

**EXISTENCE AND SMOOTHNESS OF TRANSITION
 DENSITY FOR JUMP-TYPE MARKOV PROCESSES:
 APPLICATIONS OF MALLIAVIN CALCULUS**

BY SEIJI HIRABA

Introduction.

Let X_t be a d -dimensional stable Lévy process with exponent $0 < \alpha < 2$, defined on a probability space (Ω, \mathcal{F}, P) . Then its Lévy measure is given by

$$(1) \quad K(A) = \int_{S^{d-1}} M(d\sigma) \int_0^\infty 1_{A \setminus \{0\}}(u\sigma) u^{-1-\alpha} du \quad (A \in \mathcal{B}(\mathbf{R}^d)),$$

where $M(d\sigma)$ is a bounded Borel measure on a $(d-1)$ -dimensional unit sphere S^{d-1} . For each $t > 0$ denote by $P_t(dx)$, the distribution of X_t . It is known that, under the following non-degeneracy condition on $M(d\sigma)$:

$$(2) \quad M(\{\sigma : x \cdot \sigma \neq 0\}) > 0 \quad \text{for every } x \in \mathbf{R}^d \setminus \{0\},$$

$P_t(dx)$ is absolutely continuous with respect to the Lebesgue measure on \mathbf{R}^d , and moreover its density $p_t(x)$ is a C^∞ -function of $x \in \mathbf{R}^d$ for all $t > 0$ (cf. [7]).

Next let Y_t be a d -dimensional Lévy process of class L (simply called an L -process). Then its Lévy measure is expressed by

$$K(A) = \int_{S^{d-1}} M(d\sigma) \int_0^\infty 1_{A \setminus \{0\}}(u\sigma) u^{-1} k(u, \sigma) du,$$

where $M(d\sigma)$ is a bounded Borel measure on S^{d-1} and $k(u, \sigma)$ is a nonnegative measurable function on $(0, \infty) \times S^{d-1}$, nonincreasing, right continuous in u satisfying that $0 < k(0+, \sigma) \leq \infty$ for each $\sigma \in S^{d-1}$ and $\int_{S^{d-1}} M(d\sigma) \int_0^\infty (u \wedge u^{-1}) \cdot k(u, \sigma) du < \infty$. Then the distribution $P_t(dx)$ of Y_t has a density $p_t(x)$ with respect to the Lebesgue measure on \mathbf{R}^d (henceforth we simply say “a density $p_t(x)$ ”) for $t > 0$ under the condition (2) ([8]). Moreover for any $r \geq 0$, $p_t(x)$ can be chosen a C^r -function of $x \in \mathbf{R}^d$, provided t is larger than some constant depending on $(M(d\sigma), k(0+, \sigma), d, r)$ (cf. [7], [10]).

On the other hand Bass [1] discussed the martingale problem for the Lévy generator L :

Received February 8, 1991.

$$(3) \quad L\varphi(x) = D_x\varphi(x) \cdot a(x) + \int [\varphi(x+y) - \varphi(x) - D_x\varphi(x) \cdot y 1_{(|y| < 1)}] K(x, dy)$$

and proved the well-posedness of the martingale problem for a fairly wide class of the Lévy generator L . In fact, his class includes the following case:

$$(4) \quad L = -(-\Delta)^{\alpha(x)/2},$$

where $\alpha(x): \mathbf{R}^d \rightarrow \mathbf{R}$ is Dini continuous and satisfies the inequalities

$$(5) \quad 0 < \zeta_1 \leq \alpha(x) \leq \zeta_2 < 2 \text{ for all } x \in \mathbf{R}^d \text{ with some constants } \zeta_1 \text{ and } \zeta_2.$$

In this case the Lévy kernel of this operator is expressed by

$$(6) \quad K(x, A) = \theta_{\alpha(x)} \int_{S^{d-1}} d\sigma \int_0^\infty 1_{A \setminus \{0\}}(u\sigma) u^{-1-\alpha(x)} du,$$

where

$$\theta_\alpha = \sin \frac{\pi\alpha}{2} \Gamma(1+\alpha) \Gamma\left(\frac{d+\alpha}{2}\right) \pi^{-(d+1)/2} \Gamma\left(\frac{1+\alpha}{2}\right)^{-1}.$$

To the Lévy generator (4) there corresponds a Markov process on \mathbf{R}^d , which is called a *stable-like process* following Bass [1].

Generalizing the Lévy kernels (1) and (6), we introduce the following class of Lévy kernels:

$$(7) \quad K(x, A) = \int_S M(d\sigma) \int_0^\infty 1_{A \setminus \{0\}}(y(x, u, \sigma)) u^{-1-\alpha(x)} m(x, u, \sigma) du,$$

where

- (i) M is a bounded Borel measure on a topological space S ,
- (ii) $y(x, u, \sigma): \mathbf{R}^d \times [0, \infty) \times S \rightarrow \mathbf{R}^d$ is measurable, of class C^1 in u on $[0, \infty)$ and $y(x, 0, \sigma) = 0$,
- (iii) $m(x, u, \sigma): \mathbf{R}^d \times (0, \infty) \times S \rightarrow \mathbf{R}$ is a bounded nonnegative measurable function,
- (iv) $\alpha(x): \mathbf{R}^d \rightarrow [0, 2)$ is a measurable function.

Our main concern in the present paper is to investigate the absolute continuity of the transition probability and the smoothness of the transition density for the Markov process on \mathbf{R}^d associated with the Lévy generator L of (3) with (7). In particular, in the case where the condition (5) holds, it would be expected that the transition probability $P_t(x, dy)$ has a density $p_t(x, y)$ which is smooth in y or in (x, y) for every $t > 0$ under a regularity condition and a non-degeneracy condition on $K(x, dy)$. Furthermore we shall also discuss the Lévy kernel (7) with $\alpha(x) \equiv 0$. In this case the result on one-dimensional L -processes ([10]) would suggest that for any $r \geq 0$ the transition density $p_t(x, y)$ would be of class C^r in y or in (x, y) only for sufficiently large t 's.

To avoid some technical difficulties, we shall mainly consider the following Lévy generator obtained by cutting off big jumps in (7):

$$(8) \quad L\varphi(x) = D_x\varphi(x) \cdot a(x) + \int [\varphi(x+y) - \varphi(x) - D_x\varphi(x) \cdot y] K(x, dy)$$

with

$$(9) \quad K(x, A) = \int_S M(d\sigma) \int_0^{u_0} 1_{A \setminus \{0\}}(y(x, u, \sigma)) g(x, u, \sigma) du,$$

where $u_0 > 0$ is a fixed constant and we assume that

(i) $y(x, u, \sigma)$ is a bounded measurable function on $\mathbf{R}^d \times [0, u_0] \times S$, of class C^1 in u on $[0, u_0]$ and $y(x, 0, \sigma) = 0$,

(ii) $g(x, u, \sigma) = u^{-1-\alpha(x)} m(x, \sigma)$

with $0 \leq \alpha(x) < 2$ and a bounded positive function $m(x, \sigma)$. In particular, if $S = S^{d-1}$, $y(x, u, \sigma) = u\sigma$ and $m(x, \sigma) = \theta_{\alpha(x)}$ then the corresponding Markov process should be called a *stable-like process with bounded jumps*, since big jumps are suppressed in (6).

Our main results in the present paper are summarized as follows: Suppose that a smoothness condition on $a(x)$, $y(x, u, \sigma)$, $m(x, \sigma)$ and $\alpha(x)$, and a non-degeneracy condition on $M(d\sigma)$ and $y(x, u, \sigma)$ are fulfilled. Then

(a) if $\alpha(x)$ satisfies (5), the transition probability $P_t(x, dy)$ has a density $p_t(x, y)$ which is smooth in y or in (x, y) for every $t > 0$.

(b) If $\alpha(x) \equiv 0$, then also $P_t(x, dy)$ has a density $p_t(x, y)$, and for a given integer $r \geq 0$ it is of class C^r in x or in (x, y) , provided t is larger than a constant depending on $(y(x, u, \sigma), M(d\sigma), m(x, \sigma), d, r)$.

The smoothness and non-degeneracy conditions are rather complicated in the general case (which are given in §3); however we can simply describe the result for a stable-like process with bounded jumps: the transition probability has a C^∞ -density if $\alpha(x)$ is a C^∞ -function whose derivatives of all order are bounded, it satisfies (5) and $|D_x\alpha(x)|$ is sufficiently small. We remark that these results should, at least partially, hold for a stable-like process itself because its Lévy generator is a perturbation from that of a stable-like process with bounded jumps by adding a bounded operator.

The proof of our results is essentially based on Malliavin calculus on a Poisson space which was developed by Bichteler, Gravereaux and Jacod [2], but we have to modify their framework to apply it to our class including stable-like processes with bounded jumps. Furthermore we have to emphasize our smoothness result for the transition density holds for every $t > 0$ when $\alpha(x)$ satisfies (5), although [2] discusses the smoothness of the density only for sufficiently large t 's under a different situation from ours.

The present paper is organized as follows: In §1 we introduce a Poisson space and a notion of derivatives of Poisson functionals with some modifications of [2], and by using it we define a Malliavin operator. In §2 we apply Malliavin calculus to stochastic differential equations and summarize the results used later. In §3 we state the main results for the generator L defined by (8) with the kernel $K(x, dy)$ defined by (9) and for a stable-like process. In §4 the proofs of our main results are given by using the results in §2.

Throughout this paper we shall use the following notations:

A^T : the transpose of a matrix A ,

$|E|$: the n -dimensional Lebesgue measure of $E \subset \mathbf{R}^n$,

C^r : the totality of r -times continuously differentiable functions,

C_0^r : the totality of C^r -functions with compact support,

C_b^r : the totality of C^r -functions with bounded derivatives of all order

between 0 and r ,

$C = C_0$, $C_0 = C_0^0$, $C_b = C_b^0$,

$C^\infty = \bigcap_{r \geq 0} C^r$, $C_0^\infty = \bigcap_{r \geq 0} C_0^r$, $C_b^\infty = \bigcap_{r \geq 0} C_b^r$,

$\partial_i = \partial / \partial x_i$.

