, nous définissons les quantités (éventuellement infinies) suivantes :

$$B_\alpha^+ = \sup_{x \geq \alpha} \left\{ \int_x^\infty \mu(dy) \int_\alpha^x \frac{1}{g_\mu(y)} dy \right\} \tag{5.4}$$

et

$$B_\alpha^- = \sup_{x \leq \alpha} \left\{ \int_{-\infty}^x \mu(dy) \int_x^\alpha \frac{1}{g_\mu(y)} dy \right\}.$$

Le théorème suivant permet de caractériser les mesures sur $\mathbb{R}$ qui vérifient une inégalité de Poincaré.

Théorème 5.1.6. *Soit μ une mesure sur $\mathbb{R}$ absolument continue par rapport à la mesure de Lebesgue, de densité strictement positive, et m sa médiane. Alors μ vérifie une inégalité de Poincaré si et seulement si B_m^+ et B_m^- sont finis. Dans ce cas, la constante optimale de Poincaré c vérifie*

$$\frac{1}{2}\big(B_m^+ \vee B_m^-\big) \leq c \leq 4\big(B_m^+ \vee B_m^-\big).$$

Ce résultat montre que seul le comportement des queues de la mesure joue un rôle dans le fait que la mesure vérifie une inégalité de Poincaré.

Exercice 5.1.7. *Quelques exemples pour fixer les idées.*

1. *Retrouver que la mesure de densité $(1/2)\exp(-|x|)$ vérifie une inégalité de Poincaré. Donner un encadrement de la constante.*
2. *Retrouver que la mesure $\mathcal{N}(0,1)$ vérifie une inégalité de Poincaré.*
3. *Plus généralement, à quelle condition sur α la mesure de densité proportionnelle à $\exp(-|x|^\alpha)$ vérifie-t-elle une inégalité de Poincaré ?*

5.2 Inégalité de Sobolev logarithmique

On définit l'entropie d'une fonction positive μ-intégrable g par rapport à μ par

$$\mathrm{Ent}_\mu(g) \stackrel{\text{déf.}}{=} \int g \log g \, d\mu - \int g \, d\mu \log \int g \, d\mu = \int g \log \left(\frac{g}{\int g \, d\mu} \right) d\mu.$$

L'entropie vérifie la propriété d'homogénéité suivante :

$$\forall \lambda > 0, \quad \mathrm{Ent}_\mu(\lambda g) = \lambda \mathrm{Ent}_\mu(g).$$

L'inégalité de Jensen pour la fonction convexe $x \mapsto x \log x$ assure la positivité de l'entropie :

$$\mathrm{Ent}_\mu(g) = \left(\int g \, d\mu \right) \int \left(\frac{g}{\int g \, d\mu} \right) \log \left(\frac{g}{\int g \, d\mu} \right) d\mu \geq 0.$$

L'entropie est donc une fonctionnelle positive qui s'annule si et seulement si f est μ-presque sûrement constante.

Proposition 5.2.1. *L'entropie peut aussi être définie par la formulation variationnelle suivante :*

$$\mathrm{Ent}_\mu(g) = \sup\left\{\mathbb{E}_\mu(gh),\ \mathbb{E}_\mu(e^h) = 1\right\}, \tag{5.5}$$

Exercice 5.2.2. *Démontrer la proposition 5.2.1. On pourra remarquer que pour $u \geq 0$ et $v \in \mathbb{R}$,*

$$uv \leq u \log u - u + e^v.$$

Définition 5.2.3. *La mesure μ sur $\mathbb{R}^d$ satisfait une inégalité de Sobolev logarithmique de constante c si pour toute fonction f régulière*

$$\mathrm{Ent}_\mu\left(f^2\right) \leq c \int |\nabla f|^2 \, d\mu. \tag{LS}$$

De même que pour l'inégalité de Poincaré, on peut généraliser cette définition. Le générateur infinitésimal A de mesure invariante μ sur E vérifie une inégalité de Sobolev logarithmique de constante c si pour toute fonction f régulière

$$\mathrm{Ent}_\mu\left(f^2\right) \leq -c \int f A f \, d\mu = c \int \mathbf{\Gamma}(f) \, d\mu.$$

L'inégalité de Sobolev logarithmique est plus forte que l'inégalité de Poincaré comme le montre le résultat suivant.

Proposition 5.2.4. *L'inégalité de Sobolev logarithmique de constante c implique l'inégalité de Poincaré de constante $c/2$.*

Preuve. Il suffit de démontrer l'inégalité de Poincaré pour g bornée d'intégrale nulle. On pose alors $f = 1 + \varepsilon g$. Comme g est d'intégrale nulle sous μ, on obtient les développements limités suivants :

$$\mathrm{Ent}_\mu\left(f^2\right) = 2\varepsilon^2 \int g^2 \, d\mu + o(\varepsilon^2) \quad \text{et} \quad \mathcal{E}_\mu(f) = \epsilon^2 \mathcal{E}_\mu(g) + o(\varepsilon^2).$$

L'inégalité de Sobolev logarithmique appliquée à f donne alors :

$$2\varepsilon^2 \int g^2 \, d\mu + o(\varepsilon^2) \leq c\varepsilon^2 \mathcal{E}_\mu(g) + o(\varepsilon^2).$$

Reste alors à diviser les deux membres par $2\varepsilon^2$ et faire tendre ε vers 0 pour obtenir l'inégalité de Poincaré recherchée. □

Proposition 5.2.5. *La mesure gaussienne standard sur $\mathbb{R}$ vérifie l'inégalité de Sobolev logarithmique suivante : pour toute fonction $f \in H^1(\gamma)$,*

$$\mathrm{Ent}_\gamma\left(f^2\right) \leq 2 \int |f'|^2 \, d\gamma. \tag{5.6}$$

Preuve. Soit f une fonction $\mathcal{C}^\infty$ sur $\mathbb{R}$ telle que $0 < a \leq f \leq b$ pour certaines constantes a et b. On pose :

$$F(t) = \int_{\mathbb{R}} \frac{1}{N_t f} \left(\frac{\partial}{\partial x} N_t f\right)^2 d\gamma,$$

où $(N_t)_{t\geq 0}$ désigne le semi-groupe d'Ornstein-Uhlenbeck. Par définition du semi-groupe d'Ornstein-Uhlenbeck (voir définition 1.2.1),

$$\frac{\partial}{\partial x}N_tf(x) = e^{-t}N_t(f')(x).$$

Cette formule permet de commuter l'opérateur de dérivation et le semi-groupe. Elle entraîne en particulier que

$$\sqrt{\mathbf{\Gamma} N_t f} = |\partial_x N_t f| \leq e^{-t}N_t(|f'|) = e^{-t}N_t\Big(\sqrt{\mathbf{\Gamma}(f)}\Big).$$

Cette propriété se révélera essentielle pour établir les inégalités de Sobolev logarithmiques dans des contextes plus généraux. D'après l'inégalité de Cauchy-Schwarz,

$$(N_t(f'))^2 = \left(N_t\left(\frac{f'}{\sqrt{f}}\sqrt{f}\right)\right)^2 \leq N_t\left(\frac{f'^2}{f}\right)N_t(f)$$

Donc, puisque γ est la mesure invariante de (N_t),

$$F(t) \leq e^{-2t}\int N_t\left(\frac{f'^2}{f}\right)d\gamma = e^{-2t}\int\frac{f'^2}{f}\,d\gamma.$$

Introduisons la fonction α : pour $t \geq 0$,

$$\alpha(t) = \int N_tf\log N_tf\,d\gamma.$$

Puisque $N_0f = f$ et $N_tf \to \mathbb{E}_\gamma(f)$, remarquons que

$$\alpha(0) - \lim_{t\to\infty}\alpha(t) = \mathrm{Ent}_\gamma(f).$$

De plus,

$$\begin{aligned}\alpha'(t) &= \int\partial_tN_tf(1+\log N_tf)\,d\gamma = \int AN_tf(1+\log N_tf)\,d\gamma\\ &= -\int dr_xN_tf\partial_x(1+\log N_tf)\,d\gamma = -F(t).\end{aligned}$$

On obtient donc

$$\mathrm{Ent}_\mu(f) = \int_0^\infty -\alpha'(t)\,dt = \int_0^\infty -F(t)\,dt \leq \int\frac{f'^2}{f}d\gamma\int_0^\infty e^{-2t}\,dt \leq \frac{1}{2}\int\frac{f'^2}{f}d\gamma.$$

En remplaçant f par g^2, on obtient l'inégalité (5.6).

Pour montrer que 2 est la constante optimale dans (5.6), il suffit de choisir f de la forme $e^{\lambda x}$ avec λ dans $\mathbb{R}$. On dit que ces fonctions sont les fonctions extrémales pour l'inégalité (5.6), ou encore qu'elles saturent l'inégalité (5.6). $\square$

5.3 Application à l'équation de Fokker-Planck

On considère l'équation aux dérivées partielles suivante

$$\begin{cases} \partial_t u(t,x) = \partial^2_{xx} u(t,x) + \partial_x(u(t,x)x) & \text{pour tout } (t,x) \in]0,+\infty[\times\mathbb{R}, \\ u(0,x) = u_0(x) & \text{pour tout } x \in \mathbb{R}, \end{cases} \tag{FP}$$

où u_0 est une fonction ■régulière■, positive, d'intégrale 1 par rapport à la mesure de Lebesgue.

On peut montrer par des méthodes d'analyse classique que l'équation de Fokker-Planck admet une solution unique positive telle que $u(t,\cdot)$ est encore une densité de probabilité par rapport à la mesure de Lebesgue. Remarquons que ce dernier point résulte du raisonnement suivant :

$$\partial_t \int u(t,x)\,dx = \int \partial_t u(t,x)\,dx = \int \partial_x(\partial_x u(t,x) + u(t,x)x)\,dx = 0,$$

par intégration par partie. Il faut bien sûr justifier l'intervertion de la dérivation et l'intégration par partie mais l'idée est là.

L'approche probabiliste fournit une représentation de u. Soit en effet X_0 de loi $u_0 dx$ et X le processus d'Ornstein-Uhlenbeck solution de l'équation différentielle stochastique :

$$X_t = X_0 + \sqrt{2} B_t - \int_0^t X_s\,ds.$$

Par définition du semi-groupe d'Ornstein-Uhlenbeck (voir définition 1.2.1), on sait que pour toute fonction f régulière,

$$\begin{aligned} \mathbb{E} f(X_t) &= \int f(X_t | X_0 = x) u_0(x)\,dx = \int N_t f(x) u_0(x)\,dx \\ &= \iint f(y) \exp\left(-\frac{(y - c_t x)^2}{2d_t^2}\right) \frac{dy}{\sqrt{2\pi}}\, u_0(x)\,dx \end{aligned}$$

où $c_t = e^{-t}$ et $d_t = \sqrt{1 - e^{-2t}}$. On peut donc écrire $\mathbb{E} f(X_t)$ comme l'intégrale sur $\mathbb{R}$ de f contre la fonction $v(t,y)$ définie par :

$$v(t,y) = \int u_0(x) \exp\left(-\frac{(y - c_t x)^2}{2d_t^2}\right) \frac{dx}{\sqrt{2\pi d_t^2}}.$$

La loi de X_t admet donc une densité par rapport à la mesure de Lebesgue et cette densité n'est rien d'autre que la fonction $x \mapsto v(t,x)$. Remarquons que l'approche probabiliste permettrait de montrer que si $u_0(\cdot)$ et toutes ses dérivées appartiennent à l'espace des fonctions à croissance lente, il en est de même pour $v(t,\cdot)$ pour tout t.

Soit à présent une fonction f de classe $\mathcal{C}^2$, la formule d'Itô assure que

$$f(X_t) - f(X_s) = \sqrt{2} \int_s^t f'(X_u)\,dB_u + \int_s^t Af(X_u)\,du,$$

où A est le générateur infinitésimal du semi-groupe d'Ornstein-Uhlenbeck. Prenons l'espérance pour faire disparaître le terme martingale et utilisons le fait que $\mathcal{L}(X_t)$ admet une densité par rapport à la mesure de Lebesgue :

$$\int f(x) v(t,x)\,dx - \int f(x) v(s,x)\,dx = \int_s^t \int Af(x) v(u,x)\,dx\,du.$$

Si A^* désigne l'opérateur adjoint de A par rapport à la mesure de Lebesgue, c'est-à-dire que pour g de classe $\mathcal{C}^2$,

$$A^*g(x) = \partial^2_{xx}g(x) + \partial_x(xg(x)),$$

alors on obtient que v vérifie la relation suivante :

$$\int f(x)v(t,x)\,dx - \int f(x)v(s,x)\,dx = \int_s^t \int f(x)A^*v(u,x)\,dx\,du,$$

on dit alors que v est une solution faible de l'équation de Fokker-Planck. Le terme ■faible■ est à comprendre au sens où, pour toute fonction régulière f,

$$\int f(x)\left\{v(t,x) - v(s,x) - \int_s^t A^*v(u,x)\,du\right\}dx = 0.$$

D'après sa forme explicite, il est clair que v est une fonction régulière et elle donc aussi la solution forte de l'équation de Fokker-Planck (FP) de condition initiale u_0. Dans cet exemple précis, on peut d'ailleurs vérifier directement que v est solution de l'équation (FP).

Les inégalité de Poincaré et de Sobolev logarithmique pour la mesure $\mathcal{N}(0,1)$ vont nous permettre de décrire quantitativement le comportement en temps grand de la solution de (FP).

Proposition 5.3.1. *On notera dans la suite $V(x) = x^2/2 + \log\sqrt{2\pi}$ de manière à ce que*

$$\gamma(dx) = e^{-V(x)}\,dx.$$

Pour tout $t \geq 0$, posons

$$\alpha(t) = \int \left(\frac{u(t,x)}{e^{-V(x)}} - 1\right)^2 e^{-V(x)}\,dx = \int \left(ue^V - 1\right)^2 e^{-V}.$$

Alors α est une fonction décroissante du temps et l'inégalité de Poincaré implique que $\alpha(t) \leq \alpha(0)e^{-2t}$.

Pour tout $t \geq 0$, posons

$$\beta(t) = \int u(t,x)\log u(t,x)\,dx + \int u(t,x)V(x)\,dx = \int u\log u + \int uV.$$

Alors β est une fonction décroissante du temps et l'inégalité de Sobolev logarithmique pour la loi $\mathcal{N}(0,1)$ implique que $\beta(t) \leq \beta(0)e^{-2t}$.

Preuve. Calculons tout d'abord la dérivée de α :

$$\alpha'(t) = 2\int \partial_t u\left(ue^V - 1\right) = -2\int [\partial_x u + ux]\partial_x\left(ue^V\right) = -2\int \left[\partial_x\left(ue^V\right)\right]^2 e^{-V}.$$

La fonction α est donc bien décroissante. De plus, l'inégalité de Poincaré appliquée à la fonction $f = ue^V$ s'écrit :

$$\int \left(ue^V - 1\right)^2 = \mathrm{var}_\gamma(ue^V) \leq \int \left[\partial_x(ue^V)\right]^2 e^{-V}.$$

Cette inégalité peut encore s'écrire de la manière suivante : $\alpha(t) \leq -\alpha'(t)/2$. La fonction $t \mapsto \alpha(t)e^{2t}$ est donc décroissante et en particulier, $\alpha(t) \leq \alpha(0)e^{-2t}$.

