Smoothness of the law of manifold-valued Markov processes with jumps

JEAN PICARD[1,2,*] and CATHERINE SAVONA[1,2,**]

[1]*Clermont Université, Université Blaise Pascal, Laboratoire de Mathématiques, BP 10448, F-63000 Clermont-Ferrand, France. E-mail:* [*]*Jean.Picard@math.univ-bpclermont.fr;*
[**]*Catherine.Savona@math.univ-bpclermont.fr; url:* [*http://math.univ-bpclermont.fr/~picard/*](http://math.univ-bpclermont.fr/~picard/)
[2]*CNRS, UMR 6620, LM, BP 80026, F-63171 Aubière, France*

Consider on a manifold the solution X of a stochastic differential equation driven by a Lévy process without Brownian part. Sufficient conditions for the smoothness of the law of X_t are given, with particular emphasis on noncompact manifolds. The result is deduced from the case of affine spaces by means of a localisation technique. The particular cases of Lie groups and homogeneous spaces are discussed.

Keywords: Lie groups and homogeneous spaces; Malliavin calculus; manifold-valued processes; processes with jumps

1. Introduction

Consider a $\mathbb{R}^m$-valued Lévy process Λ_t without Brownian part, a d-dimensional manifold M, and the M-valued solution X_t of an equation

$$X_{t+dt} = a(X_t, d\Lambda_t) + b(X_t)\,dt, \qquad X_0 = x_0 \tag{1}$$

for coefficients a and b such that $a(x, 0) = x$. The precise meaning of this equation will be given later, as well as conditions implying the existence and uniqueness of a solution. The aim of this article is to give sufficient conditions ensuring the smoothness of the law of X_t at any time $t > 0$. We also study more precisely the case of Lévy processes on Lie groups, and of some classes of processes on homogeneous spaces.

Proving the smoothness of the law of a random variable has motivated, in the case of continuous diffusions, the introduction of Malliavin's calculus. When $M = \mathbb{R}^d$, Bismut's approach to this calculus has proved to be useful for processes with jumps which are solutions of equations of type (1), and this topic has been intensively studied since [5] and [4]. Different techniques, each of them having its own domain of applicability, have been introduced afterwards. These techniques can be roughly divided into two classes.

The first class relies, as in [5], on some infinitesimal perturbations (in space or in time) on the jumps of the Lévy process Λ. A differential calculus can be based on these perturbations, and the associated integration by parts formula enables to study the smoothness of the law of X_t.

The second class of techniques (also when $M = \mathbb{R}^d$) has been worked out in [24]; rather than a differential calculus, one uses a finite difference calculus consisting in appending and removing jumps. This is not a differential calculus so there is no integration by parts formula in the usual

sense, but there is still a duality formula which can be written on the Poisson space of jumps, and this formula can be interpreted as a duality between appending and removing jumps, see [22,23]; this calculus has been applied to the smoothness of the law of X_t by [24]. Its advantage is that no smoothness is required for the Lévy measure of Λ; in particular this measure may have a countable support. The proof of the smoothness of the law is based on an estimation of the characteristic function (Fourier transform) of X_t; if this function is proved to decrease rapidly at infinity, then X_t has a C^∞ density. In [24] (as it will be the case in this article), it is assumed that Λ has no Brownian part and that the coefficient a satisfies a nondegeneracy condition similar to the ellipticity condition for continuous diffusions; notice that the class of so-called canonical equations with a Brownian part was studied with this finite difference calculus by [14], and Hörmander type conditions were considered by [17], also for canonical equations.

The extension of [24] to the case where M is a manifold is not immediate. The basic tool of the case $M = \mathbb{R}^d$ was the Fourier transform. When M is a symmetric space, the Fourier transform can be replaced by the so-called spherical transform, and this has been used by [20] in order to prove the smoothness of the law for a class of processes (but the Lévy measure was not allowed to be purely atomic in the case without Brownian part). However, adapting the Fourier transform technique of $\mathbb{R}^d$ seems difficult for general manifolds. When M is diffeomorphic to $\mathbb{R}^d$, we can of course apply the result for $\mathbb{R}^d$, but the assumptions needed for this result are generally not translated into canonical assumptions on M: they depend on the choice of the diffeomorphism, and they may be hard to verify, see in Section 5.1 the example of the hyperbolic space. Actually, even the case where M is an open subset of $\mathbb{R}^d$ is not trivial.

In order to get around these difficulties, our aim here is to apply some localisation techniques; we consider an atlas of M and use the results of the affine case on each local chart. However, it is known that localisation is made difficult by the presence of jumps. In [26], we have applied such a technique in order to study the smoothness of harmonic functions on some domain D of $\mathbb{R}^d$, and it appears that these functions are not always C^∞, even for processes with a C^∞ density; their order of regularity depends on the number of jumps needed to exit D, so that the smaller the jumps are, the smoother the function is. Our plan is therefore, first, to apply the localisation technique and obtain the C^ℓ regularity of the density when big jumps are removed, and then check that adding big jumps does not destroy this smoothness. In order to conclude, when the manifold is not compact, we have to make an assumption on the size of these big jumps, namely that for any relatively compact subset U of M, the set of points from which the process can jump into U is relatively compact (one cannot jump from a very distant location). This condition can be viewed as dual to the condition of [26] for the smoothness of harmonic functions. The result (Theorem 2) is given in Section 2, after the particular case of canonical equations (Theorem 1); it is proved in Section 3.

Then we relax this "big jumps" condition in three cases:

- when the Lévy kernel for these big jumps is smooth (Theorem 3),
- when M is a Lie group G and X_t is a Lévy process on it,
- when M is an homogeneous space G/H and X_t is obtained by projecting on M a Lévy process on G.

Lie groups and homogeneous spaces are the purpose of Section 4.

In Section 5, we give some examples, and also some counterexamples where the "big jumps" condition is not satisfied, and the smoothness of the law fails.

2. The smoothness result on manifolds

In this section, we give a precise meaning for equation (1), but before that we introduce the manifold M and the Lévy process Λ_t. Then we give the assumptions on the equation, give the smoothness results about the law of the solution, and make some comments on our model. Actually, before explaining the general case, we state the main result in the particular case of so-called canonical equations; assumptions are indeed much easier to write in this case. The proofs are postponed to Section 3.

2.1. The manifold

The manifold M is supposed to be Hausdorff, separable, paracompact, C^∞ and of dimension d; we do not suppose that it is connected; it may have an at most countable number of connected components, and our processes will be allowed to jump from a component to another. The tangent bundle is denoted by $TM = \bigcup_x T_x M$. If M is not compact, we consider its one-point compactification $M \cup \{\infty\}$; if M is compact, ∞ is a point which is disconnected from M. This additional point will be viewed as a cemetery point; this means that real functions f on M are extended by putting $f(\infty) = 0$, and that a process on $M \cup \{\infty\}$ which hits ∞ or tends to ∞ stays at that point forever.

Under these assumptions, one can embed M into an affine space $\mathbb{R}^N$ by means of Whitney's theorem, and one can consider Riemannian metrics on M by using a partition of unity; Riemannian distances are of course only defined between points in the same connected component, and they otherwise take the value $+\infty$. We would like to have intrinsic assumptions, which do not depend on a particular embedding or a particular metric. This will be possible because Riemannian distances are equivalent on compact subsets, so a metric condition which is supposed to hold uniformly on compact subsets will actually not depend on the choice of the metric. For instance, if f is a C^ℓ function and K is a compact subset of M, if we denote by $D^j f(x)$ the iterated differential of f (this is a multi-linear form on $(T_x M)^j$), we can define

$$|f|_{K,\ell} = \sum_{j=0}^{\ell} \sup_{x \in K} |D^j f(x)|$$

if we have chosen a Riemannian metric on M, but changing the metric will lead to an equivalent semi-norm; by allowing K to vary, we obtain a Fréchet space $C^\ell(M)$. A signed measure ν on M is said to be absolutely continuous, respectively, C^ℓ, if its restrictions to local charts are absolutely continuous, respectively, have C^ℓ densities with respect to the Lebesgue measure; it is said to have positive density if it has (strictly) positive density with respect to the Lebesgue measure on charts; this does not depend on the atlas; let $\mathcal{M}^\ell(M)$ be the set of C^ℓ measures. If we choose a C^∞ reference measure dx with positive density, we can define the family of semi-norms

$$|\nu|_{K,\ell} = |d\nu/dx|_{K,\ell}.$$

Changing the reference measure and the Riemannian metric will not change the topology of $\mathcal{M}^\ell(M)$. We also let

$$\mathcal{M}_K^\ell = \left\{ \nu \in \mathcal{M}^\ell(M); \nu(K^c) = 0 \right\}$$

if K is a compact subset of M.

The notation $U \Subset V$ for open subsets of M will mean that U is relatively compact in V.

2.2. The Lévy process

Let us first recall some basic facts about Lévy processes Λ_t with values in $\mathbb{R}^m$. A Lévy measure is a measure μ on $\mathbb{R}^m \setminus \{0\}$ such that $\int (|\lambda|^2 \wedge 1) \mu(d\lambda) < \infty$, and a Lévy process Λ_t without Brownian part and with Lévy measure μ is a process which can be written by means of the Lévy–Itô representation formula

$$\Lambda_t = \kappa t + \int_0^t \int_{\{|\lambda| \leq 1\}} \lambda \widetilde{N}(ds, d\lambda) + \int_0^t \int_{\{|\lambda| > 1\}} \lambda N(ds, d\lambda) \tag{2}$$

for some $\kappa \in \mathbb{R}^m$, where the random measure $N(dt, d\lambda) = \sum_t \delta_{(t, \Delta \Lambda_t)}$ is a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}^m$ with intensity $dt\, \mu(d\lambda)$, and $\widetilde{N}(dt, d\lambda) = N(dt, d\lambda) - dt\, \mu(d\lambda)$ is the compensated Poisson measure. When $\int (|\lambda| \wedge 1) \mu(d\lambda) < \infty$, then Λ_t has finite variation, and (2) can be simplified as

$$\Lambda_t = \kappa_0 t + \int_0^t \int_{\mathbb{R}^m} \lambda N(ds, d\lambda).$$

The relation between κ and κ_0 is easily written, and Λ_t is a pure jump process when $\kappa_0 = 0$. We will assume an approximate self-similarity condition and a non-degeneracy condition on the Lévy process written as follows.

Assumption 1. *There exist some $0 < \alpha < 2$ and some positive c and C such that*

$$c\rho^{2-\alpha} |u|^2 \leq \int_{\{|\lambda| \leq \rho\}} \langle \lambda, u \rangle^2 \mu(d\lambda) \leq C\rho^{2-\alpha} |u|^2$$

for $u \in \mathbb{R}^m$, $0 < \rho \leq 1$, and where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product. If $\alpha = 1$, we suppose moreover that

$$\limsup_{\varepsilon \downarrow 0} \left| \int_{\{\varepsilon \leq |\lambda| \leq 1\}} \lambda \mu(d\lambda) \right| < \infty.$$

If $\alpha < 1$ (finite variation case), we suppose that Λ_t is pure jump ($\kappa_0 = 0$).

The process Λ has finite variation if and only if $\alpha < 1$. Notice that no smoothness is assumed on μ; it can, for instance, have a countable support. The additional conditions of the cases $\alpha = 1$ and $\alpha < 1$ are needed to apply the results of [25,26].

As an example, we can consider stable Lévy processes of index α, or more generally semi-stable processes (see Chapter 3 of [27] for definitions and properties concerning these processes).

If Λ_t is a semi-stable process and if μ is not supported by a strict subspace of $\mathbb{R}^m$, then Assumption 1 is satisfied when $1 < \alpha < 2$. When $0 < \alpha \leq 1$, the additional conditions mean that Λ_t should be a strictly semi-stable process. Actually, Assumption 1 is only concerned with small jumps, so that we can also consider for instance truncated semi-stable processes (where jumps greater than some value are removed).

2.3. The equation

Let us now introduce the process X_t, solution of (1). When $M = \mathbb{R}^d$, the meaning of this equation is

$$
\begin{aligned}
X_t = X_u &+ \int_u^t \overline{a}(X_{s-}) \, \mathrm{d}\Lambda_s + \int_u^t b(X_s) \, \mathrm{d}s \\
&+ \sum_{u < s \leq t} \big(a(X_{s-}, \Delta\Lambda_s) - X_{s-} - \overline{a}(X_{s-})\Delta\Lambda_s\big)
\end{aligned}
\tag{3}
$$

for $u \leq t$, where $\overline{a}(x)$ is the differential at 0 of $\lambda \mapsto a(x, \lambda)$. Under convenient smoothness conditions, the Itô integral with respect to Λ is well defined, the sum converges, and the equation has a unique solution for any initial condition $X_0 = x_0$. Notice that the jumps are given by $X_t = a(X_{t-}, \Delta\Lambda_t)$. Itô's formula enables to write

$$
\begin{aligned}
f(X_t) = f(x_0) &+ \int_0^t (Df\overline{a})(X_{s-}) \, \mathrm{d}\Lambda_s + \int_0^t (Dfb)(X_s) \, \mathrm{d}s \\
&+ \sum_{0 < s \leq t} \big((f \circ a)(X_{s-}, \Delta\Lambda_s) - f(X_{s-}) - (Df\overline{a})(X_{s-})\Delta\Lambda_s\big)
\end{aligned}
\tag{4}
$$

for smooth functions f, and where Df denotes the differential of f.

This formula can be used to give a meaning to (1) when M is a manifold. We consider a coefficient a and a vector field b,

$$
a : M \times \mathbb{R}^m \to M \cup \{\infty\}, \qquad b : M \to TM.
$$

We again let $\overline{a}(x)$ be the differential of $\lambda \mapsto a(x, \lambda)$ at 0, so that $\overline{a}(x)$ is a linear map from $\mathbb{R}^m$ into $T_x M$:

$$
\overline{a} : M \to L\big(\mathbb{R}^m, TM\big), \qquad x \mapsto \partial_\lambda|_{\lambda=0} a(x, \lambda).
\tag{5}
$$

We say that $(X_t; t \geq 0)$ is a solution of (1) if X_t is a càdlàg (right continuous with left limits) process with values in $M \cup \{\infty\}$, which is adapted with respect to the completed filtration of Λ_t, and such that for any smooth function f and any compact subset K of M, equation (4) holds true up to the first exit time from K. We say that the process dies at time t if $X_s \in M$ for $s < t$ and $X_t = \infty$. This occurs if either $X_{t-} = \infty$ (the solution of the equation explodes at time t), or if $X_{t-} \in M$ and $X_t = a(X_{t-}, \Delta\Lambda_t) = \infty$ (the process is killed at a jump of Λ).

If we use a proper embedding $\mathcal{I} : M \to \mathbb{R}^N$, we can apply (4) to the components of the embedding $\mathcal{I}$ and deduce the equation that should be satisfied on $\mathcal{I}(M)$ by $\mathcal{I}(X_t)$; in order to obtain

an equation on $\mathbb{R}^N$, we have to extend the coefficients out of $\mathcal{I}(M)$; it is then sufficient to solve the equation in $\mathbb{R}^N$ and prove that the solution remains in M. This is this point of view which is generally used in order to prove the existence and uniqueness of a solution, see [8] in the case without killing ($a(x, \lambda) \in M$ for any $x \in M$). We are here in a slightly different framework (possible killing), and shall prefer to give a proof for the existence and uniqueness by means of local charts; this is because the proof based on local charts is then used for studying the smoothness of the law of X_t (it is clear that the embedded process does not have a density in $\mathbb{R}^N$, so proving the existence of a density on M cannot be made through a Malliavin calculus on $\mathbb{R}^N$). The equation on a local chart can also be deduced from (4).

The infinitesimal generator of X_t is

$$
\mathcal{L}f(x) = Df(x)\big(b(x) + \overline{a}(x)\kappa\big) + \int_{\{|\lambda| \leq 1\}} \big(f\big(a(x, \lambda)\big) - f(x) - Df(x)\overline{a}(x)\lambda\big)\mu(\mathrm{d}\lambda)
$$
$$
+ \int_{\{|\lambda| > 1\}} \big(f\big(a(x, \lambda)\big) - f(x)\big)\mu(\mathrm{d}\lambda)
$$

(6)

for f bounded and C^∞. In particular, if f is constant on a neighbourhood of x, then

$$
\mathcal{L}f(x) = \int_M \big(f(y) - f(x)\big)\mu_x(\mathrm{d}y),
$$

where μ_x is the image of the Lévy measure μ by $\lambda \mapsto a(x, \lambda)$. This measure describes the intensity of jumps; it is the Lévy kernel of the process X.

2.4. Canonical equations

Up to now, we have been given an equation (1) on the manifold M, and have explained the rigourous meaning of this equation; in this explanation we need the function $\overline{a}$ given by (5). It is clear that many functions a are associated to the same $\overline{a}$. However, there is a particular class of equations, called canonical equations, and which were introduced by [21] (see also [18]), for which a and $\overline{a}$ are in one-to-one correspondence.

Let us first consider a smooth field $\overline{a}(x) \in L(\mathbb{R}^m, T_x M)$, and let

$$
a(x, \lambda) = x_\lambda(1) \qquad \text{for } x_\lambda(t) = x + \int_0^t \overline{a}\big(x_\lambda(s)\big)\lambda \,\mathrm{d}s, \tag{7}
$$

assuming that the solution of this ordinary differential equation does not explode. Then it is easily seen that a and $\overline{a}$ are related to each other by (5), and $x \mapsto a(x, \lambda)$ is a diffeomorphism of M onto itself with inverse $x \mapsto a(x, -\lambda)$. Notice that x and $a(x, \lambda)$ are in the same connected component, so the study can be reduced to connected manifolds.