1. Differential calculus on a Poisson space.

Let E be an open subset in the n -dimensional Euclidean space \mathbf{R}^n with $|E| = \infty$ and let \mathcal{E} be the Borel σ -field on E . Fix $T > 0$ and let $(S, \mathcal{S}, M(d\sigma))$ be a topological finite measure space with $\mathcal{S} = \mathcal{B}(S)$ (the Borel σ -field of S). Next we take a Poisson space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ associated with a Poisson random measure distributing on $[0, T] \times E \times S$ with the intensity measure $\nu(dsdzd\sigma) = dsdzM(d\sigma)$ in the following way. Let Ω denote the collection of point measures $\mu = \sum_n \delta_{(t_n, z_n, \sigma_n)}$ on $[0, T] \times E \times S$ such that $\mu([0, T] \times K \times S) < \infty$ for any compact set K in E , $\mu(\{t\} \times E \times S) \leq 1$ for all $t \in [0, T]$ and $\mu(\{0\} \times E \times S) = 0$. Set $\mathcal{F}_t^0 = \sigma(\mu(A) : A \in \mathcal{B}([0, t]) \otimes \mathcal{E} \otimes \mathcal{S})$ for $t \in [0, T]$, $\mathcal{F}_t = \bigcap_{u > t} \mathcal{F}_u^0$ for $t \in [0, T]$, and $\mathcal{F} = \mathcal{F}_T^0$. Then there exists a unique probability measure P on (Ω, \mathcal{F}) such that, under P , μ is a Poisson random measure distributing on $[0, T] \times E \times S$ with the intensity measure $\nu(dsdzd\sigma)$. As usual we extend \mathcal{F}_t by adding all negligible events, and we use the same notation \mathcal{F}_t as for the original. Then $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ is our basic Poisson space. We define a σ -finite measure Q on $\Omega \times [0, T] \times E \times S$ by $Q(d\mu dsdzd\sigma) = P(d\mu)\mu(dsdzd\sigma)$.

Let $\Phi(\mu)$ be a Poisson functional, that is, a random variable on the Poisson space (Ω, \mathcal{F}, P) . For $\mu \in \Omega$ and $(s, z, \sigma) \in \text{supp } \mu$, we set $\mu^\theta(s, z, \sigma) = \mu - \delta_{(s, z, \sigma)} + \delta_{(s, z + \theta, \sigma)}$, which belongs to Ω if $z + \theta \in E$.

DEFINITION 1.1. We define an m -th derivative of Φ by

$$\mathcal{D}^m \Phi(\mu; s, z, \sigma) = D_\theta^m \Phi(\mu^\theta(s, z, \sigma))|_{\theta=0} \quad (n^m\text{-vector})$$

if it is well defined for Q -a. e. (μ, s, z, σ) , where D_θ^m is the m -th differential operator in θ (in the usual sense). Moreover define

$$\mathcal{D}^k(\mathcal{D}^m \Phi)(\mu; s, z, \sigma) = D_\theta^k(\mathcal{D}^m \Phi(\mu^\theta(s, z, \sigma); s, z + \theta, \sigma))|_{\theta=0};$$

then $\mathcal{D}^k \mathcal{D}^m = \mathcal{D}^{k+m}$.

Next we introduce basic functional spaces $\mathcal{R}^{(m)}$ ($m=0, 1, 2, \dots, \infty$). Let $C_{0,E}^m$ denote the totality of real-valued measurable functions $f(s, z, \sigma)$ on $[0, T] \times E \times S$ those which vanish outside $[0, T] \times K \times S$ for some compact set

K in E , of class C^m in z and their derivatives in z up to order m are bounded in (s, z, σ) . We set

$$\mathcal{R}^{(m)} = \{\Phi(\mu) = F(\mu(f_1), \dots, \mu(f_k)) : F \in C_p^m(\mathbf{R}^k), f_i \in C_{0,E}^m, i=1, \dots, k, k=1, 2, \dots\},$$

where $\mu(f) = \int_{[0,T] \times E \times S} f(s, z, \sigma) \mu(ds dz d\sigma)$ and $C_p^m(\mathbf{R}^k)$ is the totality of real-valued C^m -functions on \mathbf{R}^k with polynomial growth derivatives up to order m . Then $\mathcal{R}^{(m)}$ is dense in $L^p = L^p(\mathcal{Q}, P)$ for all $1 \leq p < \infty$, stable under C_p^m -composition (i.e. if $\Phi = (\Phi^1, \dots, \Phi^k)$, $\Phi^i \in \mathcal{R}^{(m)}$ and $F \in C_p^m(\mathbf{R}^k)$ then $F(\Phi) \in \mathcal{R}^{(m)}$) and $\mathcal{F}_T^0 = \sigma(\mathcal{R}^{(m)})$. Clearly every element of $\mathcal{R}^{(m)}$ is m -times differentiable in the sense of Definition 1.1.

Now we fix an auxiliary function $\rho(z) \in C_b^\infty(E \rightarrow (0, \infty))$ and introduce some higher order symmetric differential operators.

DEFINITION 1.2. (i) For $m \geq 1$ and $\Phi \in \mathcal{R}^{(2m)}$, set

$$\mathcal{L}^{(m)}\Phi = \mu \left\{ \sum_{k=0}^m \binom{m}{k} D_z^k \rho(\mathcal{D}^2(\text{trace } \mathcal{D}^2)^{m-k} \Phi)^\top \right\},$$

where $\text{trace } \mathcal{D}^2 \Phi(\mu; s, z, \sigma) = \Delta_\theta \Phi(\mu^\theta(s, z, \sigma))|_{\theta=0}$ (Δ_θ is the Laplacian with respect to the variable θ). $\mathcal{L}^{(m)}\Phi$ is called the $\mathcal{L}^{(m)}$ -derivative of Φ .

(ii) For $m \geq 1$ and $\Phi, \Psi \in \mathcal{R}^{(2m)}$, set

$$\langle \mathcal{D}^m \Phi, \mathcal{D}^m \Psi \rangle = \mu(\rho \mathcal{D}^m \Phi \mathcal{D}^m \Psi^\top).$$

Then we have the following proposition by a similar argument to Proposition 9-3 of [2].

PROPOSITION 1.3. For $m \geq 1$, $\mathcal{L}^{(m)} : \mathcal{R}^{(2m)} \rightarrow \bigcap_{p < \infty} L^p$ is a linear operator such that for $\Phi, \Psi \in \mathcal{R}^{(2m)}$

$$(1.1) \quad E[\Phi \mathcal{L}^{(m)}\Psi] = E[\Psi \mathcal{L}^{(m)}\Phi] = (-1)^m E[\langle \mathcal{D}^m \Phi, \mathcal{D}^m \Psi \rangle].$$

Set $\mathcal{L} = \mathcal{L}^{(1)}$ and $\mathcal{R} = \mathcal{R}^{(2)}$; then it holds that for $\Phi = F(\mu(f_1), \dots, \mu(f_k))$ and $\Psi = H(\mu(h_1), \dots, \mu(h_q)) \in \mathcal{R}$

$$\begin{aligned} \mathcal{D}\Phi &= \sum_{i=1}^k \partial_i F(\mu(f_1), \dots) D_z f_i, \\ \mathcal{L}\Phi &= \mu(\rho \text{trace } \mathcal{D}^2 \Phi + D_z \rho \mathcal{D}\Phi^\top) \\ &= \sum_{i=1}^k \partial_i F(\mu(f_1), \dots) \mu(\rho \Delta_z f_i + D_z \rho D_z f_i^\top) \\ &\quad + \sum_{i,j=1}^k \partial_{ij}^2 F(\mu(f_1), \dots) \mu(\rho D_z f_i D_z f_j^\top), \\ \langle \mathcal{D}\Phi, \mathcal{D}\Psi \rangle &= \mu(\rho \mathcal{D}\Phi \mathcal{D}\Psi^\top) \\ &= \sum_{i=1}^k \sum_{j=1}^q \partial_i F(\mu(f_1), \dots) \partial_j H(\mu(h_1), \dots) \mu(\rho D_z f_i D_z h_j^\top). \end{aligned}$$

Moreover if $\Phi = (\Phi^1, \dots, \Phi^d)$, $\Phi^i \in \mathcal{R}$ and $F \in C_p^2(\mathbf{R}^d)$, then

$$\begin{aligned} \mathcal{D}(F \circ \Phi) &= \sum_{i=1}^d \partial_i F(\Phi) \mathcal{D}\Phi^i, \\ \mathcal{L}(F \circ \Phi) &= \sum_{i=1}^d \partial_i F(\Phi) \mathcal{L}\Phi^i + \sum_{i,j=1}^d \partial_{ij}^2 F(\Phi) \langle \mathcal{D}\Phi^i, \mathcal{D}\Phi^j \rangle. \end{aligned}$$

In particular, for $\Phi, \Psi \in \mathcal{R}$

$$\mathcal{L}(\Phi\Psi) = \Phi \mathcal{L}\Psi + 2\langle \mathcal{D}\Phi, \mathcal{D}\Psi \rangle + \Psi \mathcal{L}\Phi.$$

Proof. We shall show only (1.1) since the proofs of the other parts are quite routine. Let $\Phi = F(\mu(f_1), \dots, \mu(f_k))$ and $\Psi = H(\mu(h_1), \dots, \mu(h_q)) \in \mathcal{R}^{(2m)}$, and K be a compact set in E such that $[0, T] \times K \times S$ contains all the supports of f_i 's and h_i 's. Now set $(S_i, Z_i, A_i)_{i \leq N} = \text{supp } \mu \cap ([0, T] \times K \times S)$ with $S_1 < \dots < S_N$, where $N = \mu([0, T] \times K \times S) \geq 0$. Then N is a Poisson random variable with a parameter $T|K|$, and the random variables $(Z_i)_{i \leq N}$ are conditionally independent on the σ -field $\mathcal{G} \equiv \sigma(N, S_1, \dots, S_N)$ with the uniform distribution over K , $U_K(dz) = dz/|K|$; that is, for each $r \geq 1$ and bounded Borel functions g on K^r , it holds

$$\begin{aligned} E[g(Z_1, \dots, Z_r) \mathbf{1}_{\{N=r\}} | \mathcal{G}] \\ = \int_{K^r} g(z_1, \dots, z_r) U_K(dz_1) \cdots U_K(dz_r) \mathbf{1}_{\{N=r\}}. \end{aligned}$$

So it is enough to show that

$$(1.2) \quad E[\langle \mathcal{D}^m \Phi, \mathcal{D}^m \Psi \rangle | \mathcal{G}] = (-1)^m E[\Phi \mathcal{L}^{(m)} \Psi | \mathcal{G}].$$