Le raisonnement est le même pour la décroissance de l'entropie. La dérivée de β vaut :

$$\beta'(t) = 2\int \partial_t u(1+\log u + V) = -\int [\partial_x u + ux]\partial_x(\log u + V) = -\int \frac{\left[\partial_x\left(ue^V\right)\right]^2}{ue^V} e^{-V}.$$

La fonction β est donc bien décroissante. De plus, l'inégalité de Sobolev logarithmique (5.6) appliquée à la fonction $f = \sqrt{ue^V}$ s'écrit :

$$\int ue^V \log(ue^V) e^{-V} = \mathrm{Ent}_\gamma\left(ue^V\right) \leq 2\int \left[\partial_x \sqrt{ue^V}\right]^2 e^{-V} = \frac{1}{2}\int \frac{\left[\partial_x\left(ue^V\right)\right]^2}{ue^V} e^{-V}.$$

Cette inégalité s'écrit encore $\beta(t) \leq -\beta'(t)/2$. La fonction $t \mapsto \beta(t)e^{2t}$ est donc décroissante et en particulier, $\beta(t) \leq \beta(0)e^{-2t}$. □

Remarquons à présent ce que signifie les contrôles obtenus. Le premier peut encore s'écrire :

$$\int \left(u - e^{-V}\right)^2 e^V \leq Ke^{-2t}.$$

Ceci signifie que $v(t,\cdot)$ converge vers e^{-V} dans $L^2(e^V)$ à vitesse exponentielle ou encore que $v(t,\cdot)e^V$ converge vers 1 dans $L^2(\gamma)$. Quant à la deuxième, elle assure que cette convergence a lieu au sens de l'entropie relative. Rappelons que l'entropie relative de μ par rapport à ν s'écrit

$$\mathrm{Ent}(\mu \,|\, \nu) = \begin{cases} \displaystyle\int \frac{d\mu}{d\nu} \log \frac{d\mu}{d\nu}\, d\nu = \int \log \frac{d\mu}{d\nu}\, d\mu & \text{si } \mu \ll \nu \\ +\infty & \text{sinon.} \end{cases}$$

La relation $\mu \ll \nu$ signifie que μ admet une densité par rapport à ν. L'inégalité de Sobolev logarithmique pour la mesure gaussienne γ implique donc que l'entropie relative de $v(t,\cdot)e^V$ par rapport à γ décroît exponentiellement vite.

5.4 Hypercontractivité

Lemme 5.4.1. *Soit $(\mathbf{P}_t)_{t\geq 0}$ un semi-groupe de Kolmogorov de générateur infinitésimal $\mathbf{L}$ et de mesure invariante μ. Soit u une fonction de classe $\mathcal{C}^2$ sur $[a,b]$ et f un élément de $L^2(\mu)$ tel que presque sûrement $a \leq f \leq b$. Alors l'application $t \mapsto u(\mathbf{P}_t f) \in L^2(\mu)$ est dérivable pour $t > 0$, continue en 0, de dérivée $-u'(\mathbf{P}_t f)\mathbf{L}u(\mathbf{P}_t f)$ et*

$$\frac{d}{dt}\int u(\mathbf{P}_t f)\,d\mu = -\frac{1}{2}\int u''(\mathbf{P}_t f)|\nabla \mathbf{P}_t f|^2\,d\mu.$$

Preuve. Par positivité du semi-groupe, $\mathbf{P}_t f \in [a,b]$ et le théorème spectral assure que $\mathbf{P}_t f$ appartient au domaine de $\mathbf{L}$ pour $t > 0$ avec

$$\frac{\partial}{\partial t}\mathbf{P}_t f = \mathbf{L}\mathbf{P}_t f.$$

Il existe une constante M telle que, pour presque tout x,

$$|u(\mathbf{P}_{t+h} f)(x) - u(\mathbf{P}_t f(x) - hu'(\mathbf{P}_t f)(x)(\mathbf{P}_{t+h} f(x) - \mathbf{P}_t f(x)))| \leq Mh^2.$$

Ceci est encore vrai en norme L^2, ce qui assure que $u(\mathbf{P}_t f)$ est dérivable (en temps) dans L^2 de dérivée égale à $u'(\mathbf{P}_t f)\mathbf{L}\mathbf{P}_t f$. Comme $\mathbf{P}_t f$ est dans le domaine de A, il est dans celui de $A^{1/2}$ qui n'est autre que $H^1(\mu)$. Un calcul immédiat sur les distributions montre que $u'(\mathbf{P}_t f)$ appartient aussi à $H^1(\mu)$ avec

$$\nabla u'(\mathbf{P}_t f) = u''(\mathbf{P}_t f)\nabla \mathbf{P}_t f.$$

Donc

$$\frac{d}{dt}\int u(\mathbf{P}_t f)\,d\mu = \int u'(\mathbf{P}_t f)\mathbf{L}\mathbf{P}_t f\,d\mu = -\mathcal{E}_\mu(u'(\mathbf{P}_t f), \mathbf{P}_t f) = -\int u''(\mathbf{P}_t f)|\nabla \mathbf{P}_t f|^2\,d\mu.$$

□

Ce lemme suggère donc que dès que u est une fonction convexe, $t \mapsto \int u(\mathbf{P}_t f)\,d\mu$ est décroissante.

Théorème 5.4.2. *(Gross 1975) Soit $(\mathbf{P}_t)_{t\geq 0}$ un semi-groupe de Kolmogorov vérifiant l'inégalité de Sobolev logarithmique de constante c. Pour tous $p > 1$ et $t \geq 0$, on a*

$$\|\mathbf{P}_t f\|_{q(t)} \leq \|f\|_p, \quad f \in L^p(\mu),$$

où $q(t) = 1 + e^{4t/c}(p-1)$.

Réciproquement, La majoration précédente implique l'inégalité de Sobolev logarithmique de constante c.

Remarque 5.4.3. *Ce théorème relie l'inégalité de Sobolev logarithmique à la propriété régularisante du semi-groupe : $\mathbf{P}_t$ envoie L^p dans $L^{q(t)}$ avec $q(t) > p$.*

Preuve. Par un argument de densité immédiat, il suffit de considérer le cas d'une fonction f à valeurs dans $[a,b]$ avec $a > 0$ (ce qui assure $a \leq \mathbf{P}_t f \leq b$). Puisque $\mathbf{P}_t f \in D(\mathbf{L}) \subset H^1(\mu)$, $(\mathbf{P}_t f)^{q/2} \in H^1(\mu)$. On applique alors à cette fonction l'inégalité de Sobolev logarithmique pour obtenir :

$$\mathrm{Ent}_\mu((\mathbf{P}_t f)^q) \leq c\mathbb{E}_\mu\Big(\big|\nabla(\mathbf{P}_t f)^{q/2}\big|^2\Big) = \frac{cq^2}{4}\mathbb{E}_\mu\big((\mathbf{P}_t f)^{q-2}|\nabla \mathbf{P}_t f|^2\big). \tag{5.7}$$

La clé de la démonstration repose sur le calcul de la dérivée de $\|\mathbf{P}_t f\|_{h(t)}$, pour une fonction $h : \mathbb{R}_+ \to [1 + \infty[$ strictement croissante et de classe C^1 qui sera choisie plus tard. Afin de simplifier les calculs, il convient de dériver le logarithme de la norme. Il vient alors :

$$\begin{aligned}
\frac{d}{dt}\Big(\log \|\mathbf{P}_t f\|_{h(t)}\Big) &= \frac{d}{dt}\left(\frac{1}{h(t)}\log \mathbb{E}_\mu\big((\mathbf{P}_t f)^{h(t)}\big)\right)\\
&= -\frac{h'(t)}{h(t)^2}\log \mathbb{E}_\mu\big((\mathbf{P}_t f)^{h(t)}\big) + \frac{1}{h(t)}\frac{1}{\mathbb{E}_\mu((\mathbf{P}_t f)^{h(t)})}\frac{d}{dt}\big(\mathbb{E}_\mu\big(e^{h(t)\log \mathbf{P}_t f}\big)\big)\\
&= \frac{h'(t)}{h(t)^2\mathbb{E}_\mu((\mathbf{P}_t f)^{h(t)})}\Big[-\mathbb{E}_\mu\big((\mathbf{P}_t f)^{h(t)}\big)\log \mathbb{E}_\mu\big((\mathbf{P}_t f)^{h(t)}\big) + \mathbb{E}_\mu\big(h(t)(\mathbf{P}_t f)^{h(t)}\log \mathbf{P}_t f\big)\\
&\quad + \frac{h(t)^2}{h'(t)}\mathbb{E}_\mu\big((\mathbf{P}_t f)^{h(t)-1}\mathbf{L}\mathbf{P}_t f\big)\Big]\\
&= \frac{h'(t)}{h(t)^2\mathbb{E}_\mu((\mathbf{P}_t f)^{h(t)})}\left[\mathrm{Ent}_\mu\big((\mathbf{P}_t f)^{h(t)}\big) - \frac{h(t)^2(h(t)-1)}{h'(t)}\mathbb{E}_\mu\big((\mathbf{P}_t f)^{h(t)-2}|\nabla \mathbf{P}_t f|^2\big)\right].
\end{aligned}$$

On obtient donc, grâce au contrôle (5.7), la majoration suivante :

$$\frac{d}{dt}\Big(\log \|\mathbf{P}_t f\|_{h(t)}\Big) \leq \frac{h'(t)\mathbb{E}_\mu\big((\mathbf{P}_t f)^{h(t)-2}|\nabla \mathbf{P}_t f|^2\big)}{\mathbb{E}_\mu((\mathbf{P}_t f)^{h(t)})}\left(\frac{c}{4} - \frac{h(t)-1}{h'(t)}\right).$$

Il reste à choisir la fonction h comme la solution de l'équation différentielle $q'/(q-1) = 4/c$ avec $q(0) = p$, c'est-à-dire $q(t) = 1 + (p-1)\exp(4t/c)$. En effet, ce choix garantit que la fonction

$$t \mapsto \|\mathbf{P}_t f\|_{q(t)}$$

est décroissante et en particulier que

$$\|\mathbf{P}_t f\|_{q(t)} \leq \|\mathbf{P}_0 f\|_{q(0)} = \|f\|_p.$$

Réciproquement, le même calcul assure que

$$\frac{d}{dt}\Big(\log \|\mathbf{P}_t f\|_{q(t)}\Big) = \frac{q'(t)}{q(t)^2\mathbb{E}_\mu((\mathbf{P}_t f)^{q(t)})}\left[\mathrm{Ent}_\mu\big((\mathbf{P}_t f)^{q(t)}\big) - \frac{cq(t)^2}{4}\mathbb{E}_\mu\big((\mathbf{P}_t f)^{q(t)-2}|\nabla \mathbf{P}_t f|^2\big)\right],$$

et l'hypercontractivité assure que

$$\left(\frac{d}{dt}\|\mathbf{P}_t f\|_{q(t)}\right)_{|t=0} \leq 0.$$

Ceci implique donc que

$$0 \geq \mathrm{Ent}_\mu(f^p) - \frac{cp^2}{4}\mathbb{E}_\mu\big(f^{p-2}|\nabla f|^2\big),$$

ce qui n'est autre que l'inégalité de Sobolev logarithmique de constante c pour $f^{p/2}$. □

Exercice 5.4.4. *Soit $(N_t)_{t\geq 0}$ le semi-groupe d'Ornstein-Uhlenbeck et γ la mesure gaussienne standard sur $\mathbb{R}$. Pour λ réel considérons la fonction f_λ définie sur $\mathbb{R}$ par $f_\lambda(x) = \exp(\lambda x)$.*

1. *Montrer que $f_\lambda \in H^1(\gamma)$, $\|f_\lambda\|_p = e^{p\lambda^2/2}$ et*

$$N_t(f_\lambda) = \exp\left(\frac{\lambda^2(1-e^{-2t})}{2}\right) f_{\lambda e^{-t}}.$$

2. *En déduire que sous l'hypothèse $q - 1 > e^{2t}(p-1)$, le rapport $\|N_t(f_\lambda)\|_q / \|f_\lambda\|_p$ n'est pas borné et que N_t n'est donc pas continu de L^p dans L^q.*

5.5 Tensorisation et perturbation

Les inégalités de Poincaré et de Sobolev logarithmique partagent une propriété fondamentale qui leur confèrent un très grand intérêt : elles sont stables par passage au produit. Plus précisément, si μ et ν vérifient une de ces inégalités avec la même constante alors c'est aussi le cas de pour la mesure $\mu \otimes \nu$.

5.5.1 Tensorisation

Soient μ_a et μ_b deux mesures de probabilité sur des espaces mesurés $(E_a, \mathcal{F}_a)$ et $(E_b, \mathcal{F}_b)$. Soit A_a (resp. A_b) un générateur infinitésimal agissant sur une algèbre standard $\mathcal{A}_a$ (resp. $\mathcal{A}_b$) de fonctions de E_a (resp. E_b) dans $\mathbb{R}$. Nous supposerons que la mesure μ_a (resp. μ_b) est invariante pour A_a (resp. A_b).

Prolongeons de façons naturelle les opérateurs A_a et A_b aux fonctions de $E = E_a \times E_b$ dans $\mathbb{R}$, en fixant une variable et en faisant agir l'opérateur sur l'autre. Notons alors $A = A_a + A_b$ la somme des deux opérateurs sur une classe de fonctions de $E_a \times E_b$ dans $\mathbb{R}$.

Nous supposerons dans la suite que l'opérateur $A_a + A_b$ est un générateur infinitésimal sur $E_a \times E_b$. Remarquons qu'alors $\mu \overset{\text{déf.}}{=} \mu_a \otimes \mu_b$ est invariante pour $A_a + A_b$. On suppose qu'il existe une algèbre standard $\mathcal{A}$ associée au générateur $A_a + A_b$ incluse dans l'ensemble des fonctions de $E_a \times E_b$ dans $\mathbb{R}$ telle que pour toute fonction $f \in \mathcal{A}$, $f(x, \cdot) \in \mathcal{A}_b$ pour $x \in E_a$ et $f(\cdot, y) \in \mathcal{A}_a$ pour $y \in E_b$.

Ces hypothèses contraignantes sont toujours vérifiées dans le cas de processus de diffusion sur $\mathbb{R}^d$ ou dans le cas d'espaces d'états finis.