For canonical equations, the assumptions needed for our main result (Theorem 2 below), or at least a sufficient condition ensuring that they are satisfied, can be written in the following simple form.

Theorem 1. *Let Λ be a Lévy process satisfying Assumption 1, and let $\overline{a}$ and b be C^∞ functions on M, with values, respectively, in $L(\mathbb{R}^m, TM)$ and TM. Consider equation (1) with coefficient a given by (7), assuming that the ordinary differential equation never explodes on M. We suppose that the jumps of Λ are bounded, and that the linear map $\overline{a}(x) : \mathbb{R}^m \to T_x M$ is surjective for any x; if $\alpha < 1$ we also suppose that $b = 0$. Then (1) has for any initial condition x_0 a unique solution X_t, and the law of X_t is C^∞ for any $t > 0$.*

2.5. Assumptions on the equation

We now return to the case of a general coefficient a. Let us give the assumptions on equation (1) which will imply the existence and uniqueness of a solution, and the smoothness of the law of this solution.

Assumption 2. *The conditions on the coefficients a and b are as follows.*

1. *Consider, for any $\varepsilon > 0$, the map $a_\star^\varepsilon$ which sends a measure ν to the measure on M*

$$\left(a_\star^\varepsilon \nu\right)(A) = \int \int 1_A\big(a(x, \lambda)\big) 1_{\{|\lambda| > \varepsilon\}} \nu(\mathrm{d}x) \mu(\mathrm{d}\lambda). \tag{8}$$

   Let K be any compact subset of M. If ε is small enough, then $a_\star^\varepsilon$ is a continuous map from $\mathcal{M}_K^\ell$ into $\mathcal{M}^\ell(M)$ (see Section 2.1 for the definition of these spaces).
2. *Let K be any compact subset of M. There exists $\varepsilon > 0$ such that $x \mapsto a(x, \lambda)$ is C^∞ on K for μ-almost any λ such that $|\lambda| \le \varepsilon$. The function $\overline{a}$ given by (5) is assumed to exist and to be C^∞. Letting D be the differentiation operator on M with respect to x, there exists some $\alpha \vee 1 < \gamma \le 2$ such that*

$$\left| D^j \big(f\big(a(x, \lambda)\big) - f(x) - Df(x)\overline{a}(x)\lambda \big) \right| \le C_{f, j, K} |\lambda|^\gamma$$

   for any C^∞ real function f, for $x \in K$, for μ-almost any $|\lambda| \le \varepsilon$, and for any j.
3. *The coefficient b is a C^∞ vector field on M with values in TM. In the case $\alpha < 1$, we suppose that $b = 0$.*
4. *The linear map $\overline{a}(x) \in L(\mathbb{R}^m, T_x M)$ is surjective for any x.*

Here are some comments about these four conditions.

First condition. This condition states that jumps preserve the smoothness of the law of the process. Suppose that $x \mapsto a(x, \lambda)$ is a C^ℓ diffeomorphism of M onto itself with inverse $y \mapsto a^{-1}(y, \lambda)$. Fix a Riemannian metric on M and the associated measure $\mathrm{d}x$; then the density $p_\star(y)$ of $a_\star^\varepsilon \nu$ is obtained from the density p of ν by means of the classical formula

$$p_\star(y) = \int_{\{|\lambda| > \varepsilon\}} \left| \det Da^{-1}(y, \lambda) \right| p\big(a^{-1}(y, \lambda)\big) \mu(\mathrm{d}\lambda)$$

(the determinant is computed for orthonormal bases on the tangent spaces). It is therefore sufficient to estimate the derivatives of a^{-1}; the condition is for instance satisfied for the canonical

equations of Theorem 1. Moreover, the process can be killed when it quits some open subset. More precisely, consider a process on a manifold M_0 associated to an equation with coefficient a_0, let M be an open subset of M_0, and kill the process when it quits M. This process is obtained by considering the equation with coefficient

$$a(x, \lambda) = \begin{cases} a_0(x, \lambda) & \text{if } a_0(x, \lambda) \in M, \\ \infty & \text{otherwise.} \end{cases} \tag{9}$$

Then $a_\star^\varepsilon$ is continuous if the same property holds true for a_0, because the map which sends a measure on M_0 to its restriction to M is continuous from $\mathcal{M}^\ell(M_0)$ to $\mathcal{M}^\ell(M)$. We shall however notice that Assumption 3 below often fails for this example. Other examples will be given in Section 2.7.

Second condition. In this condition, the smoothness of $a(x, \lambda)$ is assumed with respect to x for λ small, but no smoothness is assumed with respect to λ except for $\lambda \to 0$. If however a is smooth in (x, λ), then the condition is satisfied for $\gamma = 2$. In particular, canonical equations of Section 2.4 satisfy the condition.

Third condition. The additional assumption on b in the finite variation case $\alpha < 1$ means that the solution X of our equation is a pure jump process; it is required in order to apply the results of [25,26].

Fourth condition. The surjectivity of $\overline{a}(x)$ is a nondegeneracy condition. This condition says that small jumps go in all the directions, and is similar to the ellipticity condition for continuous diffusions.

Let us now give the additional assumption concerning big jumps. It is stating, roughly speaking, that the process cannot come from a very distant point by jumping.

Assumption 3. *If U is relatively compact, then*

$$a^{-1}(U) = \{x; \mu\{\lambda; a(x, \lambda) \in U\} > 0\}$$

is also relatively compact.

This assumption is trivially satisfied when M is compact. It is also satisfied in Theorem 1 because the jumps of Λ are in some ball B of $\mathbb{R}^m$, so

$$a^{-1}(U) \subset a(U \times B),$$

and the relative compactness of this set follows from the relative compactness of $U \times B$ and the continuity of a. The assumption often fails when the process is obtained by killing as explained in (9); difficulties generally arise when the original process can jump from $M_0 \setminus M$ into M; a counterexample, showing that the law of X_t is not always smooth in this case, will be given in Section 5.3.

2.6. The results

The main smoothness result for the solution of (1) is the following one.

Theorem 2. *Under Assumptions* 1 *and* 2, *the equation* (1) *has a unique solution* X_t *for any initial condition* $X_0 = x_0$, *and the law of* X_t *is absolutely continuous for any* $t > 0$. *If moreover Assumption* 3 *holds true, the law of* X_t *is* C^∞ *for* $t > 0$. *More precisely, if* K *is a compact subset of* M, *if* $t_0 > 0$, *and if* $p(t, x_0, x)$ *is the density of* X_t *with respect to some* C^∞ *reference measure with positive density, then the derivatives of* p *with respect to* x *satisfy*

$$\left| D^j p(t, x_0, x) \right| \le C_{j,K} t^{-(d+j)/\alpha} \tag{10}$$

uniformly for $0 < t \le t_0$, x_0 *in* M *and* x *in* K.

In the case of canonical equations, we see that Theorem 1 is a corollary of this result. In this result, jumps of Λ were supposed to be bounded. When M is compact but jumps of Λ are unbounded, we see that the condition which may cause problems is the first condition of Assumption 2. More precisely, if the vector field $\overline{a}(\cdot)\lambda$ has for some $\lambda = \lambda_1$ a stable equilibrium x_1, then jumps $\Delta\Lambda = c\lambda_1$ for large c concentrate the mass near x_1, so that the density of $a^\varepsilon_\star \nu$ may be unbounded near x_1.

When the big jumps condition (Assumption 3) is not satisfied, the conclusion of the theorem still holds if these big jumps are smooth; some other cases will be studied in Section 4, and on the other hand, counterexamples will be given in Section 5.

Theorem 3. *Under Assumptions* 1 *and* 2, *suppose that there exists a decomposition* $\mu = \mu^\flat + \mu^\sharp$ *of the Lévy measure such that only* $\mu^\flat$ *satisfies Assumption* 3, *and* $\mu^\sharp$ *is finite. Let* $\mu^\sharp_x$ *be the image of* $\mu^\sharp$ *by the map* $\lambda \mapsto a(x, \lambda)$; *suppose that* $\mu^\sharp_x$ *is* C^∞, *and that* $x \mapsto |\mu^\sharp_x|_{K,\ell}$ *is bounded on* M *for any compact* K *and any* ℓ (*see the definition of this family of semi-norms in Section* 2.1). *Then the solution* X_t *of* (1) *has a* C^∞ *law for* $t > 0$, *satisfying* (10).

The kernel $\mu^\sharp_x$ is the part of the Lévy kernel μ_x coming from $\mu^\sharp$. The assumption of Theorem 3 therefore requires the part of the Lévy kernel for jumps coming from distant locations to be smooth, whereas the Lévy kernel could be purely atomic in Theorem 2.

2.7. About the jump coefficient

In this article, we study Markov processes X_t with infinitesimal generator of the form (6), with coefficients a and b satisfying some smoothness assumptions. This covers a class of Markov processes, but not all of them. In particular, in this model, the set of (times of) jumps of X is contained in the set of jumps of the driving Lévy process Λ; the inclusion may be strict, since one may have $a(x, \lambda) = x$ for some $\lambda \neq 0$, but introducing such a behaviour generally destroys the smoothness of a. Thus, X and Λ have more or less the same times of jumps, and the rate of jump of X is not allowed to depend on its present state x. This is a drawback of this approach, as well as of other approaches based on different versions of Malliavin's calculus. A weak dependence of the rate of jumps with respect to x can however be obtained through Girsanov transformation, see [26], and cases of more general dependence have been studied (under frameworks which are different from ours) by [3,11,15]. We now verify that such a dependence is possible if we

drop the assumption of smoothness of a, and that this is compatible with our assumptions if only finitely many jumps are concerned by this behaviour.

Suppose that the generator of X is $\mathcal{L} + \mathcal{L}'$, where $\mathcal{L}$ satisfies the assumptions of Theorem 2 or 3, and

$$\mathcal{L}'f(x) = \int_{M\cup\{\infty\}} \big(f(y) - f(x)\big)\mu'_x(\mathrm{d}y),$$

where $x \mapsto \mu'_x(M \cup \{\infty\})$ is finite and C^∞. We also have to assume that the kernel μ'_x is Borel measurable, that $\nu \mapsto \nu'$ with

$$\nu'(A) = \int \mu'_x(A)\nu(\mathrm{d}x) \tag{11}$$

is continuous from $\mathcal{M}_K^\ell$ into $\mathcal{M}^\ell(M)$ (as in Assumption 2), and that jumps corresponding to this kernel satisfy Assumption 3 (the extension with the assumptions of Theorem 3 is also possible).

In order to prove that this case enters our framework, we apply the fact that the measurable space $M \cup \{\infty\}$ can be viewed as a Borel subset of $\mathbb{R}_+$ (it is a Lusin space, see [9]); thus we can view μ'_x as a measure on $\mathbb{R}_+$, and if we define

$$\widetilde{a}(x, u) = \inf\big\{v \geq 0;\ \mu'_x([0, v]) \geq u\big\} \in \mathbb{R}_+ \cup \{+\infty\},$$

the image of the Lebesgue measure on $\mathbb{R}_+$ by $\widetilde{a}(x, \cdot)$ is the measure μ'_x on $\mathbb{R}_+$, plus an infinite mass at $+\infty$; notice in particular that $\widetilde{a}(x, u) = +\infty$ if u is large enough. Let us then introduce a real symmetric Lévy process Λ'_t with Lévy measure $\mu'(\mathrm{d}\lambda') = \alpha|\lambda'|^{-\alpha-1}\mathrm{d}\lambda'$, independent of Λ; then (Λ, Λ') is a Lévy process on $\mathbb{R}^{m+1}$ with Lévy measure $\mu(\mathrm{d}\lambda)\delta_0(\mathrm{d}\lambda') + \delta_0(\mathrm{d}\lambda)\mu'(\mathrm{d}\lambda')$ satisfying Assumption 1. We then consider the equation driven by (Λ, Λ'), with coefficient $a(x, \lambda, \lambda')$, where $a(x, \lambda, 0)$ is the coefficient associated to the part $\mathcal{L}$ of the generator, and

$$a\big(x, 0, \lambda'\big) = \begin{cases} \widetilde{a}\big(x, \lambda'^{-\alpha}\big) & \text{if } \lambda' > 0 \text{ and this quantity is finite,} \\ x & \text{otherwise.} \end{cases}$$

In particular, $a(x, 0, \lambda') = x$ if $\lambda' < \mu'_x(M \cup \{\infty\})^{-1/\alpha}$. The image of μ' by $a(x, 0, \cdot)$ is μ'_x plus an infinite mass at x, so the solution of the equation has generator $\mathcal{L} + \mathcal{L}'$ as required. On the other hand, if K is a compact subset of M and if

$$\varepsilon < \Big(\sup_{x\in K} \mu'_x\big(M \cup \{\infty\}\big)\Big)^{-1/\alpha},$$

then, with the notations (8) and (11),

$$\big(a(\cdot, 0, \cdot)^\star_\varepsilon \nu\big)(\mathrm{d}x) = \nu'(\mathrm{d}x) + \big(\varepsilon^{-\alpha} - \mu'_x\big(M \cup \{\infty\}\big)\big)\nu(\mathrm{d}x).$$

Thus these jumps satisfy the first part of Assumption 2, though $x \mapsto a(x, 0, \lambda')$ is generally not continuous.

We have already seen in (9) that we can consider hard killing of a process (we kill it when it hits an obstacle), but this may cause difficulties with Assumption 3. With the construction we

have just described, we can also consider soft killing where the process is killed at some rate $h(x) \geq 0$ depending smoothly on x; this means that we add the term $\mathcal{L}' f(x) = -h(x) f(x)$ to $\mathcal{L} f(x)$, and μ'_x is the mass $h(x)$ at ∞; in this case, the measure ν' of (11) is the zero measure on M, so $\nu \mapsto \nu'$ is trivially continuous.

3. Proof of Theorems 2 and 3

The two theorems are proved in several steps.

3.1. Construction of the solution

In order to prove the existence and uniqueness of a solution of (1), we first write the equation on a local chart (U, Φ), where Φ is a diffeomorphism from an open subset U of M onto an open subset V of $\mathbb{R}^d$. We can restrict ourselves to atlases such that Φ is the restriction to U of a smooth map on M. If τ is the exit time of X from U, then (4) applied to the components of Φ shows that $Y_t = \Phi(X_t)$, $t < \tau$, should be a solution of an equation (3) on V with new coefficients

$$a_\Phi(y, \lambda) = \Phi\big(a\big(\Phi^{-1}(y), \lambda\big)\big), \qquad b_\Phi(y) = D\Phi\big(\Phi^{-1}(y)\big)b\big(\Phi^{-1}(y)\big),$$
$$\overline{a}_\Phi(y) = D\Phi\big(\Phi^{-1}(y)\big)\overline{a}\big(\Phi^{-1}(y)\big). \tag{12}$$

More precisely, X is solution of (1) if it is a càdlàg process on $M \cup \{\infty\}$ with initial condition $X_0 = x_0$, satisfying the conditions:

- For any local chart (U, Φ) and for any time u, if τ is the exit time after u of X from U, the process $Y_t = \Phi(X_t)$ satisfies the equation (3) with coefficients (a_Φ, b_Φ) on $\{u \leq t < \tau\}$.
- The jumps of X are given by $X_t = a(X_{t-}, \Delta \Lambda_t)$.
- If X_{t-} or X_t is at ∞, then $X_s = \infty$ for any $s \geq t$ (∞ is a cemetery point).

In order to solve the equation (1) up to the first exit time from U, we shall have to extend coefficients (a_Φ, b_Φ) out of V, and solve the resulting equation (3) on $\mathbb{R}^d$.

Lemma 1. *The equation* (1) *has a unique solution* X_t *for any initial condition* x_0.

Proof. Consider open subsets of M with relatively compact inclusions $U_1 \Subset U_2 \Subset U_3$. We suppose that there exists a diffeomorphism Φ from U_3 onto an open subset V_3 of $\mathbb{R}^d$, so that (U_3, Φ) is a local chart. We define $V_1 = \Phi(U_1)$, $V_2 = \Phi(U_2)$, and let $h : \mathbb{R}^d \to [0, 1]$ be a smooth function such that $h = 1$ on V_1 and $h = 0$ on V_2^c. The coefficient $b_\Phi(y)$ of (12) is defined on V_3. On the other hand, from Assumption 2, there exists $\varepsilon > 0$ such that

$$|\lambda| \leq \varepsilon \quad \Rightarrow \quad a(U_2, \lambda) \subset U_3.$$

Then $a_\Phi(y, \lambda)$ is well defined for $y \in V_2$ and $|\lambda| \leq \varepsilon$, and takes its values in V_3. Thus, for $y \in \mathbb{R}^d$ and $|\lambda| \leq \varepsilon$, we can define

$$\big(\widetilde{a}_\Phi(y, \lambda), \widetilde{b}_\Phi(y)\big) = \begin{cases} \big(h(y)a_\Phi(y, \lambda) + \big(1 - h(y)\big)y, h(y)b_\Phi(y)\big) & \text{if } y \in V_2, \\ (y, 0) & \text{otherwise} \end{cases} \tag{13}$$

which is an interpolation between (a_Φ, b_Φ) on V_1 and the motionless process on the complement of V_2. Notice that from Assumption 2,

$$\widetilde{a}_\Phi(y, \lambda) = y + h(y)\overline{a}_\Phi(y)\lambda + O\big(|\lambda|^\gamma\big)$$

for $\overline{a}_\Phi$ given by (12), and similarly for its derivatives. We can consider on $\mathbb{R}^d$ the equation (3) with coefficients $(\widetilde{a}_\Phi, \widetilde{b}_\Phi)$, driven by

$$\Lambda_t^\varepsilon = \Lambda_t - \sum_{s \leq t} \Delta\Lambda_s 1_{\{|\Delta\Lambda_s| > \varepsilon\}}. \tag{14}$$

Our smoothness assumptions on a and b imply that it has a unique solution Y_t for $Y_0 = \Phi(x_0)$ fixed. Let

$$\tau = \inf\big\{t \geq 0; Y_t \notin V_1 \text{ or } |\Delta\Lambda_t| > \varepsilon\big\}.$$

Defining $X_t = \Phi^{-1}(Y_t)$ for $t < \tau$, and $X_\tau = a(X_{\tau-}, \Delta\Lambda_\tau)$, the process X is solution of our equation (1) up to the time τ; if $|\Delta\Lambda_\tau| > \varepsilon$ and $X_\tau \in U_1$, then we can solve again the equation from this time τ and point X_τ; since jumps greater than ε are in finite number on any finite time interval, we deduce the existence of a solution X_t up to its first exit time from U_1. Conversely, we can go from X to Y and deduce the uniqueness of X from the uniqueness of Y. Thus, equation (1) has a unique solution up to the first exit time from U_1.

