ON LONG TIME ALMOST SURE ASYMPTOTICS OF RENORMALIZED BRANCHING DIFFUSION PROCESSES

Nicolas FOURNIER, Bernard ROYNETTE

Institut Elie Cartan, campus scientifique, BP 239, 54506 Vandoeuvre-lès-Nancy cedex, France

ABSTRACT. – We are concerned with the long time behavior of branching diffusion processes. We give a partial answer to the following question: given a smooth density g_0, a branching rate and a spatial motion, does there exist a (nonspatially homogeneous) binary offspring distribution such that the corresponding renormalized branching process tends a.s. to $g_0(y)\,dy$ as time grows to infinity?
© 2003 Éditions scientifiques et médicales Elsevier SAS

MSC: 60J80; 60G57

Keywords: Branching diffusion processes; Long time behaviour

RÉSUMÉ. – On s'intéresse au comportement en temps long de certains processus de branchement-diffusion. On donne une réponse partielle à la question suivante : étant donnés une densité de probabilité régulière g_0, un taux de branchement, et un mouvement spatial, existe-t-il une loi de reproduction binaire (inhomogène en espace) telle que le processus de branchement-diffusion renormalisé correspondant converge presque sûrement, en temps infini, vers $g_0(y)\,dy$.
© 2003 Éditions scientifiques et médicales Elsevier SAS

1. Introduction, notations and result

Consider a compact Riemannian manifold M of class C^∞, and denote by dy the Riemannian measure on M, by ∇ and Δ the Riemannian gradient and Laplacian operators. Consider the diffusion process $\{X_t^x\}_{t \geqslant 0}$ on M with generator

$$Lf(x) = \frac{1}{2}\big[\Delta f(x) - \nabla u(x) \cdot \nabla f(x)\big], \tag{1.1}$$

where u is a C^∞ map from M into $\mathbb{R}$. It is well-known (see Ikeda and Watanabe [5], p. 235) that strong existence and uniqueness holds for such a diffusion. We will also need its stationary measure ν, whose explicit expression is

$$\nu(dy) = \exp\big(-u(y)\big)\,dy. \tag{1.2}$$

E-mail addresses: fournier@iecn.u-nancy.fr (N. Fournier), roynette@iecn.u-nancy.fr (B. Roynette).

We will denote by $\langle f, h\rangle_{L^2(\nu)} = \int_M f(y)h(y)\nu(dy)$. Then it is well known that L is symmetric with respect to ν, in the sense that for all f and h in $C^2(M)$,

$$\langle Lf, h\rangle_{L^2(\nu)} = \langle Lh, f\rangle_{L^2(\nu)} = -\frac{1}{2}\int_M \nabla f(y) \cdot \nabla h(y)\nu(dy). \tag{1.3}$$

Consider now a family of probability measures on $\{0, 2\}$, $\{p_0(y), p_2(y)\}_{y\in M}$ indexed by the points of M. This family is called the "offspring distributions".

We also need a "branching rate" $\lambda > 0$, which is simply a strictly positive real number.

Then, we consider the branching process $\{Y_t^x\}_{t\geqslant 0}$, which is a Markov process with values in

$$\mathcal{A} = \left\{\sum_{i=1}^n \delta_{x_i};\ n \in \mathbb{N},\ x_1, \ldots, x_n \in M\right\} \tag{1.4}$$

(where $\mathbb{N} = \{0, 1, 2, \ldots\}$) starting from δ_x, associated with the motion L, with the branching rate $\lambda > 0$, and with the branching probabilities $\{p_0(x), p_2(x)\}_{x\in M}$.

We refer to Dynkin, [2] for the rigorous definition of $\{Y_t^x\}_{t\geqslant 0}$ (using the notation of Dynkin, we are in the case where $\xi_t = X_t$, $K(s, t) = \lambda(t - s)$ and $\phi^t(x, z) = p_0(x) + z^2 p_2(x)$).

Roughly speaking, Y_t^x stands for the point measure describing the positions at time t of some particles which follow the dynamic below:
  (i) Each particle has a random exponential clock with parameter λ, independent of the others.
 (ii) Each particle moves independently of the others, according to the dynamic of the diffusion process with generator L.
(iii) When the exponential clock of one particle (located at y) rings, the particle dies, and the number of offspring is a random variable with law $(p_0(y), p_2(y))$.
(iv) The only interaction between the particles is that birth time and position of offspring coincide with the death time and position of their parent.

For $\mu = \sum_{i=1}^n \delta_{x_i}$ in $\mathcal{A}$ and h a function on M, we set $\langle \mu, h\rangle = \sum_{i=1}^n h(x_i)$.

We finally will use the extinction event, which we denote by

$$\mathcal{E}_x = \{\exists t,\ \forall s \geqslant t,\ \langle Y_s^x, 1\rangle = 0\} = \{\exists t,\ \langle Y_t^x, 1\rangle = 0\}. \tag{1.5}$$

Our main result is the following.