Proposition 5.5.1. *Avec les notations ci-dessus, on a*

$$\mathrm{Var}_\mu(f) \le \mathbb{E}_\mu(\mathrm{Var}_{\mu_a}(f)) + \mathbb{E}_\mu(\mathrm{Var}_{\mu_b}(f))$$

et

$$\mathrm{Ent}_\mu(f) \le \mathbb{E}_\mu(\mathrm{Ent}_{\mu_a}(f)) + \mathbb{E}_\mu(\mathrm{Ent}_{\mu_b}(f)),$$

où les notations $\mathrm{Var}_{\mu_i}(f)$ *et* $\mathrm{Ent}_{\mu_i}(f)$ *signifient que seule la* i^{eme} *variable est concernée par les intégrations.*

Preuve. L'inégalité de tensorisation de la variance est équivalente à

$$\mathbb{E}_\mu(f^2) - \mathbb{E}_\mu\big[(\mathbb{E}_{\mu_a} f)^2\big] - \mathbb{E}_\mu\big[(\mathbb{E}_{\mu_b} f)^2\big] + (\mathbb{E}_\mu f)^2 \ge 0,$$

qui peut encore s'écrire

$$\mathbb{E}_\mu\big[(f - \mathbb{E}_{\mu_a} f - \mathbb{E}_{\mu_b} f + \mathbb{E}_\mu f)^2\big] \ge 0,$$

ce qui est toujours vrai.

La tensorisation de l'entropie est un peu plus subtile et fait appel à la formulation variationnelle de l'entropie (5.5). Soit g définie sur E telle que $\mathbb{E}_\mu(e^g) = 1$. On écrit

$$g = \underbrace{g - \log \int e^g \, d\mu_a}_{=g_a} + \underbrace{\log \frac{\int e^g \, d\mu_a}{\int e^g \, d\mu}}_{=g_b}.$$

Remarquons que $\mathbb{E}_{\mu_a}(e^{g_a}) = \mathbb{E}_{\mu_b}(e^{g_b}) = 1$, d'où, en utilisant la formule variationnelle de l'entropie (5.5) pour les mesures μ_a et μ_b,

$$\mathbb{E}_{\mu_a}(f g_a) + \mathbb{E}_{\mu_b}(f g_b) \le \mathrm{Ent}_{\mu_b}(f) + \mathrm{Ent}_{\mu_b}(f).$$

On a donc
$$\mathbb{E}_\mu(fg) = \mathbb{E}_\mu(fg_a) + \mathbb{E}_\mu(fg_b) \le \mathbb{E}_\mu(\mathrm{Ent}_{\mu_b}(f)) + \mathbb{E}_\mu(\mathrm{Ent}_{\mu_b}(f)).$$
La proposition est donc démontrée. □

Théorème 5.5.2. *Supposons que les mesures μ_a et μ_b vérifient l'inégalité de Poincaré de constante c_a (resp. c_b) d'opérateur carré du champ $\mathbf{\Gamma}_a$ associé à A_a (resp. $\mathbf{\Gamma}_b$ associé A_b), alors il en est de même pour μ et $\mathbf{\Gamma} = \mathbf{\Gamma}_a + \mathbf{\Gamma}_b$ avec la constante $c = \max(c_a, c_b)$: pour toute fonction $f \in \mathcal{A}$,*
$$\mathrm{Var}_{\mu_a \otimes \mu_b}(f) \le c\mathbb{E}_{\mu_a \otimes \mu_b}(\mathbf{\Gamma} f).$$

Théorème 5.5.3. *Supposons que les mesures μ_a et μ_b vérifient l'inégalité de Sobolev logarithmique de constante c_a (resp. c_b) pour l'opérateur carré du champ $\mathbf{\Gamma}_a$ associé à A_a (resp. $\mathbf{\Gamma}_b$ associé à A_b), alors il en est de même pour μ et $\mathbf{\Gamma} = \mathbf{\Gamma}_a + \mathbf{\Gamma}_b$ avec la constante $c = \max(c_a, c_b)$: pour toute fonction $f \in \mathcal{A}$,*
$$\mathrm{Ent}_{\mu_a \otimes \mu_b}(f^2) \le c\mathbb{E}_{\mu_a \otimes \mu_b}(\mathbf{\Gamma} f).$$

Exercice 5.5.4. *Démontrer les théorèmes 5.5.2 et 5.5.3.*

5.5.2 Perturbation

Soit μ une mesure de probabilité sur E et U une fonction bornée de E dans $\mathbb{R}$. On définit la mesure ν ■perturbée■ de μ par U comme la mesure de probabilité suivante :
$$d\nu = \frac{e^U}{Z} d\mu \quad \text{où} \quad Z = \int e^U \, d\mu.$$

Les inégalités de Poincaré et Sobolev logarithmiques sont stables (à une constante près) par cette perturbation.

Théorème 5.5.5. *Si μ vérifie l'inégalité de Poincaré (resp. de Sobolev logarithmique) de constante c alors ν vérifie l'inégalité de Poincaré (resp. de Sobolev logarithmique) de constante $c\exp(\mathrm{osc}(U))$, avec $\mathrm{osc}(U) = \sup(U) - \inf(U)$.*

Exercice 5.5.6. *Démontrer la stabilité par perturbation de l'inégalité de Poincaré.*

Preuve. Soit ρ une probabilité sur E. La fonction
$$t \mapsto \int (f^2 \log(f^2) - f^2 \log(t^2) - f^2 + t^2) \, d\rho$$
atteint son minimum sur $\mathbb{R}_+$ pour $t = \|f\|_{L^2(\rho)}$, ce qui assure que, pour tout $t > 0$,
$$\int f^2 \log\left(\frac{f^2}{\|f\|^2_{L^2(\rho)}}\right) d\rho \le \int (f^2 \log(f^2) - f^2 \log(t^2) - f^2 + t^2) \, d\rho. \tag{5.8}$$
En particulier, pour $\rho = \delta_x$, on obtient, pour tous t et x,
$$f(x)^2 \log(f(x)^2) - f(x)^2 \log(t^2) - f(x)^2 + t^2 \ge 0. \tag{5.9}$$

Appliquons l'inégalité (5.8) pour $\rho = \nu$ et $t = \|f\|_{L^2(\mu)}$:

$$\int f^2 \log\left(\frac{f^2}{\|f\|^2_{L^2(\nu)}}\right) d\nu \leq \int \left\{ f^2 \log(f^2) - f^2 \log(\|f\|^2_{L^2(\mu)}) - f^2 + \|f\|^2_{L^2(\mu)} \right\} d\nu.$$

Comme d'après (5.9) l'intégrande du membre de droite est positif, on peut écrire :

$$\begin{aligned}
\int f^2 \log\left(\frac{f^2}{\|f\|^2_{L^2(\nu)}}\right) d\nu &\leq \frac{1}{Z} e^{-\inf(U)} \int \left\{ f^2 \log(f^2) - f^2 \log(\|f\|^2_{L^2(\mu)}) - f^2 + \|f\|^2_{L^2(\mu)} \right\} d\mu \\
&= \frac{1}{Z} e^{-\inf(U)} \int f^2 \log\left(\frac{f^2}{\|f\|^2_{L^2(\mu)}}\right) d\mu \\
&\leq \frac{c}{Z} e^{-\inf(U)} \int \mathbf{\Gamma}(f)\, d\mu \\
&\leq c e^{\operatorname{osc}(U)} \int \mathbf{\Gamma}(f)\, d\nu.
\end{aligned}$$

□

Exercice 5.5.7. *Soit $n \in \mathbb{N}$. Soit Φ_n l'application de $\mathbb{R}^n$ dans $\mathbb{R}$ définie par*

$$\Phi_n(x) = \sum_{i=1}^n \left[\frac{x_i^2}{2} + U(x_i) \right],$$

où U est une fonction bornée de $\mathbb{R}$ dans $\mathbb{R}$. Soit μ_n la mesure de probabilité sur $\mathbb{R}^n$ définie par

$$\mu_n(dx) = \frac{1}{Z_n} e^{-\Phi_n(x)}\, dx, \quad \text{où} \quad Z_n = \int e^{-\Phi_n(x)}\, dx.$$

1. *Montrer (en utilisant successivement la perturbation puis la tensorisation) que la mesure μ_n vérifie l'inégalité de Sobolev logarithmique*

$$\operatorname{Ent}_{\mu_n}(f^2) \leq 2 e^{\operatorname{osc}(U)} \int |\nabla f|^2\, d\mu_n.$$

   Remarquez que la constante ne dépend pas de la dimension n.

2. *Expliquer pourquoi il faut ■perturber puis tensoriser■ et non ■tensoriser puis perturber■.*

5.6 Le critère de convexité

Nous allons établir une nouvelle fois l'inégalité de Sobolev logarithmique pour la mesure gaussienne canonique sur $\mathbb{R}^d$. Toutefois, la méthode employée ici, dite de semi-groupe, se généralisera au cas d'une mesure de Boltzmann. Le semi-groupe d'Ornstein-Uhlenbeck vérifie la propriété de commutation suivante :

$$\nabla N_t f(x) = e^{-t} N_t \nabla f(x). \tag{5.10}$$

De plus, il est clair que $N_t f(x)$ converge vers $\mathbb{E}_\gamma(f)$ quand t tend vers $+\infty$. On a donc

$$\begin{aligned}
\mathrm{Ent}_\gamma(f) &= -\int\int_0^\infty \partial_s[N_s f \log N_s f]\, ds\, d\gamma \\
&= -\int\int_0^\infty (\log N_s f + 1) A N_s f\, ds\, d\gamma \\
&= \int_0^\infty \int \nabla \log N_s f \cdot \nabla N_s f\, d\gamma\, ds = \int_0^\infty \int \frac{|\nabla N_s f|^2}{N_s f}\, d\gamma\, ds \qquad (5.11)\\
&\leq \int_0^\infty \int e^{-2s} \frac{(N_s |\nabla f|)^2}{N_s f}\, d\gamma\, ds \\
&\leq \int_0^\infty \int e^{-2s} N_s \left(\frac{|\nabla f|^2}{f} \right) d\gamma\, ds \qquad (5.12)\\
&\leq \int \frac{|\nabla f|^2}{f}\, d\gamma \int_0^\infty e^{-2s}\, ds = \frac{1}{2} \int \frac{|\nabla f|^2}{f}\, d\gamma,
\end{aligned}$$

où (5.11) est obtenue par réversibilité de γ et (5.12) découle de l'inégalité de Cauchy-Schwarz. En remplaçant f par f^2, on retrouve l'inégalité de Sobolev logarithmique de constante 2 pour la mesure gaussienne

$$\mathrm{Ent}_\gamma(f^2) \leq 2 \int |\nabla f|^2\, d\gamma.$$

Ainsi la clef du raisonnement est la relation (5.10) de commutation de N_t et de ∇. Nous allons voir que le critère de Bakry-Emery permet de caractériser les situations où

$$\sqrt{\mathbf{\Gamma} N_t f} \leq e^{-\rho t} N_t \sqrt{\mathbf{\Gamma} f}$$

est vraie pour un certain ρ réel et tout t positif.

Définition 5.6.1. *On appelle* opérateur $\mathbf{\Gamma_2}$ *la forme bilinéaire symétrique suivante : pour f et g dans $\mathcal{A}$,*

$$\mathbf{\Gamma_2}(f,g) \overset{\text{déf.}}{=} \frac{1}{2}[\mathbf{L}\mathbf{\Gamma}(f,g) - \mathbf{\Gamma}(f,\mathbf{L}g) - \mathbf{\Gamma}(\mathbf{L}f,g)]. \qquad (5.13)$$

On notera $\mathbf{\Gamma}(f)$ et $\mathbf{\Gamma_2}(f)$ pour $\mathbf{\Gamma}(f,f)$ et $\mathbf{\Gamma_2}(f,f)$. Dans le cadre des processus de Kolmogorov, rappelons que $\mathbf{\Gamma}(f,g) = \nabla f \cdot \nabla g$. Calculons à présent $\mathbf{\Gamma_2}(f)$:

$$\mathbf{L}\mathbf{\Gamma}(f) = 2\sum_{i,j} \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 + 2\sum_{i,j} \frac{\partial f}{\partial x_j} \frac{\partial^3 f}{\partial x_i^2 \partial x_j} - 2\sum_{i,j} \frac{\partial U}{\partial x_j} \frac{\partial f}{\partial x_i} \frac{\partial^2 f}{\partial x_i \partial x_j}$$

et

$$2\mathbf{\Gamma}(f,\mathbf{L}f) = 2\sum_{i,j} \frac{\partial f}{\partial x_j} \frac{\partial^3 f}{\partial x_i^2 \partial x_j} - 2\sum_{i,j} \frac{\partial U}{\partial x_i} \frac{\partial f}{\partial x_j} \frac{\partial^2 f}{\partial x_i \partial x_j} - 2\sum_{i,j} \frac{\partial f}{\partial x_j} \frac{\partial f}{\partial x_i} \frac{\partial^2 U}{\partial x_i \partial x_j}.$$

Ceci nous donne, par définition de l'opérateur $\mathbf{\Gamma_2}$,

$$\mathbf{\Gamma_2}(f) = \|\mathrm{Hess}\ f\|_2^2 + (\nabla f)^\top (\mathrm{Hess}\ U)(\nabla f),$$

où

$$\|\mathrm{Hess}\ f\|_2^2 = \sum_{i,j=1}^n \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2.$$

Définition 5.6.2. *Nous dirons que l'opérateur* $\mathbf{L}$ *vérifie une* inégalité de courbure $C(\rho)$, *avec* $\rho \in \mathbb{R}$ *si*

$$\mathbf{\Gamma_2}(f) \geq \rho \mathbf{\Gamma}(f) \tag{5.14}$$

Exercice 5.6.3. *Montrer que le générateur infinitésimal d'un processus de Kolmogorov vérifie le critère* $C(\rho)$ *si et seulement si* ρ *minore le spectre de* Hess U *sur tout* $\mathbb{R}^d$, i.e., *pour tout* $x \in \mathbb{R}^d$,

$$\forall y \in \mathbb{R}^d, \quad \langle y, \mathrm{Hess}\, U(x) y\rangle \geq \rho \langle y, y\rangle.$$

5.6.1 Inégalité de Poincaré

Soit $(\mathbf{P}_t)_{t\geq 0}$ le semi-groupe de Markov associé à $\mathbf{L}$. Nous appellerons inégalité locale une inégalité fonctionnelle vérifiée par les mesures $\mathbf{P}_t(\cdot)(x)$ uniformément en x.

Proposition 5.6.4. *Les assertions suivantes sont équivalentes :*

(i) $\mathrm{C}(\rho)$ *est vérifiée,*

(ii) $\forall f \in \mathcal{A}, \forall t > 0, \quad \mathbf{\Gamma}(\mathbf{P}_t f) \leq e^{-2\rho t}\mathbf{P}_t(\mathbf{\Gamma}(f))$,

(iii) $\forall f \in \mathcal{A}, \forall t > 0, \quad \mathbf{P}_t(f^2) - (\mathbf{P}_t f)^2 \leq \dfrac{1-e^{-2\rho t}}{\rho}\mathbf{P}_t(\mathbf{\Gamma}(f))$.

Preuve.