Let us deduce the existence of a solution $(X_t; 0 \leq t \leq t_0)$ for a fixed $t_0 > 0$. We consider a locally finite atlas of M made of subsets $U_0(k)$ (any compact subset intersects finitely many of these $U_0(k)$) such that $U_0(k)$ is relatively compact in an open subset $U(k)$ of the type of the set U_1 of the first part of the proof. We can choose for any x an index $k(x)$ such that k is measurable and $x \in U_0(k(x))$, and we solve the equation starting from x up to the first exit time τ_1 of the set $U(k(x))$ (apply the first part of the proof). Let $\mathbb{P}_x$ be the law of this solution. If δ is some Riemannian distance on M, one can check that $\sup_{t \leq u} \delta(x, X_t)$ under $\mathbb{P}_x$ converges in probability to 0 as $u \downarrow 0$, uniformly for x such that $k(x) = k$ (this follows from the similar property satisfied by $Y = \Phi(X)$); thus there exists $u_k > 0$ such that

$$\mathbb{P}_x[\tau_1 \geq u_k] \geq \mathbb{P}_x\Big[\sup_{t \leq u_k} \delta(x, X_t) < \delta\big(U_0(k), U(k)^c\big)\Big] \geq 1/2 \tag{15}$$

if $k = k(x)$.

The equation is then solved by means of the following iterative procedure. For a fixed initial condition x_0, we solve the equation from time $\tau_0 = 0$ up to the exit time τ_1 from $U(k(x_0))$. If $\tau_1 \geq t_0$ (the time up to which we want to solve the equation), we have obtained the solution up to time t_0 and we can stop the procedure; if $X_{\tau_1} = \infty$ the process stays at ∞ and the procedure can also be stopped; otherwise, starting at time τ_1 from X_{τ_1}, we solve the equation up to the exit time τ_2 from $U(k(X_{\tau_1}))$, and so on. We stop the procedure either when $\tau_j \geq t_0$, or when the process has been killed (jump to ∞). Thus, the procedure goes on forever when $\tau_j < t_0$ and $X_{\tau_j} \neq \infty$ for any j. On the other hand, denoting by $\mathcal{F}_t$ the filtration of Λ_t, we have from (15) and the strong Markov property that

$$\mathbb{P}[\tau_{j+1} - \tau_j \geq u_k | \mathcal{F}_{\tau_j}] \geq 1/2 \qquad \text{on } A_j^k = \{\tau_j < t_0\} \cap \big\{k(X_{\tau_j}) = k\big\}.$$

We deduce that

$$\sum_{j\geq 0}\mathbb{P}\big[A_j^k\big] \leq 2\sum_{j\geq 0}\mathbb{P}\big[A_j^k \cap \{\tau_{j+1}-\tau_j \geq u_k\}\big] \leq 2(1+t_0/u_k)$$

because there are at most t_0/u_k disjoint intervals of length $\geq u_k$ included in $[0,t_0]$. Thus, for k fixed, A_j^k cannot be satisfied infinitely many times. We deduce that if the procedure goes on forever, then $k(X_{\tau_j})$ tends to infinity, so $X_{\tau-} = \infty$ for $\tau = \lim \tau_j$ (the solution explodes at time τ). In this case the solution is obtained by putting $X_t = \infty$ for $t \geq \tau$.

The uniqueness can be proved by considering the first time τ at which two solutions X_t^1 and X_t^2 differ, and by supposing that $\tau < \infty$ with positive probability; then $X_{\tau-}^1 = X_{\tau-}^2$, so $X_\tau^1 = X_\tau^2$ and the uniqueness of the solution in $U(k)$ for $k = k(X_\tau^1)$ leads to a contradiction. $\square$

3.2. The case with only small jumps

In Lemma 1, we have worked out a construction of the process by means of local charts. We now verify that this construction also provides a smoothness result for the law on these local charts. In all the proofs, we choose a Riemannian metric on M, and the associated C^∞ measure dx with respect to which we will consider the densities of M-valued random variables. We shall study the solution X^ε of equation (1) driven by the Lévy process without its jumps greater than ε (the process Λ^ε defined in (14)). The parameter ε will be fixed and will be assumed to be small enough; notice however that the constants involved in the calculations will not be uniform in ε; this does not cause any difficulty because we shall never take the limit as $\varepsilon \downarrow 0$.

Lemma 2. *Consider open subsets of M with relatively compact inclusions $U_0 \Subset U_1 \Subset U_2 \Subset U_3$, such that U_3 is diffeomorphic to an open subset of $\mathbb{R}^d$. Moreover, let B be a neighbourhood of the diagonal in $M \times M$. We consider the process X^ε solution of equation (1) driven by Λ^ε. The surjectivity of $\overline{a}(x)$ is only assumed on U_2. Let $\ell \geq 0$. Then the following properties hold true if ε is small enough.*

1. *The law of the process X^ε starting from $x_0 \in U_1$ and killed at the exit from U_1 has a C^ℓ density on U_0.*
2. *The density $x \mapsto q^\varepsilon(t,x_0,x)$ of this killed process and more generally its derivatives of order $j \leq \ell$ are uniformly dominated by $t^{-(d+j)/\alpha}$ for $0 < t \leq t_0$, $x_0 \in U_1$ and $x \in U_0$.*
3. *The density and its derivatives up to order ℓ are uniformly bounded for $0 < t \leq t_0$ and $(x_0,x) \in (U_1 \times U_0) \setminus B$.*

Proof. We apply the construction given in the proof of Lemma 1, denote $V_i = \Phi(U_i)$, and obtain the process Y_t solution of (3) driven by Λ^ε, with coefficients $(\widetilde{a}_\Phi, \widetilde{b}_\Phi)$ given by (13). Then X^ε can be written as $X_t^\varepsilon = \Phi^{-1}(Y_t)$ strictly before the first exit from U_1, so the killed processes X^ε and $\Phi^{-1}(Y)$ coincide. The nonkilled process Y_t has of course not a smooth law for any initial condition, since it is motionless out of V_2. It is however possible to modify Y by adding extra independent noise in its equation, without modifying the killed process. For $\lambda \in \mathbb{R}^m$, $|\lambda| \leq \varepsilon$ and

$\lambda' \in \mathbb{R}^d$, we can replace $\widetilde{a}_\Phi$ by

$$\widetilde{a}_\Phi(y, \lambda, \lambda') = \begin{cases} h(y)a_\Phi(y, \lambda) + (1 - h(y))(y + \lambda') & \text{if } y \in V_2, \\ y + \lambda' & \text{otherwise,} \end{cases}$$

so that

$$\widetilde{a}_\Phi(y, \lambda, \lambda') = y + h(y)\overline{a}_\Phi(y)\lambda + (1 - h(y))\lambda' + \mathrm{O}(|\lambda|^\gamma).$$

Then, letting Λ' be a d-dimensional Lévy process independent of Λ, satisfying Assumption 1 and with jumps bounded by ε, we can solve the equation (3) with coefficients $(\widetilde{a}_\Phi, \widetilde{b}_\Phi)$ and driven by $(\Lambda^\varepsilon, \Lambda')$. The advantage is that now the differential of $\widetilde{a}_\Phi$ with respect to (λ, λ') at $(0, 0)$ is now surjective onto $\mathbb{R}^d$, uniformly in y; moreover, if ε has been chosen small enough and if $|\lambda|$ and $|\lambda'|$ are $\leq \varepsilon$, the map $y \mapsto \widetilde{a}_\Phi(y, \lambda, \lambda')$ is a diffeomorphism from $\mathbb{R}^d$ onto itself, and its Jacobian determinant is uniformly positive. Consequently, we can apply Theorem 1 of [25], and deduce that Y_t has a smooth density $y \mapsto p_Y(t, y_0, y)$, with derivatives satisfying

$$\sup_{y_0, y} |D^j p_Y(t, y_0, y)| \leq C_j t^{-(d+j)/\alpha}. \tag{16}$$

Moreover, it is proved in Lemma 2 of [26] that p_Y and its derivatives up to order ℓ are actually bounded as $t \downarrow 0$ if the number of jumps necessary to go from y_0 to y is large enough; thus, for $c > 0$ fixed and if ε has been chosen small enough,

$$\sup\{|D^j p_Y(t, y_0, y)|; 0 < t \leq t_0, |y - y_0| \geq c\} \leq C_j. \tag{17}$$

If τ is the first exit time of Y from V_1, we have from

$$\mathbb{E}_{y_0}[f(Y_t)1_{\{t < \tau\}}] = \mathbb{E}_{y_0}[f(Y_t)] - \mathbb{E}_{y_0}[f(Y_t)1_{\{t \geq \tau\}}]$$

and the strong Markov property that the process Y killed at τ has density

$$q_Y(t, y_0, y) = p_Y(t, y_0, y) - \mathbb{E}_{y_0}[p_Y(t - \tau, Y_\tau, y)1_{\{t \geq \tau\}}].$$

The first term is estimated from (16) and (17), and for the second one, we notice that $Y_\tau \notin V_1$, and that the number of jumps necessary to go from V_1^c into V_0 is large if ε is small enough, so $y \mapsto p_Y(t - \tau, Y_\tau, y)$ and its derivatives up to order ℓ are bounded on V_0. We deduce the smoothness of q_Y, and, by applying Φ^{-1}, the smoothness of the law of the process X^ε killed as well as the estimates claimed in the lemma. □

Lemma 3. *Let U be a relatively compact open subset of M. The surjectivity of $\overline{a}(x)$ is only assumed on the closure of U. Consider again the solution X^ε of (1) driven by Λ^ε. If ε is small enough, then X_t^ε has a C^ℓ density $p^\varepsilon(t, x_0, x)$ on U for any $x_0 \in M$; the density and more generally its derivatives of order $j \leq \ell$ are uniformly dominated by $t^{-(d+j)/\alpha}$, for $0 < t \leq t_0$, x_0 in M and x in U.*

Proof. It is sufficient to prove the result for $U = U_0$, for open subsets $U_0 \Subset U_1 \Subset U_2 \Subset U_3 \Subset U_4 \Subset M$ such that U_4 is diffeomorphic to an open subset of $\mathbb{R}^d$, and $\overline{a}(x)$ is surjective on U_4; the subset U of the lemma can indeed be covered by a finite number of such sets U_0. Put $\tau_0 = 0$, and

$$\tau'_k = \inf\{t \geq \tau_k; X^\varepsilon_t \notin U_3\}, \qquad \tau_{k+1} = \inf\{t \geq \tau'_k; X^\varepsilon_t \in U_2\}.$$

We can associate to this sequence of stopping times an expansion for the law of X^ε_t on U_0,

$$\mathbb{P}_{x_0}\big[X^\varepsilon_t \in dx\big] = \sum_{k=0}^{\infty} \mathbb{P}_{x_0}\big[X^\varepsilon_t \in dx, \tau_k \leq t < \tau'_k\big]$$

$$= \sum_{k=0}^{\infty} \mathbb{E}_{x_0}\big[\mathbb{P}_{x_0}\big[X^\varepsilon_t \in dx, t < \tau'_k | \mathcal{F}_{\tau_k}\big] 1_{\{t \geq \tau_k\}}\big]$$

$$= \sum_{k=0}^{\infty} \mathbb{E}_{x_0}\big[Q^\varepsilon\big(t - \tau_k, X^\varepsilon_{\tau_k}, dx\big) 1_{\{t \geq \tau_k\}}\big],$$

where $\mathcal{F}_t$ is the filtration of Λ and Q^ε is the transition kernel of the process X^ε killed at the exit from U_3. From Lemma 2, this kernel has for any ℓ a C^ℓ density q^ε on U_0 if ε is small enough, so the law of X^ε_t is absolutely continuous on U_0 with density

$$p^\varepsilon(t, x_0, x) = \sum_{k=0}^{\infty} \mathbb{E}_{x_0}\big[q^\varepsilon\big(t - \tau_k, X^\varepsilon_{\tau_k}, x\big) 1_{\{t \geq \tau_k\}}\big]$$

$$= q^\varepsilon(t, x_0, x) + \sum_{k=1}^{\infty} \mathbb{E}_{x_0}\big[q^\varepsilon\big(t - \tau_k, X^\varepsilon_{\tau_k}, x\big) 1_{\{t \geq \tau_k\}}\big].$$

$$(18)$$

We already know from Lemma 2 that the first term (which is 0 if $x_0 \notin U_3$) and its derivatives are dominated by $t^{-(d+j)/\alpha}$. Moreover, we can choose ε small enough so that the process cannot jump from U_2^c into U_1, and in this case $X^\varepsilon_{\tau_k} \notin U_1$ for $k \geq 1$; in particular $(X^\varepsilon_{\tau_k}, x)$ remains out of a neighbourhood of the diagonal of $M \times M$ for $x \in U_0$. Thus, $q^\varepsilon(t - \tau_k, X^\varepsilon_{\tau_k}, x)$ and its derivatives are uniformly bounded on U_0 for $k \geq 1$ (third assertion of Lemma 2). Thus, the proof of the lemma is complete from

$$\big|D^j p^\varepsilon(t, x_0, x)\big| \leq C_j\left(1_{U_3}(x_0) t^{-(d+j)/\alpha} + \sum_{k=1}^{\infty} \mathbb{P}_{x_0}[t \geq \tau_k]\right) \tag{19}$$

as soon as we prove that the series converges and is bounded. We have similarly to (15) that $\mathbb{P}_x[\tau'_0 \geq u] \geq 1/2$ for $x \in U_2$ if u is small enough, so by applying the strong Markov property of X,

$$\mathbb{P}_{x_0}\big[\tau'_k > u | \mathcal{F}_{\tau_k}\big] \geq \mathbb{P}_{x_0}\big[\tau'_k - \tau_k \geq u | \mathcal{F}_{\tau_k}\big] \geq 1/2$$

for $k \geq 1$ on $\{\tau_k < \infty\}$. Thus,

$$\mathbb{P}_{x_0}\big[X_t^\varepsilon \in U_3\big] \geq \sum_{k=1}^{\infty} \mathbb{P}_{x_0}\big[\tau_k \leq t < \tau_k'\big] = \sum_{k=1}^{\infty} \mathbb{E}_{x_0}\big[\mathbb{P}_{x_0}\big[\tau_k' > t|\mathcal{F}_{\tau_k}\big]1_{\{t \geq \tau_k\}}\big]$$

$$\geq \frac{1}{2}\sum_{k=1}^{\infty} \mathbb{P}_{x_0}[t \geq \tau_k]$$

for $t \leq u$, so the series in (19) is bounded, and we have

$$\big|D^j p^\varepsilon(t, x_0, x)\big| \leq C_j\big(1_{U_3}(x_0)t^{-(d+j)/\alpha} + 2\mathbb{P}_{x_0}\big[X_t^\varepsilon \in U_3\big]\big)$$

for $t \leq u$. If $t > u$, we use the Markov property, write

$$p^\varepsilon(t, x_0, x) = \mathbb{E}_{x_0}\big[p^\varepsilon\big(u, X_{t-u}^\varepsilon, x\big)\big]$$

and deduce

$$\begin{aligned}\big|D^j p^\varepsilon(t, x_0, x)\big| &\leq C_j\big(\mathbb{P}_{x_0}\big[X_{t-u}^\varepsilon \in U_3\big]t^{-(d+j)/\alpha} + 2\mathbb{P}_{x_0}\big[X_t^\varepsilon \in U_3\big]\big) \\ &\leq C_j't^{-(d+j)/\alpha}\mathbb{P}_{x_0}\big[X_t^\varepsilon \in U_4\big]\end{aligned} \tag{20}$$

by using the fact that $\mathbb{P}[X_t^\varepsilon \in U_4|X_{t-u}^\varepsilon] \geq 1/2$ on $\{X_{t-u}^\varepsilon \in U_3\}$ if u has been chosen small enough. $\qquad\square$

We have proved the estimate (20) which is more precise than the statement of the lemma. This property will be used in Section 4.