THEOREM 1.1. – *Consider a C^∞ strictly positive probability density function g_0 on M. Denote by $g(y) = g_0(y)\exp(u(y))$. Assume that $\lambda > C$, where $C = \sup_{x\in M} |Lg(x)/g(x)|$. Assume that for some $C_0 \in (0, 1 - C/\lambda)$,*

$$p_0(y) = \frac{1}{2}\left(1 - C_0 + \frac{Lg(y)}{\lambda g(y)}\right), \qquad p_2(y) = \frac{1}{2}\left(1 + C_0 - \frac{Lg(y)}{\lambda g(y)}\right). \tag{1.6}$$

Then $P[\mathcal{E}_x] < 1$, and the following convergence results hold:

(i) *there exists a nonnegative random variable $\widetilde{Z}^x$, strictly positive on $\mathcal{E}_x^c$, such that a.s., for all $h \in C(M)$,*

$$\lim_{t \to \infty} e^{-\lambda C_0 t} \langle Y_t^x, h \rangle = \widetilde{Z}^x \int_M h(y) g_0(y) \, dy; \qquad (1.7)$$

(ii) *almost surely, for all $h \in C(M)$,*

$$\lim_{t \to \infty} \mathbf{1}_{\mathcal{E}_x^c} \frac{\langle Y_t^x, h \rangle}{\langle Y_t^x, 1 \rangle} = \mathbf{1}_{\mathcal{E}_x^c} \int_M h(y) g_0(y) \, dy. \qquad (1.8)$$

One easily checks that for all $y \in M$, $p_0(y) + p_2(y) = 1$ and that $p_2(y) \in (0, 1)$. Notice that (ii) is an immediate corollary of (i). Remark also that Theorem 1.1 is quite natural. Indeed, our choice for p_0 and p_2 ensures that $Lg + \lambda(p_2 - p_0)g = \lambda C_0 g$. Intuitively, this means that g is a sort of "stable state" (with an exponentially growing number of particles): for each x in M, the particles which come at x (thanks to the spatial motion) are represented by $Lg(x)$, the appearance (or disappearance) of new particles at x by branching can be found in $\lambda(p_2(x) - p_0(x))g(x)$, and this leads to the new state $\lambda C_0 g(x)$.

Let us mention some literature on related topics. In the case of discrete-time multi-type Galton–Watson processes, with a finite number of types, Kesten and Stigum [6] proved a result similar to (1.7). However, the result in that paper which would correspond to the result $\{\widetilde{Z}^x \neq 0\} = \mathcal{E}_x^c$ in our paper was not proved. Still in the context of [6], Kurtz et al. [7] obtained a result of type (1.8), on the whole non-extinction set.

In the case of superprocesses on $\mathbb{R}^d$, and in a quite general context, Pinsky [8] obtained a result of type (1.7), when taking expectations. It has been recently improved by Engländer and Turaev [3], who replaced convergence "in expectation" by a (stronger) convergence in law. Let us notice that these authors do not need a symmetry asumption of type (1.3).

Let us now give two motivations for this problem. First of all, consider the following problem in economics: suppose that the diffusion process $\{X_t\}_{t \geqslant 0}$ corresponds to a "natural" spatial motion (on the sphere for example) of some commercial products. As time tends to infinity, the distribution of each of these products naturally tends to the probability measure ν. Assume that some institution (political, or commercial), wants to modify this distribution, in order that it becomes $g_0 = g \, d\nu$, for some g fixed. To this end, the institution can act by using taxes (very high at some places, and low or inexistent in other places), or something else. Then, the independent spatial motions may be replaced by a branching process $\{Y_t\}_{t \geqslant 0}$, with values in $\mathcal{A}$. Theorem 1.1 shows how to choose the offspring distributions (i.e., the tax rate as a function of the position), in order to obtain asymptotically the desired distribution.

Consider now the following biological problem. Suppose that the diffusion process $\{X_t\}_{t \geqslant 0}$ corresponds to the dynamics of a "natural hereditary quality" of a population. As time tends to infinity, the distribution of this quality tends to the probability measure ν. Assume now that by selecting the offspring, a biologist wants to transform this asymptotical distribution. Theorem 1.1 shows how to select the offspring (as a

function of the quality of the parent) in order to obtain asymptotically a given distribution $g_0 = g \, d\nu$ of the qualities.

Let us now comment on our result. First remark that it was not *a priori* obvious that $P[\mathcal{E}_x] < 1$. Indeed, the choice of p_0 and p_2 yield that $\int_M (p_2(x) - p_0(x)) g_0(x) \, dx = C_0 > 0$, but it is possible that at some places, $p_2(y) < p_0(y)$. The fact that $P[\mathcal{E}_x] < 1$ comes from the fact that where g_0 is large, i.e., where the (expected) asymptotic distribution of our process is large, p_2 is greater than p_0. A discussion on this point is given in Engländer and Turaev [3] in the case of superprocesses (see Appendix B and Lemma 8, the eigenvalue λ_c is given here by λC_0).

Let us explain our assumptions. Our main requirement is that the branching rate is greater, in some sense, than the speed of the spatial motion. Our condition $(\lambda > C)$ is probably not necessary, but it seems that if the motion is "much faster" than the branching rate, then one may not obtain any desired density: even if we force particles to be born at the locations where g_0 is large, then they will quickly move away from these points.

The next assumption is that g_0 is strongly equivalent to ν, in the sense that $g_0(y) \, dy = g(y)\nu(dy)$, for some bounded from above and from below density function g (since M is compact). Although this condition is probably too strong, it seems that at least the equivalence of $g_0(y) \, dy$ and of $\nu(dy)$ is necessary (this is also the case in [3]). Indeed, if $\nu(A) = 0$ for some subset $A \subset M$, we can not force particles to go in the region A using only branching. On the other hand, if $\nu(A) > 0$ for some subset $A \subset M$ we can not hope for the complete disappearence of particles in A, even by setting $p_2(y) = 0$ for all y in A.

Finally, it would probably be possible to treat the more general case where the offspring distribution charges $\mathbb{N}$ (instead of $\{0, 2\}$). In such a case and if this distribution admits a second order moment, the conclusion of Theorem 1.1 might hold, replacing (1.6) by $\sum_{n \geqslant 0} n p_n(y) = 1 + C_0 - Lg(y)/\lambda g(y)$.