(i)⇒(ii)

Soit $f \in \mathcal{A}$ et $t > 0$. On considère la fonction $\Psi : [0,t] \to \mathbb{R}$ définie par :

$$\Psi(s) = \mathbf{P}_s(\mathbf{\Gamma}(\mathbf{P}_{t-s}f)).$$

Remarquons que c'est pour que Ψ soit bien définie et pour que les calculs qui suivent soient licites que nous nous plaçons sur l'algèbre $\mathcal{A}$. On notera dans la suite g pour $\mathbf{P}_{t-s}f$. On a alors, d'après la formule de dérivation des fonctions composées,

$$\begin{aligned}\Psi'(s) &= \mathbf{P}_s\left[\mathbf{L}\mathbf{\Gamma}(g) - 2\mathbf{\Gamma}(g, \mathbf{L}g)\right] \\ &= 2\mathbf{P}_s\mathbf{\Gamma_2}(g) \geq 2\rho\mathbf{P}_s\mathbf{\Gamma}(g) = 2\rho\Psi(s).\end{aligned}$$

Donc, d'après le lemme de Gronwall, $\Psi(t) \geq \Psi(0)\exp(2\rho t)$, ce qui n'est autre que (ii).

(ii)⇒(iii)

Posons à présent $\Psi(s) = \mathbf{P}_s[(\mathbf{P}_{t-s}f)^2]$ et encore $g = \mathbf{P}_{t-s}f$. Alors

$$\Psi'(s) = \mathbf{P}_s[\mathbf{L}(g^2) - 2g\mathbf{L}g] = 2\mathbf{P}_s(\mathbf{\Gamma}(g)).$$

Donc

$$\begin{aligned}\mathbf{P}_t(f^2) - (\mathbf{P}_t f)^2 = \Psi(t) - \Psi(0) &= 2\int_0^t \mathbf{P}_s\mathbf{\Gamma}(\mathbf{P}_{t-s}f)ds \\ &\leq 2\int_0^t e^{-2\rho s}\mathbf{P}_s\mathbf{P}_{t-s}\mathbf{\Gamma}(f)ds \qquad (5.15) \\ &= \left(2\int_0^t e^{-2\rho s}ds\right)\mathbf{P}_t\mathbf{\Gamma}(f) \qquad (5.16) \\ &= \frac{1-e^{-2\rho t}}{\rho}\mathbf{P}_t(\mathbf{\Gamma}(f)).\end{aligned}$$

(iii)⇒(i)
Par la formule de Taylor appliquée à $t \mapsto \mathbf{P}_t f$ en 0, il vient :

$$\mathbf{P}_t f = f + t\mathbf{L}f + \frac{t^2}{2}\mathbf{L}(\mathbf{L}f) + o(t^2).$$

Ceci donne

$$\begin{aligned}\mathbf{P}_t(f^2) - (\mathbf{P}_t f)^2 &= t\left[\mathbf{L}(f^2) - 2f\mathbf{L}f\right] \\ &\quad + \frac{t^2}{2}\left[\mathbf{L}\big(\mathbf{L}(f^2)\big) - 2f\mathbf{L}(\mathbf{L}f) - 2(\mathbf{L}f)^2\right] + o(t^2).\end{aligned}$$

En utilisant deux fois la relation

$$\mathbf{L}(gh) = 2\mathbf{\Gamma}(g,h) + g\mathbf{L}h + h\mathbf{L}g,$$

on remarque que

$$\mathbf{L}(\mathbf{L}(f^2)) = 2\mathbf{L}[\mathbf{\Gamma}f + f\mathbf{L}f] = 2\left[\mathbf{L}\mathbf{\Gamma}f + 2\mathbf{\Gamma}(f,\mathbf{L}f) + (\mathbf{L}f)^2 + f\mathbf{L}(\mathbf{L}f)\right].$$

Il vient donc

$$\mathbf{P}_t(f^2) - (\mathbf{P}_t f)^2 = 2t\mathbf{\Gamma}f + t^2[\mathbf{L}(\mathbf{\Gamma}f) + 2\mathbf{\Gamma}(f,\mathbf{L}f)] + o(t^2).$$

D'autre part,

$$\frac{1-e^{-2\rho t}}{\rho}\mathbf{P}_t(\mathbf{\Gamma}(f)) = 2t\mathbf{\Gamma}f + 2t^2[\mathbf{L}(\mathbf{\Gamma}f) - \rho\mathbf{\Gamma}f] + o(t^2).$$

En comparant les deux développements grâce à l'assertion (iii), on obtient exactement la propriété (i). Remarquons que l'on n'utilise ici qu'un développement limité de $(1-e^{-2\rho t})/\rho$ à l'ordre 2 pour obtenir (i). On peut donc remplacer (iii) par une version affaiblie, i.e. remplacer $(1-e^{-2\rho t})/\rho$ par une fonction ϕ définie sur $[0,t_0]$ et à valeurs dans $\mathbb{R}^+$ telle que

$$\phi(t) = 2t - 2\rho t^2 + o(t^2) \text{ en } 0.$$

□

5.6.2 L'inégalité de Sobolev logarithmique locale

On peut améliorer la proposition 5.6.4 lorsque $(\mathbf{P}_t)_{t\geq 0}$ est un semi-groupe de diffusion. Il existe des définitions abstraites de ce qu'est une diffusion dans une variété mais nous nous cantonnons ici au cas réel. Pour fixer les idées, on pourra penser à une diffusion comme un opérateur différentiel du second ordre sans terme constant.

$$\mathbf{L}f(x) = \sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{n} b_i(x)\frac{\partial f}{\partial x_i}(x),$$

de matrice $(a_{ij})_{i,j}$ symétrique positive.

Si $\mathbf{L}$ est un opérateur de diffusion alors pour toute fonction Φ de classe $\mathcal{C}^\infty$ de $\mathbb{R}$ dans $\mathbb{R}$,

$$\mathbf{L}(\Phi(f)) = \Phi'(f)\mathbf{L}f + \Phi''(f)\mathbf{\Gamma}(f). \tag{5.17}$$

Lemme 5.6.5. *Si* $\mathrm{C}(\rho)$ *est vérifiée et* $\mathbf{L}$ *est un générateur de diffusion, on a*

$$\forall f \in \mathcal{A}, \quad 4\mathbf{\Gamma}(f)[\mathbf{\Gamma_2}(f) - \rho\mathbf{\Gamma}(f)] \geq \mathbf{\Gamma}(\mathbf{\Gamma}(f)). \tag{5.18}$$

Preuve. La preuve dans le cas général est un peu pénible. Nous montrons ce qu'il en ait pour les semi-groupes de Kolmogorov. D'après l'exercice 5.6.3,

$$\mathbf{\Gamma_2}(f) - \rho\mathbf{\Gamma}(f) \geq \|\mathrm{Hess}\ f\|_2^2.$$

On a donc

$$4\mathbf{\Gamma}(f)[\mathbf{\Gamma_2}(f) - \rho\mathbf{\Gamma}(f)] \geq 4\sum_{i=1}^{d}(\partial_i f)^2 \sum_{j,k=1}^{d}(\partial_{jk}^2 f)^2.$$

D'autre part,

$$\mathbf{\Gamma}(\mathbf{\Gamma}f) = \sum_{i=1}^{d}\left(\partial_i \sum_{j=1}^{d}(\partial_j f)^2\right)^2 = 4\sum_{i=1}^{d}\left(\sum_{j=1}^{d}\partial_j f \partial_{ij} f\right)^2.$$

L'inégalité de Cauchy-Schwarz assure alors le résultat. □

Ce lemme constitue une amélioration très important du critère de courbure. Il entraîne en particulier le résultat suivant.

Proposition 5.6.6. *Si* $\mathbf{L}$ *est une diffusion, les assertions suivantes sont équivalentes :*

(i) $\mathrm{C}(\rho)$ *est vérifiée,*

(ii) $\forall f \in \mathcal{A}, \forall t > 0, \sqrt{\mathbf{\Gamma}(\mathbf{P}_t f)} \leq \exp(-\rho t)\mathbf{P}_t(\sqrt{\mathbf{\Gamma} f})$.

(iii) pour tout $f \in \mathcal{A}$ *et tout* $t > 0$,

$$\mathbf{P}_t(f^2 \log f^2) - \mathbf{P}_t(f^2)\log \mathbf{P}_t(f^2) \leq \frac{2}{\rho}(1 - e^{-2\rho t})\mathbf{P}_t(\mathbf{\Gamma}(f)).$$

Remarque 5.6.7. *L'inégalité de Cauchy-Schwarz assure que*

$$(\mathbf{P}_t h)^2 \leq \mathbf{P}_t(h^2),$$

donc l'assertion (ii) ci-dessus renforce bien la propriété (ii) de la proposition 5.6.4.

Remarquons encore que dans le cas où $\mathbf{L}$ *est de la forme* $\Delta - \nabla\Phi \cdot \nabla$ *sur* $\mathbb{R}^n$, $\mathbf{\Gamma}$ *est tout simplement le carré de la norme du gradient et donc l'assertion (ii) s'écrit aussi*

$$|\nabla \mathbf{P}_t f| \leq \exp(-\rho t)\mathbf{P}_t|\nabla f|.$$

Preuve de la proposition 5.6.6. **(i)⇒(ii)**
D'après le lemme 5.6.5, il suffit de montrer que (5.18) implique (ii). Pour cela, considérons la fonction $\Psi(s) = \exp(-\rho s)\mathbf{P}_s\sqrt{\mathbf{\Gamma}(\mathbf{P}_{t-s}f)}$. On écrira encore $g = \mathbf{P}_{t-s}f$.

$$\Psi'(s) = -\rho\Psi(s) + \exp(-\rho s)\mathbf{P}_s[\mathbf{L}\sqrt{\mathbf{\Gamma}(g)}] - \exp(-\rho s)\mathbf{P}_s\left[\frac{\mathbf{\Gamma}(g, \mathbf{L}g)}{\sqrt{\mathbf{\Gamma}(g)}}\right].$$

D'après la formule de changement de variables (5.17), on a

$$\mathbf{L}\sqrt{\mathbf{\Gamma}(g)} = \frac{\mathbf{L}(\mathbf{\Gamma}(g))}{2\sqrt{\mathbf{\Gamma}(g)}} - \frac{\mathbf{\Gamma}(\mathbf{\Gamma}(g))}{4\mathbf{\Gamma}(g)^{3/2}}.$$

Donc

$$\begin{aligned}\Psi'(s) &= \exp(-\rho s)\mathbf{P}_s\left[\frac{\mathbf{L}(\mathbf{\Gamma}(g)) - 2\mathbf{\Gamma}(g,\mathbf{L}g)}{2\sqrt{\mathbf{\Gamma}(g)}} - \frac{\mathbf{\Gamma}(\mathbf{\Gamma}(g))}{4(\mathbf{\Gamma}(g))^{3/2}} - \rho\sqrt{\mathbf{\Gamma}(g)}\right] \\ &= \exp(-\rho s)\mathbf{P}_s\left[\frac{4\mathbf{\Gamma}(g)(\mathbf{\Gamma}_2(g) - \rho\mathbf{\Gamma}(g)) - \mathbf{\Gamma}(\mathbf{\Gamma}(g))}{4(\mathbf{\Gamma}(g))^{3/2}}\right] \geq 0,\end{aligned}$$

donc Ψ est croissante et $\Psi(t) \geq \Psi(0)$, ce qui n'est autre que (ii).

(ii)⇒(iii)

Soit Φ une fonction de classe $\mathcal{C}^2$. On note $\Psi(s) = \mathbf{P}_s(\Phi(\mathbf{P}_{t-s}f))$. D'après la formule de changement de variables (5.17),

$$\Psi'(s) = \mathbf{P}_s[\mathbf{L}\Phi(\mathbf{P}_{t-s}f) - \Phi'(\mathbf{P}_{t-s}f)\mathbf{L}\mathbf{P}_{t-s}f] = \mathbf{P}_s[\Phi''(\mathbf{P}_{t-s}f)\mathbf{\Gamma}\mathbf{P}_{t-s}f].$$

En choisissant $\Phi(x) = x\log x$, on obtient, grâce à (ii),

$$\Psi'(s) = \mathbf{P}_s\left[\frac{\mathbf{\Gamma}(\mathbf{P}_{t-s}f)}{\mathbf{P}_{t-s}f}\right] \leq e^{-2\rho(t-s)}\mathbf{P}_s\left[\frac{\left[\mathbf{P}_{t-s}\sqrt{\mathbf{\Gamma}f}\right]^2}{\mathbf{P}_{t-s}f}\right].$$

De plus, par l'inégalité de Cauchy-Schwarz, pour toutes fonctions h et k positives,

$$\frac{(\mathbf{P}_{t-s}k)^2}{\mathbf{P}_{t-s}h} \leq \mathbf{P}_{t-s}\left(\frac{k^2}{h}\right). \tag{5.19}$$

Avec $h = f$ et $k = \sqrt{\mathbf{\Gamma}(f)}$, on obtient

$$\Psi'(s) \leq e^{-2\rho(t-s)}\mathbf{P}_s\mathbf{P}_{t-s}\left(\frac{\mathbf{\Gamma}(f)}{f}\right) = e^{-2\rho(t-s)}\mathbf{P}_t\left(\frac{\mathbf{\Gamma}(f)}{f}\right).$$

Il ne reste plus qu'à écrire

$$\begin{aligned}\mathrm{Ent}_{\mathbf{P}_t(\cdot)}(f) &= \Psi(t) - \Psi(0) \\ &\leq \left(\int_0^t e^{-2\rho(t-s)}ds\right)\mathbf{P}_t\left(\frac{\mathbf{\Gamma}(f)}{f}\right) \\ &\leq \frac{1}{2\rho}(1 - e^{-2\rho t})\mathbf{P}_t\left(\frac{\mathbf{\Gamma}(f)}{f}\right).\end{aligned}$$

On applique l'inégalité précédente à f^2 en utilisant $\mathbf{\Gamma}(f^2)/f^2 = 4\mathbf{\Gamma}(f)$.

(iii)⇒(i)

On utilise ici l'argument qui permet de comparer les inégalités de Poincaré et de Sobolev logarithmique. Dans (iii), on remplace f par $1 + \varepsilon g$. Pour toute mesure de probabilité ν, quand $\varepsilon \to 0$,

$$\mathrm{Ent}_\nu(f)\big((1 + \varepsilon g)^2\big) = 2\varepsilon^2\mathrm{Var}_\nu(g) + o(\varepsilon^2).$$

Comme de plus $\mathbf{P}_t(\mathbf{\Gamma}(1 + \varepsilon g)) = \varepsilon^2\mathbf{P}_t(\mathbf{\Gamma}g)$, il vient

$$\mathbf{P}_t(g^2) - (\mathbf{P}_t g)^2 \leq \frac{c(t)}{2}\mathbf{P}_t(\mathbf{\Gamma}(g)),$$

qui est une inégalité de Poincaré locale qui est équivalente à C(ρ) d'après la proposition 5.6.4. □

Exemple 5.6.8. *Nous allons retrouver ici de manière très rapide que la mesure gaussienne standard de $\mathbb{R}^n$ satisfait une inégalité de Sobolev logarithmique de constante 2 dont nous savons déjà qu'elle est optimale.*

Le laplacien dans $\mathbb{R}^n$ vérifie le critère $C(0)$. D'autre part, le semi-groupe associé au laplacien (dit semi-groupe de la chaleur*) admet la représentation suivante : pour toute fonction f de classe $\mathcal{C}^\infty$ à dérivées bornées,*

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}^n} f(y) e^{-\frac{|x-y|^2}{4t}} \frac{dy}{(4\pi t)^{n/2}}.$$

Remarquons que $(\mathbf{P}_t)_{t \geq 0}$ est aussi, à une renormalisation en temps près, le semi-groupe du mouvement brownien standard sur $\mathbb{R}^n$.

