MICHEL ÉMERY
RÉMI LÉANDRE

Sur une formule de Bismut

Séminaire de probabilités (Strasbourg), tome 24 (1990), p. 448-452
<http://www.numdam.org/item?id=SPS_1990__24__448_0>

Université de Strasbourg
Séminaire de Probabilités

SUR UNE FORMULE DE BISMUT

par M. Emery et R. Léandre

Soit V une variété riemannienne compacte. Désignons par $r(dx)$ la mesure riemannienne normalisée sur V, par $P_t(x,dy) = p_t(x,y)r(dy)$ les probabilités de transition du mouvement brownien sur V et par Π^x la loi du pont brownien issu de x (c'est-à-dire du mouvement brownien sur V, issu de x et conditionné pour revenir en x à l'instant 1; on ne s'intéresse qu'à l'intervalle de temps $[0,1]$). Dans ses travaux sur le théorème de l'indice[1], Bismut munit l'espace des lacets (applications continues de $[0,1]$ dans V, qui prennent la même valeur en 0 et 1) de la probabilité

$$\mathbb{\Pi} = \frac{\int \Pi^x \, p_1(x,x)\, r(dx)}{\int p_1(x,x)\, r(dx)} ,$$

qui s'interprète comme la loi du mouvement brownien X conditionné par l'événement $\{X_1 = X_0\}$. (En réalité, dans le travail de Bismut, Π^x n'est pas exactement la loi du pont en x, mais cela ne change rien à la discussion qui suit). Pourquoi la «trace» $p_1(x,x)$ apparaît-elle, alors que la probabilité $\int \Pi^x \, r(dx)$ semblerait à première vue un choix parfaitement raisonnable? Nous allons tenter de l'expliquer à l'aide de quelques arguments heuristiques.

Remarquons tout d'abord que conditionner un processus de Markov Y par $\{Y_1 = Y_0\}$ est une opération encore moins évidente[2] que la construction des ponts, c'est-à-dire le conditionnement du processus issu de y par $\{Y_1 = y\}$. En effet, en notant $\mathbb{L}^\lambda(Y)$ la loi du processus sous une mesure initiale λ, l'égalité classique $\int \mathbb{L}^y(Y)\,\lambda(dy) = \mathbb{L}^\lambda(Y)$ n'entraîne pas la relation analogue entre lois conditionnelles : en général, si A est une partie de l'espace d'états,

$$\int \mathbb{L}^y[Y|Y_1 \in A]\,\lambda(dy) \neq \mathbb{L}^\lambda[Y|Y_1 \in A] .$$

1. Index Theorem and Equivariant Cohomology on the Loop Space, Comm. Math. Phys. 98 (1985).
2. Sur la difficulté de définir les ponts, voir L. Schwartz, Le mouvement brownien sur $\mathbb{R}^N$ en tant que semimartingale dans S_N, Ann. I.H.P. 21 (1985).

C'est pourquoi, alors que seules les probabilités de transition du processus interviennent dans la construction des ponts, il faut également faire intervenir la loi initiale pour conditionner par $\{Y_1 = Y_0\}$.

Si l'espace d'états de Y est fini (ce qui simplifie les notations et évite bien des difficultés techniques), en notant λ la loi initiale et $q_t(y,z)$ les probabilités de transition, la loi conditionnelle du processus sachant $\{Y_1 = Y_0\}$ est donnée pour $0 < t_1 < \ldots < t_n < 1$ par

$$\mathbb{L}^\lambda[Y_0 = y_0, \ldots, Y_{t_n} = y_{t_n} \mid Y_1 = Y_0]$$
$$= \frac{\lambda(y_0) q_{t_1}(y_0, y_{t_1}) q_{t_2 - t_1}(y_{t_1}, y_{t_2}) \cdots q_{t_n - t_{n-1}}(y_{t_{n-1}}, y_{t_n}) q_{1-t_n}(y_{t_n}, y_0)}{\sum_{y \in E} \lambda(y) q_1(y,y)}$$
$$= \frac{\lambda(y_0) q_1(y_0, y_0) \Pi^{y_0}[Y_0 = y_0, \ldots, Y_{t_n} = y_{t_n}]}{\sum \lambda(y) q_1(y,y)}$$