To end this introduction, we would like to give the main intuition of our result (1.8). On $\mathcal{E}_x^c$, it is quite clear that the number of particles $\langle Y_t^x, 1 \rangle$ grows to infinity with t. Applying the Markov property, we may write, for each $s, t \geqslant 0$, for all $h \in C(M)$,

$$\frac{\langle Y_{t+s}^x, h \rangle}{\langle Y_{t+s}^x, 1 \rangle} = \frac{\langle Y_t^x, 1 \rangle}{\sum_{i=1}^{\langle Y_t^x, 1 \rangle} \langle Y_s^{X_t^i}, 1 \rangle} \cdot \frac{\sum_{i=1}^{\langle Y_t^x, 1 \rangle} \langle Y_s^{X_t^i}, h \rangle}{\langle Y_t^x, 1 \rangle}, \tag{1.9}$$

where X_t^i (for $i \in \{1, \ldots, \langle Y_t^x, 1 \rangle\}$) are the locations of the particles at t, and where conditioned on $\mathcal{F}_t$, $\{Y_s^{X_t^i}\}_{s \geqslant 0}$ are independent branching processes with initial conditions $\delta_{X_t^i}$. Hence, expecting that a law of large numbers holds, we might take the limit as t tends to infinity, and obtain that for some measure μ on M,

$$\lim_{t \to \infty} \frac{\langle Y_{t+s}^x, h \rangle}{\langle Y_{t+s}^x, 1 \rangle} = \frac{\int_M E(\langle Y_s^x, h \rangle) \mu(dx)}{\int_M E(\langle Y_s^x, 1 \rangle) \mu(dx)} = \langle \xi_s^\mu, h \rangle, \tag{1.10}$$

where the last equality stands for a definition. One can check that ξ_s^μ satisfies the partial differential equation: for all $\phi \in C^2(M)$,

$$\partial_t \langle \xi_t^\mu, \phi \rangle = \langle \xi_t^\mu, L\phi + \lambda(p_2 - p_0)\phi \rangle - \lambda \langle \xi_t^\mu, p_2 - p_0 \rangle \langle \xi_t^\mu, \phi \rangle. \tag{1.11}$$

But p_2 and p_0 have been chosen in such a way that the only stationary solution of this P.D.E. is $g_0(x)\,dx$. Hence, we expect that for any μ, ξ_s^μ goes, as s tends to infinity, to $g_0(x)\,dx$. Letting s tend to infinity in (1.10) gives us the desired conclusion.

2. Proof

In this section, we give the proof of our result. The assumptions of the previous section will always be supposed to hold.

We will use the notation A for a constant whose value changes from line to line.

We will first obtain some martingale properties of our branching process, which will in particular allow us to control the speed of growth of the size of the population.

Then, we will use a well-chosen orthonormal basis of $L^2(\nu)$. We will first prove (1.7) for each element h of this basis, and then extend the result to any continuous function h on M.

We first of all recall the generator of Y^x.

LEMMA 2.1. – *The process $\{Y_t^x\}_{t \geqslant 0}$ is a Markov process with values in $\mathcal{A}$. We denote by $\mathcal{L}$ its generator. For $f \in C^2(M)$ and $\mu \in \mathcal{A}$, we set*

$$F^f(\mu) = \langle \mu, f \rangle \quad and \quad G^f(\mu) = \langle \mu, f \rangle^2. \tag{2.1}$$

Then

$$\mathcal{L}F^f(\mu) = \left\langle \mu, Lf + \left(\lambda C_0 - \frac{Lg}{g} \right) f \right\rangle, \tag{2.2}$$

$$\mathcal{L}G^f(\mu) = \langle \mu, \lambda f^2 + |\nabla f|^2 \rangle + 2\langle \mu, f \rangle \left\langle \mu, Lf + \left(\lambda C_0 - \frac{Lg}{g} \right) f \right\rangle. \tag{2.3}$$

Proof. – Consider an element $\mu = \sum_{i=1}^n \delta_{x_i}$ of $\mathcal{A}$. Consider also $\phi \in C^2(\mathbb{R})$ and $f \in C^2(M)$. Then one easily checks that

$$\mathcal{L}\big[\phi(\langle \cdot, f \rangle)\big](\mu) = \partial_t E_\mu \big[\phi(\langle Y_t, f \rangle)\big]_{t=0}$$

$$= \partial_t E \left[\phi \left(\sum_{i=1}^n f(X_t^{x_i}) \right) \right]_{t=0} + \lambda \sum_{i=1}^n \big\{ p_2(x_i)\big[\phi(\langle \mu, f \rangle + f(x_i)) - \phi(\langle \mu, f \rangle)\big]$$

$$+ p_0(x_i)\big[\phi(\langle \mu, f \rangle - f(x_i)) - \phi(\langle \mu, f \rangle)\big] \big\}, \tag{2.4}$$

where X^{x_i} are independent diffusion processes on M with generator L, starting from x_i. Using the explicit expression of p_0 and p_2, we first obtain (2.2) by setting $\phi(x) = x$, and then (2.3) by setting $\phi(x) = x^2$. □

Then we deduce some martingale properties of our branching process.

