MALLIAVIN CALCULUS FOR LÉVY PROCESSES WITH ARBITRARY LÉVY MEASURES

UDC 519.21

A. M. KULIK

ABSTRACT. A new method is proposed to prove the absolute continuity of distributions of solutions of stochastic differential equations with jumps. The method is based on the differentiation in time in the space of functionals of the Poisson point measure. In contrast to the known Bismut and Picard methods, our approach can be applied to point measures with arbitrary Lévy measures. We obtain sufficient conditions for the absolute continuity of the solutions expressed in terms of the coefficients of the equation; the conditions do not involve assumptions on properties of the Lévy measure.

1. INTRODUCTION

We present an approach that allows one to introduce a Malliavin type calculus for functionals of general Lévy processes and to obtain sufficient conditions for the absolute continuity of solutions of stochastic differential equations with jumps (we do not pose any assumptions about regularity of the intensity of the jumps). Our investigations are motivated by a pioneering idea due to Bismut [1] and developed further by many authors. The idea is to extend the Malliavin approach to regularity of Wiener functionals to more general probability spaces by introducing a smooth structure in these spaces in terms of a “differentiation rule”, integration-by-parts formula, and by further applications of the stochastic calculus of variations to smooth functionals with nondegenerate derivatives. Another approach for solving this problem, the Davydov “stratification method”, also requires the existence of some smooth structure on a probability space; namely, it requires a measurable group of transformations of the space.

The first question arising in the realization of the above idea is how to introduce an appropriate smooth structure on the space of trajectories of Lévy processes. The answer to this question depends essentially on properties of the Lévy measure of the initial process. The most studied is the case where the Lévy measure has the (regular) density with respect to the Lebesgue measure in $\mathbb{R}^d$. A necessary integration-by-parts formula can be obtained in this case by changing the value of the jumps and by introducing the compensating transformation of the Girsanov measure. This case is studied by many authors: among them are J. M. Bismut, K. Bichteler, J. Jacod, J. Norris, R. Leandre, H. Kunita, T. Komatsu, just to mention a few (see [1]–[4] and the references therein). Conditions for a solution to have a smooth distribution are obtained in this case for stochastic differential equations governed by an appropriate Poisson random point measure.

2000 *Mathematics Subject Classification.* Primary 60H07; Secondary 60G51.

Key words and phrases. Malliavin calculus, Lévy process, stratification method, admissible time-stretching transformations.

Supported by the Ministry of Science and Education of Ukraine, Project # 01.07/103.

The case of an arbitrary Lévy measure (or, more generally, the case of an arbitrary phase space) is studied much less. Generally speaking, the transformations whose jumps change their values are not admissible in this case, and therefore the approach mentioned above does not work.

The main idea exploited throughout the paper is that admissible transformations that change the moments of their jumps still exist. This idea looks quite simple if the initial process is treated as a Poisson random measure on the time-space product space. Then the corresponding measure of intensity is smooth with respect to time. Nevertheless, here only a few results are known. Elliott and Tsoi [5] obtained the integration-by-parts formula for a pure (one point) Poisson process for a tight class of functionals (roughly speaking, this is the class of C^1 -functions of a finite number of moments of jumps). Elliott and Tsoi [5] obtained also an interesting analog of the Ocone–Hausmann formula for the integrand in the martingale representation of Poisson functionals in terms of their stochastic derivatives.

The author obtained in [6] an integration-by-parts formula for general Lévy processes and for a more general class of functionals. Some results on the regularity of the distributions of integral functionals and solutions of one-dimensional stochastic differential equations are also obtained in [6].

It turns out that differentiation with respect to time has some nice properties absent in the case of differentiation with respect to space or in the Wiener case. Admissible time-stretching transformations (the exact construction is given in Section 2) do not commute in general; thus the corresponding vector fields have a nontrivial curvature. This is reflected in the structure of the stochastic integral (defined as the adjoint operator to the differentiation) and in solutions of anticipating stochastic differential equations. These problems are studied by Privault [7, 8] in the one-point case and by the author [9, 10] in the general case.

We also study in detail the properties of the distributions of solutions of stochastic differential equations with jumps that can be expressed in terms of differentiation with respect to time.