puisque Π^{y_0} est donné par la même formule où λ est remplacé par une masse unité en y_0. La trace $q_1(y,y)$ apparaît donc naturellement, avec une interprétation bien intuitive : elle vient modifier la mesure λ de manière à mettre plus de poids sur les positions initiales pour lesquelles la trajectoire a le plus de chances de se refermer en un lacet (non continu ici).

Ceci explique la formule de Bismut $\mathbb{\Pi} = \int \Pi^x \, p_1(x,x) \, r(dx) \,/ \int p_1(x,x) \, r(dx)$, mais en partie seulement, parce qu'il nous faut aussi choisir, non seulement une mesure initiale pour le processus (λ dans l'exemple ci-dessus), mais une mesure de référence servant à définir les densités $p_t(x,y)$ des probabilités de transition $P_t(x,dy)$ (la mesure de comptage dans l'exemple ci-dessus). Changer de mesure de référence multiplie en effet les $p_t(x,y)$ par une fonction positive $g(y)$ arbitraire et modifie donc $\mathbb{\Pi}$. Or, si le choix de la mesure riemannienne r paraît s'imposer comme mesure initiale pour le brownien (si V est connexe c'est l'unique loi invariante et elle est réversible), il est bien moins clair que c'est aussi elle que nous devons prendre comme référence; après tout, dans ces questions de lacets, la mesure $p_1(x,x) \, r(dx)$ apparaît, nous allons le voir, de façon parfaitement naturelle, alors pourquoi ne pas baser tous les calculs sur elle? Ou encore, pourquoi ne pas choisir comme référence la mesure $p_1(x,x)^{-1} r(dx)$, qui présente l'avantage de fournir la formule plus simple $\mathbb{\Pi} = \int \Pi^x r(dx)$?

Un premier argument consiste à remarquer qu'un changement de mesure de référence peut aussi s'interpréter comme un changement de mesure initiale, tous deux se traduisant par l'introduction d'une fonction $g(x)$ dans la formule. En particulier, la formule obtenue n'est pas modifiée si l'on multiplie les deux mesures (initiale et de référence) par la même fonction, et l'on obtiendrait le même résultat que Bismut en remplaçant r par une mesure équivalente quelconque, pourvu que les densités p_t soient calculées à l'aide de cette même mesure. En d'autres termes, la mesure $p_1(x,x) \, r(dx)$ est intrinsèquement liée au processus non conditionné, elle ne dépend d'aucun choix arbitraire. (On pourrait la noter $P_1(x,dx)$, mais cette notation est dangereuse : elle masque le fait que l'on a utilisé de bonnes versions des densités.)

Mais il y a aussi une autre raison, bien plus importante par ses conséquences : *Si V est connexe, le choix fait par Bismut est le seul pour lequel toutes les v. a. X_t ont même loi sous* $\mathbb{\Pi}$. *En outre, le processus* $(X_t)_{t\in\mathbb{R}/\mathbb{Z}}$ *est alors stationnaire et réversible.*

La stationnarité signifie que la loi du processus est invariante sous l'action du groupe des rotations $S^1 \simeq \mathbb{R}/\mathbb{Z}$; elle va résulter de l'égalité entre les mesures initiales et de référence et elle n'est absolument pas liée à l'invariance de la mesure riemannienne, mais seulement au caractère markovien homogène du brownien. La réversibilité signifie que la loi est invariante par le changement de t en $1-t$, et découlera de celle de la mesure riemannienne pour le brownien. Si V n'est pas connexe, on peut pondérer chaque composante à l'aide d'un facteur arbitraire, et on perd donc l'unicité.

Tout ceci reste probablement vrai pour des processus de Markov plus généraux, pourvu que le semi-groupe ait de bonnes versions et que les ponts existent; l'hypothèse de connexité sera bien sûr remplacée par une condition d'ergodicité.