For simplicity of notation we set $f_i^j(z) = f_i(S_j, z, A_j)$ and $h_i^j(z) = h_i(S_j, z, A_j)$. For a fixed $\mu \in \Omega$, set $N = N(\mu)$ and

$$\begin{aligned} g(z_1, \dots, z_N) &= \sum_{r=1}^N \rho(z_r) D_{z_r}^m \left(F \left(\sum_{s=1}^N f_i^s(z_s), \dots, \sum_{s=1}^N f_k^s(z_s) \right) \right) \\ &D_{z_r}^m \left(H \left(\sum_{s=1}^N h_1^s(z_s), \dots, \sum_{s=1}^N h_q^s(z_s) \right) \right)^\top; \end{aligned}$$

then we have $\langle \mathcal{D}^m \Phi, \mathcal{D}^m \Psi \rangle = \mu(\rho \mathcal{D}^m \Phi \mathcal{D}^m \Psi^\top) = g(Z_1, \dots, Z_N)$. Moreover

$$\begin{aligned} E[\langle \mathcal{D}^m \Phi, \mathcal{D}^m \Psi \rangle | \mathcal{G}] &= \int_{K^N} U_K(dz_1) \cdots U_K(dz_N) g(z_1, \dots, z_N) \\ &= \sum_{r=1}^N \sum_{|\alpha|=m} \frac{m!}{\alpha!} \int_{K^N} U_K(dz_1) \cdots U_K(dz_N) \rho(z_r) \\ &\quad \times D_{z_r}^\alpha \left(F \left(\sum_{s=1}^N f_i^s(z_s), \dots \right) \right) D_{z_r}^\alpha \left(H \left(\sum_{s=1}^N h_1^s(z_s), \dots \right) \right), \end{aligned}$$

where $\alpha=(\alpha_1, \dots, \alpha_n)$ is a multi-index and $\alpha!=\alpha_1! \cdots \alpha_n!$. Since $\text{supp} f_i^j$ and $\text{supp} h_i^j$ are contained in K , $D_{z_r}^\alpha(F(\cdots))$ and $D_{z_r}^\alpha(H(\cdots))$ vanish in K^c if $|\alpha| \geq 1$. Hence by using the classical integration-by-parts formula we get

$$\begin{aligned} E[\langle \mathcal{D}^m \Phi, \mathcal{D}^m \Psi \rangle | \mathcal{G}] \\ = (-1)^m \sum_{r=1}^N \sum_{|\alpha|=m} \frac{m!}{\alpha!} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \int_{K^N} U_K(dz_1) \cdots U_K(dz_N) \\ F\left(\sum_{s=1}^N f_i^s(z_s), \dots\right) D_{z_r}^\beta \rho(z_r) D_{z_r}^{2\alpha-\beta} \left(H\left(\sum_{s=1}^N h(z_s), \dots\right)\right). \end{aligned}$$

On the other hand it is easy to see that

$$\begin{aligned} \mathcal{L}^{(m)} \Psi &= \mu \left\{ \sum_{k=0}^m \binom{m}{k} \mathcal{D}_z^k \rho(\mathcal{D}^k (\text{trace } \mathcal{D}^2)^{m-k} \Psi)^\top \right\} \\ &= \sum_{r=1}^N \sum_{|\alpha|=m} \frac{m!}{\alpha!} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D_{z_r}^\beta \rho(Z_r) D_{z_r}^{2\alpha-\beta} \left(H\left(\sum_{s=1}^N h(Z_s), \dots\right)\right). \end{aligned}$$

Therefore (1.2) holds, and this completes the proof of (1.1). \square

Remark 1.4. $(\mathcal{L}, \mathcal{R})$ is essentially the same one as in [2]. Note that $\mathcal{L}^{(i)}: \mathcal{R}^{(2m)} \rightarrow \mathcal{R}^{(2m-2i)}$, $i=1, \dots, m$, but $\mathcal{L}^{(2)} \neq \mathcal{L}^2 \equiv \mathcal{L} \circ \mathcal{L}$.

Next we want to extend the domains of \mathcal{D}^m and $\mathcal{L}^{(m)}$, which are of interest for themselves, although they will be later used only in the case $m=1$.

Let \mathcal{H} be the completion of $\mathcal{F} \otimes \mathcal{B}([0, T]) \otimes \mathcal{E} \otimes \mathcal{S}$ by Q . For each $k \geq 1$ and $p \in (0, \infty)$, denote by $L_{(k)}^p$ the set of all $(\mathbf{R}^n)^k$ -valued \mathcal{H} -measurable functions F on $\Omega \times [0, T] \times E \times S$ with the norm $\|F\|_p = \|\mu(\rho |F|^2)^{1/2}\|_{L^p} < \infty$, where $\|\cdot\|_{L^p}$ denotes the L^p -norm in $L^p = L^p(\Omega, P)$. For each $k \geq 1$ and $p \in [1, \infty)$, $(L_{(k)}^p, |\cdot|_p)$ is a Banach space by identifying two functions which are equal Q -a. e. In particular, $L_{(k)}^2$ is a Hilbert space with the inner product $(F, G)_L = E[\langle F, G \rangle]$, where $\langle F, G \rangle = \mu(\rho FG^\top)$.

Let us introduce several spaces of Poisson functionals in a similar way to the case of the Wiener space ([9]).

DEFINITION 1.5. For $p_0, p_1, \dots, p_m \geq 2$, it is said that Φ belongs to $H(p_0, p_1, \dots, p_m)$ (resp. $H(p_0 | p_1, \dots, p_m)$) if there is a sequence $\{\Phi_k\}$ in $\mathcal{R}^{(m)}$ (resp. $\mathcal{R}^{(2m)}$) such that $\Phi_k \rightarrow \Phi$ in L^{p_0} and, for each i ($i=1, \dots, m$), $\{\mathcal{D}^i \Phi_k\}_{k=1}^\infty$ (resp. $\{\mathcal{L}^{(i)} \Phi_k\}_{k=1}^\infty$) is a Cauchy sequence in $L_{(i)}^{p_i}$ (resp. L^{p_i}). Denote the limit by $\mathcal{D}^i \Phi$ (resp. $\mathcal{L}^{(i)} \Phi$), which is called the i -th weak derivative (resp. the $\mathcal{L}^{(i)}$ -weak derivative) of Φ . Moreover set, for $p_0, p_1, \dots, p_m, q_1, \dots, q_r \geq 2$,

$$H(p_0, p_1, \dots, p_m | q_1, \dots, q_r) = H(p_0, p_1, \dots, p_m) \cap H(p_0 | q_1, \dots, q_r).$$

By using (1.1) we shall show that $\mathcal{D}^m \Phi$ and $\mathcal{L}^{(m)} \Phi$ are well defined for every $\Phi \in H(p_0, p_1, \dots, p_m)$ and $\Phi \in H(p_0 | p_1, \dots, p_m)$ respectively. It is easy to see that the approximating sequences can be chosen in $\mathcal{R}^{(\infty)}$. Suppose that

$\Phi_k \in \mathcal{R}^{(\infty)} \rightarrow 0$ in L^2 and $\mathcal{D}^i \Phi_k \rightarrow F_i$ in $L^2_{(i)}$; then for any $\Psi \in \mathcal{R}^{(\infty)}$

$$\begin{aligned} (F_i, \mathcal{D}^i \Psi)_L &= \lim_{k \rightarrow \infty} (\mathcal{D}^i \Phi_k, \mathcal{D}^i \Psi)_L \\ &= (-1)^i \lim_{k \rightarrow \infty} E[\Phi_k \mathcal{L}^{(i)} \Psi] \quad (\text{by (1.1)}) \\ &= 0. \end{aligned}$$

Thus $\|F_i\|_2^2 = (F_i, F_i)_L = \lim_{k \rightarrow \infty} (F_i, \mathcal{D}^i \Phi_k)_L = 0$, i. e., $F_i = 0$ Q -a. e. Also suppose that $\mathcal{L}^{(i)} \Phi_k \rightarrow \theta_i$ in L^2 ; then for all $\Psi \in \mathcal{R}^{(\infty)}$

$$E[\theta_i \Psi] = \lim_{k \rightarrow \infty} E[\Psi \mathcal{L}^{(i)} \Phi] = \lim_{k \rightarrow \infty} E[\Phi_k \mathcal{L}^{(i)} \Psi] = 0.$$

Since $\mathcal{R}^{(\infty)}$ is dense in L^2 , we have $\theta_i = 0$ a. s.

Clearly the above spaces in Definition 1.5 are Banach spaces with the following norms respectively:

$$\|\Phi\|_{H(p_0, \dots, p_m)} = \|\Phi\|_{L^{p_0}} + \sum_{i=1}^m \|\mathcal{D}^i \Phi\|_{p_i},$$

$$\|\Phi\|_{H(p_0 | q_1, \dots, q_m)} = \|\Phi\|_{L^{p_0}} + \sum_{i=1}^m \|\mathcal{L}^{(i)} \Phi\|_{L^{q_i}}$$

and

$$\|\Phi\|_{H(p_0, p_1, \dots, p_m | q_1, \dots, q_r)} = \|\Phi\|_{L^{p_0}} + \sum_{i=1}^m \|\mathcal{D}^i \Phi\|_{p_i} + \sum_{i=1}^r \|\mathcal{L}^{(i)} \Phi\|_{L^{q_i}}.$$

In particular, from (1.3) of the following proposition it follows that

$$H_{\frac{1}{2}}^{(m)} \equiv H(2, \underbrace{2, \dots, 2}_m | 2, \dots, 2) = H(2 | \underbrace{2, \dots, 2}_m),$$

PROPOSITION 1.6. a) $\mathcal{L}^{(m)} : H_{\frac{1}{2}}^{(m)} \rightarrow L^2$ is a symmetric linear operator;
 b) if we set $\Gamma^m(\Phi, \Psi) = \mu(\rho \mathcal{D}^m \Phi \mathcal{D}^m \Psi^T)$, then $\Gamma^m : H_{\frac{1}{2}}^{(m)} \times H_{\frac{1}{2}}^{(m)} \rightarrow L^1$ is a continuous bilinear symmetric nonnegative operator satisfying

$$(1.3) \quad E[\Phi \mathcal{L}^{(m)} \Psi] = E[\Psi \mathcal{L}^{(m)} \Phi] = (-1)^m E[\Gamma^m(\Phi, \Psi)].$$

The proof is obvious.

By using them we can define a Malliavin operator (\mathcal{L}, H_∞) as in [2].

DEFINITION 1.7. For $q \in [2, \infty)$, denote by H_q the Banach space $H(q, q | q)$ with the norm $\|\Phi\|_{H_q} = \|\Phi\|_{L^q} + \|\mathcal{L} \Phi\|_{L^q} + \|\mathcal{D} \Phi\|_q$, and set $H_\infty \equiv \bigcap_{q \geq 2} H_q$, which is a Fréchet space. Then (\mathcal{L}, H_∞) (resp. $\Gamma(\Phi, \Psi) \equiv \Gamma^1(\Phi, \Psi)$) is called a Malliavin operator (resp. a Malliavin covariance of (Φ, Ψ)).