COROLLARY 2.2. – *Let $f \in C^2(M)$. Then the process*

$$M_t^{x,f} = e^{-\lambda C_0 t} \langle Y_t^x, f \rangle - f(x) - \int_0^t e^{-\lambda C_0 s} \left\langle Y_s^x, Lf - f \frac{Lg}{g} \right\rangle ds \tag{2.5}$$

is a càdlàg (with an a.s. finite number of jumps on each compact time interval) martingale starting from 0 with (predictable) quadratic variation

$$\langle M^{x,f} \rangle_t = \int_0^t e^{-2\lambda C_0 s} \langle Y_s^x, \lambda f^2 + |\nabla f|^2 \rangle \, ds. \tag{2.6}$$

The proof of this corollary is immediate: it suffices to apply Itô's formula, (2.2) and (2.3).

We thus obtain some results about the growth of the population as t increases.

COROLLARY 2.3. –
(i) *There exists a constant A such that for all $t \geqslant 0$, all x in M,*

$$E\left[e^{-\lambda C_0 t} \langle Y_t^x, g \rangle\right] = g(x), \tag{2.7}$$

$$E\left[e^{-2\lambda C_0 t} \langle Y_t^x, g \rangle^2\right] \leqslant A. \tag{2.8}$$

(ii) *Hence, for any bounded nonnegative measurable function h on M, there exists a constant A_h such that for all $t \geqslant 0$, all x in M*

$$E\left[e^{-\lambda C_0 t} \langle Y_t^x, h \rangle\right] + E\left[e^{-2\lambda C_0 t} \langle Y_t^x, h \rangle^2\right] \leqslant A_h. \tag{2.9}$$

Proof. – We first prove (i). Using Corollary 2.2 with $f = g$, it is obvious that

$$M_t^{x,g} = e^{-\lambda C_0 t} \langle Y_t^x, g \rangle - g(x) \tag{2.10}$$

is a martingale starting from 0. Hence (2.7) holds. Next, using the expression of the bracket of $M^{x,g}$, we deduce that

$$E\left[e^{-2\lambda C_0 t} \langle Y_t^x, g \rangle^2\right] = g^2(x) + \int_0^t e^{-\lambda C_0 s} E\left[e^{-\lambda C_0 s} \langle Y_s^x, \lambda g^2 + |\nabla g|^2 \rangle\right] ds$$

$$\leqslant g^2(x) + \left\|\frac{\lambda g^2 + |\nabla g|^2}{g}\right\|_\infty \int_0^t e^{-\lambda C_0 s} E\left[e^{-\lambda C_0 s} \langle Y_s^x, g \rangle\right] ds$$

$$\leqslant A + A \int_0^t e^{-\lambda C_0 s} \, ds \leqslant A, \tag{2.11}$$

the value of the constant A changing from line to line. Point (ii) is an immediate consequence, since for any h bounded, there exists a constant $\tilde{A}_h$ such that for all $y \in M$, $h(y) \leqslant \tilde{A}_h g(y)$. □

In order to build a suitable basis of $L^2(\nu)$, we introduce some operators.

LEMMA 2.4. – *For $f \in C^2(M)$, and $t \geqslant 0$, we define*

$$R_t f(x) = E\left[e^{-\lambda C_0 t} \langle Y_t^x, f \rangle\right]. \tag{2.12}$$

Then R_t is a semi-group with generator K defined by

$$Kf(x) = Lf(x) - f(x)\frac{Lg(x)}{g(x)}. \tag{2.13}$$

Consider also the h-transform of R_t defined by

$$Q_t f(x) = \frac{1}{g(x)} R_t[fg](x). \tag{2.14}$$

Then Q_t is a Markovian semi-group, and the associated generator is given by

$$\widetilde{K} f(x) = \frac{1}{g(x)} \big(L[fg](x) - f(x)Lg(x)\big)$$

$$= \frac{1}{2}\left[\Delta f(x) + \nabla f(x) \cdot \left(2\frac{\nabla g(x)}{g(x)} - \nabla u(x)\right)\right]. \tag{2.15}$$

The stationary measure of Q_t is $g^2\nu$, and for all h and f in $C^2(M)$,

$$\big\langle \widetilde{K}h, f\big\rangle_{L^2(g^2\nu)} = \big\langle h, \widetilde{K}f\big\rangle_{L^2(g^2\nu)} = -\frac{1}{2}\int_M \nabla f(y) \cdot \nabla h(y)g^2(y)\nu(dy). \tag{2.16}$$

Proof. – We first prove that R_t is a semi-group. Let thus $f \in C^2(M)$, and let $s, t \geqslant 0$. We first write Y_{t+s}^x as $\sum_{i=1}^{\langle Y_t^x, 1\rangle} Y_s^{X_t^i}$, where X_t^i (for $i \in \{1, \dots, \langle Y_t^x, 1\rangle\}$) are the locations of the particles at t, and where conditioned on $\mathcal{F}_t$, $\{Y_s^{X_t^i}\}_{s\geqslant 0}$ are independent branching processes with initial conditions $\delta_{X_t^i}$. We obtain

$$R_{t+s}f(x) = E\left[e^{-\lambda C_0 t}\sum_{i=1}^{\langle Y_t^x, 1\rangle} E\{e^{-\lambda C_0 s}\langle Y_s^{X_t^i}, f\rangle | \mathcal{F}_t\}\right] = E\left[e^{-\lambda C_0 t}\sum_{i=1}^{\langle Y_t^x, 1\rangle} R_s f(X_t^i)\right]$$

$$= E\big[e^{-\lambda C_0 t}\langle Y_t^x, R_s f\rangle\big] = R_t[R_s f](x). \tag{2.17}$$

Next, it is clear from Corollary 2.2 that the generator of R_t is given by K. It is immediately deduced that Q_t is a semi-group and that (2.15) and (2.16) hold. The fact that Q_t is Markov (i.e., $Q_t 1 = 1$) is deduced from Corollary 2.3(i). □

Next we present useful properties of R_t.