Appliquons à présent le théorème 5.6.6. La constante $2(1 - e^{-2\rho t})/\rho$ se prolonge en $4t$ pour ρ nulle. Enfin, remarquons que pour $t = 1/2$ et $x = 0$, on retrouve l'inégalité de Sobolev logarithmique pour la gaussienne standard, ce qui achève notre propos.

5.7 Inégalités pour la mesure réversible et critères intégrés

Lorsque $\rho > 0$, le critère de courbure permet d'obtenir les inégalités de Poincaré et de Sobolev logarithmique pour la mesure invariante et le carré du champ associés à l'opérateur $\mathbf{L}$. Il suffit pour cela de supposer que le semi-groupe est ergodique :

Définition 5.7.1. *Un semi-groupe $(\mathbf{P}_t)_{t \geq 0}$ de mesure invariante μ est dit* ergodique, *au sens $\mathbf{L}^2(\mu)$, si*

$$\lim_{t \to \infty} \mathbf{P}_t f = \mathbf{E}_\mu(f) \text{ dans } \mathbf{L}^2(\mu).$$

On peut montrer qu'une condition suffisante d'ergodicité est que l'opérateur $\mathbf{\Gamma}$ ne s'annule que pour les fonctions constantes. Les semi-groupes de Kolmogorov sont toujours ergodiques puisqu'alors $\mathbf{\Gamma}(f)$ est égal à $|\nabla f|^2$.

Si $(\mathbf{P}_t)_{t \geq 0}$ est ergodique et $\rho > 0$, on peut faire tendre t vers l'infini dans la proposition 5.6.4 pour montrer que μ vérifie l'inégalité de Poincaré de constante $1/\rho$ et dans le théorème 5.6.6 pour montrer que μ vérifie l'inégalité de Sobolev logarithmique de constante $2/\rho$.

Une question naturelle est alors de déterminer à quelle condition la mesure invariante μ d'un semi-groupe de Kolmogorov de densité $\exp(-\Phi)$ par rapport à la mesure de Lebesgue vérifie une inégalité de Sobolev logarithmique. Une condition suffisante évidente est alors de supposer Φ strictement uniformément convexe c'est-à-dire que le spectre de la matrice hessienne de Φ est minoré par un nombre strictement positif.

Corollaire 5.7.2. *Si la fonction Φ sur $\mathbb{R}^n$ s'écrit comme la somme d'une fonction strictement uniformément convexe U et d'une fonction bornée B, alors la mesure de probabilité μ de densité $Z_\Phi^{-1} \exp(-\Phi)$ vérifie une inégalité de Sobolev logarithmique.*

Preuve. L'opérateur défini pour f dans $\mathcal{A}$ par

$$\mathbf{L}f = \Delta f - \nabla U \cdot \nabla f$$

a pour mesure réversible la mesure de probabilité $Z_U^{-1} \exp(-U)$ et vérifie le critère $\mathrm{C}(\rho)$ pour un certain $\rho > 0$ car

$$\mathbf{\Gamma}_2(f) = \|\mathrm{Hess} f\|_2^2 + (\mathrm{Hess} U \nabla f) \cdot \nabla f \geq \rho |\nabla f|^2 = \rho \mathbf{\Gamma}(f).$$

La mesure $Z_U^{-1}\exp(-U)$ vérifie donc une inégalité de Sobolev logarithmique de constante $2/\rho$. On utilise ensuite le théorème de perturbation pour conclure. □

Toutefois, ce critère peut être affaibli lorsque l'on s'intéresse uniquement à la mesure réversible, comme le montrent les résultats qui suivent.

5.7.1 L'inégalité de Poincaré pour la mesure réversible

Dans toute la suite du chapitre, nous nous placerons sous l'hypothèse de réversibilité de la mesure invariante. Nous regroupons dans la remarque suivante les conséquences de cette hypothèse auxquelles nous aurons recours de manière intensive.

Remarque 5.7.3. *Si la mesure invariante est réversible, on a en particulier :*

$$\mathbf{E}_\mu(\mathbf{\Gamma}(f,g)) = -\mathbf{E}_\mu(f\mathbf{L}g) \text{ et } \mathbf{E}_\mu(\mathbf{\Gamma_2}f) = -\mathbf{E}_\mu(\mathbf{\Gamma}(f,\mathbf{L}f)) = \mathbf{E}_\mu\big((\mathbf{L}f)^2\big). \tag{5.20}$$

Proposition 5.7.4. *Si $(\mathbf{P}_t)_{t\geq 0}$ admet μ pour mesure réversible, est ergodique et si $\rho > 0$, les assertions suivantes sont équivalentes :*

(i) $\forall f \in \mathcal{A},\quad \mathbf{E}_\mu(\mathbf{\Gamma_2}(f)) \geq \rho\mathbf{E}_\mu(\mathbf{\Gamma}(f))$

(ii) $\forall f \in \mathcal{A},\quad \mathrm{Var}_\mu(f) \leq \dfrac{1}{\rho}\mathbf{E}_\mu(\mathbf{\Gamma}(f)).$

Preuve.

(ii)⇒(i)

L'invariance de la mesure et l'inégalité de Cauchy-Schwarz assurent que

$$-\int f\mathbf{L}f d\mu = \int(\mathbf{E}_\mu(f)-f)\mathbf{L}f d\mu \leq (\mathrm{Var}_\mu(f))^{1/2}\Big(\int(\mathbf{L}f)^2 d\mu\Big)^{1/2},$$

donc, par (5.20) et (ii), il vient

$$\int \mathbf{\Gamma}(f)d\mu \leq \frac{1}{\sqrt{\rho}}\Big(\int \mathbf{\Gamma}(f)d\mu\Big)^{1/2}\Big(\int \mathbf{\Gamma_2}(f)d\mu\Big)^{1/2}.$$

(i)⇒(ii)

La clé est d'écrire

$$\begin{aligned}\mathrm{Var}_\mu(f) &= -\int\int_0^\infty \frac{d}{ds}(\mathbf{P}_s f)^2 ds d\mu \\ &= -2\int_0^\infty\int(\mathbf{P}_s f)\mathbf{L}\mathbf{P}_s f d\mu ds = 2\int_0^\infty\int \mathbf{\Gamma}(\mathbf{P}_s f)d\mu ds.\end{aligned}$$

Posons $\Psi(s) = \mathbf{E}_\mu(\mathbf{\Gamma}(\mathbf{P}_s f))$. On a alors

$$\begin{aligned}\Psi'(s) &= 2\int \mathbf{\Gamma}(\mathbf{P}_s f, \mathbf{L}\mathbf{P}_s f)d\mu \\ &= -2\int \mathbf{\Gamma_2}(\mathbf{P}_s f)d\mu \leq -2\rho\int \mathbf{\Gamma}(\mathbf{P}_s f)d\mu = -2\rho\Psi(s).\end{aligned}$$

On a donc $\Psi(s) \leq \exp(-2\rho s)\Psi(0)$, d'après le lemme de Gronwall. Ceci donne

$$2\int_0^\infty \Psi(s)ds \leq \Big(2\int_0^\infty \exp(-2\rho s)ds\Big)\int \mathbf{\Gamma}(f)d\mu,$$

ou encore

$$\mathrm{Var}_\mu(f) \leq \frac{1}{\rho} \int \mathbf{\Gamma}(f) d\mu.$$

□

5.7.2 L'inégalité de Sobolev logarithmique pour la mesure réversible

Sous l'hypothèse supplémentaire de diffusion nous allons établir un critère suffisant mais non nécessaire pour l'inégalité de Sobolev logarithmique.

Proposition 5.7.5. *Soit $(\mathbf{P}_t)_{t\geq 0}$ un semi-groupe de diffusion réversible et ergodique par rapport à μ et soit $\rho > 0$, satisfaisant l'assertion suivante, dite critère super intégral,*

$$\forall f \in \mathcal{A}, \quad \mathbf{E}_\mu\big(e^f \mathbf{\Gamma_2}(f)\big) \geq \rho \mathbf{E}_\mu\big(e^f \mathbf{\Gamma}(f)\big). \tag{5.21}$$

Alors la mesure μ vérifie l'inégalité de Sobolev logarithmique

$$\forall f \in \mathcal{A}, \quad \mathrm{Ent}_\mu\big(f^2\big) \leq \frac{2}{\rho} \mathbf{E}_\mu(\mathbf{\Gamma}(f)).$$

5.8 Concentration de la mesure

5.8.1 Cas de la mesure gaussienne

Théorème 5.8.1. *Soit F une fonction lipschitzienne sur $\mathbb{R}$ de norme de Lipschitz $\|F\|_{Lip} \leq 1$. Alors, pour tout $r \geq 0$,*

$$\gamma(F \geq \mathbb{E}_\gamma(F) + r) \leq e^{-r^2/2}. \tag{5.22}$$

Preuve. Soit F une fonction de norme $\|F\|_{Lip} \leq 1$. La fonction F admet des moments exponentiels de tous ordres par rapport à γ car pour tout $\lambda > 0$,

$$\mathbb{E}_\gamma\big(e^{\lambda F}\big) \leq e^{\lambda F(0)} \mathbb{E}_\gamma\big(e^{\lambda |x|}\big) < \infty.$$

Afin de majorer $\gamma(F \geq \mathbb{E}_\gamma(F) + r)$, appliquons l'inégalité de Markov (exponentielle). On obtient ainsi, pour tout réel strictement positif λ,

$$\gamma(F \geq \mathbb{E}_\gamma(F) + r) = \gamma\big(e^{\lambda(F - \mathbb{E}_\gamma(F))} \geq e^{\lambda r}\big) \leq e^{-\lambda r} \mathbb{E}_\gamma\big(e^{\lambda(F - \mathbb{E}_\gamma(F))}\big).$$

Nous sommes ramenés à contrôler la transformée de Laplace de $F - \mathbb{E}_\gamma(F)$ sous γ. Pour ce faire, le point clé est de montrer l'estimation sous-gaussienne suivante :

$$\mathbb{E}_\gamma\big(e^{\lambda(F - \mathbb{E}_\gamma(F))}\big) \leq \exp(\lambda^2/2). \tag{5.23}$$

Ce contrôle assure alors que, pour tout $\gamma > 0$,

$$\gamma(F - \mathbb{E}_\gamma(F) \geq r) \leq \exp(\lambda r + \lambda^2/2).$$

Une optimisation en λ conduit alors à l'inégalité de déviation (5.22).

Prouvons à présent l'estimation (5.23). Dans un premier temps, nous supposerons que F est de classe $\mathcal{C}^\infty$ bornée. De plus, on peut considérer que F est de moyenne nulle. Soit $(N_t)_{t\geq 0}$

le semi-groupe d'Ornstein-Uhlenbeck et $\lambda > 0$. Notons H la transformée de Laplace de $N_t F$ prise au point λ : pour tout $t \geq 0$,

$$H(\lambda) \stackrel{\text{déf.}}{=} \mathbb{E}_\gamma(e^{\lambda N_t F}).$$

La fonction H est dérivable et sa dérivée vaut

$$H'(t) = \lambda \mathbb{E}_\gamma\big(e^{\lambda N_t F} \mathbf{L} N_t F\big).$$

Rappelons que $N_t F$ converge, quand t tend vers l'infini, vers la fonction constante égale à $\mathbb{E}_\gamma(F) = 0$ et $|N_t F| \leq \sup F$ donc, par convergence dominée, $H(t)$ tend vers 1 quand t tend vers l'infini. Ainsi, nous pouvons écrire que

$$H(t) = 1 - \int_t^\infty H'(s)\, ds.$$

Or

$$H'(s) = \lambda \int e^{\lambda N_s F} \mathbf{L} N_s F \, d\gamma = -\lambda \int \nabla e^{\lambda N_s F} \nabla N_s F \, d\gamma = -\lambda^2 \int e^{\lambda N_s F} |\nabla N_s F|^2 \, d\gamma.$$

Par définition de N_t et grâce à l'inégalité de Cauchy-Schwarz, on a

$$|\nabla N_s F|^2 = e^{-2s} |N_s(\nabla F)|^2 \leq e^{-2s} N_s\big(|\nabla F|^2\big).$$

Remarquons enfin que puisque $\|F\|_{Lip} \leq 1$ et F est régulière, $|\nabla F| \leq 1$. On obtient donc

$$H'(s) \geq -\lambda^2 e^{-2s} \int e^{\lambda N_s F} \, d\gamma = -\lambda^2 H(s),$$

et par suite, H vérifie l'inéquation intégrale suivante :

$$H(t) \leq 1 + \lambda^2 \int_t^\infty e^{-2s} H(s)\, ds$$

Le lemme de Gronwall entraîne alors la majoration

$$H(t) \leq \exp\left(\lambda^2 \int_t^\infty e^{-2s}\, ds\right) = \exp\left(\frac{\lambda^2}{2} e^{-2t}\right).$$

On aboutit finalement à l'inégalité souhaitée : $\mathbb{E}_\gamma(e^{\lambda F}) = H(0) \leq e^{\lambda^2/2}$.

L'extension de l'estimation (5.23) à toute fonction F telle que $\|F\|_{Lip} \leq 1$ découle du lemme d'approximation suivant.