We here list several properties of the Malliavin operator and the Malliavin covariance which can be checked by the same arguments as in Theorem 8-18 of [2].

THEOREM 1.8. Let (\mathcal{L}, H_∞) be the Malliavin operator.

a) $\mathcal{L} : H_\infty \subset \bigcap_{p < \infty} L^p \rightarrow \bigcap_{p < \infty} L^p$ is a linear operator and satisfies the following properties:

- (i) H_∞ is stable under C_p^2 -composition and $\mathfrak{F}_T^0 = \sigma(H_\infty)$;
- (ii) \mathcal{L} is symmetric in L^2 , i.e., $E[\Phi \mathcal{L} \Psi] = E[\Psi \mathcal{L} \Phi]$;
- (iii) $\Gamma : H_\infty \times H_\infty \rightarrow \bigcap_{p < \infty} L^p$ is a continuous bilinear symmetric nonnegative operator;
- (iv) if $\Phi = (\Phi^1, \dots, \Phi^d)$, $\Phi^i \in H_\infty$ and $F \in C_p^2(\mathbf{R}^d)$ then

$$\mathcal{D}(F \circ \Phi) = \sum_{i=1}^d \partial_i F(\Phi) \mathcal{D} \Phi^i,$$

$$\mathcal{L}(F \circ \Phi) = \sum_{i=1}^d \partial_i F(\Phi) \mathcal{L} \Phi^i + \sum_{i,j=1}^d \partial_{ij}^2 F(\Phi) \Gamma(\Phi^i, \Phi^j).$$

In particular, for $\Phi, \Psi \in H_\infty$

$$\mathcal{L}(\Phi \Psi) = \Phi \mathcal{L} \Psi + 2\Gamma(\Phi, \Psi) + \Psi \mathcal{L} \Phi.$$

- b) (i) If $\Phi = (\Phi^1, \dots, \Phi^d)$, $\Phi^i \in H_\infty$, $F \in C_p^2(\mathbf{R}^d)$ and $\Psi \in H_\infty$ then

$$\Gamma(F \circ \Phi, \Psi) = \sum_{i=1}^d \partial_i F(\Phi) \Gamma(\Phi^i, \Psi);$$

- (ii) $\mathcal{L} 1 = 0$, $E[\mathcal{L} \Phi] = 0$ and $E[\Phi \mathcal{L} \Psi] = E[\Psi \mathcal{L} \Phi] = -E[\Gamma(\Phi, \Psi)]$.

c) For $\Phi = (\Phi^1, \dots, \Phi^d) \in (H_\infty)^d$, $F \in C_p^2(\mathbf{R}^d)$ and $\Psi \in H_\infty$, the integration-by-parts formula holds:

$$E\left(\Psi \sum_{i=1}^d \partial_i F(\Phi) \Gamma(\Phi^i, \Phi^i)\right) = -E\left(F(\Phi) \{\Psi \mathcal{L} \Phi^j + \Gamma(\Phi^j, \Psi)\}\right).$$

§ 2. Applications of Malliavin calculus to stochastic differential equations.

Now exactly in the same way as in [2], we can apply the Malliavin operator to a solution of the following stochastic differential equation driven by the Poisson random measure $\mu(dsdzd\sigma)$:

$$(2.1) \quad X_t^x = x + \int_0^t a(X_s^x) ds + \int_0^t \int_{E \times S} f(X_{s-}^x, z, \sigma) \tilde{\mu}(dsdzd\sigma) \quad (x \in \mathbf{R}^d),$$

where $\tilde{\mu} = \mu - \nu$ is a compensated Poisson random measure and

$$a(x) = (a^i(x)) \quad : \mathbf{R}^d \rightarrow \mathbf{R}^d,$$

$$f(x, z, \sigma) = (f^i(x, z, \sigma)) : \mathbf{R}^d \times E \times S \rightarrow \mathbf{R}^d$$

are measurable functions. We also write (2.1) in the differential form

$$dX^x = a(X^x)dt + f(X^x)d\tilde{\mu}, \quad X_0^x = x.$$

We give an assumption on $a(x)$ and $f(x, z, \sigma)$. Let $r \geq 1$.

- Assumption (A-r). (i) $a(x)$ is a C_b^r -function on \mathbf{R}^d ;
(ii) $f(x, z, \sigma)$ is r -times differentiable in (x, z) on $\mathbf{R}^d \times E$ and moreover

$$\begin{aligned} f(0, \cdot, \cdot) &\in \bigcap_{2 \leq p \leq \infty} L^p(E \times S, dzM(d\sigma)), \\ \sup_x |D_x^i f(x, \cdot, \cdot)| &\in \bigcap_{2 \leq p \leq \infty} L^p(E \times S, dzM(d\sigma)) \quad \text{for } 1 \leq i \leq r, \\ \sup_{x, z, \sigma} |D_x^i D_z^j f(x, z, \sigma)| &< \infty \quad \text{for } 1 \leq i+j \leq r \text{ and } j \geq 1. \end{aligned}$$

We shall always assume at least (A-1). Then (2.1) has a unique solution X_t^x , which satisfies $\sup_{t \leq T} |X_t^x| \in \bigcap_{p < \infty} L^p$ and defines a standard strong Markov process X^x with the generator L :

$$L\varphi(x) = D_x \varphi(x) \cdot a(x) + \int [\varphi(x+y) - \varphi(x) - D_x \varphi(x) \cdot y] K(x, dy)$$

where K is given by

$$K(x, A) = \int_{E \times S} 1_{A \setminus \{0\}}(f(x, z, \sigma)) dzM(d\sigma).$$

From now on, we fix a Malliavin operator with an auxiliary function ρ satisfying

Condition (ρ). $\rho: E \rightarrow (0, \infty)$ is a C_b^∞ -function such that

- (i) $\rho(z)/d(z)^2, |D_z \rho(z)|/d(z) \in \bigcap_{1 \leq p < \infty} L^p(E, dz)$, where $d(z) \equiv \inf_{y \in E^c} |z-y| \wedge 1$ (if $E^c \equiv \mathbf{R}^n \setminus E \neq \emptyset$), $\equiv 1$ (if $E = \mathbf{R}^n$);
(ii) for $r \geq 0, |D_z^r \rho| \in L^1(E, dz)$.

For simplicity we consider the equation (2.1) with a fixed initial value $x \in \mathbf{R}^d$ and we denote the solution by $X_t = X_t^x$. Then we have

THEOREM 2.1 (cf. [2], Theorem 10-3). Assume (A-3). Then for each $t \leq T$ the components X_t^i of the solution X_t belong to H_∞ . Moreover set $U^{ij} = \Gamma(X_t^i, X_t^j)$ and $V^i = \mathcal{L} X_t^i$; then the processes $U = (U^{ij})_{i, j \leq d}$ and $V = (V^i)_{i \leq d}$ are solutions to the following equations, respectively:

$$\left\{ \begin{aligned} dU &= [U D_x a(X)^\top + D_x a(X) U] dt \\ &\quad + [U \cdot D_x f(X_-)^\top + D_x f(X_-) U \cdot] d\tilde{\mu} \\ &\quad + [\rho D_z f(X_-) D_z f(X_-)^\top + D_x f(X_-) U \cdot D_x f(X_-)^\top] d\mu, \\ U_0 &= 0, \end{aligned} \right.$$

and

$$\left\{ \begin{aligned} dV &= D_x a(X) V dt + D_x f(X_-) V \cdot d\tilde{\mu} \\ &\quad + \sum_{i, j=1}^d [\partial_{i,j}^2 a(X) U^{ij} dt + \partial_{i,j}^2 f(X_-) U^{ij} d\tilde{\mu}] \\ &\quad + [\rho \Delta_z f(X_-) + D_z \rho D_z f(X_-)^\top] d\mu, \\ V_0 &= 0, \end{aligned} \right.$$

where $\partial_{i_j}^2 = \partial^2 / (\partial x_i \partial x_j)$.

Finally we modify the results of [2] according to the present situation, which will be used to prove our main results of §3.

Denote by ∇X_t^x the Fréchet derivative of the map $x \in \mathbf{R}^d \rightarrow X_t^x \in L^p$. Then under (A-2) ∇X_t^x exists for each $x \in \mathbf{R}^d$ and $t \in [0, T]$, and it is a unique solution to the following linear stochastic differential equation:

$$d\nabla X^x = D_x a(X^x) \nabla X^x dt + D_x f(X^x) \nabla X^x d\bar{\mu}, \quad \nabla X_0^x = I$$

([2] Theorem 6-29). Let us introduce the $\mathbf{R}^d \otimes \mathbf{R}^d$ -valued process K^x defined by

$$dK^x = D_x a(X^x) dt + D_x f(X^x) d\bar{\mu}, \quad K_0^x = 0.$$

Then

$$d\nabla X^x = dK^x \nabla X^x, \quad \nabla X_0^x = I.$$

Define the sequence of stopping times:

$$T_0^x = 0, \quad T_n^x = \inf\{t > T_{n-1}^x : \det(I + \Delta K_t^x) = 0\} \quad (= T \text{ if } \{\cdot\} = \phi),$$

where $\Delta K_t^x = K_t^x - K_{t-}^x$. We further define the process $\nabla X^x(n)$ by

$$d\nabla X^x(n) = dK^x(n) \nabla X^x(n)_-, \quad \nabla X^x(n)_0 = I,$$

where $K^x(n)_t = K_t^x - K_{t \wedge T_{n-1}^x}^x$. Clearly $K^x(1) = K^x$ and $\nabla X^x(1) = \nabla X^x$. Since K^x is a right continuous process with left-hand limits and $\{0 < t < T : \det(I + \Delta K_t^x) = 0\}$ is a finite set, $T_n^x \uparrow T$. It is also known that ([5])

$$\nabla X^x(n)_t = I \text{ if } t \leq T_{n-1}^x,$$

$$\nabla X^x(n)_t \text{ is invertible if } t < T_n^x,$$

$$\nabla X^x(n)_t \text{ is not invertible if } t \geq T_n^x,$$

$$\nabla X^x(n)_{t-} \text{ is invertible if } t \leq T_n^x.$$

Moreover we define the process H^x by

$$dH^x = \rho D_x f D_x f^\top(X^x) d\mu \quad \text{and} \quad H_0^x = 0.$$

Then it is shown in [2] (Proposition 10-22) that under (A-3)

$$U_t^x = \nabla X^x(n)_t \left(U_{T_{n-1}^x}^x + \int_{T_{n-1}^x}^t \int_{E \times S} \{ \nabla X^x(n)_{s-}^{-1} (I + D_x f)^{-1} D_x f D_x f^\top \right.$$

$$\left. (I + D_x f)^{-1, \top} (X_{s-}^x, z, \sigma) \nabla X^x(n)_{s-}^{-1, \top} \rho(z) \} \mu(ds dz d\sigma) \right) \nabla X^x(n)_t^\top$$

$$\text{if } T_{n-1}^x \leq t < T_n^x,$$

$$U_{\bar{x}}^x = \Delta H_{\bar{x}}^x + (I + \Delta K_{\bar{x}}^x) U_{\bar{x}}^x - (I + \Delta K_{\bar{x}}^x)^T \quad \text{and} \quad U_0^x = 0.$$

Therefore, by Theorem 10-30 (a) of [2], we see that, under (A-4), if $U_{\bar{x}}$ is invertible a.s. then the distribution of $X_{\bar{x}}$ has a density. Furthermore, using the above formulae and the arguments in §11 of [2], we can obtain the following lemma, which will be used to prove Theorem 3.1.