The absolute continuity of the law of X_t claimed in Theorem 2 follows easily from Lemma 3, by conditioning on the last jump of Λ before time t greater than ε. If τ is this last jump (we put $\tau = 0$ when there is no big jump), then we deduce that X has a density given by

$$p(t, x_0, x) = \mathbb{E}_{x_0}\big[p^\varepsilon(t - \tau, X_\tau, x)\big]. \tag{21}$$

However, this formula is not sufficient to obtain the smoothness and even the local boundedness of p, because $p^\varepsilon(t - \tau, X_\tau, x)$ is of order $(t - \tau)^{-d/\alpha}$, at least when X_τ and x are close to each other, and $(t - \tau)^{-d/\alpha}$ is not integrable if $d \geq \alpha$.

3.3. The case with big jumps

In Lemma 3, we have proved the smoothness of the law when big jumps of Λ have been removed. We now have to take into account the effect of these big jumps. Notice that the following lemma completes the proof of Theorem 2 when M is compact (choose $U = M$).

Lemma 4. *Let U be a relatively compact open subset of M. There exists a Markov process Y_t such that the laws of X and Y killed at the exit from U coincide, and Y_t has on U a C^ℓ density $p_Y(t, y_0, y)$, the derivatives of which satisfy*

$$\left| D^j p_Y(t, y_0, y) \right| \le C_j t^{-(d+j)/\alpha} \tag{22}$$

for $0 < t \le t_0$, y_0 in M and y in U.

Proof. For any $\varepsilon > 0$, we have a decomposition $\mathcal{L} = (\mathcal{L} - \mathcal{L}^\varepsilon) + \mathcal{L}^\varepsilon$ of the infinitesimal generator, where $\mathcal{L} - \mathcal{L}^\varepsilon$ is the generator of the process X^ε driven by Λ^ε, and

$$\mathcal{L}^\varepsilon f(x) = \int_{\{|\lambda| > \varepsilon\}} \left(f\big(a(x, \lambda)\big) - f(x) \right) \mu(\mathrm{d}\lambda).$$

Choose $U \Subset U' \Subset U'' \Subset M$. Fix a Riemannian metric on M and consider the Riemannian exponential function; then its inverse $\exp_x^{-1} y \in T_x M$ is well defined and smooth if y is close to x. Thus, if ε is small enough, we can consider $a_0(x, \lambda) = \exp_x^{-1} a(x, \lambda)$ for $x \in U''$ and $|\lambda| \le \varepsilon$. Let $h_0 : M \to [0, 1]$ be a smooth function such that $h_0 = 1$ on U' and $h_0 = 0$ on U''^c, and let

$$\widetilde{a}(x, \lambda) = \exp_x \big(h_0(x) a_0(x, \lambda)\big), \qquad \widetilde{b}(x) = h_0(x) b(x)$$

for $|\lambda| \le \varepsilon$. Let $\widetilde{\mathcal{L}}^\varepsilon$ be the infinitesimal generator $\mathcal{L} - \mathcal{L}^\varepsilon$ where coefficients (a, b) have been replaced by $(\widetilde{a}, \widetilde{b})$; this corresponds to a process $\widetilde{X}$ driven by Λ^ε, which is interpolated between X^ε and the motionless process (this is similar to the construction of Lemma 1 but we are here on the manifold instead of $\mathbb{R}^d$); it satisfies Assumption 2, except the surjectivity condition which does not hold out of U'' but holds on the closure of U'. On the other hand, let $h : M \to [0, 1]$ be a smooth function such that $h = 1$ on U and $h = 0$ on U'^c, and define

$$\overline{\mathcal{L}}^\varepsilon f(x) = \int_{\{|\lambda| > \varepsilon\}} \left(f\big(a(x, \lambda)\big) h\big(a(x, \lambda)\big) h(x) - f(x) \right) \mu(\mathrm{d}\lambda).$$

This means that we consider the jumps λ of Λ greater than ε; if the process is at a point x before this jump, we kill it with probability $1 - h(x)$; if it is not killed, it jumps to $x_1 = a(x, \lambda)$ and is killed with probability $1 - h(x_1)$. We let Y be the process with generator $\widetilde{\mathcal{L}}^\varepsilon + \overline{\mathcal{L}}^\varepsilon$; this is the process $\widetilde{X}$ interlaced with jumps described by $\overline{\mathcal{L}}^\varepsilon$. This process enters our framework from Section 2.7, and X and Y coincide when killed at the exit from U.

From Lemma 3, if ε is small enough, the process $\widetilde{X}$ has on U' a C^ℓ density $\widetilde{p}(t, x_0, x)$ with respect to the Riemannian measure. On the other hand, there are N_t jumps $\Delta \Lambda_{\tau_k}$ greater than ε on the time interval $[0, t]$, for a random $\tau = (\tau_1, \tau_2, \ldots)$; we also let $\tau_0 = 0$ and append a last $\tau_{N_t+1} = t$. Let K be the random index k such that $\tau_{k+1} - \tau_k$ is maximal. Then

$$\mathbb{P}[Y_{\tau_{K+1}-} \in \mathrm{d}y | \tau; \Lambda_s, 0 \le s \le \tau_K] = \widetilde{p}(\tau_{K+1} - \tau_K, Y_{\tau_K}, y) \, \mathrm{d}y$$

on U', with

$$\left| D^j \widetilde{p}(\tau_{K+1} - \tau_K, Y_{\tau_K}, y) \right| \le C_j (\tau_{K+1} - \tau_K)^{-(d+j)/\alpha}. \tag{23}$$

On $\{K = N_t\}$, we obtain the conditional density of Y_t. Otherwise, we have to apply the jump at τ_{K+1} to this distribution; we first kill the process with probability $1 - h(y)$ and therefore get a C^ℓ law on M supported by U'; from Assumption 2, this law is then transformed by $a_\star^\varepsilon$ into a C^ℓ law on M, which is restricted into a C^ℓ law supported by U' by the second killing. We therefore obtain

$$\mathbb{P}[Y_{\tau_{K+1}} \in dy | \tau; \Lambda_s, 0 \le s \le \tau_K] = p_\star(y) \, dy$$

for a conditional density $p_\star$ which is C^ℓ, with derivatives dominated as in (23).

This density is then propagated from τ_{K+1} to $(\tau_{K+2})-$ by means of the semigroup of $\widetilde{X}$ with generator $\widetilde{\mathcal{L}}^\varepsilon$; if ε is small enough, then $x \mapsto \widetilde{a}(x, \lambda)$ are diffeomorphisms of M onto itself for $|\lambda| \le \varepsilon$, and the process $\widetilde{X}$ can be written as $\widetilde{X}_t = \Phi_t(\widetilde{X}_0)$ for a flow of diffeomorphisms Φ_t of M onto itself; the technique of [12] for compact manifolds can be adapted to our case since the process is motionless out of a compact part of M. We choose a copy of Φ which is independent of Λ, and obtain

$$\mathbb{P}[Y_{\tau_{K+2}-} \in dy | \tau; \Lambda_s, 0 \le s \le \tau_K] \tag{24}$$
$$= \widetilde{\mathbb{E}}\big[p_\star\big(\Phi_{\tau_{K+2}-\tau_{K+1}}^{-1}(y)\big) \big|\det\big(D\Phi_{\tau_{K+2}-\tau_{K+1}}^{-1}\big)(y)\big|\big] \, dy,$$

where $\widetilde{\mathbb{E}}$ is the expectation only with respect to Φ, and the determinant is computed relatively to orthonormal bases on the tangent spaces. The differential of Φ_t is solution of

$$D\Phi_{t+dt}(x) = D\widetilde{a}\big(\Phi_t(x), d\widetilde{\Lambda}_t\big)D\Phi_t(x) + D\widetilde{b}\big(\Phi_t(x)\big)D\Phi_t(x) \, dt, \qquad D\Phi_0(x) = I.$$

If M is embedded in $\mathbb{R}^N$, this can be transformed into an equation on $\mathbb{R}^N$ and it is a standard procedure to prove that $\sup_{t \le T} |D\Phi_t(x)|$ has bounded moments, uniformly in x: use the technique of [13]. By differentiating this equation, the same property holds true for higher order derivatives, so actually $\sup_{t,x} |D^j \Phi_t(x)|$, $t \le t_0$, has bounded moments. The same property can be verified for the derivatives of the inverse map Φ_t^{-1}, by looking at the equation of its derivative. Thus, in (24), we obtain an estimate on the derivatives of the conditional density of $Y_{\tau_{K+2}-}$ similar to (23). By iterating this procedure on all the subsequent jumps τ_k, we prove that Y_t has a conditional density; the unconditioned density p_Y is then obtained by taking the expectation and satisfy

$$\big|D^j p_Y(y)\big| \le C\mathbb{E}\big[e^{CN_t}(\tau_{K+1} - \tau_K)^{-(d+j)/\alpha}\big]$$

on U', because each jump at τ_k and each use of the flow Φ between τ_k and τ_{k+1} appends a multiplicative constant in the estimation, and they are at most N_t of these jumps. The number N_t of big jumps is a Poisson variable so has finite exponential moments. Moreover,

$$\tau_{K+1} - \tau_K \ge \frac{t}{N_t + 1},$$

so

$$(\tau_{K+1} - \tau_K)^{-(d+j)/\alpha} \le (N_t + 1)^{(d+j)/\alpha} t^{-(d+j)/\alpha}$$

and we can conclude. $\qquad\square$

We now give an estimation of the probability for X_t to be in some relatively compact open subset when one needs many jumps to come from the initial condition.

Lemma 5. *Consider open subsets of M with relatively compact inclusions $U_n \Subset V_n \Subset U_{n+1}$, and suppose that X cannot jump from U_{n+1}^c into V_n. Let $n \geq 1$. Then $\mathbb{P}_{x_0}[X_t \in U_0]$ is $O(t^n)$ as $t \downarrow 0$ uniformly for x_0 in V_{n-1}^c.*

Proof. A similar result was proved in Lemma 1 of [26] for $\mathbb{R}^d$; the proof of this variant is much simpler. Let h_k, $0 \leq k \leq n-1$, be a smooth function with values in $[0, 1]$, which is 1 on U_k and 0 on V_k^c. Let $C_k = \sup |\mathcal{L}h_k|$ for the generator $\mathcal{L}$ of X. Then $h_k(x_0) = 0$ for $x_0 \in V_{n-1}^c \subset V_k^c$, and $\mathcal{L}h_k(x) = 0$ for $x \in U_{k+1}^c$, so

$$\mathbb{P}_{x_0}[X_t \in U_k] \leq \mathbb{E}_{x_0}\big[h_k(X_t)\big] = \mathbb{E}_{x_0}\int_0^t \mathcal{L}h_k(X_s)\,\mathrm{d}s \leq C_k t \sup_{0 \leq s \leq t} \mathbb{P}_{x_0}[X_s \in U_{k+1}].$$

Applying this inequality for $k = 0, 1, \ldots, n-1$ completes the proof. $\qquad\square$

Proof of Theorem 2. The absolute continuity has already been proved in (21), the smoothness has been obtained in Lemma 4 in the compact case, so we now have to study the smoothness of the density in the non compact case. Let U_0 be a relatively compact open subset of M which is diffeomorphic to an open subset of $\mathbb{R}^d$, and let $K \subset U_0$ be compact. Under Assumption 3, there exists a sequence (U_n, V_n) satisfying the conditions of Lemma 5 starting from the given U_0. For ε and n which will be chosen respectively small and large enough, there also exists a process Y constructed in Lemma 4 for $U = U_{n+1}$, with a C^ℓ density $p_Y(t, y_0, y)$ on U_{n+1}. The law of Y killed at the exit τ from U_{n+1} has on U_0 the density

$$q_Y(t, y_0, y) = p_Y(t, y_0, y) - \mathbb{E}_{y_0}\big[p_Y(t - \tau, Y_\tau, y)1_{\{\tau < t\}}\big]. \tag{25}$$

Consider the process Y when it is out of U_{n+1}; its big jumps (coming from $\overline{\mathcal{L}}^\varepsilon$) are included in jumps of X, so by construction of the sequence (U_n, V_n), it is not possible to jump directly from U_{n+1}^c into V_n; the small jumps (coming from $\widetilde{\mathcal{L}}^\varepsilon$) have been modified, but they are small, so if ε is small enough, it is again not possible to jump directly from U_{n+1}^c into V_n. On the other hand, Y coincides with X on U_{n+1} and has therefore the same jumps, so it cannot jump from U_{k+1}^c into V_k for $k < n$. Thus, we can apply Lemma 5 to Y and deduce that $\mathbb{P}_{y_0}[Y_t \in U_0]$ is $O(t^n)$ for $y_0 \notin V_{n-1}$; we also have the uniform estimates (22), and we can deduce as in Lemma 2 of [26] that if n has been chosen large enough, then $D^j p_Y(t, y_0, y)$ is uniformly bounded for $0 < t \leq t_0$, $y \in K$ and $y_0 \notin V_{n-1}$ (this result was proved on $\mathbb{R}^d$ but U_0 can be viewed as a subset of $\mathbb{R}^d$). These estimates on p_Y imply by means of (25) that $D^j q_Y(t, y_0, y)$ is $O(t^{-(d+j)/\alpha})$ uniformly for $0 < t \leq t_0$, $y_0 \in U_{n+1}$, $y \in K$, and is bounded if $y_0 \in U_{n+1} \setminus V_{n-1}$.

The killed processes X and Y coincide, and the smoothness of the non killed process X is then deduced as in the proof of Lemma 3 by considering successive exits from U_{n+1} and entrances into V_n; we obtain an expansion similar to (18). When the process enters V_n, it is at a point out of V_{n-1}, so we have the uniform boundedness of the derivatives of q_Y starting from this point, and we can proceed and estimate the series as in Lemma 3. $\qquad\square$

Proof of Theorem 3. The decomposition $\mu = \mu^\flat + \mu^\sharp$ of the Lévy measure corresponds to a decomposition $\Lambda_t = \Lambda_t^\flat + \Lambda_t^\sharp$ of the Lévy process into independent Lévy processes, where $\Lambda^\sharp$ is of pure jump type. We can apply Theorem 2 and deduce that the process $X^\flat$ driven by $\Lambda^\flat$ has a smooth density $p^\flat$. Let τ be the time of the last jump of $\Lambda^\sharp$ before t, with $\tau = 0$ when $\Lambda^\sharp$ has no jump on $[0, t]$; this last event has probability $\exp -t\mu^\sharp(\mathbb{R}^m)$. Similarly to (21), the density of X_t can be written as

$$p(t, x_0, x) = \mathbb{E}_{x_0}\big[p^\flat(t - \tau, X_\tau, x)\big]$$
$$= p^\flat(t, x_0, x)\exp\big(-t\mu^\sharp(\mathbb{R}^m)\big) + \mathbb{E}_{x_0}\big[p^\flat(t - \tau, X_\tau, x)1_{\{\tau>0\}}\big].$$

The first term is smooth. On the event $\{\tau > 0\}$, it follows from the assumption of the theorem about $\mu_x^\sharp$ that the conditional law of X_τ given $\Lambda^\flat$ has a smooth density p_0 on M, so

$$\mathbb{E}_{x_0}\big[p^\flat(t - \tau, X_\tau, x)1_{\{\tau>0\}}\big] = \mathbb{E}_{x_0}\bigg[1_{\{\tau>0\}}\int p^\flat(t - \tau, z, x)p_0(z)\,\mathrm{d}z\bigg].$$

We want to verify that the smoothness of p_0 is preserved by $p^\flat(t - \tau, \cdot, \cdot)$. We know from the proof of Theorem 2 that $p^\flat(t - \tau, z, \cdot)$ is C_b^ℓ on K if the initial condition z is out of a large enough subset, so it is sufficient to consider the case of a smooth p_0 with compact support. It also follows from previous proof that it is sufficient to prove the result for the modified process Y, and the propagation of the smoothness of the law by the semigroup of Y has been obtained in the proof of Lemma 4. $\qquad\square$

4. Lie groups and homogeneous spaces

In Theorem 2, we have assumed that the process cannot come from infinity by jumping (except in the case of smooth jumps of Theorem 3). This assumption is not needed in the affine case $M = \mathbb{R}^d$, but other conditions are required in this case, and assumptions made for instance in [24] are not intrinsic if $\mathbb{R}^d$ is viewed as a differentiable manifold (they are not invariant by diffeomorphisms). We can of course apply the affine theorem when M is diffeomorphic to $\mathbb{R}^d$, but again the conditions on the coefficients will depend on the diffeomorphism.

Thus, if jumps are not bounded, we need additional structure on M (as this is the case on $\mathbb{R}^d$ where the affine structure is used). We consider here the case of Lévy processes on Lie groups; more generally, we consider the case where M is an homogeneous space on which a Lie group G acts, and X_t is the projection on M of a Lévy process on G.

Exposition about Lévy processes on Lie groups can be found in [19]. We shall use the theory of integration on Lie groups or homogeneous spaces (Haar measures, invariant and relatively invariant measures), which is explained in [6], see also for instance [28]. We recall here the points which are useful for our study.