Lemme 5.8.2. *Soit (X, d) un espace métrique et μ une mesure de probabilité sur la tribu borélienne de X. Soit $(F_n)_{n \in \mathbb{N}}$ une suite de fonctions intégrables telles que, pour tout $n \in \mathbb{N}$, $\|F_n\|_{Lip} \leq 1$, tendant presque sûrement vers F, et telle que, pour tout entier n et tout réel positif λ,*

$$\mathbb{E}_\mu\big(e^{\lambda(F_n - \mathbb{E}_\mu(F_n))}\big) \leq e^{\lambda^2/2}. \tag{5.24}$$

Alors F est intégrable et vérifie l'inégalité (5.24)

Preuve. Montrons, dans un premier temps, l'uniforme intégrabilité de la suite $(F_n)_n$. Pour ce faire, nous allons contrôler uniformément les moments d'ordre 2 en les décomposant de manière classique :

$$\mathbb{E}_\mu(F_n^2) = \mathbb{E}_\mu\big((F_n - \mathbb{E}_\mu(F_n))^2\big) + (\mathbb{E}_\mu(F_n))^2.$$

Pour contrôler le terme de variance, remarquons que l'inégalité (5.24) entraîne l'inégalité de concentration

$$\mu(|F_n - \mathbb{E}_\mu(F_n)| \geq r) \leq 2e^{-r^2/2}.$$

En conséquence, on obtient la majoration suivante :

$$\mathbb{E}_\mu\big((F_n - \mathbb{E}_\mu(F_n))^2\big) = \int_0^\infty 2r\mu(|F_n - \mathbb{E}_\mu(F_n)| \geq r)\,dr \leq 4\int_0^\infty re^{-r^2/2}\,dr = 4.$$

D'autre part on peut trouver une borne uniforme des espérances des $(F_n)_n$. En effet, soient m assez grand pour que $\mu(|F| \leq m) \geq 3/4$ et n_0 tel que, pour tout $n \geq n_0$, $\mu(|F_n - F| \geq 1) \leq 1/4$. Alors $\mu(|F_n| \leq m+1) \geq 1/2$, pour $n \geq n_0$. Soit maintenant r_1 tel que $2e^{r_1^2/2} < 1/2$. L'ensemble $\{|F_n| \leq m+1\}\backslash\{|F_n - \mathbb{E}_\mu(F_n)| \geq r_1\}$ est non vide puisque de mesure strictement positive. Pour $n \geq n_0$, il existe donc ω tel que

$$F_n(\omega) - \mathbb{E}_\mu(F_n) \leq r_1 \quad \text{et} \quad |F_n(\omega)| \leq m+1.$$

Nous avons ainsi obtenu la majoration $|\mathbb{E}_\mu(F_n)| \leq r_1 + m + 1$, pourvu que $n \geq n_0$.

On vient ainsi de montrer que $\sup_n \mathbb{E}_\mu(F_n^2) < \infty$ et donc que la suite $(F_n)_n$ est uniformément intégrable. Il en découle alors que $\mathbb{E}_\mu(|F|) < \infty$ et que F_n converge vers F dans $L^1(\mu)$. Par conséquent, grâce au lemme de Fatou,

$$\mathbb{E}_\mu\big(e^{\lambda F}\big) \leq \liminf_{n\to\infty} \mathbb{E}_\mu\big(e^{\lambda F_n}\big) \leq \liminf_{n\to\infty} e^{\lambda \mathbb{E}_\mu(F_n)+\lambda^2/2} = e^{\lambda\mathbb{E}_\mu(F)+\lambda^2/2},$$

ce qui achève la preuve du lemme. □

Soient G une fonction lipschitzienne bornée telle que $\|G\|_{Lip} \leq 1$et $(\rho_\varepsilon)_\varepsilon$ une famille régularisante. Alors $G_\varepsilon \stackrel{\text{déf.}}{=} \rho_\varepsilon * G$ est une fonction de classe $\mathcal{C}^\infty$, bornée et $\|G_\varepsilon\|_{Lip} \leq \|G\|_{Lip} \leq 1$. Ainsi, G_ε vérifie (5.23) et, d'après le lemme ci-dessus, il en est de même pour G, puisque $(G_\varepsilon)_\varepsilon$ converge presque sûrement vers G. L'extension à toutes les fonctions F lipschitziennes telles que $\|F\|_{Lip} \leq 1$ est obtenue de la même manière à l'aide de l'approximation suivante : si F est une fonction lipschitzienne telle que $\|F\|_{Lip} \leq 1$, alors $F_n \stackrel{\text{déf.}}{=} \max(-n, \min(F, n))$ est bornée, $\|F_n\|_{Lip} \leq \|F\|_{Lip} \leq 1$ et $(F_n)_n$ tend vers F presque sûrement. □

5.8.2 Concentration gaussienne et inégalité de Sobolev logarithmique

Il existe un lien très fort entre la mesure gaussienne et l'inégalité de Sobolev logarithmique. Le résultat suivant souligne ce fait en établissant qu'une mesure qui satisfait à une inégalité de Sobolev logarithmique vérifie la propriété de concentration gaussienne.

Théorème 5.8.3. *(Herbst) Soit μ une mesure de probabilité sur $\mathbb{R}^d$ vérifiant l'inégalité de Sobolev logarithmique*

$$\forall f \in \mathcal{C}_b^\infty(\mathbb{R}^d), \quad \mathrm{Ent}_\mu\big(f^2\big) \leq c\mathbb{E}_\mu(|\nabla f|^2),$$

alors, pour toute fonction lipschitzienne F telle que $\|F\|_{Lip} \leq 1$,

$$\mu(F \geq \mathbb{E}_\mu(F) + r) \leq e^{-r^2/c} \quad \textit{et} \quad \mu(|F - \mathbb{E}_\mu(F)| \geq r) \leq 2e^{-r^2/c}.$$

Preuve. Comme dans la preuve précédente, on suppose dans un premier temps que F est de classe $cC^i nfty$ bornée telle que $\|F\|_{_{Lip}} \leq 1$.

Soit $H(\lambda) \stackrel{\text{déf.}}{=} \mathbb{E}_\mu(e^{\lambda F})$ la transformée de Laplace de F. L'inégalité de Sobolev logarithmique appliquée à $f^2 = e^{\lambda F}$ entraîne l'inégalité suivante :

$$\int \lambda F e^{\lambda F}\, d\mu - H(\lambda)\log H(\lambda) = \lambda H'(\lambda) - H(\lambda)\log H(\lambda) \leq \frac{c\lambda^2}{4}\int e^{\lambda F}|\nabla F|^2\, d\mu = \frac{c\lambda^2}{4} H(\lambda), \tag{5.25}$$

puisque

$$\left|\nabla e^{\lambda F/2}\right|^2 = \frac{\lambda^2}{4} e^{\lambda F}|\nabla F|^2 \leq \frac{\lambda^2}{4} e^{\lambda F}.$$

De plus, H est une fonction strictement positive donc (5.25) s'écrit encore

$$\frac{H'(\lambda)}{\lambda H(\lambda)} - \frac{\log H(\lambda)}{\lambda^2} \leq \frac{c}{4}.$$

En posant $K(\lambda) \stackrel{\text{déf.}}{=} (1/\lambda)\log H(\lambda)$ pour $\lambda > 0$, cela est équivalent à $K'(\lambda) \leq c/4$. Un développement à l'ordre un nous assure de la convergence de $K(\lambda)$ vers $\mathbb{E}_\mu(F)$, quand λ tend vers 0. Pour $\lambda > 0$, comme K est continue sur $[0, \lambda]$ et dérivable sur $]0, \lambda[$,

$$K(\lambda) - K(0) = \int_0^\lambda K'(s)\, ds \leq \frac{c\lambda}{4},$$

ce qui correspond à l'estimation sous-gaussienne de la transformée de Laplace :

$$H(\lambda) \leq e^{\lambda \mathbb{E}_\mu(F) + (c/4)\lambda^2}.$$

L'inégalité de Markov permet aisément d'en déduire l'inégalité de déviation. L'extension à toutes les fonctions lipschitziennes s'effectue par régularisation et troncature. □

5.8.3 Concentration pour la mesure exponentielle

Dans toute la section la mesure μ désignera la mesure exponentielle (symétrique) sur $\mathbb{R}$, de densité $(1/2)e^{-|x|}$. Il paraît clair qu'il est totalement illusoire d'espérer obtenir pour μ des inégalités de Sobolev logarithmique et de concentration gaussienne. Nous allons toutefois montrer que μ vérifie une inégalité de Poincaré et une inégalité de Sobolev logarithmique modifiée dont nous déduirons une propriété de concentration appropriée.

Exercice 5.8.4. *Soit ϕ régulière sur $\mathbb{R}$. Montrer que*

$$\int \phi\, d\mu = \phi(0) + \int \operatorname{sgn}(x)\phi'(x)\, \mu(dx). \tag{5.26}$$

Lemme 5.8.5. *La mesure exponentielle (double) sur $\mathbb{R}$ vérifie une inégalité de Poincaré de constante 4 : pour toute fonction f régulière sur $\mathbb{R}$,*

$$\operatorname{Var}_\mu(f) \leq 4\int f'^2 d\mu.$$

Preuve. Soit $g(x) = f(x) - f(0)$. Alors, d'après (5.26) et l'inégalité de Cauchy-Schwarz,

$$\int g^2\,d\mu = 2\int \mathrm{sgn}(x)g'(x)g(x)\,\mu(dx) \le 2\Big(\int g'^2\,d\mu\Big)^{1/2}\Big(\int g^2\,d\mu\Big)^{1/2}.$$

Il ne reste plus qu'à remarquer que $\mathrm{Var}_\mu(f) \le \int g^2\,d\mu$ et $g' = f'$. □

Théorème 5.8.6. *Pour tout $0 < \rho < 1$ et toute fonction lipschitzienne f sur $\mathbb{R}$ telle que $|f'| \le \rho$ presque partout,*

$$\mathrm{Ent}_\mu(e^f) \le \frac{2}{1-\rho}\int f'^2 e^f\,d\mu.$$

Preuve. Quitte à changer f en $f + a$, on peut supposer que $f(0) = 0$. Comme, pour tout $u \ge 0$, $-u\log u \le -u + 1$, nous avons

$$\mathrm{Ent}_\mu(f) = \int fe^f\,d\mu - \int e^f\,d\mu \log\int e^f\,d\mu \le \int (fe^f - e^f + 1)\,d\mu.$$

La relation (5.26) assure

$$\int (fe^f - e^f + 1)\,d\mu = \int \mathrm{sgn}(x)f'(x)f(x)e^{f(x)}\,\mu(dx).$$

Donc, en utilisant l'inégalité de Cauchy-Schwarz, il vient

$$\mathrm{Ent}_\mu(e^f) \le \Big(\int f'^2e^f\,d\mu\Big)^{1/2}\Big(\int f^2e^f\,d\mu\Big)^{1/2}$$

Encore par (5.26), nous pouvons écrire

$$\int f^2e^f\,d\mu = 2\int \mathrm{sgn}(x)f'(x)f(x)e^{f(x)}\,\mu(dx) + \int \mathrm{sgn}(x)f'(x)f(x)^2e^{f(x)}\,\mu(dx).$$

D'après l'inégalité de Cauchy-Schwarz et le contrôle $|f'| \le \rho$, nous obtenons donc

$$\int f^2e^f\,d\mu \le 2\Big(\int f'^2e^f\,d\mu\Big)^{1/2}\Big(\int f^2e^f\,d\mu\Big)^{1/2} + \rho\int f^2e^f\,d\mu,$$

ou encore,

$$\int f^2e^f\,d\mu \le \Big(\frac{2}{1-\rho}\Big)^2\int f'^2e^f\,d\mu.$$

On a donc montré que

$$\mathrm{Ent}_\mu(e^f) \le \frac{2}{1-\rho}\int f'^2e^f\,d\mu,$$

ce qui achève la preuve. □

Théorème 5.8.7. *Soit F de $\mathbb{R}^n$ dans $\mathbb{R}$ telle que*

$$\max_{1\le i\le n}|\partial_i F| \le 1 \quad \textit{et} \quad \sum_{i=1}^n (\partial_i F)^2 \le a^2$$

presque partout. Alors, pour tout $r \ge 0$,

$$\mu^n\Big(\Big\{F \ge \int F\,d\mu^n + r\Big\}\Big) \le \exp\Big(-\frac{1}{4}\min\Big(r, \frac{r^2}{4a^2}\Big)\Big).$$

Preuve. Pour toute fonction F sur $\mathbb{R}^n$ telle que $\max_{1\leq i\leq n}|\partial_i F| \leq 1$ presque partout et tout λ tel que, $|\lambda| \leq \rho < 1$,

$$\mathrm{Ent}_{\mu^n}\left(e^{\lambda F}\right) \leq \frac{2\lambda^2}{1-\rho}\int\sum_{i=1}^n (\partial_i F)^2 e^{\lambda F}\, d\mu^n. \tag{5.27}$$

Fixons $\rho = 1/2$ dans un premier temps. Supposons de plus que $\sum_{i=1}^n(\partial_i F)^2 \leq a^2$ presque partout. Alors, d'après (5.27),

$$\mathrm{Ent}_{\mu^n}\left(e^{\lambda F}\right) \leq 4a^2\lambda^2\int e^{\lambda F}\, d\mu^n,$$

pour tout $|\lambda| \leq 1/2$. Soit H définie sur $[-1/2, 1/2]$ par

$$H(\lambda) = \int e^{\lambda F - 4a^2\lambda^2}\, d\mu^n.$$

L'inégalité précédente peut se reformuler ainsi : pour $\lambda \in [-1/2, 1/2]$,

$$\lambda H'(\lambda) \leq H(\lambda)\log H(\lambda).$$

On obtient alors, en intégrant cette relation que, pour $\lambda \in [-1/2, 1/2]$,

$$\int e^{\lambda F}\, d\mu^n \leq \exp\left(\lambda\int F\, d\mu^n + 4a^2\lambda^2\right).$$

L'inégalité de Markov exponentielle fournit l'inégalité de déviation habituelle :

$$\mu^n\left(\left\{F \geq \int F\, d\mu^n + r\right\}\right) \leq e^{-\lambda r + 4a^2\lambda^2},$$

que l'on minimise en $\lambda \in [-1/2, 1/2]$. On prendra ainsi $\lambda = r/8a^2$ si $r \leq 4a^2$ et $\lambda = 1/2$ si $r \geq 4a^2$. Ceci assure alors que, pour tout $r \geq 0$,

$$\mu^n\left(\left\{F \geq \int F\, d\mu^n + r\right\}\right) \leq \exp\left(-\frac{1}{4}\min\left(r, \frac{r^2}{4a^2}\right)\right).$$

□

Corollaire 5.8.8. *Soit f de $\mathbb{R}$ dans $\mathbb{R}$ telle que $|f'| \leq 1$. Soient $X_1, \ldots, X_n$ des v.a. i.i.d. de loi μ. Alors, pour tout $r \geq 0$,*

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n f(X_i) - \mathbb{E}f(X)\right| \geq r\right) \leq 2\exp\left(-\frac{n}{4}\min(r, r^2/4)\right).$$

Remarque 5.8.9. *À r fixé, on reconnaît dans cette expression la queue exponentielle pour les petites valeurs de n et la queue gaussienne (qui vient du théorème limite central) pour des grandes valeurs de n.*

L'exemple de la mesure exponentielle est en fait généralisable à toutes les mesures vérifiant une inégalité de Poincaré.

Théorème 5.8.10. *Soit μ une mesure de probabilité sur $\mathbb{R}^d$ qui vérifie l'inégalité de Poincaré de constante c : pour toute fonction f localement lipschitzienne,*

$$\mathrm{Var}_\mu(f) \leq c \int |\nabla f|^2 \, d\mu.$$

Alors pour toute fonction f telle que $\|\nabla f\|_\infty \leq \rho < 2/\sqrt{c}$,

$$\mathrm{Ent}_\mu(e^f) \leq \frac{c}{2}\left(\frac{2+\rho\sqrt{c}}{2-\rho\sqrt{c}}\right)^2 e^{\rho\sqrt{5c}} \int |\nabla f|^2 e^f \, d\mu.$$

Corollaire 5.8.11. *Soit F de $(\mathbb{R}^d)^n$ dans $\mathbb{R}$ telle que*

$$\max_{1\leq i\leq n} |\nabla_i F| \leq b \quad et \quad \sum_{i=1}^n (\nabla_i F)^2 \leq a^2$$

presque partout. Alors, pour tout $r \geq 0$,

$$\mu^n\left(\left\{F \geq \int F \, d\mu^n + r\right\}\right) \leq \exp\left(-\frac{1}{K}\min\left(\frac{r}{b}, \frac{r^2}{4a^2}\right)\right),$$

où K s'exprime en fonction de la constante de Poincaré c.