LEMMA 2.2. *Let $t > 0$. Under (A-4) if, for each $n \geq 1$,*

$$\int_{T_{n-1}^x} \int_{E \times S} [\nabla X^x(n)_{s-}^{-1} (I + D_x f)^{-1} D_x f D_x f^T (I + D_x f)^{-1, T} (X_{s-}^x, z, \sigma) \nabla X^x(n)_{s-}^{-1, T} \rho(z)] \mu(ds dz d\sigma)$$

is invertible a.s. on $\{\mu: T_{n-1}^x(\mu) < t < T_n^x(\mu)\}$, then the distribution of $X_{\bar{x}}$ has a density.

To obtain the smoothness results, we introduce a function $h(z)$ on E following [2], which is related to the integrability of $(U_{\bar{x}})^{-1}$.

DEFINITION 2.3. Let $\zeta, \theta > 0$ be given. A measurable function $h: E \rightarrow [0, \infty)$ is called (ζ, θ) -broad if

$$\int_0^\infty ds s^{\zeta-1} \exp\left\{-\theta \int_E (1 - e^{-sh(z)}) dz\right\} < \infty.$$

Example 2.4. Let $E = (\eta, \infty)$ for some $\eta \in \mathbf{R}$. Let $\zeta, \theta > 0$ be given and set $h(z) = |z|^\gamma e^{-\delta|z|}$ ($z \in E$) for some constants $\gamma, \delta \in \mathbf{R}$. If $\delta < \theta/\zeta$ then h is (ζ, θ) -broad for all $\gamma \in \mathbf{R}$. If $\delta = \theta/\zeta$ then h is also (ζ, θ) -broad for $\gamma > \delta/\theta = 1/\zeta$. Otherwise h is not (ζ, θ) -broad. In particular, if $\delta = 0$, that is, $h(z) = |z|^\gamma$ then h is (ζ, θ) -broad for all $\zeta, \theta > 0$ and $\gamma \in \mathbf{R}$.

The assumption (SB- (ζ, θ)) of [2] is modified as follows:

Assumption (B- (ζ, θ)). There exist constants $\varepsilon > 0, \delta \geq 0$, a (ζ, θ) -broad function h , a function ρ satisfying the condition (ρ) and a family of disjoint sets $\{S_i\}_{i=1}^N \subset \mathcal{S}$ with $M(S_i) = M_0 > 0$ such that for each $x, y \in \mathbf{R}^d$

$$\sum_{i=1}^N \inf_{z: h(z) > 0, \sigma \in S_i} \frac{\rho(z)}{h(z)} y^T F(x, z, \sigma) y \geq \varepsilon \frac{|y|^2}{1 + |x|^\delta},$$

where F is the function with values in the set of symmetric nonnegative definite $d \times d$ -matrices defined by

$$(2.2) \quad F(x, z, \sigma) = \begin{cases} (I + D_x f)^{-1} D_x f D_x f^T (I + D_x f)^{-1, T} (x, z, \sigma) & \text{if } \det(I + D_x f(x, z, \sigma)) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Now we can modify the results of [2] (Theorems 2-27, 2-28 and 2-29) in the following way.

THEOREM 2.5. Let $r \geq 0$ and $t > 0$. Assume (A-j) and (B-(ζ, θ)) for some j and $\zeta, \theta > 0$. Assume also $\inf_{x, z, \sigma} |\det(I + D_x f(x, z, \sigma))| > 0$. Then

a) the distribution of X_t^x has a density $p_t(x, \cdot)$ of class C^r for each $x \in \mathbf{R}^d$, provided either

(i) $j \geq r + d + 5$, $\theta \leq M_0 t$ and $\zeta > 2d(r + d + 1)/[M_0 t / \theta]$ (where $[a]$ denotes the integer part of a), or

(ii) $j \geq r + 5$, $\theta \leq M_0 t$ and $\zeta > 2d^2(r + 1)/[M_0 t / \theta]$.

b) Moreover the function: $(x, y) \rightarrow p_t(x, y)$ is of class C^r , provided that $j \geq r + 2d + 5$, $\theta \leq M_0 t$ and $\zeta > 2d(r + 2d + 1)/[M_0 t / \theta]$ hold.

c) Furthermore assume $\inf_{v \in [0, 1], x, z, \sigma} |\det(I + v D_x f(x, z, \sigma))| > 0$. If $j \geq 2r + 4d + 8$, $\theta \leq M_0 t$ and $\zeta > 4d(r + 2d + 2)/[M_0 t / \theta]$ then the function: $(s, x, y) \rightarrow p_s(x, y)$ is of class C^r on $(t, T) \times \mathbf{R}^d \times \mathbf{R}^d$.

Remark 2.6. In the above theorem, we only give the result corresponding to the first assertion (i) of Theorem 2-28 of [2]; the proof of the second assertion (ii) of Theorem 2-28 includes trivial mistakes (see the equations (4-24) and (4-37) in [2]) and it is hard to correct them. Here we need two more degrees of differentiability on the coefficients (a, f) than that in [2], because we use Malliavin's approach (see Remark 10-33 of [2]). However if we take Bismut's approach the regularity assumption on (a, f) would be improved.

§ 3. Main results.

By using Theorem 2.5 and Lemma 2.2 of § 2 we give some results on the existence and smoothness of a density for the transition probability $P_t(x, dy)$ of the jump Markov process associated with the Lévy generator L defined by (8). Here its Lévy kernel $K(x, dy)$ is given by (9) with the following functions:

$$y(x, u, \sigma) : \mathbf{R}^d \times [0, u_0] \times S \rightarrow \mathbf{R}^d, \quad y(x, 0, \sigma) = 0 \quad (\text{curves});$$

$$g(x, u, \sigma) : \mathbf{R}^d \times (0, u_0] \times S \rightarrow (0, \infty) \quad (\text{weight}).$$

We restrict the function $g(x, u, \sigma)$ as follows:

$$g(x, u, \sigma) = u^{-1 - \alpha(x)} m(x, \sigma),$$

where $m(x, \sigma) > 0$ and $0 \leq \alpha(x) < 2$.

In order to relate L and the equation (2.1) we introduce the following functions:

$$G(x, u, \sigma) = \int_u^{u_0} g(x, v, \sigma) dv \quad \text{for } u \in (0, u_0],$$

(3.1) $\hat{G}(x, \cdot, \sigma)$ is the inverse function of $G(x, \cdot, \sigma)$ and

$$(3.2) \quad f(x, z, \sigma) = y(x, \hat{G}(x, z, \sigma), \sigma).$$

Then K is expressed by

$$K(x, A) = \int_S M(d\sigma) \int_0^\infty 1_{A \setminus \{0\}}(f(x, z, \sigma)) dz.$$

We here choose $E = (0, \infty)$ and a Poisson space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ over $[0, T] \times E \times S$ with the intensity measure $dt dz M(d\sigma)$. Then we can consider the stochastic differential equation (2.1) associated with $a(x)$ and $f(x, z, \sigma)$ of (3.2) on the Poisson space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$; its solution solves the martingale problem for the Lévy generator L .

Now we give an assumption on (a, y, m, α) . Let $r \geq 1$.

Assumption $(L-r)$. (i) $a(x)$ is a C_b^r -function on \mathbf{R}^d ;

(ii) $y(x, u, \sigma)/u$ is r -times differentiable in (x, u) on $\mathbf{R}^d \times [0, u_0]$, $D_x^i D_u^j (y(x, u, \sigma)/u)$ is bounded in (x, u, σ) for all i, j with $0 \leq i + j \leq r$ and $D_u y(x, 0, \sigma)$ is continuous at every $\sigma \in \text{supp } M(d\sigma)$;

(iii) $m(x, \sigma)$ is r -times differentiable in x such that $0 < \xi_1 \leq m(x, \sigma) \leq \xi_2 < \infty$ and $D_x^k m(x, \sigma)$ is bounded in (x, σ) for all k with $0 \leq k \leq r$;

(iv) $\alpha(x) \equiv 0$, or $\alpha(x)$ is a C_b^r -function on \mathbf{R}^d and satisfies the inequality $0 < \zeta_1 \leq \alpha(x) \leq \zeta_2 < 2$ with some constants ζ_1, ζ_2 .

Under the assumption $(L-1)$ the equation (2.1) has a unique solution X_t^x for each initial condition $X_0^x = x$. Since $X^x = (X_t^x)$ defines a standard Markov process on \mathbf{R}^d , we denote its transition probability by $P_t(x, dy)$.

The next assumption is a non-degeneracy condition for $K(x, dy)$:

Assumption (N) .

$$\text{span}\{D_u y(x, 0, \sigma) : \sigma \in \text{supp } M(d\sigma)\} = \mathbf{R}^d$$

for every $x \in \mathbf{R}^d$.

Note that under $(L-1)$ and (N) , $K(x, dy)$ satisfies $K(x, \mathbf{R}^d) = \infty$ and $\int |y|^2 K(x, dy) < \infty$.

The first result of this paper is the following.

THEOREM 3.1 (Existence of transition density). *Assume that $(L-4)$ and (N) hold. Then $P_t(x, dy)$ has a density $P_t(x, \cdot)$ for all $t > 0$, $x \in \mathbf{R}^d$.*

Remark 3.2. When S is a finite set, this assertion was shown in Theorem 3-8 of [2]. Léandre [6] also discussed this problem under a weaker non-degeneracy condition than (N) , in the case where S is finite and $g(x, u, \sigma)$ is independent of x . We here generalize S and $g(x, u, \sigma)$, which enables us to treat a stable-like process. We shall prove the above theorem by combining two methods of [2] and [6].