4.1. Lie groups

Let $M = G$ be a d-dimensional Lie group with neutral element e and Lie algebra $\mathfrak{g}$; as a vector space, $\mathfrak{g}$ is the tangent space T_eG; it can be identified to the space of left invariant vector fields

on G; the Lie bracket of two elements of $\mathfrak{g}$, as well as the exponential map $\exp\colon \mathfrak{g} \to G$ can be constructed from this identification. We can choose as a smooth reference measure on G a left Haar measure $\mathcal{H}^G_\leftarrow$, or a right Haar measure $\mathcal{H}^G_\rightarrow$; each of them is unique modulo a multiplicative constant; they satisfy

$$\mathcal{H}^G_\leftarrow(gA) = \mathcal{H}^G_\leftarrow(A), \qquad \mathcal{H}^G_\leftarrow(Ag^{-1}) = \chi_G(g)\mathcal{H}^G_\leftarrow(A),$$

$$\mathcal{H}^G_\rightarrow(Ag) = \mathcal{H}^G_\rightarrow(A), \qquad \mathcal{H}^G_\rightarrow(gA) = \chi_G(g)\mathcal{H}^G_\rightarrow(A) \tag{26}$$

for a group homomorphism $\chi_G\colon G \to \mathbb{R}^\star_+$ which is the modulus of G. If we are given $\mathcal{H}^G_\leftarrow$, we can define $\mathcal{H}^G_\rightarrow$ by

$$\mathcal{H}^G_\rightarrow(dg) = \chi_G(g)\mathcal{H}^G_\leftarrow(dg). \tag{27}$$

The group G is said to be unimodular if $\chi_G \equiv 1$; this holds true when G is compact because any group homomorphism from G into $\mathbb{R}^\star_+$ must be equal to 1. Let $\mathrm{Ad}\colon G \to GL(\mathfrak{g})$ be the adjoint representation of G; this means that Ad_g is the differential at e of the inner automorphism $x \mapsto gxg^{-1}$. Then

$$\chi_G(g) = |\det \mathrm{Ad}_g|. \tag{28}$$

On G there are two differential calculi, a left invariant one and a right invariant one. The left and right invariant derivatives are the linear forms

$$D_\leftarrow f(g)u = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} f\big(g\exp(\varepsilon u)\big), \qquad D_\rightarrow f(g)u = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} f\big(\exp(\varepsilon u)g\big)$$

for smooth functions $f\colon G \to \mathbb{R}$ and $u \in \mathfrak{g}$. The invariance means that $D_\leftarrow L_h f = L_h D_\leftarrow f$ and $D_\rightarrow R_h f = R_h D_\rightarrow f$, with the notations $L_h f(g) = f(hg)$ and $R_h f(g) = f(gh)$. The left and right invariant derivatives are related to each other by

$$D_\rightarrow f(g)u = D_\leftarrow f(g)\mathrm{Ad}^{-1}_g u. \tag{29}$$

If we choose an inner product on $\mathfrak{g}$, we can consider the norms $|D_\leftarrow f(g)|$ and $|D_\rightarrow f(g)|$ of the linear forms, and norms corresponding to different inner products are of course equivalent. We can also consider classes $C^\ell_{b,\leftarrow}$ or $C^\ell_{b,\rightarrow}$ of functions for which the left or right invariant derivatives are bounded (without boundedness we can simply use the notation C^ℓ since the classes for the left and right calculi coincide).

Let X_t be a left Lévy process on G with initial condition $X_0 = e$. This is a strong Markov process which is invariant under left multiplication, so that its semigroup satisfies $P_t L_h = L_h P_t$; equivalently, the infinitesimal generator should satisfy $\mathcal{L}L_h = L_h\mathcal{L}$. For $0 \leq s \leq t$, the variable $X_s^{-1}X_t$ must be independent of $(X_u; 0 \leq u \leq s)$ and must have the same law as X_{t-s}. We consider here the subclass of Lévy processes without Brownian part. Let V be a relatively compact neighbourhood of e in G which is diffeomorphic to a neighbourhood U of 0 in $\mathfrak{g}$ by means of the Lie exponential function. Then a left Lévy process without Brownian part is characterised by a drift $\kappa \in \mathfrak{g}$ and a Lévy measure μ_X on $G \setminus \{e\}$ which integrates smooth bounded functions f

such that $f(e) = D_{\leftarrow}f(e) = 0$; the infinitesimal generator of X can be written in the Hunt form as

$$\mathcal{L}f(g) = D_{\leftarrow}f(g)\kappa + \int_V \left(f(gx) - f(g) - D_{\leftarrow}f(g)\exp^{-1}x\right)\mu_X(dx) \qquad (30)$$

$$+ \int_{V^c} \left(f(gx) - f(g)\right)\mu_X(dx).$$

It is explained in [2] that X can be viewed as the solution of an equation driven by a Poisson measure on $\mathbb{R}_+ \times G$; by means of a technique similar to Section 2.7, it is also the solution of an equation of type (1) driven by a $\mathfrak{g}$-valued Lévy process. More precisely, let μ be the measure on $U \setminus \{0\}$ which is the image of $\mu_X|_V$ by $\exp^{-1}$; on the other hand, there exists a bi-measurable bijection i from V^c onto a Borel subset U' of U^c, so we can let μ be on U^c the image of $\mu_X|_{V^c}$ by i. Then

$$\mathcal{L}f(g) = D_{\leftarrow}f(g)\kappa + \int_U \left(f(g\exp\lambda) - f(g) - D_{\leftarrow}f(g)\lambda\right)\mu(d\lambda)$$

$$+ \int_{U'} \left(f\left(gi^{-1}(\lambda)\right) - f(g)\right)\mu(d\lambda).$$

Thus X_t can be viewed as the solution of $X_t = a(X_{t-}, d\Lambda_t)$, where Λ_t is a Lévy process in the vector space $\mathfrak{g}$ with Lévy measure μ supported by $U \cup U'$, and

$$a(g, \lambda) = \begin{cases} g\exp\lambda & \text{if } \lambda \in U, \\ gi^{-1}(\lambda) & \text{if } \lambda \in U'. \end{cases}$$

Then $\overline{a}(g): \mathfrak{g} \to T_g(G)$ is the map which sends u to the value at g of the left invariant vector field associated to u, so it is bijective, and it is also not difficult to verify the second part of Assumption 2. The nondegeneracy condition of Assumption 1 on μ is immediately transferred to an assumption on μ_X as

$$c\rho^{2-\alpha}|u|^2 \le \int_{\{|\exp^{-1}x|\le\rho\}} \left\langle\exp^{-1}x, u\right\rangle^2 \mu_X(dx) \le C\rho^{2-\alpha}|u|^2 \qquad (31)$$

for $u \in \mathfrak{g}$ and $\rho \le \rho_0$ small enough so that $\exp^{-1}$ is well defined; the additional condition in the case $\alpha = 1$ is written as

$$\limsup_{\varepsilon\downarrow0}\left|\int_{\{\varepsilon\le|\exp^{-1}x|\le\rho_0\}} \exp^{-1}x\,\mu_X(dx)\right| < \infty. \qquad (32)$$

These two conditions do not depend on the choice of the inner product on $\mathfrak{g}$. If $\alpha < 1$, we also suppose that the process is pure jump, so that (30) becomes

$$\mathcal{L}f(g) = \int_G \left(f(gx) - f(g)\right)\mu_X(dx).$$

For the first part of Assumption 2, we have to study the propagation of the smoothness of the measure by a right translation; this is easy if we choose a right Haar measure as a reference measure.

Thus, under these conditions, we can apply Theorem 2 and deduce that X_t has a smooth density if μ_X has compact support. If the support is not compact, a possibility is to use Theorem 3. Otherwise, we can use the following result.

Theorem 4. *Let X_t be a left Lévy process on G with $X_0 = e$, the Lévy measure μ_X of which satisfies* (31), *and the additional condition* (32) *if $\alpha = 1$. If $\alpha < 1$, suppose moreover that X_t is a pure jump process. Then the law of X_t, $t > 0$, is absolutely continuous with respect to the left Haar measure $\mathcal{H}_{\leftarrow}^G$. Let $\ell \geq 0$. If*

$$\int_{V^c} \chi_G(g) |\mathrm{Ad}_g|^j \mu_X(\mathrm{d}g) < \infty \tag{33}$$

for a relatively compact neighbourhood V of e and for $j \leq \ell$, then the density is in $C_{b,\leftarrow}^\ell$. In particular, if

$$\int_{V^c} |\mathrm{Ad}_g|^j \mu_X(\mathrm{d}g) < \infty \tag{34}$$

for any j, the density is in $C_{b,\leftarrow}^\infty$.

Proof. The absolute continuity follows from Theorem 2. Assume now (33) for $j = 0$. Let V be a relatively compact neighbourhood of e, and let $\mu_X^\flat$ and $\mu_X^\sharp$ be the restrictions of μ_X to V and V^c; as usually, X can be written on $[0, t]$ as a Lévy process $X^\flat$ with Lévy measure $\mu_X^\flat$ interlaced with N_t big jumps at times τ_k described by $\mu_X^\sharp$. We proceed as in the proof of Lemma 4 and let τ_{K+1} be the end of the longest subinterval of $[0, t]$ without big jump. Then, from Theorem 2, conditionally on (τ_k), the variable $Y = X_{\tau_{K+1}-}$ has with respect to the left Haar measure a C^ℓ density p_Y satisfying

$$\left| D_{\leftarrow}^j p_Y(y) \right| \leq C_j (\tau_{K+1} - \tau_K)^{-(d+j)/\alpha} \tag{35}$$

on V, and the right-hand side is integrable as in Lemma 4. But Theorem 2 also states that the estimate is uniform with respect to the initial condition, so the same estimate holds true for any gY, and one deduces that (35) holds uniformly on G. Then, we have $X_t = YZ$ where $Z = X_{\tau_{K+1}-}^{-1} X_t$ is, conditionally on (τ_k), independent of Y. Conditionally on Z and (τ_k), the variable X_t is therefore absolutely continuous with density

$$p(x|Z) = \chi_G(Z) p_Y(x Z^{-1}) \tag{36}$$

(this follows from (26)). Thus, by integrating this formula and applying (35), we deduce that the density of X_t is bounded and continuous as soon as $\chi_G(Z)$ is integrable. On the other hand,

$$\chi_G(Z) = \chi_G\left((X_{\tau_{K+1}}^\flat)^{-1} X_t^\flat \right) \prod_{k=K+1}^{N_t} \chi_G\left(X_{\tau_k-}^{-1} X_{\tau_k} \right),$$

where the different terms are conditionally independent given (τ_k); the process $\chi_G(X_t^\flat)$ is a geometric Lévy process with bounded jumps, so the first term has bounded conditional expectation, and we deduce from (33) for $j = 0$ that the conditional expectation of $\chi_G(Z)$ is bounded by some exponential of N_t, so $\chi_G(Z)$ is integrable and the case $\ell = 0$ is proved.

We can differentiate (36) and get

$$D_\leftarrow p(x|Z)u = \chi_G(Z)D_\leftarrow p_Y\big(xZ^{-1}\big)\mathrm{Ad}_Z(u)$$

for $u \in \mathfrak{g}$. For higher order derivatives, we have

$$D_\leftarrow^j p(x|Z)(u_1,\ldots,u_j) = \chi_G(Z)D_\leftarrow^j p_Y\big(xZ^{-1}\big)\big(\mathrm{Ad}_Z(u_1),\ldots,\mathrm{Ad}_Z(u_j)\big).$$

We deduce the smoothness of the law of X_t and the boundedness of its derivatives if we prove that $\chi_G(Z)|\mathrm{Ad}_Z|^j$ is integrable. To this end, we notice that

$$\chi_G(Z)|\mathrm{Ad}_Z|^j \leq \prod_{k=K+1}^{N_t} \big(\chi_G\big(\big(X_{\tau_k}^\flat\big)^{-1}X_{\tau_{k+1}}^\flat\big)|\mathrm{Ad}_{((X_{\tau_k}^\flat)^{-1}X_{\tau_{k+1}}^\flat)}|^j\big)$$

$$\times \prod_{k=K+1}^{N_t} \big(\chi_G\big(X_{\tau_k-}^{-1}X_{\tau_k}\big)|\mathrm{Ad}_{(X_{\tau_k-}^{-1}X_{\tau_k})}|^j\big)$$

where the different terms are again conditionally independent given (τ_k); the integrability is deduced similarly to the case $j = 0$ by using (33) for the terms of the second product. The last claim of the theorem follows from (28). $\square$

Condition (33) is trivially satisfied for $j = 0$ when G is unimodular, so in this case we obtain the existence of a continuous bounded density without any assumption on the big jumps.

Notice also that (34) is related to the existence of exponential moments for big jumps. Assume that μ_X is the image of a measure μ on $\mathfrak{g}$ by the Lie exponential (this holds for instance when the exponential is surjective). For $\lambda \in \mathfrak{g}$, let $\mathrm{ad}_\lambda : \mathfrak{g} \to \mathfrak{g}$ be the adjoint action given by $\mathrm{ad}_\lambda(u) = [\lambda, u]$ for the Lie bracket $[\cdot, \cdot]$. We have $\mathrm{Ad}_{\exp(\lambda)} = \exp(\mathrm{ad}_\lambda)$, so

$$|\mathrm{Ad}_{\exp(\lambda)}| \leq \exp\big(|\mathrm{ad}_\lambda|\big) \leq \exp\big(c|\lambda|\big),$$

and (34) is satisfied if

$$\int_{\{|\lambda|>1\}} \exp\big(C|\lambda|\big)\mu(\mathrm{d}\lambda) < \infty$$

for any C. This is however the worst case. If G is nilpotent, then the expansion of $\exp(\mathrm{ad}_\lambda)$ is finite, and exponential moments can be replaced by ordinary moments; notice however that the class of stable processes introduced by [16] on simply connected nilpotent groups does not enter our framework, because (31) is not satisfied.

Theorem 4 can of course be translated into the case of right Haar measure, right invariant derivatives, and right Lévy processes (invariant by right multiplication). Then the conditions

(33) and (34) are replaced by conditions on $\chi_G(g)^{-1}$ and $|\mathrm{Ad}_g^{-1}|$. On the other hand, left and right Lévy processes with the same infinitesimal generator at e have the same law at any fixed time t, because the right Lévy process Y can be deduced from the left process X on $[0, t]$ by the formula $Y_s = X_{t-s}^{-1} X_t$. Thus, in order to study the law of $X_t = Y_t$, one can choose between left and right calculi. Notice however that if for instance we apply the result for the right Lévy process, we obtain that the density $p_\rightarrow$ of X_t with respect to $\mathcal{H}_\rightarrow^G$ is of class $C_{b,\rightarrow}^\ell$; from (27), the relation between the densities with respect to the left and right Haar measures is $p_\rightarrow = \chi_G^{-1} p_\leftarrow$; the density $p_\leftarrow$ is of class C^ℓ, but not necessarily of class $C_{b,\leftarrow}^\ell$ (the relation between left and right invariant derivatives is given in (29)).

4.2. Homogeneous spaces

We now consider the case where the manifold M is an homogeneous space $M = G/H$. More precisely, G is a m-dimensional Lie group of transformations acting transitively and smoothly on the left on M, and $H = \{h \in G; h(o) = o\}$ is the isotropy group of some fixed point o of M; the projection $\pi : G \to M$ is given by $\pi(g) = g(o)$. We can choose a Lebesgue measurable section S of π^{-1} (which exists from the measurable section theorem), and any g in G can then be uniquely written as $g = S_x h$ for $x = \pi(g) = g(o)$ and some h in H. The action of G on M can be written as $g(y) = \pi(g S_y)$. We will denote by $\mathfrak{g}$ and $\mathfrak{h}$ the Lie algebras of G and H.

We can look for a measure on M which would be invariant under the action of G, but such a measure does not always exist. We therefore weaken the invariance into a relative invariance property, see [6,28] for an introduction to the topic and some of the properties which are given below; we say that a Radon non-identically zero measure $\mathcal{H}^M$ is relatively invariant under the action of G with multiplier χ if $\chi : G \to \mathbb{R}_+^\star$ is a group homomorphism and

$$\mathcal{H}^M\big(g(A)\big) = \chi(g)\mathcal{H}^M(A).$$

Then the measure is invariant if $\chi \equiv 1$; this is necessarily the case when G is compact. For instance, we have seen in (26) that the right Haar measure on G is relatively invariant under the left multiplication with multiplier χ_G. A relatively invariant measure also does not always exist, but it exists in more general situations than invariant measures. It exists on $M = G/H$ if and only if $\chi_G/\chi_H : H \to \mathbb{R}_+^\star$ can be extended to a group homomorphism $\chi : G \to \mathbb{R}_+^\star$; in this case there exist a relatively invariant measure with multiplier χ, and this measure is unique modulo a multiplicative constant. In particular, an invariant measure exists if and only if $\chi_G = \chi_H$ on H; this property holds in particular when H is compact.

The relationship between left Haar measures on G and H and relatively invariant measures on M is the following one. If $\mathcal{H}_\leftarrow^H$ is a left Haar measure on H and if $\chi : G \to \mathbb{R}_+^\star$ is a group homomorphism, then a Radon measure $\mathcal{H}^M$ on M is relatively invariant with multiplier χ if and only if the measure $\mathcal{H}_\leftarrow^G$ defined on G by

$$\int_G f(g)\mathcal{H}_\leftarrow^G(\mathrm{d}g) = \int\int_{M \times H} \frac{f}{\chi}(S_x h)\mathcal{H}^M(\mathrm{d}x)\mathcal{H}_\leftarrow^H(\mathrm{d}h) \tag{37}$$

is a left Haar measure. For right Haar measures, (37) becomes (use (27))

$$\int_G f(g)\mathcal{H}^G_{\rightarrow}(dg) = \int\int_{M\times H} f(S_x h)\frac{\chi_G}{\chi}(S_x)\mathcal{H}^M(dx)\mathcal{H}^H_{\rightarrow}(dh). \tag{38}$$

If Ξ is a G-valued variable with densities $p_{\leftarrow}$ and $p_{\rightarrow}$ with respect to $\mathcal{H}^G_{\leftarrow}$ and $\mathcal{H}^G_{\rightarrow}$, we deduce from (37) and (38) that $\pi(\Xi)$ has density

$$p(x) = \int_H \frac{p_{\leftarrow}}{\chi}(S_x h)\mathcal{H}^H_{\leftarrow}(dh) = \frac{\chi_G}{\chi}(S_x)\int_H p_{\rightarrow}(S_x h)\mathcal{H}^H_{\rightarrow}(dh) \tag{39}$$

with respect to $\mathcal{H}^M$. Estimates on p can therefore be deduced from estimates on $p_{\leftarrow}$ or $p_{\rightarrow}$, but this is clearly simpler when H is compact.