5.9 Inégalités de transport

5.9.1 Distances de transport

Soit (E,d) un espace métrique muni de la tribu borélienne $\mathcal{F}$. Notons $\mathcal{P}_0$ l'ensemble des mesures de probabilité sur $(E,\mathcal{F})$. Pour tout $k \geq 1$, soit $\mathcal{P}_k$ l'ensemble des mesures de probabilité μ sur $(E,\mathcal{F})$ qui admettent un moment d'ordre k, *i.e.* telles que, pour tout $y \in E$,

$$\int d(x,y)^k \, \mu(dx) < \infty.$$

Remarque 5.9.1. *La mesure $\mu \in \mathcal{P}_0$ appartient à $\mathcal{P}_k$ si et seulement si, il existe $z \in E$ tel que*

$$\int d(x,z)^k \, \mu(dx) < \infty.$$

Définition 5.9.2. *Soit μ et ν deux éléments de $\mathcal{P}_0$ et soit $P(\mu,\nu)$ l'ensemble des mesures de probabilité sur $E \times E$ admettant comme marges μ et ν. On définit l'application T_0 sur $\mathcal{P}_0 \times \mathcal{P}_0$ par*

$$T_0(\mu,\nu) = \inf\left\{\int \mathbf{1}_{x\neq y} \, \pi(dx,dy) \; ; \; \pi \in P(\mu,\nu)\right\}.$$

Pour $k \geq 1$, on définit, de plus, l'application T_k sur $\mathcal{P}_k \times \mathcal{P}_k$ par

$$T_k(\mu,\nu) = \left(\inf\left\{\int \frac{d(x,y)^k}{k} \, \pi(dx,dy) \; ; \; \pi \in P(\mu,\nu)\right\}\right)^{1/k}.$$

Proposition 5.9.3. *Pour $k = 0$ ou $k \geq 1$, l'application T_k est une distance sur $\mathcal{P}_k$.*

La distance T_0 coïncide, à une constante près à la notion de variation totale. Soit μ une mesure signée sur $(E, \mathcal{F})$, la norme en variation totale de μ est définie par :

$$\|\mu\|_{VT} = \sup\left\{\int f\,d\mu,\ f \text{ mesurable bornée par } 1\right\}.$$

Proposition 5.9.4. *Pour toutes mesures μ et ν de $\mathcal{P}_0$,*

$$T_0(\mu,\nu) = \frac{1}{2}\|\mu-\nu\|_{VT}.$$

Preuve. Montrons une inégalité. Soit μ et ν dans $\mathcal{P}_0$. Par définition,

$$T_0(\mu,\nu) = \inf\{\mathbb{P}(X \neq Y)\},$$

où l'infimum est pris sur les variables aléatoires X et Y à valeurs dans E de lois respectives μ et ν. Soit f mesurable et bornée par 1 et $\pi \in P(\mu,\nu)$,

$$\begin{aligned}\int f\,d\mu - \int f\,d\nu &= \int (f(x)-f(y))\,\pi(dx,dy) = \int (f(x)-f(y))\mathbf{1}_{x\neq y}\,\pi(dx,dy)\\ &\leq 2\int \mathbf{1}_{x\neq y}\,\pi(dx,dy) \leq 2T_0(\mu,\nu).\end{aligned}$$

Donc $\|\mu-\nu\|_{TV} \leq 2T_0(\mu,\nu)$. □

Exercice 5.9.5. *Montrer que si ν admet f pour densité par rapport à μ alors*

$$\|\mu-\nu\|_{VT} = \int |f-1|d\mu.$$

Dualité de Monge-Kantorovitch

Il est possible d'obtenir une formulation duale analogue à celle de la variation totale pour les distances T_k, $k \geq 1$.

Théorème 5.9.6. *On a la représentation duale suivante des distance de transport :*

$$\inf \iint T(x,y)\pi(dx,dy) = \sup\left[\int g\,d\nu - \int f\,d\mu\right],$$

où l'infimum est pris sur les mesures de probabilité sur $E\times E$ muni de la tribu produit tandis que le supremum est pris sur les fonctions mesurables bornées telles que pour tous x et y,

$$g(x) \leq f(y) + T(x,y).$$

En particulier,

$$T_1(\mu,\nu) = \sup\left[\int f\,d\nu - \int f\,d\mu\right],$$

où le supremum est pris sur les fonctions f 1-lipschitziennes

De même

$$T_2(\mu,\nu)^2 = \sup\left[\int g\,d\nu - \int f\,d\mu\right],$$

pour f et g telles que, pour tous x, y,

$$g(x) \leq f(y) + \frac{1}{2}|x-y|^2.$$

Remarquons que le choix optimal pour g s'écrit

$$g(x) = \inf_{y \in \mathbb{R}^d} \left[f(y) + \frac{1}{2}|x-y|^2 \right].$$

5.9.2 Inégalité de Csiszár-Kullback

Il s'agit de montrer que, de manière générique, la distance T_0 est contrôlée par l'entropie relative.

Théorème 5.9.7. *Soit μ et ν dans $\mathcal{P}_0$ tels que ν admet la densité f par rapport à μ.*

$$\|\nu - \mu\|_{VT} \leq \sqrt{2\mathrm{Ent}(\nu \,|\, \mu)} = \sqrt{2\mathrm{Ent}_\mu(f)}.$$

On dira alors que la mesure μ vérifie l'inégalité de transport T_0 de constante 2.

Preuve. La démonstration repose sur une inégalité astucieuse due à Pinsker : pour tout $u \geq 0$,

$$3(u-1)^2 \leq (2u+4)(u \log u - u + 1).$$

D'après les inégalités de Cauchy-Schwarz et de Pinsker,

$$\begin{aligned}
\int |f-1| \, d\mu &\leq \int \sqrt{\frac{2f+4}{3}} \sqrt{f \log f - f + 1} \, d\mu \\
&\leq \sqrt{\int \frac{2f+4}{3} \, d\mu} \sqrt{\int f \log f \, d\mu} \\
&\leq \sqrt{2 \int f \log f \, d\mu} = \sqrt{2\mathrm{Ent}_\mu(f)}.
\end{aligned}$$

D'après l'exercice 5.9.5, le résultat s'ensuit. □

5.9.3 Inégalité de transport pour le coût linéaire

Il est possible de caractériser les mesures μ qui vérifie une inégalité de transport T_1 : ce sont les mesures pour lesquelles la transformée de Laplace de toute fonction lipschitzienne est sous-gaussienne.

Théorème 5.9.8. *La mesure μ vérifie l'inégalité de transport T_1 de constante $c > 0$ si*

$$T_1(\mu, \nu) \leq \sqrt{\mathrm{cEnt}_\mu\left(\frac{d\nu}{d\mu}\right)}, \qquad \text{IT1(c)}$$

si et seulement si, pour tout $\lambda > 0$,

$$\int e^{\lambda f} \, d\mu \leq e^{\lambda \int f \, d\mu + c\lambda^2/4}. \qquad \text{IE(c/2)}$$

Preuve. Notons φ la densité de ν par rapport à μ. L'inégalité de transport IT1(c) assure que pour toute fonction f lipschitzienne telle que $\|f\|_L \leq 1$,

$$\int f\,d\nu - \int f\,d\mu \leq \sqrt{c\mathrm{Ent}_\mu(\varphi)},$$

où, de manière équivalente, pour tout $\lambda > 0$,

$$\int f\,d\nu - \int f\,d\mu \leq \frac{c\lambda}{4} + \frac{1}{\lambda}\mathrm{Ent}_\mu(\varphi).$$

L'inégalité précédente peut encore s'écrire

$$\int \psi\varphi\,d\mu \leq \mathrm{Ent}_\mu(\varphi),$$

où l'on a noté

$$\psi = \lambda f - \lambda \int f\,d\mu - \frac{c\lambda^2}{4}.$$

En choisissant alors φ égale à $e^\psi / \int e^\psi\,d\mu$, il vient que $\log \int e^\psi\,d\mu \leq 0$ ce qui n'est rien d'autre que IE(c/2).

La formule variationnelle de l'entropie (5.5) assure que la réciproque est vraie. □

Corollaire 5.9.9. *Si la mesure de probabilité μ sur $\mathbb{R}^d$ vérifie une inégalité de Sobolev logarithmique*

$$\mathrm{Ent}_\mu(f^2) \leq c \int |\nabla f|^2\,d\mu,$$

alors elle vérifie l'inégalité de transport IT1(c).

Preuve. D'après l'argument de Herbst (théorème 5.8.3), l'inégalité de Sobolev logarithmique de constante c implique l'intégrabilité exponentielle IC(c/2) et donc l'inégalité de transport IT1(c). □

Corollaire 5.9.10. *La mesure de probabilité μ sur $\mathbb{R}^d$ vérifie l'inégalité de transport IT1(c) si et seulement si elle vérifie la propriété de concentration gaussienne pour les fonctions lipschitziennes : si F est de (semi-)norme de Lipschitz inférieure à 1,*

$$\mu(|F - \mathbb{E}_\mu(F)| \geq r) \leq 2e^{-r^2/c}.$$

Preuve. L'inégalité IT1(c) implique IE(c/2) qui implique la concentration gaussienne des fonctions lipschitziennes. Réciproquement, on applique la formule

$$\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X \geq t)\,dt$$

à e^F pour obtenir IE(c/2). □

5.9.4 Inégalité de transport pour le coût quadratique

On souhaite étudier à présent l'inégalité de transport IT2. Il faut pour cela avoir recours à quelques notions de la théorie des équations aux dérivées partielles. L'équation aux dérivées partielles non linéaire suivante, dite de Hamilton-Jacobi, de condition initiale f supposée lipschitzienne bornée,

$$\begin{cases} \partial_t v(t,x) + \frac{1}{2}|\nabla v(t,x)|^2 = 0 & \text{si } (t,x) \in]0,\infty[\times\mathbb{R}^d, \\ v(0,x) = f(x) \text{ si } x \in \mathbb{R}^d, \end{cases} \tag{5.28}$$

admet pour solution la fonction Q définie sur $]0,\infty[\times\mathbb{R}^d$ par

$$Q_t f(x) = \inf_{y\in\mathbb{R}^d} \left[f(y) + \frac{1}{2t}|x-y|^2 \right].$$

La famille $(Q_t)_t$ définit un semi-groupe de générateur infinitésimal (non-linéaire) $-(1/2)|\nabla f|^2$.

Rappelons que $Q_1 f$, que nous noterons Qf, apparaît comme le choix optimal de la fonction g dans la version duale de la distance T_2, c'est-à-dire que

$$T_2(\mu,\nu)^2 = \sup_f \left\{ \int Qf\, d\nu - \int f\, d\mu \right\}.$$

Hypercontractivité des solutions de Hamilton-Jacobi

Il s'agit ici de relier l'inégalité de Sobolev logarithmique à une propriété d'hypercontractivité du semi-groupe de Hamilton-Jacobi dans l'esprit du théorème de Gross 5.4.2.

Théorème 5.9.11. *Soit μ une mesure de probabilité sur $\mathbb{R}^d$ absolument continue par rapport à la mesure de Lebesgue. S'il existe c tel que pour toute fonction régulière f sur $\mathbb{R}^d$,*

$$\mathrm{Ent}_\mu(f^2) \le c \int |\nabla f|^2\, d\mu, \tag{5.29}$$

alors, pour toute fonction mesurable bornée f sur $\mathbb{R}^d$, tout $t \ge 0$ et tout $a \in \mathbb{R}$,

$$\left\| e^{Q_t f} \right\|_{a+2t/c} \le \left\| e^f \right\|_a. \tag{5.30}$$

Réciproquement, si (5.30) est vérifiée pour tout $t \ge 0$ et un réel a non nul, alors μ vérifie l'inégalité de Sobolev logarithmique (5.29).

Remarque 5.9.12. *L'application $a \mapsto \left\| e^f \right\|_a$ est bien définie pour tout $a \in \mathbb{R}_*$ et se prolonge par continuité en 0 par $e^{\int f\, d\mu}$.*

Preuve. Soit F la fonction sur $\mathbb{R}_+$ par

$$F(t) = \left\| e^{Q_t f} \right\|_{\lambda(t)},$$

où $\lambda(t) = a + 2t/c$. La fonction F est dérivable en tout point où $\lambda(t) \ne 0$ et l'on a alors :

$$\lambda(t)^2 F(t)^{\lambda(t)-1} F'(t) = \frac{2}{c}\mathrm{Ent}_\mu\left(e^{\lambda(t)Q_t f}\right) + \int \lambda(t)^2 \partial_t Q_t f e^{\lambda(t)Q_t f}\, d\mu.$$

Puisque

$$\partial_t Q_t f(x) = -\frac{1}{2}|\nabla Q_t f(x)|^2,$$

on obtient donc,

$$\lambda(t)^2 F(t)^{\lambda(t)-1} F'(t) = \frac{2}{c}\mathrm{Ent}_\mu\big(e^{\lambda(t)Q_t f}\big) - \int \frac{\lambda(t)^2}{2}|\nabla Q_t f|^2 e^{\lambda(t)Q_t f}\, d\mu.$$

L'inégalité (5.29) appliquée à $e^{(\lambda(t)/2)Q_t f}$ assure que $F'(t) \leq 0$ pour tout $t > 0$ (sauf peut-être en un point lorsque $a < 0$). La fonction F est décroissante, ce qui fournit l'hypercontractivité.