Assumption (SN) . There exist a constant $\varepsilon > 0$ and a finite number of points $\sigma_i \in \text{supp } M(d\sigma)$, $i=1, \dots, N$ and disjoint neighborhoods S_i of σ_i , $i=1, \dots, N$ such that for all $x, y \in \mathbf{R}^d$

$$\sum_{i=1}^N \inf_{\sigma \in S_i} |y^\top D_u y(x, 0, \sigma)|^2 \geq \varepsilon |y|^2.$$

This is stronger than (N). The following assumptions are rather artificial, but they are required by technical reasons. Let $f(x, z, \sigma)$ be defined by (3.2).

Assumption (C). $\inf_{x, z, \sigma} |\det(I + D_x f(x, z, \sigma))| > 0$.

Assumption (SC). $\inf_{v \in [0, 1], x, z, \sigma} |\det(I + v D_x f(x, z, \sigma))| > 0$.

These assumptions hold, for instance, if either $|y/u|$, $|D_x y/u|$, $|D_u(y/u)|$, or $|D_x y/u|$, $|D_x m|$, $|D_x \alpha|$ are sufficiently small, or if u_0 is sufficiently small. These conditions will be discussed in Remark 3.6.

Now we can state the main result of this paper.

THEOREM 3.3 (Smoothness of transition density). *Assume that (SN) and (C) hold. Let $r \geq 1$ and let $M_0 = \min_{1 \leq i \leq N} M(S_i)$ where S_i is found in (SN).*

a) Let $0 < \zeta_1 \leq \alpha(x) \leq \zeta_2 < 2$.

- (i) Under $(L - (r+5))$, $P_t(x, dy)$ has a density $P_t(x, y)$ of class C^r in y for all $x \in \mathbf{R}^d$, $t > 0$.
- (ii) Under $(L - (r+2d+5))$, $p_t(x, y)$ is of class C^r in (x, y) for all $t > 0$.
- (iii) Assume further that (SC) holds. Then, under $(L - (2r+4d+8))$, $p_t(x, y)$ is of class C^r in (t, x, y) of $(0, \infty) \times \mathbf{R}^d \times \mathbf{R}^d$.
- (iv) In the cases (i), (ii) and (iii), if $(L-r)$ is fulfilled for all $r \geq 1$ then the term C^r can be replaced by C^∞ .

b) Let $\alpha(x) \equiv 0$, and set $\eta = 1/\xi_1$.

- (i) Under $(L - (r+d+5))$ (resp. $(L - (r+5))$), $P_t(x, dy)$ has a density $p_t(x, y)$ of class C^r in y , if $t > \{4d(r+d+1)\eta + 2\eta^2\}/M_0$ (resp. $t > \{4d^2(r+1)\eta + 2\eta^2\}/M_0$).
- (ii) Under $(L - (r+2d+5))$, $p_t(x, y)$ of class C^r in (x, y) , if $t > \{4d(r+2d+1)\eta + 2\eta^2\}/M_0$.
- (iii) Assume further that (SC) holds. Then, under $(L - (2r+4d+8))$, $p_s(x, y)$ is of class C^r in (s, x, y) of $(t, \infty) \times \mathbf{R}^d \times \mathbf{R}^d$, if $t > \{8d(r+2d+2)\eta + 2\eta^2\}/M_0$.

Remark 3.4. (i) The result corresponding to (b) of Theorem 3.3 is found in Theorem 3-12 of [2], but their assumption involves $f(x, z, \sigma)$. However it would be desirable to give the assumption in terms of the Lévy generator (8) with (9), without using $f(x, z, \sigma)$.

(ii) As mentioned in Remark 2.6 it is possible to drop one or two degrees of differentiability on the coefficients (a, y, m, α) in our results. However it is not interest for us.

Finally we shall apply our results to a stable-like process with bounded jumps and a stable-like process itself. Recall that the generator of a stable-like process is decomposed into $L = A^0 + A^1$ where

$$A^0 \varphi(x) = \int_{|y| < u_0} [\varphi(x+y) - \varphi(x) - D_x \varphi(x) \cdot y] K(x, dy),$$

$$A^1 \varphi(x) = \int_{|y| \geq u_0} [\varphi(x+y) - \varphi(x)] K(x, dy)$$

and $K(x, dy)$ is given by (6).

Assume that $\alpha(x)$ is a C_b^1 -function satisfying that the inequalities (5). Since A^0 is a special type of (8), there corresponds a Markov process on \mathbf{R}^d which is defined by the unique solution of the SDE (2.1) with $f(x, z, \sigma)$ of (3.2) on a Poisson space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ over $[0, T] \times (0, \infty) \times S^{d-1}$. We denote by $P_t^0(x, dy)$ its transition probability.

On the other hand Bass [1] establishes the well-posedness of the martingale problem for L , so that L determines a Markov process on \mathbf{R}^d uniquely. We denote by $P_t(x, dy)$ its transition probability. Then we obtain

THEOREM 3.5. *Suppose that $\alpha(x)$ is a C_b^j -function on \mathbf{R}^d for some j , and satisfies (5). Let $r \geq 0$.*

- a) *(Stable-like process with bounded jumps)*
 - (i) *If $j \geq 4$ then $P_t^0(x, dy)$ has a density $p_t^0(x, y)$ for all $t > 0$ and $x \in \mathbf{R}^d$.*
 - (ii) *Assume (C). If $j \geq r + 5$ then $p_t^0(x, y)$ is of class C^r in y for all $t > 0$, $x \in \mathbf{R}^d$. If $j \geq r + 2d + 5$ then $p_t^0(x, y)$ is of class C^r in (x, y) for all $t > 0$.*
 - (iii) *Assume (SC). If $j \geq 2r + 4d + 8$ then $p_t^0(x, y)$ is of class C^r in (t, x, y) of $(0, \infty) \times \mathbf{R}^d \times \mathbf{R}^d$.*

b) *(Stable-like process)*

If $j \geq 4$ then $P_t(x, dy)$ has a density $p_t(x, y)$ given by

$$(3.5) \quad p_t(x, y) = p_t^0(x, y) + \int_0^t ds \int P_{t-s}(x, dy_1) A^1 p_s^0(y_1, y)$$

for all $t > 0$ and $x \in \mathbf{R}^d$.

Proof. a) It is easy to check that $m(x, \sigma) = \theta_{\alpha(x)}$ satisfies $(L-j)$ -(iii). Hence the claims follow from Theorem 3.1 and Theorem 3.3(a).

b) Let $C_\infty = C_\infty(\mathbf{R}^d)$ be the Banach space of continuous functions on \mathbf{R}^d vanishing at infinity with supremum norm. It is easy to see that $P_t^0(x, dy)$ induces an s -continuous conservative Markov semi-group $\{T_t^0\}$ acting on C_∞ , and that A^1 is a bounded generator of a Markov semi-group acting on C_∞ . Let $(G^0, \mathcal{D}(G^0))$ be the generator of $\{T_t^0\}$, and set $G = G^0 + A^1$ and $\mathcal{D} \equiv \mathcal{D}(G) = \mathcal{D}(G^0)$. Then $(G, \mathcal{D}(G))$ generates an s -continuous conservative Markov semi-group $\{T_t\}$ acting on C_∞ . Furthermore the following perturbation formula holds:

$$(3.6) \quad T_t = T_t^0 + \int_0^t T_{t-s} A^1 T_s^0 ds \quad \text{on } C_\infty$$

(see Theorem 3.1 of [3], also Chap. 1 and 4 of [4]). Then there is a Markov process (X_t, P_x) associated with the Markov semi-group $\{T_t\}$. Since $\mathcal{D} \supset C_b^2(\mathbf{R}^d)$ and $G = L$ on $C_b^2(\mathbf{R}^d)$, the probability law of (X_t) under P_x is a solution of the martingale problem for $(L, C_b^2(\mathbf{R}^d))$. Hence by the well-posedness of the martingale problem, $\{T_t\}$ coincides with the semi-group associated with $P_t(x, dy)$. Therefore (3.5) follows from (3.6). \square

Remark 3.6. a) To obtain the smoothness results of $P_t(x, dy)$ for a

stable-like process, the formula (3.5) might be useful. But we need more information on integrability of derivatives of $P_t^y(x, y)$, which seems rather difficult.

b) For the assumptions (C) and (SC) we can give a sufficient condition as follows: Set $n(x, u, \sigma) = y(x, u, \sigma)/u$. Then, under (L-r), $n(x, u, \sigma)$ is r -times differentiable in (x, u) on $\mathbf{R}^d \times [0, u_0]$, its derivatives are bounded in (x, u, σ) up to order r , and

$$f(x, z, \sigma) = \hat{G}(x, z, \sigma)n(x, \hat{G}(x, z, \sigma), \sigma).$$

(i) In the case of $\alpha(x) \equiv 0$, $\hat{G}(x, z, \sigma) = u_0 \exp[-z/m(x, \sigma)]$. Then we have

$$\begin{aligned} |D_x f(x, z, \sigma)| &\leq u_0 (|n| + |D_u n|)(x, \hat{G}(x, z, \sigma), \sigma) |D_x m(x, \sigma)| / \xi_1 \\ &\quad + |D_x n(x, \hat{G}(x, z, \sigma), \sigma)|. \end{aligned}$$

From this we deduce that if either $|n|$, $|D_u n|$, $|D_x n|$ or $|D_x m|$, $|D_x n|$ are sufficiently small then (SC) holds.

For instance, suppose that all the functions α, n, m are independent of (x_2, \dots, x_d) . Then $\det(I + vD_x f) = 1 + v\partial_1 f$ for $v \in [0, 1]$. In this case (C) and (SC) hold if $u_0 < \xi_1 / (\|n\| + \|D_u n\| \|D_x m\| + \xi_1 \|D_x n\|)$, where $\|\cdot\|$ denotes the supremum norm with respect to the variables x, u and σ .