We now explain how left and right invariant differential calculi on G can be transported to M. For the left invariant calculus, notice that if F is a smooth function on M, we can differentiate $f = F \circ \pi$ which is defined on G and consider

$$D_{\leftarrow} F(x)u = D_{\leftarrow} f(S_x)u \tag{40}$$

for $u \in \mathfrak{g}$; notice that the result is 0 if u is in $\mathfrak{h}$, so the differential is actually a linear form on the vector space $\mathfrak{g}/\mathfrak{h}$. The problem is that it depends on the choice of the section S; if S' is another section and if we fix x, then $S'_x = S_x h$ for some $h \in H$, and

$$D_{\leftarrow} f(S'_x)u = D_{\leftarrow} f(S_x)\mathrm{Ad}_h u.$$

If however H is compact, then we can choose on $\mathfrak{g}$ an $\mathrm{Ad}(H)$-invariant inner product (this means that $|\mathrm{Ad}_h u| = |u|$ for h in H), we can consider $D_{\leftarrow} F(x)$ on the orthogonal $\mathfrak{p}$ of $\mathfrak{h}$ in $\mathfrak{g}$, and the norm $|D_{\leftarrow} F(x)|$ will not depend on the choice of S. It is invariant under the action of G in the sense

$$\left|D_{\leftarrow}(F \circ g)(x)\right| = \left|(D_{\leftarrow} F)(g(x))\right|.$$

Higher order derivatives have a similar behaviour, and we can consider the classes of functions $C^{\ell}_{b,\leftarrow}$ on M.

For the right differential calculus, we only consider the case where $(M, \cdot)$ is itself a Lie group, and left translations of M form a normal subgroup of G. In this case, a canonical choice for the section S is to let S_x be the left translation by x, so that $S: M \to G$ is an injective group homomorphism. The group G is a semi-direct product

$$G = S(M) \rtimes H \tag{41}$$

satisfying the commutation property $h S_x = S_{h(x)} h$ for $h \in H$. All elements of G can be written in the form $S_x h$ for some $x \in M$ and $h \in H$, and we have the product rule

$$S(x_1)h_1 S(x_2)h_2 = S(x_1 \cdot h_1(x_2))h_1 h_2.$$

The vector space $\mathfrak{g}$ can be written as $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ for the Lie algebras $\mathfrak{h}$ and $\mathfrak{m}$ of H and $M \sim S(M)$. We can consider the differential $D_\rightarrow F(x)u$ computed on the Lie group M for $u \in \mathfrak{m}$, and therefore the classes of functions $C_{b,\rightarrow}^\ell$. The relation with the differential on G is given by

$$D_\rightarrow F(x)u = D_\rightarrow f(S_x)u \qquad \text{for } f = F \circ \pi$$

similar to (40), because $\exp(\varepsilon u)S_x = S_{\exp(\varepsilon u)x}$. The behaviour under the action of G is

$$D_\rightarrow(F \circ g)(x)u = (D_\rightarrow F)\big(g(x)\big)\mathrm{Ad}_g u. \tag{42}$$

If $\mathcal{H}_\rightarrow^H$ and $\mathcal{H}_\rightarrow^M$ are right Haar measures on H and M, a right Haar measure can be defined on G by

$$\int_G f(g)\mathcal{H}_\rightarrow^G(\mathrm{d}g) = \int\int_{M \times H} f(S_x h)\mathcal{H}_\rightarrow^M(\mathrm{d}x)\mathcal{H}_\rightarrow^H(\mathrm{d}h) \tag{43}$$

because the right-hand side is invariant under the right action of H and $S(M)$ (use $S_x h S_y = S_{x \cdot h(y)}h$). By comparing with (38), we see that $\mathcal{H}_\rightarrow^M$ is a measure on M with multiplier

$$\chi(S_x h) = \chi_G(S_x h)/\chi_H(h), \tag{44}$$

since this is a group homomorphism and $\chi_G/\chi(S_x) = 1$. In particular, the Haar modulus of M is $\chi_M(x) = \chi(S_x) = \chi_G(S_x)$ and (39) becomes

$$p(x) = \int_H p_\rightarrow(S_x h)\mathcal{H}_\rightarrow^H(\mathrm{d}h). \tag{45}$$

In the next subsections, we study some classes of Markov processes on M. In the first case, we assume that H is compact, and consider the class of Markov processes on M which are invariant under the action of G; they can be written as the projection of left Lévy processes on G, and we apply the left differential calculus. We also consider the case where X is the projection of a right Lévy process. In the second case, we assume that G is a semi-direct product of type (41), and we let X be the projection of a right Lévy process on G; we then apply the right differential calculus.

4.3. Case 1: Compact isotropy subgroup

We here assume that $M = G/H$ for a compact Lie subgroup H of G; in particular, there is on H a unique probability measure $\mathcal{H}^H$ which is both a left and right Haar measure, and there is on M a measure $\mathcal{H}^M$ which is invariant under the action of G, and which is related to a left Haar measure on G by (37) with $\chi \equiv 1$. We choose an $\mathrm{Ad}(H)$-invariant inner product on $\mathfrak{g}$; it induces a H-invariant inner product on the tangent space $T_o M$, and a G-invariant Riemannian metric on M. We consider the class of Markov processes X_t which are invariant under the action of G; this means that $P_t(F \circ g) = (P_t F) \circ g$ for any $g \in G$; in particular, the law of $(X_t; t \geq 0)$ with initial condition $X_0 = o$ is invariant under the action of H. From Theorem 2.2 of [19], these processes are obtained as $X_t = \pi(\Xi_t)$, where Ξ_t is a left Lévy process on G which is invariant

under the right action of H. The invariance of X implies in particular that its Lévy kernel μ_x can be deduced from μ_o which will be simply called the Lévy measure μ_X of X; it is a H-invariant measure on $M \setminus \{o\}$ which integrates $\delta^2(o, x) \wedge 1$ (for the Riemannian distance δ). The Lévy process Ξ on G can be obtained by taking a Lévy measure given by

$$\mu_\Xi(A) = \int \int_{H \times M} 1_A\left(h S_x h^{-1}\right) \mathcal{H}^H(\mathrm{d}h) \mu_X(\mathrm{d}x), \tag{46}$$

where the H-invariance of μ_X implies that μ_Ξ does not depend on the choice of the section S (see [19]). Then the generator of Ξ can be written by means of (30), for a neighbourhood $V_G = \exp U_G$, where U_G is a small enough ball of $\mathfrak{g}$ (in particular U_G is $\mathrm{Ad}(H)$-invariant), and for a drift $\kappa \in \mathfrak{g}$ which is $\mathrm{Ad}(H)$-invariant.

For $0 \le s \le t$ fixed, notice that Ξ_t has the same law as $\Xi_s \Xi'_{t-s}$, for Ξ' an independent copy of Ξ; moreover, the variable Ξ_s can be written as $S_{X_s} h_s$, for h_s a H-valued variable. Thus, letting $X' = \pi(\Xi')$,

$$X_t = \Xi_t(o) \sim \Xi_s \Xi'_{t-s}(o) = S_{X_s} h_s\left(X'_{t-s}\right) \sim S_{X_s}\left(X'_{t-s}\right) \tag{47}$$

because the law of X'_{t-s} is H-invariant. More generally, if μ_1 and μ_2 are two H-invariant measures on M, one can define the convolution $\mu_1 * \mu_2$ as the image of the product measure by $(x, y) \mapsto S_x(y)$; it does not depend on S and is again H-invariant. The relation (47) shows that the law ν_t of X_t satisfies $\nu_t = \nu_s * \nu_{t-s}$. If the convolution product is commutative on the set of H-invariant measures, then (G, H) is said to be a Gelfand pair.

An example of space M is the hyperbolic space viewed as a subspace of the Minkowski space, namely

$$\mathbb{H}^d = \left\{x = (x_0, x_1, \ldots, x_d) \in \mathbb{R}^{1,d}; |x| = 1, x_0 > 0\right\}, \qquad |x|^2 = x_0^2 - \sum_{i=1}^{d} x_i^2,$$

with $o = (1, 0, \ldots, 0)$. It can be viewed as G/H, where $G = \mathrm{SO}^+(1, d)$ is the restricted Lorentz group of linear transformations of $\mathbb{R}^{1,d}$ which preserve the pseudo-norm, the time direction and the space orientation, and $H = \{h \in G; h(o) = o\} \sim \mathrm{SO}(d)$. Then it is known that (G, H) is a Gelfand pair.

For our result, we need the functions

$$\chi_M(x) = \chi_G(S_x), \qquad \mathcal{A}(x) = |\mathrm{Ad}_{S_x}|$$

which do not depend on S because $\chi_G = 1$ on H and the inner product of $\mathfrak{g}$ is $\mathrm{Ad}(H)$-invariant.

Theorem 5. *On $M = G/H$ for H compact, let X_t be a G-invariant Markov process without Brownian part, with $X_0 = o$, the Lévy measure μ_X of which satisfies the non degeneracy assumption (31) where $\exp^{-1} = \exp_o^{-1}$ denotes the inverse Riemannian exponential function based at o, and the additional condition (32) if $\alpha = 1$. If $\alpha < 1$ suppose moreover that X_t is a pure jump process. Then the law of X_t, $t > 0$, is absolutely continuous with respect to the G-invariant*

measure $\mathcal{H}^M$. *Let* $\ell \geq 0$. *If*

$$\int_{V^c} \chi_M(x)\mathcal{A}(x)^j \mu_X(\mathrm{d}x) < \infty \tag{48}$$

for a relatively compact neighbourhood V *of* o *and for* $j \leq \ell$, *then the density is in* $C^\ell_{b,\leftarrow}$ *(see Section* 4.2 *for the definition). In particular, if*

$$\int_{V^c} \mathcal{A}(x)^j \mu_X(\mathrm{d}x) < \infty \tag{49}$$

for any j, *the density is in* $C^\infty_{b,\leftarrow}$. *If* (G, H) *is a Gelfand pair, conditions* (48) *or* (49) *are not needed.*

Proof. We write X as $X_t = \pi(\Xi'_t)$ for a left Lévy process Ξ' with Lévy measure $\mu_{\Xi'}$ given by (46); denote by $\mathcal{L}'$ its infinitesimal generator. The question is to know whether the non degeneracy condition (31) for X can be translated into the similar condition for Ξ' (and similarly for (32) when $\alpha = 1$). Recall that $\mathfrak{g}$ is written as the orthogonal sum of $\mathfrak{h}$ and $\mathfrak{p}$, and we can choose the section S such that $S_x \in \exp \mathfrak{p}$ for x in a neighbourhood of o; then S is uniquely determined and smooth on a maybe smaller neighbourhood; the measure μ_X on a neighbourhood of o is therefore transported to a measure μ' on a neighbourhood of 0 in $\mathfrak{p}$, and μ' satisfies Assumption 1. If $S_x = \exp \lambda$, then

$$\exp^{-1}(h S_x h^{-1}) = \mathrm{Ad}_h \lambda + \mathrm{O}(|\lambda|^2).$$

On the other hand, we have from (46) that

$$I(\rho) = \int_{\{|\exp^{-1} \xi| \leq \rho\}} \langle \exp^{-1} \xi, u \rangle^2 \mu_{\Xi'}(\mathrm{d}\xi)$$

$$= \int \int_{\{|\exp^{-1}(h S_x h^{-1})| \leq \rho\}} \langle \exp^{-1}(h S_x h^{-1}), u \rangle^2 \mathcal{H}^H(\mathrm{d}h) \mu_X(\mathrm{d}x),$$

so obtaining the lower and upper bounds (31) for $I(\rho)$ is equivalent to estimating

$$I'(\rho) = \int \int_{H \times \{|\lambda| \leq \rho\}} \langle \mathrm{Ad}_h \lambda, u \rangle^2 \mathcal{H}^H(\mathrm{d}h) \mu'(\mathrm{d}\lambda).$$

The upper bound follows easily since μ' satisfies Assumption 1, and for the lower bound, we can restrict the domain of integration H to a small neighbourhood of the unity on which $|\mathrm{Ad}_h \lambda - \lambda| \leq c|\lambda|$ for c arbitrarily small. We deduce that (31) holds true for the process Ξ' but the lower bound is only for u in $\mathfrak{p}$. In order to obtain all of $\mathfrak{g}$, we add extra independent noise in Ξ'. Let Ξ'' be a left Lévy process on H with generator $\mathcal{L}''$ satisfying the conditions of Theorem 4; it is therefore associated to a measure μ'' on a neighbourhood of 0 in $\mathfrak{h}$, satisfying (31); we have $\mathcal{L}''(F \circ \pi) = 0$ for any smooth F on M. Let Ξ be the process with generator $\mathcal{L} = \mathcal{L}' + \mathcal{L}''$. Then $\mathcal{L}(F \circ \pi) = \mathcal{L}'(F \circ \pi)$, so $X_t = \pi(\Xi'_t)$ can also be written as $X_t = \pi(\Xi_t)$, and Ξ_t is a left Lévy

process associated to the measure $\mu' + \mu''$ which now satisfies (31) for any u in $\mathfrak{g}$. Condition (32) is similarly extended to Ξ when $\alpha = 1$.

Thus, we deduce from Theorem 4 that Ξ_t has a density $p_\leftarrow(t, \cdot)$ with respect to $\mathcal{H}_\leftarrow^G$. We have from (39) that the density of X_t with respect to $\mathcal{H}^M$ is given by

$$p(t, x) = \int_H p_\leftarrow(t, S_x h) \mathcal{H}^H(\mathrm{d}h).$$

Condition (48) for X implies (33) for Ξ' because $\chi_G(h) = |\mathrm{Ad}_h| = 1$ for h in H, so that $\chi_G(h S_x h^{-1}) = \chi_M(x)$ and $|\mathrm{Ad}_{(h S_x h^{-1})}| = \mathcal{A}(x)$, and (33) always holds for Ξ'' for the same reason. Thus, Theorem 4 can also be applied to Ξ for the smoothness of the law. In the case $\ell = 0$, the continuity of p follows from the continuity of $p_\leftarrow$ and the fact that we can choose a S which is smooth in a neighbourhood of x. In the case $\ell = 1$, we have

$$p\big(t, \pi(g)\big) = \int_H p_\leftarrow(t, S_{\pi(g)} h) \mathcal{H}^H(\mathrm{d}h) = \int_H p_\leftarrow(t, gh) \mathcal{H}^H(\mathrm{d}h)$$

from the left invariance of $\mathcal{H}^H$, and we deduce from the definition (40) that

$$D_\leftarrow p(t, x) u = \int_H D_\leftarrow p_\leftarrow(t, S_x h) \mathrm{Ad}_h^{-1} u \mathcal{H}^H(\mathrm{d}h),$$

so

$$\big|D_\leftarrow p(t, x)\big| \leq \int_H \big|D_\leftarrow p_\leftarrow(t, S_x h)\big| \mathcal{H}^H(\mathrm{d}h).$$

The study of higher order derivatives is similar.

If (G, H) is a Gelfand pair, we write a decomposition $\mu_X = \mu_X^\flat + \mu_X^\sharp$ where $\mu_X^\flat$ is the restriction to a H-invariant relatively compact neighbourhood of o; this corresponds to a decomposition $\mathcal{L} = \mathcal{L}^\flat + \mathcal{L}^\sharp$, and X can be viewed as the process with generator $\mathcal{L}^\flat$ interlaced with big jumps described by $\mathcal{L}^\sharp$. Conditionally on the times of the N_t big jumps before t, the law of X_t is therefore the convolution of $(2N_t + 1)$ H-invariant laws. From the commutativity of the convolution, all the big jumps can be put together, and we can write X_t in law for t fixed as $S_{X_t^\sharp}(X_t^\flat)$. The law of $X_t^\flat$ is in $C_{b, \leftarrow}^\infty$, and this smoothness is preserved under the action of $S_{X_t^\sharp}$. $\square$

If (G, H) is a Gelfand pair, the technique of previous proof for putting together big jumps can be extended to other cases. For instance, we can obtain the smoothness of the law if the Lévy measure is the sum of two measures, and only one of them satisfies the assumptions. This implies that the upper bound in (31) can be weakened, as in [20].

We can also consider the class of processes $X_t = \pi(\Xi_t)$, for right Lévy processes on G. There are Markov processes with semigroup $P_t f(x) = \mathbb{E} f(\Xi_t(x))$. Noticing that Ξ_t can also be viewed at fixed time as the value of a left Lévy process, we again apply left differential calculus and immediately obtain from the above proof the following result.

Theorem 6. *On $M = G/H$ for H compact, let $X_t = \pi(\Xi_t) = \Xi_t(o)$ for a right Lévy process Ξ_t on G. We suppose that the left Lévy process having the same generator at e as Ξ satisfies*

the assumptions of Theorem 4 for some ℓ. Then X_t has a $C^\ell_{b,\leftarrow}$ density with respect to the G-invariant measure $\mathcal{H}^M$.