The paper is organized as follows. To make the paper self-contained we provide in Section 2 a brief summary of differentiation with respect to time. A general result on the absolute continuity of differentiable functionals on the Poisson space is given in Section 3. We apply this result in Section 4 and obtain sufficient conditions for the regularity of distributions of a solution of a stochastic differential equation with jumps. It turns out that the density is not smooth in the general case. This is due to some specific features of the problem caused by both the structure of the initial vector fields and “bad” local properties of the initial noise. A corresponding counterexample is given in Section 5.

2. ADMISSIBLE TRANSFORMATIONS AND STOCHASTIC DERIVATIVE FOR A GENERAL LÉVY PROCESS

In this section, we briefly describe the construction of the stochastic derivative in the space of trajectories of a general Lévy process. More detail and proofs can be found in [9]. We touch only a few aspects of this topic and do not discuss others (such as the extended stochastic integral with respect to the compensated Poisson point measure or the joint stochastic derivative), since we do not need them in what follows (an interested reader can find necessary definitions in [9, 10]).

Let ν be a Poisson point measure on $\mathbb{R}^+ \times \mathbb{R}^d$ with the Lévy measure Π (that is, the measure of intensity of ν is equal to $\lambda^1 \times \Pi$), $\tilde{\nu}$ the corresponding compensated measure,

and

$$\xi(t) = \int_{[0,t] \times \{\|u\| \geq 1\}} u \nu(ds, du) + \int_{[0,t] \times \{\|u\| < 1\}} u \tilde{\nu}(ds, du)$$

the canonical representation of the Lévy process ξ (necessary details of this construction can be found in [11]). We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is such that $\mathcal{F} = \sigma(\nu)$; that is, any random variable is a functional of ν (or of ξ).

Put $H = L_2(\mathbb{R}^+)$, $H_0 = L_\infty(\mathbb{R}^+) \cap L_2(\mathbb{R}^+)$, and

$$jh(\cdot) = \int_0^\cdot h(s) ds, \quad h \in H.$$

For a fixed $h \in H_0$ we introduce the family $\{T_h^t, t \in \mathbb{R}\}$ of transformations of $\mathbb{R}^+$ by defining $T_h^t x$, $x \in \mathbb{R}^+$, as the value of the solution of the Cauchy problem

$$z'_{x,h}(s) = jh(z_{x,h}(s)), \quad s \in \mathbb{R}, \quad z_{x,h}(0) = x,$$

at the point $s = t$. One can check that

$$\text{a) } T_h^{s+t} = T_h^s \circ T_h^t, \quad \text{b) } \frac{d}{dt} T_h^t x|_{t=0} = jh(x), \quad \text{c) } T_h^t = T_{th}^1.$$

Put $T_h \equiv T_h^1$ and $\Pi_{\text{fin}} = \{\Gamma \in \mathcal{B}(\mathbb{R}^d), \Pi(\Gamma) < +\infty\}$. For $h \in H_0$ and $\Gamma \in \Pi_{\text{fin}}$, consider the following transformations T_h^Γ of the measure ν :

$$[T_h^\Gamma \nu]([0, t] \times \Delta) = \nu([0, T_{-h}t] \times (\Delta \cap \Gamma)) + \nu([0, t] \times (\Delta \setminus \Gamma)), \quad t \in \mathbb{R}^+, \Delta \in \Pi_{\text{fin}}.$$

Put

$$(\rho^\Gamma, h) = - \int_0^\infty h(t) \tilde{\nu}(dt, \Gamma), \quad p_h^\Gamma = \exp \left\{ \int_{\mathbb{R}^+} r_h(t) \nu(dt, \Gamma) - \lim_{t \rightarrow +\infty} [T_h t - t] \Pi(\Gamma) \right\},$$

where $r_h(t) = \int_0^1 h(T_{sh}t) ds$, $t \in \mathbb{R}^+$.

Lemma 2.1. a) *The transformation T_h^Γ is admissible for the distribution ν with the density p_h^Γ ; that is,*

$$\begin{aligned} & \mathbb{E} \varphi([T_h^\Gamma \nu]([0, t_1] \times \Delta_1), \dots, [T_h^\Gamma \nu]([0, t_n] \times \Delta_n)) \\ &= \mathcal{E} p_h^\Gamma \varphi(\nu([0, t_1] \times \Delta_1), \dots, \nu([0, t_n] \times \Delta_n)) \end{aligned}$$

for all $\{t_1, \dots, t_n\} \subset \mathbb{R}^+$, $\{\Delta_1, \dots, \Delta_n\} \subset \Pi_{\text{fin}}$, and an arbitrary Borel function

$$\varphi: \mathbb{R}^n \rightarrow \mathbb{R};$$

b) L_p - $\lim_{\varepsilon \rightarrow 0} (p_{\varepsilon h}^\Gamma - 1)/\varepsilon = -(\rho^\Gamma, h)$, $p \in (1, +\infty)$.