LEMMA 2.5. –
 (i) *There exist some C^∞ functions $q(t, x, y)$ and $r(t, x, y)$ on $(0, \infty) \times M \times M$ such that for all measurable function f on M,*

$$Q_t f(x) = \int_M q(t, x, y)f(y)\,dy, \qquad R_t f(x) = \int_M r(t, x, y)f(y)\,dy. \tag{2.18}$$

 (ii) *The semi-group $(R_t)_{t\geqslant 0}$ is strongly continuous, this is, for any $h \in C(M)$,*

$$\lim_{t\to 0} \|R_t h - h\|_\infty = 0. \tag{2.19}$$

Proof. – The generator $\widetilde{K}$ of Q_t is the generator of a M-valued uniformly elliptic diffusion process, so that it is well-known, see, e.g., Fefferman and Sanchez-Calle [4], p. 268, that there exists a C^∞ function $q(t, x, y)$ on $(0, \infty) \times M \times M$ such that for all measurable function f on M, $Q_t f(x) = \int_M q(t, x, y) f(y) \, dy$. Hence,

$$R_t f(x) = g(x) Q_t \left[\frac{f}{g}\right](x) = \int_M g(x) \frac{q(t, x, y)}{g(y)} f(y) \, dy \qquad (2.20)$$

and thus (i) holds with $r(t, x, y) = q(t, x, y) g(x)/g(y)$, which has the same regularity as r since g is C^∞, bounded from below and from above.

It is also standard that $(Q_t)_{t \geqslant 0}$ is strongly continuous, from which point (ii) it is immediately deduced. □

We now build a basis of $L^2(\nu)$, of which the elements are eigenfunctions of K.

LEMMA 2.6. – *There exists an nondecreasing sequence $\{\rho_n\}_{n \geqslant 0}$ of nonnegative real numbers and a sequence $\{\psi_n\}_{n \geqslant 0}$ of real-valued functions on M such that:*
  (i) *$\rho_0 = 0$, and $\rho_n > 0$ for all $n \geqslant 1$,*
  (ii) *$\{\psi_n\}_{n \geqslant 0}$ is an orthonormal basis of $L^2(M, \nu)$, and $\psi_0 = g/\sqrt{\int_M g^2 \, d\nu}$,*
  (iii) *for all $n \geqslant 0$, $K \psi_n = -\rho_n \psi_n$ and for all $t \geqslant 0$, $R_t \psi_n = e^{-\rho_n t} \psi_n$,*
  (iv) *for all n, $\psi_n \in C^\infty(M)$,*
  (v) *for all $t > 0$, $\sum_{n \geqslant 0} e^{-\rho_n t} < \infty$.*

Proof. – Consider the eigenvalues and eigenfunctions of $\widetilde{K}$: $\widetilde{K} \phi_n = -\rho_n \phi_n$, for all $n \geqslant 0$, where ρ_n are nonnegative real numbers. (It is clear from (2.16) that the eigenvalues of $\widetilde{K}$ are nonpositive.) One easily deduces that for each $n \geqslant 0$, any $t > 0$, $Q_t \phi_n = e^{-\rho_n t} \phi_n$.

Next, notice that for each $t > 0$ fixed, Q_t is a Hilbert–Schmidt operator. Indeed, this is a straightforward consequence of Lemma 2.5, and of Ex 49-b p. 1086 of [1]. This implies that its spectrum is discrete, and that (v) holds (since $Q_{t/2}$ is Hilbert–Schmidt). We number the ρ_n in nondecreasing order and note that it follows classically that $\rho_0 < \rho_1$. We choose $\{\phi_n\}$ an orthonormal basis of $L^2(M, g^2 \nu)$. Finally notice that $\rho_0 = 0$, and that the associated (renormalized) eigenfunction is $\phi_0 = 1/\sqrt{\int_M g^2 \, d\nu}$.

Then the sequence $\psi_n = g \phi_n$ satisfies the conclusion of the lemma. Indeed, points (i) and (ii) are straightforward. Next, it is clear that for each $n \geqslant 0$, $K \psi_n = g \widetilde{K} \phi_n = -g \rho_n \phi_n = -\rho_n \psi_n$. In the same spirit, $R_t \psi_n = g Q_t \phi_n$. Hence $R_0 \psi_n = \psi_n$, and $\partial_t R_t \psi_n = g \partial_t Q_t \phi_n = -\rho_n g Q_t \phi_n = -\rho_n R_t \psi_n$, and (iii) follows.

Point (iv) is an immediate consequence of the previous lemma: for any $t > 0$ fixed, $\psi_n = e^{\rho_n t} R_t \psi_n = e^{\rho_n t} \int_M \psi_n(y) r(t, \cdot, y) \, dy$. Since $r(t, \cdot, \cdot)$ is C^∞ on $M \times M$, the usual Lebesgue theorem allows to conclude the proof. Finally, point (v) has already been proved. □

We now study the asymptotic behaviour of $e^{-\lambda C_0 t} \langle Y_t^x, \psi_n \rangle$ for each n. This will be sufficient, since $\{\psi_n\}$ is a basis of $L^2(M, \nu)$.