4.4. Case 2: Semi-direct product

We consider as in Theorem 6 processes $X_t = \pi(\Xi_t) = \Xi_t(o)$ where Ξ_t is a right Lévy process on G, but do not assume that H is compact. Instead, we suppose that G is a semi-direct product as described in (41), and we apply the right differential calculus; recall that the right Haar measure $\mathcal{H}^M_\rightarrow$ of M is relatively invariant under the action of G with multiplier χ given by (44).

A typical example is when M is the additive group $\mathbb{R}^d$ and G is the affine group. Then $H = GL(d)$.

Theorem 7. *On $M = G/H$ for $G = S(M) \rtimes H$, let $X_t = \pi(\Xi_t) = \Xi_t(o)$ for a right Lévy process Ξ_t on G. We suppose that the Lévy measure μ_Ξ of Ξ satisfies (31), and the additional condition (32) if $\alpha = 1$, and that Ξ is a pure jump process if $\alpha < 1$. Assume also that*

$$\int_{V^c} \frac{1}{\chi(g)} \left| \mathrm{Ad}_g^{-1} \right|^j \mu_\Xi(\mathrm{d}g) < \infty \tag{50}$$

for $j \leq \ell$, where V is a relatively compact neighbourhood of e, and χ is given by (44). Then X_t has a $C^\ell_{b,\rightarrow}$ density with respect to the right Haar measure $\mathcal{H}^M_\rightarrow$.

Proof. Let us consider Ξ_t. We know that it can be viewed as the solution of an equation driven by a $\mathfrak{g}$-valued Lévy process Λ. We denote by Ξ^ε or $\Xi(\varepsilon)$ the same process when jumps of Λ greater than some ε (for some norm) have been removed, and by X^ε or $X(\varepsilon)$ its projection on M. We have seen in (20) that we have an estimate for the density of Ξ^ε and its derivatives at e involving $\mathbb{P}[\Xi^\varepsilon_t \in V_G]$ for a relatively compact neighbourhood V_G of e, uniformly in the initial condition. From the right invariance of the process, we also have an estimate for the density at g involving $\mathbb{P}[\Xi^\varepsilon_t \in V_G g]$, uniform in g. In particular, we can write

$$\left| D^j_\rightarrow p^{\Xi(\varepsilon)}_\rightarrow \left(t, S_x h g_0^{-1} \right) \right| \leq C_j t^{-(d+j)/\alpha} \mathbb{P}\left[\Xi^\varepsilon_t \in V_G S_x h g_0^{-1} \right] \tag{51}$$

for any $g_0 \in G$, $x \in M$, $h \in H$. We want to estimate the integral of this quantity with respect to h. Let V_M be a relatively compact neighbourhood of o; then $V'_G = V_G S(V_M)^{-1}$ is a relatively compact neighbourhood of e. For x in M, let $V^x_M = x^{-1} \cdot V_M \cdot x$. We have $V_G \subset V'_G S_x S_y S_x^{-1}$ for $y \in V^x_M$, so

$$\mathbb{P}\left[\Xi^\varepsilon_t \in V_G S_x h g_0^{-1} \right] \leq \mathbb{P}\left[\Xi^\varepsilon_t \in V'_G S_x S_y h g_0^{-1} \right]$$

for any $y \in V^x_M$ and $h \in H$. By taking the mean value on V^x_M, we obtain

$$\mathbb{P}\left[\Xi^\varepsilon_t \in V_G S_x h g_0^{-1} \right] \leq \frac{1}{\mathcal{H}^M_\rightarrow(V^x_M)} \int_{V^x_M} \mathbb{P}\left[\Xi^\varepsilon_t \in V'_G S_x S_y h g_0^{-1} \right] \mathcal{H}^M_\rightarrow(\mathrm{d}y),$$

and $\mathcal{H}_\rightarrow^M(V_M^x) = \mathcal{H}_\rightarrow^M(V_M)/\chi_M(x)$. Thus,

$$
\int_H \mathbb{P}\big[\Xi_t^\varepsilon \in V_G S_x h g_0^{-1}\big]\mathcal{H}_\rightarrow^H(\mathrm{d}h)
$$

$$
\leq C\chi_M(x) \int\int_{V_M^x \times H} \mathbb{P}\big[\Xi_t^\varepsilon \in V_G' S_x S_y h g_0^{-1}\big]\mathcal{H}_\rightarrow^M(\mathrm{d}y)\mathcal{H}_\rightarrow^H(\mathrm{d}h)
$$

$$
\leq C\chi_M(x) \int_G \mathbb{P}\big[\Xi_t^\varepsilon \in V_G' S_x g g_0^{-1}\big]\mathcal{H}_\rightarrow^G(\mathrm{d}g) \tag{52}
$$

$$
= C\chi_M(x)\mathbb{E}\big[\mathcal{H}_\rightarrow^G\big(S_x^{-1}(V_G')^{-1}\Xi_t^\varepsilon g_0\big)\big] = C\mathcal{H}_\rightarrow^G\big((V_G')^{-1}\big),
$$

where we have used (43) in the second inequality and $\chi_G(S_x^{-1}) = 1/\chi_G(S_x) = 1/\chi_M(x)$ in the last equality. The density of $\Xi_t^\varepsilon g_0$ at g is $p_\rightarrow^{\Xi(\varepsilon)}(t, gg_0^{-1})$, and similarly for its right invariant derivatives. The law of the variable $\pi(\Xi_t^\varepsilon g_0)$ is the law of X_t^ε with initial condition $x_0 = \pi(g_0)$, and (45) becomes

$$
p_\rightarrow^{X(\varepsilon)}(t, x_0, x) = \int_H p_\rightarrow^{\Xi(\varepsilon)}\big(t, S_x h g_0^{-1}\big)\mathcal{H}_\rightarrow(\mathrm{d}h).
$$

By differentiating, we have

$$
\big|D_\rightarrow^j p_\rightarrow^{X(\varepsilon)}(t, x_0, x)\big| \leq \int_H \big|D_\rightarrow^j p_\rightarrow^{\Xi(\varepsilon)}\big(t, S_x h g_0^{-1}\big)\big|\mathcal{H}_\rightarrow^H(\mathrm{d}h) \leq C_j t^{-(d+j)/\alpha}
$$

from (51) and (52). Thus, the smoothness of the law of X^ε is proved. We now have to take into account big jumps with the technique of Lemma 4, by considering X^ε interlaced with the big jumps, and letting (τ_K, τ_{K+1}) be the longest interval without small jumps. Previous argument shows that, conditionally on the times of big jumps, the variable $X_{\tau_{K+1}-}$ has a density $p_\star$ which is in $C_{b,\rightarrow}^\ell$, and its derivatives are of order $(\tau_{K+1} - \tau_K)^{-(d+j)/\alpha}$. The variable X_t is then obtained from the action of $\Upsilon = \Xi_t \Xi_{\tau_{K+1}-}^{-1}$, so its density is

$$
p_\rightarrow^X(t, x_0, x) = \mathbb{E}\big[p_\star\big(\Upsilon^{-1}(x)\big)/\chi(\Upsilon)\big].
$$

The variable $1/\chi(\Upsilon)$ is conditionally integrable (given the times of big jumps) if (50) holds for $j = 0$, so the theorem can be proved for $\ell = 0$ by the technique of Lemma 4; the case $\ell \geq 1$ is similar by applying (42) for the derivatives of $p_\star(\Upsilon^{-1}(x))$. $\qquad\square$

5. Examples

We here give some examples, and also some counterexamples where the "big jumps" condition (Assumption 3) does not hold, and the smoothness of the density fails.

5.1. Isotropic jumps

We have assumed that X is solution of an equation driven by some Lévy process Λ, but usually, jumps are often described by the Lévy kernel μ_x, the image of μ by $\lambda \mapsto a(x, \lambda)$. It is not easy to know when some Markov process with some Lévy kernel can be represented as the solution of an equation of our type. We have already seen in Section 2.7 how it is possible to deal with a finite part of the Lévy kernel. We now give the example of [1] where this is globally possible. Let M be a complete Riemannian manifold, and suppose that μ_x is the image by $(r, u) \mapsto \exp_x(ru)$ of $\mu_R \otimes \nu_x$, where μ_R is a measure on $(0, \infty)$ (radial part), and ν_x is the uniform probability measure on the unit sphere of $T_x M$ (angular part); this means that we choose a direction uniformly in the unit sphere, then go along a geodesic in that direction, at a distance chosen according to μ_R. Such a μ_x is singular if μ_R is singular.

   In order to construct an equation for this process, as explained in [1], we lift it to the bundle $O(M)$ of orthonormal frames, as this is classically done in the Eells–Elworthy–Malliavin construction of the Brownian motion. Points of this bundle can be written as $\xi = (x, g)$ for $x \in M$ and $g : \mathbb{R}^d \to T_x M$ is an orthogonal linear map; we put $\pi(\xi) = x$. Then, for $\lambda \in \mathbb{R}^d$, we can define $a(\xi, \lambda)$ for $\xi = (x, g)$ by

$$\pi\big(a(\xi, \lambda)\big) = \exp_x(g\lambda),$$

and the frame at $\pi(a(\xi, \lambda))$ is deduced from the frame g at x by parallel translation along the geodesic $(\exp_x(g\lambda t); 0 \le t \le 1)$. Let Ξ be the solution of the equation on $O(M)$ with this coefficient a and with $b = 0$, driven by a symmetric Lévy process Λ with Lévy measure $\mu = \mu_R \otimes \nu$, where ν is the uniform measure on the unit sphere of $\mathbb{R}^d$. Then $X = \pi(\Xi)$ is the process that we are looking for. The process Ξ can be viewed as the horizontal process above X. Notice that if (e_i) is the canonical basis of $\mathbb{R}^d$, the vector fields $\overline{a}(\cdot)e_i$ are the canonical horizontal vector fields on $O(M)$, and the equation of Ξ is a canonical equation, since a is obtained from $\overline{a}$ by means of (7).

   However, the surjectivity of $\overline{a}(\xi)$ cannot be satisfied, since the dimension of $O(M)$ which is $d(d + 1)/2$ is greater than the dimension d of $\mathbb{R}^d$. Nevertheless, we can add extra noise in Ξ without modifying the law of X and get this non degeneracy condition; more precisely, the extra noise acts vertically on the process Ξ. We enlarge the space $\mathbb{R}^d$ of the Lévy process as $\mathbb{R}^d \times O(d)$, put $a(x, \lambda, e) = a(x, \lambda)$, and let $a(x, 0, g)$ be the vertical transformation which modifies the frame by composing it with g. We let Λ' be an independent Lévy process on $O(d)$, and consider the equation driven by (Λ, Λ').

   We choose the reference measure on $O(M)$ which, when projected on M is the Riemannian measure, and which is on each fibre the uniform measure (normalised measure invariant under the action of $O(d)$). We consider the process Ξ on $O(M)$ the initial condition Ξ_0 of which has uniform law on the fibre above some x_0. We deduce from Theorem 2 the smoothness of the law of Ξ_t on $O(M)$ if μ_R has bounded support, and $\int_{\{|\lambda| \le \rho\}} |\lambda|^2 \mu_R(d\lambda)$ is bounded below and above by constants times $\rho^{2-\alpha}$ as $\rho \downarrow 0$. Conditionally on $X_t = \pi(\Xi_t)$, the variable Ξ_t has uniform law on the fibre above X_t, so the density of Ξ_t is a function of x only; this is also the density of X_t, so X_t has a smooth law too.

   When M is the sphere or the hyperbolic space, then M is a Riemannian symmetric space, and we are in the framework of a G-invariant process on $M = G/H$, where (G, H) is a Gelfand

pair. Thus, the smoothness of the law holds on $\mathbb{H}^d$ also in the case of unbounded jumps (Theorem 5). The hyperbolic space is diffeomorphic to $\mathbb{R}^d$; however, if we use normal coordinates, we cannot apply the theorem of [24] for $\mathbb{R}^d$ because the ellipticity condition is not satisfied. The matrix $(\overline{aa}^\star)^{-1}$, which is the hyperbolic Riemannian metric, explodes indeed exponentially fast at infinity, whereas at most polynomial growth was assumed in [24].

5.2. Affine transformations

Let G be the group of affine transformations of $\mathbb{R}^d$; it enters the framework of (41) as $G = \mathbb{R}^d \rtimes GL(d)$, where $g = (g_2, g_1)$ is the map $g(x) = g_1 x + g_2$. The Lie algebra is the set $\mathbb{R}^d \times \mathfrak{gl}(d)$, and if $u = (u_2, u_1)$, we have

$$g \exp(\varepsilon u) g^{-1}(x) = x + \varepsilon \big(g_1 u_1 g_1^{-1} x + g_1 \big(u_2 - g_1 u_1 g_1^{-1} g_2 \big) \big) + \mathrm{O}\big(\varepsilon^2\big),$$

so

$$\chi_G(g) = |\det g_1|, \qquad \mathrm{Ad}_g u = \big(g_1 u_1 g_1^{-1}, g_1 \big(u_2 - g_1 u_1 g_1^{-1} g_2 \big) \big).$$

We deduce from Theorem 4 the uniform smoothness of the laws of a left Lévy process Ξ_t and its projection $X_t = \pi(\Xi_t)$ on $\mathbb{R}^d$, under the nondegeneracy condition on the small jumps (with additional conditions if $\alpha \le 1$), and if the moments of $|\mathrm{Ad}_g|$ for the big jumps part are finite. Under the similar condition on $|\mathrm{Ad}_g^{-1}|$, we can also consider the right Lévy process, and deduce the smoothness of its $\mathbb{R}^d$ component from Theorem 7; in this case $H = GL(d)$ is unimodular, so $\chi = \chi_G$.

Notice that G is not unimodular, so without the assumption on big jumps, we are not even sure of the local boundedness of the density of X_t, except when we can apply Theorem 3 (smooth Lévy measure). In order to find a counterexample, we are going, for simplicity, to consider a subgroup of G.

So let now G be the group of transformations of $\mathbb{R}^d$ generated by translations and by dilations of rate e^n, $n \in \mathbb{Z}$. Thus $G = \mathbb{R}^d \rtimes \mathbb{Z}$, where (y, n) corresponds to the transformation $z \mapsto \mathrm{e}^n z + y$. The composition law of this group is

$$(y_2, n_2).(y_1, n_1) = \big(\mathrm{e}^{n_2} y_1 + y_2, n_1 + n_2 \big).$$

Its Lie algebra is Abelian and can be identified to $\mathbb{R}^d$. Denoting by $\mathrm{d}n$ and $\mathrm{d}y$ the counting measure on $\mathbb{Z}$ and the Lebesgue measure on $\mathbb{R}^d$, the Haar measures on G, the modulus and the adjoint representation are

$$\mathcal{H}_\leftarrow^G(\mathrm{d}y\,\mathrm{d}n) = \mathrm{e}^{-nd}\,\mathrm{d}y\,\mathrm{d}n, \qquad \mathcal{H}_\rightarrow^G(\mathrm{d}y\,\mathrm{d}n) = \mathrm{d}y\,\mathrm{d}n, \qquad \chi_G(y, n) = \mathrm{e}^{nd}, \qquad \mathrm{Ad}_{(y,n)}u = \mathrm{e}^n u.$$

We consider on G left Lévy processes $X_t = (Y_t, N_t)$ such that N_t is a random walk on $\mathbb{Z}$, and Y_t is deduced from a Lévy process Λ_t on $\mathbb{R}^d$, independent of N, by means of

$$Y_t = \int_0^t \mathrm{e}^{N_s}\,\mathrm{d}\Lambda_s.$$

We want to study the density near $(0,0)$ of X_1 with respect to $\mathcal{H}_{\leq}^G$, or equivalently the density near 0 of Y_1 restricted to $\{N_1 = 0\}$. Suppose that Λ is an isotropic stable process. Then

$$Y_1 \sim \Lambda\left(\int_0^1 e^{\alpha N_s}\, ds\right) \sim \left(\int_0^1 e^{\alpha N_s}\, ds\right)^{1/\alpha} \Lambda_1,$$

and its conditional density at 0 given N is

$$p^Y(0|N) = c_\Lambda\left(\int_0^1 e^{\alpha N_s}\, ds\right)^{-d/\alpha}, \tag{53}$$

where c_Λ is the density of Λ_1 at 0. Let p_n be the mass of the Lévy measure of N at n, with $\sum p_n < \infty$. Let A_n, $n \geq 1$, be the following event: the process N has exactly two jumps before time 1, one jump of size $-n$ followed by a jump of size $+n$. Then the probability of A_n is $c_N p_{-n} p_n$ with $c_N = \frac{1}{2}\exp - \sum p_j$; conditionally on A_n, the times of the two jumps are obtained from two independent variables uniformly distributed on $[0, 1]$, and by letting the negative jump be the smallest one, and the positive jump the largest one. The law of X_1 restricted to A_n is smooth; by applying (53) and by denoting $c = c_\Lambda c_N$, its density at $(0,0)$ is

$$p_{A_n}(0,0) = 2c p_{-n} p_n \int_0^1 \int_0^t \left(s + (t-s)e^{-n\alpha} + 1 - t\right)^{-d/\alpha}\, ds\, dt$$

$$= 2c p_{-n} p_n \int_0^1 (1-s)\left(1 + s\left(e^{-n\alpha} - 1\right)\right)^{-d/\alpha}\, ds$$

$$\geq c p_{-n} p_n e^{-2n\alpha}\left(1 - \left(1 - e^{-n\alpha}\right)^2\right)^{-d/\alpha}$$

$$\geq c p_{-n} p_n e^{n(d-2\alpha)}.$$

The law of X_1 restricted to $\bigcup_{1 \leq k \leq n} A_k$ is smooth with density $\sum_{k=1}^n p_{A_k}$, so for any neighbourhood U of $(0,0)$, the density p of X_1 satisfies

$$\operatorname*{ess\,sup}_U p \geq c \sum_{k=1}^n p_{-k} p_k e^{k(d-2\alpha)}.$$

If the series diverges, then X_1 does not have a locally bounded density, and examples can be constructed as soon as $d > 2\alpha$.