Réciproquement, l'hypercontractivité assure que $F'(0) \leq 0$, ce qui implique

$$\frac{2}{c}\mathrm{Ent}_\mu\big(e^{af}\big) \leq \frac{1}{2}\int |a\nabla f|^2 e^{af}\, d\mu.$$

Puisque a est supposé non nul, l'inégalité précédente est bien (5.29). □

Inégalité de Sobolev logarithmique et inégalité IT2

Théorème 5.9.13. *Supposons que μ soit absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. Si μ vérifie l'inégalité*

$$\mathrm{Ent}_\mu\big(f^2\big) \leq c\int |\nabla f|^2\, d\mu,$$

toute fonction régulière f sur $\mathbb{R}^d$, alors pour toute mesure de probabilité ν absolument continue par rapport à μ,

$$T_2(\mu,\nu) \leq \sqrt{c\mathrm{Ent}(\nu\,|\,\mu)}.$$

Preuve. Soit g une fonction lipschitzienne bornée sur $\mathbb{R}^d$. Puisque, pour $\lambda > 0$, $Q(\lambda g) = \lambda Q_\lambda g$, on a

$$Q(\lambda g) = \lambda \partial_\lambda Q(\lambda g) + \frac{1}{2}|\nabla Q(\lambda g)|^2.$$

Appliquons l'inégalité de Sobolev logarithmique (5.29) à la fonction $f = e^{(1/c)Q(\lambda g)}$. La fonction $G(\lambda) = \int e^{(2/c)Q(\lambda g)}\, d\mu$ vérifie donc l'inéquation différentielle

$$\lambda G'(\lambda) \leq G(\lambda)\log G(\lambda).$$

Puisque $G'(0) = (1/c)\int g\, d\mu$, il vient que

$$\int e^{(1/c)Qg}\, d\mu \leq e^{(1/c)\int g\, d\mu}.$$

Cette inégalité est équivalente à une inégalité de transport IT2. En effet, en vertu du théorème 5.9.6 (on note φ la densité de ν par rapport à μ),

$$T_2(\mu,\nu)^2 \leq c\mathrm{Ent}_\mu(\varphi)$$

est équivalent à

$$\sup_f \left[\int Qf\, d\nu - \int f\, d\mu\right] \leq c\mathrm{Ent}_\mu(\varphi)$$

puisque Qf est le choix optimal pour g. De l'inégalité précédente, on déduit que

$$\int \psi\varphi \, d\mu \leq c\mathrm{Ent}_\mu(\varphi),$$

où $\psi = Qf - \int f \, d\mu$, qui est valable pour tout φ, ce qui est équivalent à dire que $\int e^{(1/c)\psi} \, d\mu \leq 1$.
$\square$

Chapitre 6

Exercices et devoirs

6.1 Le processus d'Ornstein-Uhlenbeck

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soit $(B_t)_{t\geq 0}$ un mouvement brownien dans $\mathbb{R}$. On note, pour tout $s \geq 0$, $\mathcal{F}_s = \sigma(B_u,\ 0 \leq u \leq s)$ et $\mathcal{F} = (\mathcal{F}_s)_{s\geq 0}$ la filtration engendrée par le mouvement brownien. Soit $\sigma > 0$ et $\beta \in \mathbb{R}$. On considère l'équation différentielle stochastique suivante (dite de Langevin) :

$$dV_t = -\beta V_t + \sigma dB_t, \tag{6.1}$$

avec condition initiale V_0, variable aléatoire réelle de carré intégrable indépendante de $\mathcal{F}$.

1. En quels sens a-t-on existence et unicité de la solution de l'équation (6.1) ?
2. Calculer le générateur infinitésimal de $(V_t)_{t\geq 0}$. Montrer que le processus $(V_t)_{t\geq 0}$ admet une mesure invariante gaussienne que l'on calculera. Est-elle réversible ?
3. Pour tout $t \geq 0$ on définit
$$M_t = \int_0^t e^{\beta s} dB_s.$$
Montrer que $(M_t)_{t\geq 0}$ est une $\mathcal{F}$-martingale dont on précisera la variation quadratique.
4. Résoudre l'équation (6.1). On pourra pour cela écrire la formule d'intégration par partie pour $d(e^{\beta t} V_t)$.
5. On suppose que V_0 est une v.a.r. gaussienne de loi $\mathcal{N}(m, \alpha^2)$ indépendante de $\mathcal{F}$. Montrer que V_t est encore de loi normale et préciser sa moyenne et sa variance. Que se passe-t-il si $V_0 = x$ p.s. ? À quelle(s) condition(s) sur les coefficients β et σ, le processus (V_t) converge-t-il en loi quand $t \to \infty$? Quelle est sa limite ? Faire le lien avec la question 2.
6. Supposons à présent que $m = 0$. Calculer, pour $0 \leq s \leq t$, $\mathrm{Cov}(V_s, V_t)$. Quelle condition faut-il imposer sur V_0 pour que $\mathrm{Cov}(V_s, V_t)$ ne dépende que de $t - s$? Que dire alors de la loi de V_t pour tout $t \geq 0$?
7. On introduit à présent le processus $(X_t)_{t\geq 0}$ en posant
$$X_t = X_0 + \int_0^t V_s ds,$$
où X_0 est une v.a. gaussienne (éventuellement dégénérée) indépendante de V_0 et de $\mathcal{F}$. Montrer que pour tous $0 \leq s \leq t$,
$$X_t = X_s + \frac{\sigma}{\beta} \int_s^t \left[1 - e^{-\beta(t-u)}\right] dB_u$$

En déduire que X_t est une v.a. gaussienne dont on précisera les paramètres. Calculer $\mathrm{Cov}(X_s, X_t)$. En déduire que pour tous $0 = t_0 < t_1 < \cdots < t_n$, le vecteur $(X_{t_0}, \ldots, X_{t_n})$ est gaussien. Donner sa matrice de covariance.

8. Montrer que si $\beta \to +\infty$ et $\beta/\sigma \to 1$ alors pour tous $0 = t_0 < t_1 < \cdots < t_n$, le vecteur $(X_{t_0}, \ldots, X_{t_n})$ converge en loi vers le vecteur $(B_{t_0}, \ldots, B_{t_n})$.
9. Montrer que la densité $f(t, x, v)$ de la loi de (X_t, V_t) est solution faible de l'équation aux dérivées partielles
$$\partial_t f + v\partial_x f = \frac{\sigma^2}{2}\partial^2_{vv} f + \beta\partial_v(vf).$$
Cette équation appartient à la classe des équations cinétiques (où intervient la vitesse). Elle décrit l'évolution de la densité de particules ayant la vitesse v au point x à l'instant t. La dynamique des particules est donnée par le principe fondamental de la dynamique : l'accélération est égale à la somme des forces qui s'exercent sur la particule : en l'occurence des forces aléatoires dues aux chocs avec les particules du milieu (d'ou le brownien et le laplacien) et un terme de friction (force qui s'oppose à la vitesse comme le vent que l'on a toujours dans le nez à vélo).

6.2 Semi-groupe de Markov

6.2.1 Exemples de semi-groupes

Exercice 4. *Exemples classiques*
Montrer que les familles de noyaux suivantes sont des fonctions de transition homogènes. Calculer pour chaque exemple le générateur infinitésimal, la (ou les) mesures invariantes (sont-elles de masse fini, réversibles,...) et l'opérateur carré du champ. Montrer que ce sont des semi-groupe de Feller.

1. Translation à vitesse constante : $E = \mathbb{R}$, $\mathcal{E} = \mathcal{B}(\mathbb{R})$. soit $v \in \mathbb{R}$,
$$P_t(x, \cdot) = \delta_{x+vt}.$$
2. Processus de Poisson. $E = \mathbb{R}$, $\mathcal{E} = \mathcal{B}(\mathbb{R})$,
$$P_t(x, dy) = \sum_{n=0}^{\infty} e^{-t}\frac{t^n}{n!}\delta_{x+n}(dy).$$
3. Processus de Poisson composé : soit π une mesure de probabilité sur un espace $(E, \mathcal{E})$. Montrer que l'on peut définir par récurrence une probabilité de transition π^n par
$$\pi^n(x, A) = \int_E \pi(x, dy)\pi^{n-1}(y, A).$$
En déduire que
$$P_t(x, dy) = \sum_{k=0}^{\infty} e^{-t}\frac{t^n}{n!}\pi^n(x, dy),$$
est une fonction de transition. Quelle est la dynamique associée ?
4. Mouvement brownien : $E = \mathbb{R}$, $\mathcal{E} = \mathcal{B}(\mathbb{R})$, $P_t(x, \cdot)$ est la mesure gaussienne $\mathcal{N}(x, t)$.

5. Mouvement brownien réfléchi : $E = \mathbb{R}^+$, $\mathcal{E} = \mathcal{B}(\mathbb{R}^+)$, $P_t(x, \cdot)$ est la mesure de densité sur $\mathbb{R}^+$,
$$\frac{1}{\sqrt{2\pi t}}\left[\exp\left(-\frac{1}{2t}(y-x)^2\right) + \exp\left(-\frac{1}{2t}(y+x)^2\right)\right].$$

Exercice 5. *Noyaux tordus*
Montrer que les deux familles de noyaux suivantes sont des fonctions de transition markoviennes sur $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$:

1. $P_t f(x) = e^{-t/x} f(x) + \displaystyle\int_x^\infty \frac{t}{y^2}\exp(-t/y) f(y)\, dy,$
2. $Q_t f(x) = \dfrac{x}{x+t} f(x+t) + \displaystyle\int_x^\infty \frac{t}{(t+y)^2} f(t+y)\, dy.$

Exercice 6. *Semi-groupes de convolution*
Soit $(\mu_t)_{t\geq 0}$ une famille de mesures de probabilité sur $\mathbb{R}^d$ telle que

(i) pour tous $s, t \geq 0$, $\mu_s * \mu_t = \mu_{s+t}$,

(ii) $\mu_0 = \delta_0$ et $\lim_{t\to 0} \mu_t = \delta_0$ au sens

On pose alors pour $t \geq 0$,
$$P_t(x, A) = \int_{\mathbb{R}^d} \mathbf{1}_A(x+y)\, \mu_t(dy).$$
Soit X un processus associé à un semigroupe de convolution.

1. Montrer que P_t est un semi-groupe de Feller.
2. Montrer les accroissements de X sont indépendants.
3. Montrer que la loi de l'incrément $X_t - X_s$ est μ_{t-s} (on dit que le processus est à accroissements stationnaires).

Un processus à accroissements indépendants et stationnaires est appelé processus de Lévy.

4. Donner des exemples de processus de Lévy.

6.2.2 Chaînes de Markov à temps continu

Exercice 7. *L'espace à deux points*
Considérons l'espace à deux points $E = \{-1, 1\}$ que nous munissons de la mesure de Bernoulli uniforme $\beta \overset{\text{déf.}}{=} (1/2)(\delta_{-1} + \delta_1)$. Toute fonction $f : E \to \mathbb{R}$ peut s'écrire $f(x) = a + bx$ avec a et b réels. On introduit alors la famille $(P_t)_t$ d'opérateur définis par
$$P_t f(x) \overset{\text{déf.}}{=} \int_E f(y)(1 + e^{-t}xy)\, \beta(dy) = a + e^{-t}bx.$$

1. Montrer que le générateur infinitésimal de P_t vaut $Af(x) = -bx$.
2. On représente les fonctions par des vecteurs de $\mathbb{R}^2$ et les opérateurs par des matrices. Que vaut A ? Quelles sont ses valeurs propres et les vecteurs propres associés ?
3. Que vaut e^{tA} ? Quelles sont ses valeurs propres et les vecteurs propres associés ?
4. Vérifier la propriété de semi-groupe de P_t.
5. Le semi-groupe admet-il une mesure invariante ? symétrique ?

6. Que vaut l'opérateur carré du champ ?

Exercice 8. *Un exemple numérique*
Soit X le processus à temps continu de générateur

$$A = \begin{pmatrix} -2 & 1 & 1 \\ 4 & -4 & 0 \\ 2 & 1 & -3 \end{pmatrix}$$

1. Calculer P_t.
2. Quelle est la probabilité invariante μ de X ?
3. Qu'observe-t-on ?

6.2.3 Diffusion sur un intervalle

Exercice 9. *Processus de Bessel*
Soit $d \geq 2$ et B un mouvement brownien sur $\mathbb{R}^d$. On définit le processus R comme la norme de B, c'est-à-dire : pour tout $t \geq 0$,

$$R_t := |B_t| = \sqrt{(B_t^{(1)})^2 + \cdots + (B_t^{(d)})^2}.$$

Si B est issu de x alors R est issu de $r = |x|$.

1. Montrer que $\{0 \leq s \leq t,\ R_s = 0\}$ est de mesure nulle.
2. Montrer que

$$W := \sum_{i=1}^{d} W^{(i)} \quad \text{avec pour } 1 \leq i \leq d, \quad W_t^{(i)} := \int_0^t \frac{B_s^{(i)}}{R_s}\, dB_s^{(i)},$$

   est un mouvement brownien.
3. Montrer que R est solution de

$$R_t = r + \int_0^t \frac{d-1}{2R_s}\, ds + B_t.$$

   On pourra appliquer la formule d'Itô pour décrire l'équation satisfaite par $Y := R^2$, puis appliquer à nouveau la formule d'Itô à la fonction

$$g_\varepsilon(y) = \begin{cases} \frac{3}{8}\sqrt{\varepsilon} + \frac{3}{4\sqrt{\varepsilon}} y - \frac{1}{8\varepsilon\sqrt{\varepsilon}} y^2, & \text{si } y < \varepsilon \\ \sqrt{y}, & \text{si } y \geq \varepsilon, \end{cases}$$

   avant de passer à la limite $\varepsilon \to 0$.

Exercice 10. *Carrés de Bessel*
On s'intéresse au processus solution de

$$dX_t = \sigma dB_t + \frac{\alpha}{X_t} dt - \lambda X_t dt.$$

1. Pour quelles valeurs des paramètres σ, α et λ, n'y a-t-il pas explosion ?

2. Quelle est la mesure invariante de X ?
3. Soient $Y^{(1)}, \dots, Y^{(d)}$ les processus réels solutions de

$$dY_t^{(i)} = dB_t^{(i)} - \lambda Y_t^{(i)} dt,$$

   issus de 1 avec $\lambda > 0$ et les $(B^{(i)})_i$ des mouvements browniens réels indépendants. Quelle équation le processus Z défini par

$$Z_t = Y_t^{(1)} + \dots + Y_t^{(d)}$$

   vérifie-t-il ?
4. Quelle est la loi de Z_t ?
5. Montrer que la loi de Z_t converge vers sa mesure invariante.

Exercice 11. *Projection*
On considère à présent le processus solution de

$$dX_t = \sqrt{1 - X_t^2} dB_t - \lambda X_t dt.$$

1. Pour quelles valeurs du paramètre λ, n'y a-t-il pas explosion ?
2. Quelle est la mesure invariante de X ?
3. Lorsque $\lambda = n - 1$, peut-on interpréter géométriquement X (voir le titre de l'exercice) ?

Bibliographie

[ABC$^+$00] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto et G. Scheffer – *Sur les inégalités de Sobolev logarithmiques*, Panoramas et Synthèses [Panoramas and Syntheses], vol. 10, Société Mathématique de France, Paris, 2000, Avec une préface de D. Bakry et M. Ledoux.

[KS] I. Karatzas et S. Shreve – *Brownian motion and stochastic calculus*, Springer-Verlag.

[Nel67] E. Nelson – *Dynamical theories of Brownian motion*, Princeton University Press, Princeton, N.J., 1967.

[Nor98] J. R. Norris – *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 1998, Reprint of 1997 original.

[Roy99] G. Royer – *Une initiation aux inégalités de Sobolev logarithmiques*, Cours Spécialisés [Specialized Courses], vol. 5, Société Mathématique de France, Paris, 1999.

[RY91] D. Revuz et M. Yor – *Continuous martingales and brownian motion*, Springer-Verlag, 1991.