Avant de vérifier l'assertion ci-dessus, remarquons que la loi des lacets vérifie aussi une autre propriété, la propriété de Markov circulaire : Si $a < b$ sont deux points de l'ensemble des temps $[0,1[$, ils découpent le cercle $\mathbb{R}/\mathbb{Z}$ en deux arcs connexes $[a,b]$ et $[b,1[\cup[0,a]$, et les comportements du processus sur ces deux arcs sont conditionnellement indépendants étant donné le couple (X_a, X_b). Mais cette propriété ne peut pas servir à caractériser la bonne loi parmi les autres, parce qu'elle est commune à toutes les probabilités de la forme $\mathbb{\Pi}^g = \int \Pi^x p_1(x,x)\, g(x)\, r(dx)$. En effet, puisque Π^x est donnée pour $0 < t_1 < \ldots < t_n < 1$ par l'intégrale

$$\Pi^x[(X_{t_1},\ldots,X_{t_n}) \in A] =$$

$$\int I_A(x_1,\ldots,x_n) P_{t_1}(x,dx_1) P_{t_2-t_1}(x_1,dx_2)\ldots P_{t_n-t_{n-1}}(x_{n-1},dx_n)\frac{p_{1-t_n}(x_n,x)}{p_1(x,x)}\,,$$

on obtient

$$\begin{aligned}\mathbb{\Pi}^g&[(X_0,\ldots,X_{t_n}) \in A]\\ &= \int g(x)\, I_A(x,x_1,\ldots,x_n)\, P_{t_1}(x,dx_1)P_{t_2-t_1}(x_1,dx_2)\ldots\\ &\qquad\qquad \ldots P_{t_n-t_{n-1}}(x_{n-1},dx_n)\, p_{1-t_n}(x_n,x)\, r(dx)\,,\end{aligned}$$

d'où, pour $0 < u_1 < \ldots < u_k < a < v_1 < \ldots < v_\ell < b < w_1 < \ldots < w_m < 1$,

$$\begin{aligned}\mathbb{\Pi}^g&[(X_0,\ldots,X_{u_k}) \in A\,,\ (X_{v_1},\ldots,X_{v_\ell}) \in B\,,\ (X_{w_1},\ldots,X_{w_m}) \in C \mid X_a, X_b]\\ &= \int I_A(x_0,\ldots,x_k)\, I_B(y_1,\ldots,y_\ell)\, I_C(z_1,\ldots,z_m)\\ &\qquad g(x_0)\, P_{u_1}(x_0,dx_1)\ldots P_{u_k-u_{k-1}}(x_{k-1},dx_k)\, p_{a-u_k}(x_k,X_a)\\ &\qquad P_{v_1-a}(a,dy_1)\ldots P_{v_\ell-v_{\ell-1}}(y_{\ell-1},dy_\ell)\, p_{b-v_\ell}(y_\ell,X_b)\\ &\qquad P_{w_1-b}(b,dz_1)\ldots P_{w_m-w_{m-1}}(z_{m-1},dz_m)\, p_{1-w_m}(z_m,x_0)\, r(dx_0)\end{aligned}$$

et la propriété de Markov circulaire a lieu car cette intégrale se factorise en le produit d'une intégrale par rapport aux y et d'une intégrale en les x et les z.

Supposant V connexe, nous allons maintenant montrer que parmi les probabilités $\mathbb{\Pi}^g$, seule celle qui correspond à g constante donne la même loi à tous les X_t. L'hypothèse $\mathbb{\Pi}^g[X_t \in A] = \mathbb{\Pi}^g[X_0 \in A]$ pour un $t \in]0,1[$ se réécrit

$$\int_A p_1(y,y)\,g(y)\,r(dy) = \int_{y\in A}\int_x \Pi^x[X_t \in dy]\,p_1(x,x)\,g(x)\,r(dx)$$
$$= \int_A r(dy)\int p_t(x,y)\,p_{1-t}(y,x)\,g(x)\,r(dx)\ ;$$