Choose for instance

$$p_n = e^{-\beta n}, \qquad p_{-n} = e^{-\sigma n} \tag{54}$$

for $n \geq 1$, $\beta > 0$, $\sigma > 0$. In this case the density is not locally bounded if $d \geq 2\alpha + \beta + \sigma$. On the other hand, Theorem 4 implies that if $d < \beta$, the density is bounded uniformly with respect to the initial condition.

Consider now the law of Y_1. We can consider the right Lévy process X' with the same generator at e as X; if N and Λ are the same independent random walk and stable isotropic processes,

X' can be obtained as $X' = (Y', N)$, where Y' the solution of $Y'_{t+dt} = a(Y'_t, dN_t, d\Lambda_t)$, and $a(y, n, \lambda) = e^n y + \lambda$. We have the equalities in law $X'_1 \sim X_1$ and $Y'_1 \sim Y_1$.

The study of the law of Y' has been the subject of Theorem 7. For the example (54), this theorem implies that $Y'_1 \sim Y_1$ has a bounded density if $d < \sigma$. On the other hand, we have just verified that this density is not locally bounded if $d \geq 2\alpha + \beta + \sigma$; actually, this condition can be improved because we now study Y_1 instead of X_1. The event A_n can be replaced by A'_n: the process N has exactly one jump of size $-n$ before time 1, and we deduce that the density is not locally bounded if $d \geq \alpha + \sigma$.

5.3. Killed processes

Consider a process which satisfies our sufficient conditions for the smoothness of the law in some manifold, say the affine space $\mathbb{R}^d$, but which is killed at the exit from some open subset M. The existence of a locally bounded density is preserved, but it appears that this killing can destroy the smoothness of this density. We have seen in (9) that Assumption 2 is preserved, but Assumption 3 may fail; a problem can occur when jumps are allowed to enter the subset (coming from the infinity of M), or at least (by applying Theorem 3) if the non-smooth part of these jumps are allowed to enter the subset. If directions of non-smooth jumps lie in some closed cone of the vector space $\mathbb{R}^d$, then smoothness is preserved if the obstacle M^c is such a cone based at some point of $\mathbb{R}^d$; in the general case however, roughly speaking, the obstacle can produce singularity behind it. Let us give a one-dimensional example.

Let $\Lambda_t = \Lambda_t^0 - \Lambda_t^1$ where Λ_t^0 and Λ_t^1 are real, independent, and are respectively, a symmetric stable Lévy process with Lévy measure $|\lambda|^{-\alpha-1} d\lambda$ and a standard Poisson process. Let X_t be the process Λ_t killed when it quits $M = (-\infty, 1)$. It is solution of the equation corresponding to $b = 0$ and $a(x, \lambda) = x + \lambda$ if $x + \lambda < 1$, equal to ∞ otherwise; the initial condition is $x_0 = 0$.

Notice that the similar process without Λ_t^1 has smooth densities from Theorem 3. If Λ_t^1 is added instead of subtracted, the law should also be smooth since the appended jumps do not enable the process to enter M. We are now going to prove that in our framework, the law of X_1 is not C^1 as soon as $\alpha \leq 2/3$.

Let τ be the lifetime of X, which is the first exit time of Λ from M. Then the density of X_1 is

$$p(x) = q(1, x) - \mathbb{E}\big[q(1 - \tau, x - \Lambda_\tau) 1_{\{\tau < 1\}}\big], \tag{55}$$

where $q(t, \cdot)$ is the density of Λ_t. We have

$$q(t, x) = e^{-t} \sum_k \frac{t^k}{k!} q_0(t, x + k),$$

where the density q_0 of Λ^0 is C^∞ on $(\mathbb{R}_+ \times \mathbb{R}) \setminus \{(0, 0)\}$, with bounded derivatives out of any neighbourhood of $(0, 0)$. Consequently, q is smooth on $(\mathbb{R}_+ \times \mathbb{R}) \setminus (\{0\} \times \mathbb{Z}_-)$, and

$$q_{(k)}(t, x) = q(t, x) - e^{-t} \frac{t^k}{k!} q_0(t, x + k) \tag{56}$$

is smooth on $\mathbb{R}_+ \times (-k-1, -k+1)$. We are going to study the derivative of $p(x)$ as $x \uparrow 0$.

We need some information on the joint law of the exit time τ and the overshoot $\Lambda_\tau - 1$. To this end, we first check that τ is almost surely a time of jump of Λ; this means that the process cannot creep upward, see [10]; a simple way of verifying this fact is to notice that Λ^0 cannot creep both upward and downward because it has no Brownian part; since it is moreover symmetric, it creeps neither upward, neither downward, and Λ which is obtained by adding a process with finitely many jumps satisfies the same property. Consider the joint law of $(\tau, \Lambda_{\tau-}, \Lambda_\tau)$, and denote by $\mathcal{T}$ the set of jumps of Λ; notice that τ is a jump of Λ^0; then

$$\mathbb{E}\big[f(\tau, \Lambda_{\tau-}, \Lambda_\tau)1_{\{\tau<\infty\}}\big] = \mathbb{E}\sum_{t\in\mathcal{T}} f(t, X_{t-}, X_{t-} + \Delta\Lambda_t)1_{\{X_{t-}\neq\infty\}}1_{\{\Delta\Lambda_t\geq 1-X_{t-}\}}$$

$$= \mathbb{E}\int_0^\infty \int_0^\infty f(t, X_t, X_t + \lambda)1_{\{X_t\neq\infty\}}1_{\{\lambda\geq 1-X_t\}}\lambda^{-1-\alpha}\,\mathrm{d}\lambda\,\mathrm{d}t$$

from a key formula of stochastic calculus on Poisson spaces. Thus,

$$\mathbb{P}[\tau \in \mathrm{d}t, \Lambda_{\tau-} \in \mathrm{d}x, \Lambda_\tau \in \mathrm{d}y] = \mathbb{P}[X_t \in \mathrm{d}x](y-x)^{-1-\alpha}\,\mathrm{d}t\,\mathrm{d}y$$

on $(0,\infty) \times (-\infty, 1) \times (1, \infty)$, and the density of (τ, Λ_τ) is

$$\zeta(t, y) = \mathbb{E}\big[(y - X_t)^{-1-\alpha}\big] \tag{57}$$

on $(0,\infty) \times (1, \infty)$; in particular $y \mapsto \zeta(t, y)$ is C_b^∞ on $[1+\varepsilon, \infty)$ for any $\varepsilon > 0$, uniformly in t. Let $h: \mathbb{R} \to [0, 1]$ be a smooth function which is 1 on $(-\infty, 4/3]$ and 0 on $[5/3, +\infty)$. From (55), we have

$$p(x) = q(1, x) - \int_1^\infty \int_0^1 q(1-t, x-y)\zeta(t, y)\,\mathrm{d}t\,\mathrm{d}y$$

$$= q(1, x) - \int_1^\infty \int_0^1 q(1-t, x-y)\big(1 - h(y)\big)\zeta(t, y)\,\mathrm{d}t\,\mathrm{d}y - \overline{p}(x)$$

with

$$\overline{p}(x) = \int_1^\infty \int_0^1 q(1-t, x-y)h(y)\zeta(t, y)\,\mathrm{d}t\,\mathrm{d}y.$$

The measure $(1 - h(y))\zeta(t, y)\,\mathrm{d}y$ is C_b^∞, uniformly in t, so its convolution with the law of Λ_{1-t} is also smooth, and we obtain that $p + \overline{p}$ is smooth. It is therefore sufficient to study the regularity of $\overline{p}$.

The function $q(1-t, \cdot)$ is smooth out of $\mathbb{Z}_-$, so we can differentiate $\overline{p}$ on $(-1/3, 0)$ and obtain

$$D\overline{p}(x) = \int_1^{5/3} \int_0^1 Dq(1-t, x-y)h(y)\zeta(t, y)\,\mathrm{d}t\,\mathrm{d}y$$

$$= \int_1^{5/3} \int_0^1 (1-t)\mathrm{e}^{t-1}Dq_0(1-t, x+1-y)h(y)\zeta(t, y)\,\mathrm{d}t\,\mathrm{d}y + O(1)$$

as $x \uparrow 0$, because the rest involves the function $q_{(1)}$ of (56) which is smooth on $(-2, 0)$. The self-similarity of Λ^0 enables to write

$$D\overline{p}(x) = \int_1^{5/3} \int_0^1 (1-t)^{1-2/\alpha} e^{t-1} Dq_0\left(1, \frac{x+1-y}{(1-t)^{1/\alpha}}\right) h(y) \zeta(t, y) \, dt \, dy + O(1).$$

The law of the stable variable Λ_1^0 is symmetric and unimodal, so Dq_0 is nonnegative on $\mathbb{R}_-$, and is bounded below by a positive constant on some $[-C_2, -C_1] \subset (-\infty, -2^{1/\alpha}]$. Thus, the double integral can be bounded below by considering only the part where the fraction in Dq_0 is in $[-C_2, -C_1]$, where $1/2 \le t \le 1$, and where $1 < y \le 4/3$. With the change $s = 1 - t$, we obtain

$$D\overline{p}(x) \ge c \int_1^{4/3} \int_{J(y)} s^{1-2/\alpha} \zeta(1-s, y) \, ds \, dy - C \tag{58}$$

with

$$J(y) = \left\{0 < s \le 1/2; C_1 \le s^{-1/\alpha}(y-1-x) \le C_2\right\}.$$

On the other hand, from (57),

$$\zeta(t, y) = \mathbb{E}\left[(y - X_t)^{-1-\alpha}\right] \ge e^{-t} \mathbb{E}\left[(y - X_t^0)^{-1-\alpha}\right],$$

where e^{-t} is the probability for Λ^1 to be 0 up to time t, and X^0 is the process Λ^0 killed at the exit of M. We have from [7] that the law of X_t^0 with initial condition $X_0^0 = 0$ is bounded below and above by some positive constants times $(1-u)^{\alpha/2} \, du$ for $0 \le u < 1$ and $1/2 \le t \le 1$, so

$$\begin{aligned}
\zeta(1-s, y) &\ge c \int_0^1 (y-u)^{-\alpha-1}(1-u)^{\alpha/2} \, du \\
&\ge c \int_{1-3(y-1)}^{1-(y-1)} \left((y-1) + (1-u)\right)^{-\alpha-1}(1-u)^{\alpha/2} \, du \ge c'(y-1)^{-\alpha/2}
\end{aligned} \tag{59}$$

for $1 < y \le 4/3$ and $0 \le s \le 1/2$. We also have

$$J(y) = \left\{s > 0; C_1 \le s^{-1/\alpha}(y-1-x) \le C_2\right\}$$

for $1 < y \le 5/3$ and $-1/3 \le x < 0$, because $0 \le y - 1 - x \le 1$ and $C_2^{-\alpha} < C_1^{-\alpha} \le 1/2$; thus

$$\int_{J(y)} s^{1-2/\alpha} \, ds = C(y-1-x)^{2\alpha-2}. \tag{60}$$

It follows from (58), (59) and (60) that

$$D\overline{p}(x) \ge c \int_1^{4/3} (y-1-x)^{2\alpha-2}(y-1)^{-\alpha/2} \, dy - C.$$

If $\alpha \le 2/3$, we obtain that $D\overline{p}(x)$ tends to $+\infty$ as $x \uparrow 0$, so $Dp(x)$ tends to $-\infty$ and the law of X_1 is not C^1.

Acknowledgement

References

[1] Applebaum, D. and Estrade, A. (2000). Isotropic Lévy processes on Riemannian manifolds. *Ann. Probab.* **28** 166–184. MR1756002

[2] Applebaum, D. and Kunita, H. (1993). Lévy flows on manifolds and Lévy processes on Lie groups. *J. Math. Kyoto Univ.* **33** 1103–1123. MR1251218

[3] Bally, V. and Clément, E. (2011). Integration by parts formula and applications to equations with jumps. *Probab. Theory Related Fields* **151** 613–657. MR2851695

[4] Bichteler, K., Gravereaux, J.B. and Jacod, J. (1987). *Malliavin Calculus for Processes with Jumps. Stochastics Monographs* **2**. New York: Gordon and Breach Science Publishers. MR1008471

[5] Bismut, J.M. (1983). Calcul des variations stochastique et processus de sauts. *Z. Wahrsch. Verw. Gebiete* **63** 147–235. MR0701527

[6] Bourbaki, N. (1963). *Éléments de Mathématique. Fascicule XXIX. Livre VI*: *Intégration. Chapitre* 7: *Mesure de Haar. Chapitre* 8: *Convolution et Représentations. Actualités Scientifiques et Industrielles* **1306**. Paris: Hermann. MR0179291

[7] Chen, Z.Q., Kim, P. and Song, R. (2010). Heat kernel estimates for the Dirichlet fractional Laplacian. *J. Eur. Math. Soc.* (*JEMS*) **12** 1307–1329. MR2677618

[8] Cohen, S. (1996). Géométrie différentielle stochastique avec sauts. I. *Stochastics Stochastics Rep.* **56** 179–203. MR1396760

[9] Dellacherie, C. and Meyer, P.A. (1975). *Probabilités et Potentiel. Chapitres I à IV, Édition entièrement refondue, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV. Actualités Scientifiques et Industrielles* **1372**. Paris: Hermann. MR0488194

[10] Doney, R.A. (2007). *Fluctuation Theory for Lévy Processes. Lecture Notes in Math.* **1897**. Berlin: Springer. MR2320889

[11] Fournier, N. (2008). Smoothness of the law of some one-dimensional jumping S.D.E.s with non-constant rate of jump. *Electron. J. Probab.* **13** 135–156. MR2375602

[12] Fujiwara, T. (1991). Stochastic differential equations of jump type on manifolds and Lévy flows. *J. Math. Kyoto Univ.* **31** 99–119. MR1093330

[13] Fujiwara, T. and Kunita, H. (1985). Stochastic differential equations of jump type and Lévy processes in diffeomorphisms group. *J. Math. Kyoto Univ.* **25** 71–106. MR0777247

[14] Ishikawa, Y. and Kunita, H. (2006). Malliavin calculus on the Wiener–Poisson space and its application to canonical SDE with jumps. *Stochastic Process. Appl.* **116** 1743–1769. MR2307057

[15] Kulik, A.M. (2011). Absolute continuity and convergence in variation for distributions of functionals of Poisson point measure. *J. Theoret. Probab.* **24** 1–38. MR2782709

[16] Kunita, H. (1994). Stable Lévy processes on nilpotent Lie groups. In *Stochastic Analysis on Infinite-dimensional Spaces* (*Baton Rouge, LA, 1994*). *Pitman Res. Notes Math. Ser.* **310** 167–182. Harlow: Longman Sci. Tech. MR1415667

[17] Kunita, H. (2009). Smooth density of canonical stochastic differential equation with jumps. *Astérisque* **327** 69–91. MR2642353

[18] Kurtz, T.G., Pardoux, É. and Protter, P. (1995). Stratonovich stochastic differential equations driven by general semimartingales. *Ann. Inst. Henri Poincaré Probab. Stat.* **31** 351–377. MR1324812

[19] Liao, M. (2004). *Lévy Processes in Lie Groups. Cambridge Tracts in Mathematics* **162**. Cambridge: Cambridge Univ. Press. MR2060091

[20] Liao, M. and Wang, L. (2007). Lévy–Khinchin formula and existence of densities for convolution semigroups on symmetric spaces. *Potential Anal*. **27** 133–150. MR2322502

[21] Marcus, S.I. (1978). Modeling and analysis of stochastic differential equations driven by point processes. *IEEE Trans. Inform. Theory* **24** 164–172. MR0487784

[22] Nualart, D. and Vives, J. (1995). A duality formula on the Poisson space and some applications. In *Seminar on Stochastic Analysis*, *Random Fields and Applications* (*Ascona*, 1993). *Progress in Probability* **36** 205–213. Basel: Birkhäuser. MR1360277

[23] Picard, J. (1996). Formules de dualité sur l'espace de Poisson. *Ann. Inst. Henri Poincaré Probab. Stat.* **32** 509–548. MR1411270

[24] Picard, J. (1996). On the existence of smooth densities for jump processes. *Probab. Theory Related Fields* **105** 481–511. Erratum: *Probab. Theory Relat. Fields* **147** (2010) 711–713. MR1402654

[25] Picard, J. (1997). Density in small time at accessible points for jump processes. *Stochastic Process. Appl.* **67** 251–279. MR1449834

[26] Picard, J. and Savona, C. (2000). Smoothness of harmonic functions for processes with jumps. *Stochastic Process. Appl.* **87** 69–91. MR1751165

[27] Sato, K.i. (1999). *Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. MR1739520

[28] Wijsman, R.A. (1990). *Invariant Measures on Groups and Their Use in Statistics. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **14**. Hayward, CA: IMS. MR1218397