LEMMA 2.7. –

(i) *There exists a nonnegative random variable Z^x such that a.s. and in $L^2(\Omega)$,*

$$\lim_{t\to\infty} e^{-\lambda C_0 t} \langle Y_t^x, g \rangle = Z^x. \tag{2.21}$$

Furthermore, there exists a constant A such that for all $x \in M$,

$$E[Z^x] = g(x), \qquad E[(Z^x)^2] \leqslant A. \tag{2.22}$$

(ii) *For all $n \geqslant 1$, a.s.,*

$$\lim_{t\to\infty} e^{-\lambda C_0 t} \langle Y_t^x, \psi_n \rangle = 0. \tag{2.23}$$

Proof. – We begin with (i). We know from Corollaries 2.2 and 2.3 that $M_t^{x,g} = e^{-\lambda C_0 t} \langle Y_t^x, g \rangle - g(x)$ is a martingale starting from 0 and bounded in L^2, from which (i) is a straightforward consequence.

We now prove (ii). We set $W_t^{n,x} = e^{-\lambda C_0 t} \langle Y_t^x, \psi_n \rangle$. Using Corollary 2.2 and the fact that $L\psi_n - \psi_n Lg/g = K\psi_n = -\rho_n \psi_n$, we deduce that

$$W_t^{n,x} = \psi_n(x) - \rho_n \int_0^t W_s^{n,x} \, ds + M_t^{x,\psi_n}, \tag{2.24}$$

where M^{x,ψ_n} is defined in (2.5). This equation can be solved explicitely:

$$W_t^{n,x} = e^{-\rho_n t} \left[\psi_n(x) + \int_0^t e^{\rho_n s} \, dM_s^{x,\psi_n} \right]. \tag{2.25}$$

Denote by $O_t^{n,x} = \int_0^t e^{\rho_n s} \, dM_s^{x,\psi_n}$. We clearly just have to check that $D_t^{n,x} = e^{-\rho_n t} O_t^{n,x}$ goes a.s. to 0 as t increases to infinity. This will be done by using the Borel–Cantelli lemma: it suffices to check that for any $\varepsilon > 0$,

$$\sum_{k \geqslant 1} P\left[\sup_{[k,k+1]} |D_t^{n,x}| \geqslant \varepsilon \right] < \infty. \tag{2.26}$$

But for all k,

$$P\left[\sup_{[k,k+1]} |D_t^{n,x}| \geqslant \varepsilon \right] \leqslant P\left[\sup_{[k,k+1]} |O_t^{n,x}| \geqslant \varepsilon \, e^{k\rho_n} \right] \leqslant A \frac{e^{-2k\rho_n}}{\varepsilon^2} E[(O_{k+1}^{n,x})^2] \tag{2.27}$$

by using Doob's inequality, since $O^{n,x}$ is a martingale. Finally, an easy computation using the quadratic variation of $O^{n,x}$ (which is obtained from that of M^{x,ψ_n}, which is given in Corollary 2.2), using Corollary 2.3 and Lemma 2.6(iv) shows that for some constant A_n,

$$E[(O_t^{n,x})^2] = E\left[\int_0^t e^{2\rho_n s} e^{-2\lambda C_0 s} \langle Y_s^x, \lambda\psi_n^2 + |\nabla\psi_n|^2 \rangle \, ds \right]$$

$$\leqslant A_n \int_0^t e^{2\rho_n s}\, e^{-\lambda C_0 s} ds \leqslant A_n(1+t)\big(e^{2\rho_n t - \lambda C_0 t} + 1\big). \qquad (2.28)$$

We furthermore deduce that, for some constant $A_{n,\varepsilon}$,

$$P\Big[\sup_{[k,k+1]} |D_t^{n,x}| \geqslant \varepsilon\Big] \leqslant A_{n,\varepsilon}(k+2)\big[e^{-\lambda C_0 k} + e^{-2\rho_n k}\big] \qquad (2.29)$$

which allows to conclude that (2.26) holds. This ends the proof.   □

We deduce the non-triviality of the extinction event.

COROLLARY 2.8. – *Recall that $\mathcal{E}_x$ is the extinction event.*
 (i) *Set $\varepsilon = \inf_{y \in M} g(y)/2 > 0$. Then*

$$\alpha = \inf_{x \in M} P\big[Z^x > \varepsilon\big] > 0. \qquad (2.30)$$

(ii) $\sup_{x \in M} P[\mathcal{E}_x] \leqslant 1 - \alpha$.

Proof. – We first check (i). Thanks to Lemma 2.7(i), we know that for all x in M, $E[Z^x] = g(x)$ and $A = \sup_{x \in M} E[(Z^x)^2] < \infty$. Thus

$$g(x) = E\big[Z^x \mathbf{1}_{Z^x \leqslant \varepsilon}\big] + E\big[Z^x \mathbf{1}_{Z^x > \varepsilon}\big] \leqslant \varepsilon + \sqrt{A}\sqrt{P\big[Z^x > \varepsilon\big]} \qquad (2.31)$$

which allows to conclude the proof. Point (ii) is now obvious, since it is clear from Lemma 2.7(i) that $\mathcal{E}_x \subset \{Z^x = 0\} \subset \{Z^x \leqslant \varepsilon\}$.   □

We now extend Lemma 2.7 to continuous functions on M.

LEMMA 2.9. – *Set $a := 1/(\int_M g^2\, d\nu)$. Then for all $h \in C(M)$, a.s.,*

$$\lim_{t \to \infty} e^{-\lambda C_0 t}\langle Y_t^x, h\rangle = a Z^x \langle h, g\rangle_{L^2(\nu)}. \qquad (2.32)$$

Proof. – We prove this result by using the previous lemma. We thus consider $h \in C(M)$. We first recall that due to Lemma 2.5(ii),

$$\lim_{t \to 0} \|R_t h - h\|_\infty = 0. \qquad (2.33)$$

Then we split h and $R_t h$ according to the orthonormal basis ψ_n:

$$h = \sum_{n \geqslant 0} \langle h, \psi_n\rangle_{L^2(\nu)} \psi_n, \qquad R_t h = \sum_{n \geqslant 0} \langle h, \psi_n\rangle_{L^2(\nu)} e^{-\rho_n t} \psi_n. \qquad (2.34)$$

Then we consider the approximation $h_p = \sum_{n=0}^p \langle h, \psi_n\rangle_{L^2(\nu)} \psi_n$.