(ii) In the case of $0 < \zeta_1 \leq \alpha(x) \leq \zeta_2 < 2$, $\hat{G}(x, z, \sigma) = (\alpha(x)/m(x, \sigma))z + u_0^{-\alpha(x)} - 1/\alpha(x)$. It is easy to see that

$$\begin{aligned} |D_x \hat{G}(x, z, \sigma)| &\leq |D_x \alpha(x)| / \zeta_1^2 (1 \vee \zeta_2 u_0 \log u_0) \\ &\quad + u_0 / \zeta_1 (|D_x \alpha(x)| \zeta_1 + |D_x m(x, \sigma)| / \xi_1) \vee |D_x \alpha(x)| |\log u_0| \\ &\hspace{15em} \text{if } u_0 \geq 1, \\ |D_x \hat{G}(x, z, \sigma)| &\leq |D_x \alpha(x)| / \zeta_1^2 u_0^{\zeta_1(1/\zeta_2 - 1/e)} \\ &\quad + u_0 / \zeta_1 (|D_x \alpha(x)| / \zeta_1 + |D_x m(x, \sigma)| / \xi_1) \vee |D_x \alpha(x)| |\log u_0| \\ &\hspace{15em} \text{if } u_0 < 1. \end{aligned}$$

Moreover

$$\begin{aligned} |D_x f(x, z, \sigma)| &\leq (|n| + |D_u n|)(x, \hat{G}(x, z, \sigma), \sigma) |D_x \hat{G}(x, z, \sigma)| \\ &\quad + |D_x n(x, \hat{G}(x, z, \sigma), \sigma)| |\hat{G}(x, z, \sigma)|. \end{aligned}$$

From these inequalities we deduce that if either $|n|$, $|D_u n|$, $|D_x n|$, or $|D_x \alpha|$, $|D_x m|$, $|D_x n|$ are sufficiently small, or if u_0 is sufficiently small, then (SC) holds. In particular, for a stable-like process with bounded jumps, we have $n(x, u, \sigma) = \sigma$ and $m(x, \sigma) = \theta_{\alpha(x)}$. Let

$$\xi_1 = \left(\sin \frac{\pi}{2} \zeta_1 \right) \wedge \left(\sin \frac{\pi}{2} \zeta_2 \right) \Gamma(1 + \zeta_1) \Gamma((d + \zeta_1)/2) \pi^{-(d+1)/2} \Gamma((1 + \zeta_2)/2)^{-1},$$

$$\xi_2 = \Gamma(1 + \zeta_2) \Gamma((d + \zeta_2)/2) \pi^{-(d+1)/2} \Gamma((1 + \zeta_1)/2)^{-1}$$

and

$$\begin{aligned} \eta = & \left(\pi \cos \frac{\pi}{2} \zeta_1 \Gamma(1 + \zeta_2) \Gamma((d + \zeta_2)/2) + \gamma_1 \Gamma((d + \zeta_2)/2) + \Gamma(1 + \zeta_2) \gamma_2 \right. \\ & \left. + \Gamma(1 + \zeta_2) \Gamma((d + \zeta_2)/2) \gamma_3 \right) \left(2\pi^{(d+1)/2} \Gamma((1 + \zeta_1)/2) \right)^{-1} \end{aligned}$$

with $\gamma_1 = 2 \sup_x \Gamma'(1 + \alpha(x))$, $\gamma_2 = \sup_x \Gamma'((d + \alpha(x))/2)$ and $\gamma_3 = \sup_x \Gamma'((1 + \alpha(x))/2)$. Then $0 < \xi_1 \leq \theta_{\alpha(x)} \leq \xi_2 < \infty$ and $|D_x \theta_{\alpha(x)}| \leq \eta |D_x \alpha(x)|$. Let $u_0 = 1$. In this case we have

$$|D_x f(x, z, \sigma)| \leq (1/\zeta_1^2 + (1/\zeta_1 + \eta/\zeta_1)\zeta_1) |D_x \alpha(x)|.$$

Therefore if $\alpha(x) = \alpha(x_1)$ and $\sup_x |D_x \alpha(x)| < \zeta_1^2 \xi_1 / (2\xi_1 + \zeta_1 \eta)$, then (SC) holds.

§ 4 Proof.

This section will be devoted to the proofs of Theorem 3.1 and Theorem 3.3. Recall that $\hat{G}(x, z, \sigma)$ and $f(x, z, \sigma)$ are defined by (3.1) and (3.2) respectively, and $E = (0, \infty)$. The following lemma is obtained by elementary calculation.

LEMMA 4.1. For $r \geq 1$, $(L-r)$ implies $(A-r)$ and that for every $0 \leq i + j \leq r$

$$\sup_{x, \sigma} |D_x^i D_z^j f(x, z, \sigma)| \in \bigcap_{2 \leq p \leq \infty} L^p(E, dz) \text{ is bounded.}$$

Furthermore, if $r \geq 2$, then $\lim_{z \rightarrow \infty} \sup_{x, \sigma} |D_x f(x, z, \sigma)| = 0$ and for each $\varepsilon > 0$ there is a constant $\delta = \delta(\varepsilon) > 0$ such that

$$\sup_{x, \sigma, z > \delta} |D_u y(x, \hat{G}(x, z, \sigma), \sigma) - D_u y(x, 0, \sigma)| \leq \varepsilon.$$

First we prove the existence theorem.

Proof of Theorem 3.1. It suffices to show that the distribution of the solution X_t^x of (2.1) with $E = (0, \infty)$ has a density for every $0 < t \leq T$ and $x \in \mathbf{R}^d$. For simplicity we drop the super script "x" for X_t^x , T_t^x , etc. Let $F(x, z, \sigma)$ be the same as in (2.2) and we define the process R_t with values in the set of all symmetric nonnegative definite $d \times d$ -matrices by

$$R_t = \int_0^t \int_{E \times S} \rho(z) F(X_{s-}, z, \sigma) \mu(ds dz d\sigma).$$

This is well-defined by Lemma 7-4 of [2] and satisfies the condition

$$y^T R_t y \geq y^T R_s y \geq 0 \quad \text{for all } y \in S^{d-1} \text{ if } 0 \leq s \leq t.$$

By Lemma 2.2 it is enough to show that, for any fixed $n \geq 1$ and $0 \leq u < t$,

$$(4.1) \quad \int_u^t \nabla X(n)_s^{-1} dR_s \nabla X(n)_s^{-1, \top} \text{ is invertible a. s. on } \{T_n > t\}.$$

Here we restrict the Poisson space and the time interval to $\Omega' = \Omega \cap \{T_n > t\}$ and $[u, t]$ respectively. That is, we take the restricted Poisson space $(\Omega', \mathcal{F}', (\mathcal{F}'_s)_{s \in [u, t]}, P')$ defined by $P'(\cdot) = P(\cdot | \Omega')$, $\mathcal{F}'_s = (\cap_{r > s} \sigma(\{\mu(A) : A \in \mathcal{B}((u, r]) \otimes \mathcal{E} \otimes \mathcal{S} \cup \mathcal{N}\}))|_{\Omega'}$ and $\mathcal{F}' = \mathcal{F}'_t$, where \mathcal{N} is the totality of the P -null sets of \mathcal{F} . Then $\{\mathcal{F}'_s\}_{s \in [u, t]}$ is right-continuous and \mathcal{F}'_u is the trivial field. The proof of (4.1) will be divided into three steps.

(Step 1) We fix $x \in \mathbf{R}^d$ and $\tilde{y} \in S^{d-1}$ arbitrarily. By the assumption (N) there is a point $\sigma_0 \in \text{supp } M(d\sigma)$ such that $D_u y^\top(x, 0, \sigma_0) \tilde{y} \neq 0$. Note that $D_z f(x, z, \sigma) = -(D_u y/g)(x, \hat{G}(x, z, \sigma), \sigma)$. Then, by Lemma 4.1 and the continuity of $D_u y(x, 0, \sigma)$ at σ_0 , there exist a constant $L > 0$ and a neighborhood S_0 of σ_0 such that $D_z f(I + D_x f)^{-1, \top}(x, z, \sigma) \tilde{y} \neq 0$ for every $z \geq L$ and $\sigma \in S_0$.

(Step 2) Next we prove that

$$(4.2) \quad \text{for each } y \in S^{d-1} \text{ and } s \in (u, t],$$

$$Q^y(s) \equiv \int_u^s y^\top \nabla X(n)_t^{-1} dR \nabla X(n)_t^{-1, \top} y > 0 \quad P'\text{-a. s.}$$

Suppose that (4.2) does not hold. That is, there exist a vector $y_0 \in S^{d-1}$ and a time point $s \in (u, t]$ such that $P'(Q^{y_0}(s) = 0) > 0$. For $\beta > 0$ we introduce the local martingale:

$$Y_\beta(s) = \exp\left(-\beta Q^{y_0}(s) + \int_u^s \int_{E \times S} \{1 - \exp[-\beta \rho(z) \tilde{F}(r, z, \sigma)]\} dr dz M(d\sigma)\right),$$

where $\tilde{F}(s, z, \sigma) = y_0^\top \nabla X(n)_s^{-1} F(X_{s-}, z, \sigma) \nabla X(n)_s^{-1, \top} y_0$. First we suppose that $Y_\beta(s)$ is a martingale. If $Q^{y_0}(s) = 0$ then

$$\lim_{\beta \rightarrow \infty} Y_\beta(s) = \exp\left(\int_u^s dr \int_{\rho(z) \tilde{F}(r, z, \sigma) \neq 0} dz M(d\sigma)\right).$$

By Fatou's lemma we have

$$E' \left[\exp\left(\int_u^s dr \int_{\rho(z) \tilde{F}(r, z, \sigma) \neq 0} dz M(d\sigma)\right) \right] \leq 1.$$

Hence it holds almost surely on $\{Q^{y_0}(s) = 0\}$ that

$$(4.3) \quad \int_{\rho(z) \tilde{F}(\mu, r, z, \sigma) \neq 0} dz M(d\sigma) \text{ is finite for almost every } r \in (u, s].$$

In the case where $Y_\beta(s)$ is not a martingale, we also have the same result by using a stopping time argument. Since $\rho > 0$ on E and $\tilde{F}(s, z, \sigma) = |D_z f^\top(I + D_x f)^{-1, \top}(X_{s-}, z, \sigma) \nabla X(n)_s^{-1, \top} y_0|^2$, (4.3) contradicts the result in Step 1. Therefore (4.2) holds.