on en déduit que pour presque tout y

$$(*)\qquad p_1(y,y)\,g(y) = \int p_t(x,y)\,p_{1-t}(y,x)\,g(x)\,r(dx)\ .$$

Supposons que g n'est pas constante. Soit $\gamma = \operatorname{ess\,inf} g \geq 0$. L'ensemble $\{g > \gamma\}$ n'est pas négligeable, donc, pour presque tout y dans $\{g = \gamma\}$,

$$\gamma\,p_1(y,y) = g(y)\,p_1(y,y) = \int p_t(x,y)\,p_{1-t}(y,x)\,g(x)\,r(dx)$$
$$> \int p_t(x,y)\,p_{1-t}(y,x)\,\gamma\,r(dx) = \gamma\,p_1(y,y)$$

(la connexité de V a été utilisée sous la forme $p_t(x,y)\,p_{1-t}(y,x) > 0$); il s'ensuit que $\{g = \gamma\}$ est négligeable. Posons $A_\varepsilon = \{\gamma \leq g \leq \gamma + \varepsilon\}$, de sorte que $r(A_\varepsilon)$ tend vers zéro avec ε. En minorant g par $\gamma + \varepsilon - \varepsilon I_{A_\varepsilon}$ dans $(*)$ on trouve, pour presque tout y,

$$g(y)\,p_1(y,y) \geq (\gamma+\varepsilon)\,p_1(y,y) - \varepsilon\int_{A_\varepsilon} p_t(x,y)\,p_{1-t}(y,x)\,r(dx)\ .$$

Si l'on pose $C = \sup_{x,y} p_t(x,y)p_{1-t}(y,x)/\inf_y p_1(y,y)$, ceci entraîne

$$g(y) \geq (\gamma+\varepsilon) - C\varepsilon\,r(A_\varepsilon)\ ;$$

mais pour ε assez petit, $Cr(A_\varepsilon)$ est plus petit que $\frac{1}{2}$, et $g(y) \geq \gamma+\varepsilon-\frac{1}{2}\varepsilon = \gamma+\frac{1}{2}\varepsilon$ pour presque tout y, ce qui est absurde : g doit être constante.

Il ne nous reste plus qu'à vérifier que, la fonction g étant choisie constante ($g = c = \left[\int p_1(x,x)\,r(dx)\right]^{-1}$), la loi des lacets est stationnaire et réversible. C'est très facile! Recopions la formule donnant $\mathbb{\Pi}$:

$$\mathbb{\Pi}[(X_{t_1},\dots,X_{t_n}) \in A]$$
$$= c\int I_A(x_1,\dots,x_n)\,P_{t_1}(x,dx_1)P_{t_2-t_1}(x_1,dx_2)\dots$$
$$\dots P_{t_n-t_{n-1}}(x_{n-1},dx_n)\,p_{1-t_n}(x_n,x)\,r(dx)$$
$$= c\int I_A(x_1,\dots,x_n)\,p_{t_1}(x,x_1)p_{t_2-t_1}(x_1,x_2)\dots$$
$$\dots p_{t_n-t_{n-1}}(x_{n-1},x_n)\,p_{1-t_n}(x_n,x)\,r(dx)r(dx_1)\dots r(dx_n)\ .$$

En se débarassant de x par intégration, on remplace $p_{1-t_n}(x_n, x)\, r(dx)\, p_{t_1}(x, x_1)$ par $p_{t_1+1-t_n}(x_n, x_1)$, et les t_i n'interviennent plus que par leurs différences deux à deux, d'où la stationnarité (on observera, comme il a été annoncé plus haut, que l'invariance de la mesure r n'a pas été utilisée). Quant à la réversibilité, elle est vraie pour toutes les Π^g, et se déduit directement par intégration de celle des ponts Π^x, qui résulte immédiatement de la symétrie $p_t(y, z) = p_t(z, y)$. ∎