Then it is clear from Lemma 2.7, since $\psi_0 = g/\sqrt{\int_M g^2\, d\nu}$ and $\rho_0 = 0$, that for any $t > 0$ and any $p \geqslant 1$, a.s.,

$$\lim_{s\to\infty} e^{-\lambda C_0 s}\langle Y_s^x, R_t h_p\rangle = \sum_{n=0}^{p} e^{-\rho_n t}\langle h, \psi_n\rangle_{L^2(\nu)} \lim_{s\to\infty} e^{-\lambda C_0 s}\langle Y_s^x, \psi_n\rangle$$

$$= \langle h, \psi_0\rangle_{L^2(\nu)}\frac{Z^x}{\sqrt{\int_M g^2\,d\nu}} = aZ^x\langle h, g\rangle_{L^2(\nu)}. \qquad (2.35)$$

Furthermore, using Lemma 2.5(i), one easily checks that for any $t > 0$, there exists a constant $A(t)$ such that for all $n \geqslant 0$,

$$\|\psi_n\|_\infty = e^{\rho_n t}\|R_t\psi_n\|_\infty \leqslant A(t)\,e^{\rho_n t}\|\psi_n\|_{L^2(\nu)} = A(t)\,e^{\rho_n t} \qquad (2.36)$$

from which we deduce that for all $t > 0$,

$$\|R_t h - R_t h_p\|_\infty = \left\|\sum_{n>p}\langle h, \psi_n\rangle_{L^2(\nu)}\,e^{-\rho_n t}\psi_n\right\|_\infty \leqslant \|h\|_{L^2(\nu)}A\left(\frac{t}{2}\right)\sum_{n>p}e^{-\rho_n t/2} \qquad (2.37)$$

(by applying (2.36) with $t/2$ instead of t) which goes to 0 as p tends to infinity, according to Lemma 2.6(v).

Also notice that since g is bounded away from 0 and since $e^{-\lambda C_0 s}\langle Y_s^x, g\rangle$ is a converging martingale, one easily obtains that the random variable

$$U = \sup_{s\in[0,\infty)} e^{-\lambda C_0 s}\langle Y_s^x, 1\rangle \qquad (2.38)$$

is a.s. finite. We may finally conclude the proof. We approximate h by $R_t h_p$, for p large and t small. Let $\varepsilon > 0$ be fixed. We have to show that there a.s. exists s_ε such that for all $s \geqslant s_\varepsilon$,

$$\left|\langle Y_s^x, h\rangle e^{-\lambda C_0 s} - aZ^x\langle g, h\rangle_{L^2(\nu)}\right| < \varepsilon. \qquad (2.39)$$

We use the previous estimates: for any $s \geqslant 0$, $t > 0$, and $p \geqslant 1$,

$$\left|\langle Y_s^x, h\rangle e^{-\lambda C_0 s} - aZ^x\langle g, h\rangle_{L^2(\nu)}\right| \leqslant \left|\langle Y_s^x, h\rangle e^{-\lambda C_0 s} - \langle Y_s^x, R_t h\rangle e^{-\lambda C_0 s}\right|$$

$$+ \left|\langle Y_s^x, R_t h\rangle e^{-\lambda C_0 s} - \langle Y_s^x, R_t h_p\rangle e^{-\lambda C_0 s}\right|$$

$$+ \left|\langle Y_s^x, R_t h_p\rangle e^{-\lambda C_0 s} - aZ^x\langle g, h\rangle_{L^2(\nu)}\right|. \qquad (2.40)$$

Using (2.37) and (2.38), we deduce that for all s, t, p,

$$\left|\langle Y_s^x, h\rangle e^{-\lambda C_0 s} - aZ^x\langle g, h\rangle_{L^2(\nu)}\right|$$

$$\leqslant U \times \left[\|h - R_t h\|_\infty + \|h\|_{L^2(\nu)}A\left(\frac{t}{2}\right)\sum_{n>p}e^{-\rho_n t/2}\right]$$

$$+ \left|\langle Y_s^x, R_t h_p\rangle e^{-\lambda C_0 s} - aZ^x\langle g, h\rangle_{L^2(\nu)}\right|. \qquad (2.41)$$

Using (2.33), (2.37), and (2.35) allows to conclude that choosing first t small enough, then p large enough, and finally s large enough yields to (2.39). This ends the proof. □

To conclude, we just have to verify that Z^x is strictly positive on the set where there is not extinction.