(Step 3) Finally we prove (4.1). Let

$$K_s = \ker \int_u^s \nabla X(n)^{-1} dR \nabla X(n)^{-1, \top} \quad \text{for } s \in (u, t]$$

and $K_{u+} = \bigcup_{s>u} K_s$ (note that $K_r \supset K_s$ if $r \leq s$). Then (4.2) implies that for each $y \in S^{d-1}$ and any $s > u$ $P'(y \in K_s) = 0$, i.e. $P'(y \in K_{u+}) = 0$. From this we can deduce that $P'(K_{u+} = \{0\}) = 1$ as follows. Since $\text{rank} \int_u^s \nabla X(n)^{-1} dR \nabla X(n)^{-1, \top}$ is \mathcal{F}'_s -measurable, it follows that $\dim K_{u+}$ is \mathcal{F}'_u -measurable. Noting that \mathcal{F}'_u is the trivial field, we have some $0 \leq k \leq d$ such that $P'(\dim K_{u+} = k) = 1$. If $k \geq 1$ then for each $n \geq 1$ we can choose an $\mathcal{F}'_{u+1/n}$ -measurable unit random vector e_n which belongs to $K_{u+1/n}$ P' -a.s. Denote by $CL(\{e_n\})$ the totality of cluster points of $\{e_m : m \geq n\}$, that is, $CL(\{e_n\}) = \bigcap_{n \geq 1} \overline{\{e_m : m \geq n\}}$. Then from the trivality of \mathcal{F}'_u we see that there is a deterministic non-empty compact subset C of S^{d-1} such that $P'(CL(\{e_n\}) = C) = 1$ (note that we can regard $C_n = \overline{\{e_m : m \geq n\}}$ as an $\mathcal{F}'_{u+1/n}$ -measurable random variable with values in the set of all non-empty compact subsets of S^{d-1} , which is a compact metric space with the Hausdorff metric). Hence for any $y \in C$, $P'(y \in K_{u+}) = 1$, which contradicts (4.2). Thus we have $K_{u+} \subset \{0\}$ P' -a.s., from which (4.1) follows. Therefore the proof of Theorem 3.1 is complete. \square

Next we proceed to the smoothness results. The following inequalities can be easily obtained.

LEMMA 4.2. (i) *In the case of $0 < \zeta_1 \leq \alpha(x) \leq \zeta_2 < 2$ there are constants $\eta, \eta' > 0$ and $\eta'' \geq 0$ such that*

$$\sup_{x, \sigma} g(x, \hat{G}(x, z, \sigma), \sigma) \leq \eta'(1+z)^\eta \quad \text{for all } z > \eta''.$$

(ii) *In the case of $\alpha(x) \equiv 0$ there are constants $\eta, \eta' > 0$ and $\eta'' \geq 0$ such that*

$$\sup_{x, \sigma} g(x, \hat{G}(x, z, \sigma), \sigma) \leq \eta' e^{\eta'' z} \quad \text{for all } z > \eta''$$

(in this case we can take $\eta = 1/\xi_1$).

Proof of Theorem 3.3. First note that for S_i of (SN) we may assume $M(S_i) = M_0 > 0$ for $i=1, \dots, N$ by modifying $M(d\sigma)$ and $m(x, \sigma)$ as follows:

$$\bar{M}(d\sigma) = M_0 \sum_{i=1}^N M(d\sigma \cap S_i) / M(S_i) + M(d\sigma \setminus S')$$

where $S' = \sum_{i=1}^N S_i$, and

$$\bar{m}(x, \sigma) = \begin{cases} m(x, \sigma) M(S_i) / M_0 & \text{if } \sigma \in S_i \\ m(x, \sigma) & \text{otherwise} \end{cases}$$

(then $\bar{m}(x, \sigma) \geq m(x, \sigma) \geq \xi_i > 0$).

Next we verify the assertion a) by using Theorem 2.5 together with

Lemma 4.1. For this end, it suffices to prove that for any $\zeta, \theta > 0$, the assumption $(B-(\zeta, \theta))$ holds (see § 2). By the same argument as in the proof of Theorem 3-12 of [2] we see that there exist positive constants ε', ϕ and γ such that if $E' = (\phi, \infty)$ then for all $x, y \in \mathbf{R}^d$

$$(4.4) \quad \sum_{i=1}^N \inf_{z \in E'} \sup_{\sigma \in S_i} y^\top F(x, z, \sigma) y (1+z)^\gamma \geq \varepsilon' |y|^2 / 2.$$

In fact, set $\phi = \sup\{|D_u y(x, 0, \sigma)| : x \in \mathbf{R}^d, \sigma \in \sum_{i=1}^N S_i\}$; then $0 < \phi < \infty$ by (L-1) and (N). Since $\lim_{z \rightarrow \infty} \sup_{x, \sigma} |D_x f(x, z, \sigma)| = 0$, for $V \equiv (I + D_x f)^{-1} - I$ there is a constant $\phi_1 > 0$ such that

$$(4.5) \quad \|V(x, z, \sigma)\| \leq \left(\frac{1}{4\phi} \sqrt{\varepsilon/N}\right) \wedge \frac{1}{4} \quad \text{for all } z > \phi_1,$$

where $\|\cdot\|$ denotes the operator norm. Let

$$U(x, z, \sigma) = (I + D_x f(x, z, \sigma))^{-1} D_z f(x, z, \sigma),$$

which is well-defined by (C). Then

$$\begin{aligned} U(x, z, \sigma) &= (I + V(x, z, \sigma)) D_z f(x, z, \sigma) \\ &= -g(x, \hat{G}(x, z, \sigma), \sigma)^{-1} (D_u y(x, 0, \sigma) + V(x, z, \sigma) D_u y(x, 0, \sigma) \\ &\quad + (I + V(x, z, \sigma)) [D_u y(x, \hat{G}(x, z, \sigma), \sigma) - D_u y(x, 0, \sigma)]). \end{aligned}$$

Thus Lemma 4.2 (i), (4.4) and (4.5) yield that if $\phi = \phi_1 \vee \delta\left(\frac{1}{5} \sqrt{\varepsilon/N}\right) \vee \eta''$ then for all $z > \phi$, $\sigma \in \sum_{i=1}^N S_i$ and $x, y \in \mathbf{R}^d$

$$|y^\top U(x, z, \sigma)| \geq \frac{1}{\eta'} (1+z)^{-\gamma} \left(|y^\top D_u y(x, 0, \sigma)| - |y| \frac{1}{2} \sqrt{\varepsilon/N} \right).$$

Now we fix $x, y \in \mathbf{R}^d$ with $y \neq 0$. From (SN) there is an index i such that $\inf_{\sigma \in S_i} |y^\top D_u y(x, 0, \sigma)|^2 \geq \varepsilon |y|^2 / N$. Hence if $z > \phi$ and $\sigma \in S_i$, then

$$|y^\top U(x, z, \sigma)| \geq \frac{1}{\eta'} (1+z)^{-\gamma} \frac{1}{2} \sqrt{\varepsilon/N} |y|.$$

Therefore, for all $x, y \in \mathbf{R}^d$, (4.4) holds (note that $y^\top F(x, z, \sigma) y = |y^\top U(x, z, \sigma)|^2$).

Now we can choose a function $\rho(z) : E = (0, \infty) \rightarrow (0, \infty)$ such that it satisfies the condition (ρ) and $\rho(z) = (1+z)^{-2}$ for $z > \phi$. If we set $h(z) = (1+z)^{-r} \rho(z) 1_{E'}(z) = (1+z)^{-r-2} 1_{E'}(z)$ then, for any $\zeta, \theta > 0$, $h(z)$ is a (ζ, θ) -broad function on E' , which turns to a (ζ, θ) -broad function on E (see Example 2.4). Since

$$\inf_{z: h(z) > 0} \frac{\rho(z)}{h(z)} y^\top F(x, z, \sigma) y = \inf_{z \in E'} y^\top F(x, z, \sigma) y (1+z)^\gamma,$$

our claim follows.

Finally we prove the assertion b). Let $\eta = 1/\xi_1$ and denote by H the constants $2d(r+d+1)/\eta$ (resp. $2d^2(r+1)/\eta$) in (i), $2d(r+2d+1)/\eta$ in (ii), and

$4d(r+2d+2)/\eta$ in (iii), respectively. By Theorem 2.5, for the proof of b) it suffices to prove that $(B-(\zeta, \theta))$ holds for $\zeta=\eta$, $\theta=M_0t/[H+1]$. This is proved by a similar argument to the proof of a). In fact, (4.4) holds with $e^{2\eta z}$ instead of $(1+z)^r$ by Lemma 4.2 (ii). Let ρ be a function satisfying the condition (ρ) and $\rho(z)=z^{-2}$ for $z>\phi$, and set $h(z)=e^{-2\eta z}\rho(z)1_{E'}(z)=z^{-2}e^{-2\eta z}1_{E'}(z)$. Since $2\eta < M_0t/(\eta[H+1])=\theta/\eta=\theta/\zeta$ provided $t > 2\eta^2(H+1)/M_0 (\geq 2\eta^2[H+1]/M_0)$, h is (ζ, θ) -broad on E . Therefore we see that the assumption $(B-(\zeta, \theta))$ holds. \square

Acknowledgments. The author wishes to thank Professor M. Tsuchiya for leading him to this problem and for valuable discussions and suggestions. He is also grateful to Professor T. Shiga who carefully read the manuscript and suggested many critical comments which greatly improved this paper.

REFERENCES

- [1] R.F. BASS, Uniqueness in law for pure jump Markov processes, *Probab. Th. Rel. Fields* **79**, 271-287 (1988).
- [2] K. BICHTLER, J.-B. GRAVEREAUX AND J. JACOD, *Malliavin Calculus for Processes with Jumps*, Gordon and Breach, 1987.
- [3] E.B. DAVIES, *One-parameter Semigroups*, Academic Press, 1980.
- [4] S.N. ETHIER AND T.G. KURTZ, *Markov Processes. Characterization and Convergence*, Wiley, New York, 1986.
- [5] J. JACOD, Equations différentielles stochastiques linéaires; La methods de variation des constantes, *Lecture Notes in Math.*, 920, 442-446, Springer-Verlag, Berlin, 1982.
- [6] R. LÉANDRE, Régularité de processus de sauts dégénérés (I), (II), *Ann. Inst. H. Poincaré, Probab. Statist.*, 21 (1985), 125-146; 24 (1988), 209-236.
- [7] K. SATO, Class L of multivariate distributions and its subclass, *J. Multivar. Anal.*, 10 (1980), 207-232.
- [8] K. SATO, Absolute continuity of multivariate distributions of class L, *J. Multivar. Anal.*, 12 (1982), 89-94.
- [9] I. SHIGEKAWA, Derivatives of Wiener functionals and absolute continuity of induced measures, *J. Math. Kyoto Univ.*, 20 (1980), 263-289.
- [10] S.J. WOLF, On the continuity properties of L functions, *Ann. Math. statist.*, 42 (1971), 2064-2073.

TOKYO INSTITUTE OF TECHNOLOGY,
DEPARTMENT OF APPLIED PHYSICS