LEMMA 2.10. – *Recall that $\mathcal{E}_x$ is the extinction set.*

(i) *On $\mathcal{E}_x^c$, $\lim_{t\to\infty}\langle Y_t^x, 1\rangle = +\infty$ a.s.*
(ii) *Furthermore,*

$$P\left[\mathcal{E}_x^c \cap \{Z^x = 0\}\right] = 0. \tag{2.42}$$

Proof. – We first prove (i). Recall that $\langle Y_t^x, 1\rangle$ is $\mathbb{N}$-valued. Hence, one easily understands that it suffices to show that for all $N \in \mathbb{N}_* = \{1, 2, 3, \ldots\}$, $P[\liminf_{t\to\infty}\langle Y_t^x, 1\rangle = N] = 0$. Let thus $N \in \mathbb{N}_*$ be fixed. First notice that

$$\tilde{p}_0 = \inf_{y\in M} p_0(y) \geqslant \frac{1}{2}\left(1 - C_0 - \frac{C}{\lambda}\right) > 0. \tag{2.43}$$

Then, set $T_0 = 0$, $S_0 = 0$, and define recursively the stopping times

$$S_n = \inf\{s > T_n; \ \Delta\langle Y_s^x, 1\rangle \neq 0\},$$
$$T_{n+1} = \inf\{s > S_n; \ \langle Y_s^x, 1\rangle = N\}, \tag{2.44}$$

where $\Delta\langle Y_s^x, 1\rangle = \langle Y_s^x, 1\rangle - \langle Y_{s-}^x, 1\rangle$. Then,

$$P\left[\liminf_{t\to\infty}\langle Y_t^x, 1\rangle = N\right] = P\left[\lim_n T_n = \infty, \sum_{n\geqslant 1}\mathbf{1}_{\Delta\langle Y_{S_n}^x, 1\rangle = -1} < \infty\right]. \tag{2.45}$$

Due to the Borel–Cantelli lemma, this last quantity equals 0, since conditionnaly on the sequence $\{T_n, S_n\}_{n\geqslant 1}$, the events $\{\Delta\langle Y_{S_n}^x, 1\rangle = -1\}$ are independent and with probability bounded below by $\tilde{p}_0$.

Next, we prove (ii). Denote by $B = \mathcal{E}_x^c \cap \{Z^x = 0\}$. On the set B, $\lim_{t\to\infty}\langle Y_t^x, 1\rangle = +\infty$ a.s., while $\lim_{t\to\infty} e^{-\lambda C_0 t}\langle Y_t^x, 1\rangle = 0$.

For any $t \geqslant 0$, denote by X_t^i, for $i \in \{1, \ldots, \langle Y_t^x, 1\rangle\}$, the points of the support of Y_t^x. Then one may write, for all $s \geqslant 0$, $t \geqslant 0$,

$$\langle Y_{t+s}^x, 1\rangle e^{-\lambda C_0(t+s)} = e^{-\lambda C_0 t}\sum_{i=1}^{\langle Y_t^x, 1\rangle} e^{-\lambda C_0 s}\langle Y_s^{X_t^i}, 1\rangle, \tag{2.46}$$

where $Y^{X_t^i}$ are independent branching processes starting from $\delta_{X_t^i}$ conditioned on $\mathcal{F}_t$. Making s tend to infinity, we deduce, using Lemma 2.9 and the fact that $\langle 1, g\rangle_{L^2(\nu)} = 1$, that on B, for all t,

$$\sum_{i=1}^{\langle Y_t^x, 1\rangle} Z^{X_t^i} = 0. \tag{2.47}$$

Using Corollary 2.8(i) and the notations therein, we thus obtain that for all $t \geqslant 0$, all $N \geqslant 1$,

$$P[B] \leqslant P\left[\sum_{i=1}^{N} Z^{X_t^i} = 0, \langle Y_t^x, 1\rangle \geqslant N\right] + P\left[\langle Y_t^x, 1\rangle < N, \mathcal{E}_x^c\right]$$
$$\leqslant (1 - \alpha)^N + P\left[\langle Y_t^x, 1\rangle \leqslant N, \mathcal{E}_x^c\right]. \tag{2.48}$$

Making t tend to infinity, we deduce (using (i)) that for each $N \geqslant 1$, $P[B] \leqslant (1 - \alpha)^N$. Since $\alpha > 0$, we conlude that $P(B) = 0$, which was our aim. $\quad \square$

The conclusion is now straightforward.

Proof of Theorem 1.1. – Let $x \in M$ be fixed. We already have proved (see Corollary 2.8(ii)) that $P[\mathcal{E}_x] < 1$. Since for any $h \in C(M)$, $\langle h, g \rangle_{L^2(\nu)} = \int_M h(y)g_0(y)\,dy$, point (i) is straightforward from Lemmas 2.9 and 2.10, while point (ii) is an immediate consequence of point (i). $\quad \square$

REFERENCES

[1] N. Dunford, J. Schwartz, Linear Operators, Part II, Spectral Theory, Selfadjoint Operators in Hilbert Space, Reprint of the 1963 original, in: Wiley Classics Library, 1988.

[2] E.B. Dynkin, Branching particle systems and superprocesses, Ann. Probab. 19 (3) (1991) 1157–1194.

[3] J. Engländer, D. Turaev, A scaling limit theorem for a class of superdiffusions, Ann. Probab. 30 (2) (2002) 683–722.

[4] C. Fefferman, A. Sanchez-Calle, Fundamental solutions for second order subelliptic operators, Ann. of Math. 124 (1986) 247–272.

[5] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, 1989.

[6] H. Kesten, B. Stigum, A limit theorem for multidimensional Galton–Watson processes, Ann. Math. Statist. 37 (1966) 1211–1223.

[7] T. Kurtz, R. Lyons, R. Pemantle, Y. Peres, A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes, in: K. Athreya, P. Jagers (Eds.), Classical and Modern Branching Processes (Minneapolis, 1994), in: IMA Vol. Math. Appl., Vol. 84, Springer, New York, 1997, pp. 181–185.

[8] R.G. Pinsky, Transience, reccurence and local extinction properties of the support for supercritical finite measure-valued diffusions, Ann. Probab. 24 (1) (1996) 237–267.