It follows from assertion a) of Lemma 2.1 that T_h^Γ generates the corresponding family of transformations of random variables, since $\mathcal{F}$ is generated by ν . We denote these transformations by the same symbol T_h^Γ .

Let $\mathcal{C}$ be the set of functionals $f \in L_\infty(\Omega)$ such that, for all $\Gamma \in \Pi_{\text{fin}}$, there exists a random element $\nabla_H^\Gamma f \in \bigcap_p L_p(\Omega, \mathbb{P}, H)$ for which

$$(\nabla_H^\Gamma f, h)_H = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [T_{\varepsilon h}^\Gamma \circ f - f]$$

for all $h \in H_0$ (the convergence holds in all spaces L_p , $p < +\infty$).

Lemma 2.2. *For all $\Gamma \in \Pi_{\text{fin}}$, the pair $(\nabla_H^\Gamma, \mathcal{C})$ is such that*

1) *for $f_1, \dots, f_n \in \mathcal{C}$ and $F \in C_b^1(\mathbb{R}^n)$,*

$$F(f_1, \dots, f_n) \in \mathcal{C} \quad \text{and} \quad \nabla_H F(f_1, \dots, f_n) = \sum_{k=1}^n F'_k(f_1, \dots, f_n) \nabla_H f_k$$

(chain rule);

2) the mapping $\rho^\Gamma: h \mapsto (\rho^\Gamma, h)$ is a weak random element in H with weak second order moments and

$$\mathbb{E}(\nabla_H^\Gamma f, h)_H = -\mathbb{E}f(\rho^\Gamma, h), \quad h \in H, f \in \mathcal{C}$$

(integration-by-parts formula);

3) there exists a countable set $\mathcal{C}_0 \subset \mathcal{C}$ such that $\sigma(\mathcal{C}_0) = \mathcal{F}$.

Remark 2.1. A pair satisfying conditions 1)–3) is called a *complete admissible differentiation* (see [9]). A discussion concerning this and some related notions can also be found in the paper [9].

Proof. Conditions 1) and 2) follow from the definition of the class $\mathcal{C}$ and Lemma 2.1, while condition 3) holds in view of the following examples. $\square$

Example 2.1. Let $\Delta \in \Pi_{\text{fin}}$ and $f = \tau_n^\Delta \equiv \inf\{t: \nu([0, t] \times \Delta) = n\}$. Then $f \in \mathcal{C}$ and

$$[\nabla_H^\Gamma f](\cdot) = -\mathbf{1}_{[0, \tau_n^\Delta]}(\cdot) \mathbf{1}_{p(\tau_n^\Delta) \in \Gamma},$$

where $p(\tau_n^\Delta)$ is the value of the point process associated with the measure ν at the moment τ_n^Δ .

Example 2.2. Let $f = \int_0^T a(s, \xi(s)) ds$ for $a \in C_b([0, T] \times \mathbb{R}^d)$. Then $f \in \mathcal{C}$ and

$$[\nabla_H^\Gamma f](\cdot) = \sum_{\cdot \leq s \leq T, \Delta \xi_s \in \Gamma} [a(s, \xi(s-)) - a(s, \xi(s))], \quad \Gamma \in \Pi_{\text{fin}}.$$

Conditions 1)–3) imply that the mapping ∇_H^Γ viewed as an unbounded operator acting from $\mathcal{C} \subset L_2(\Omega, \mathbb{P})$ to $L_2(\Omega, \mathbb{P}, H)$ is closable.

Definition 2.1. The closure D^Γ of the mapping ∇_H^Γ is called the stochastic derivative (or Γ -stochastic derivative). The adjoint operator $I^\Gamma \equiv D^{\Gamma*}$ is called the extended stochastic integral with respect to $\tilde{\nu}(\cdot, \Gamma)$.

The following example exhibits another property that appears neither in the case of the Skorokhod derivative in the Wiener space nor in the case of the spatial derivative of Lévy processes.

Example 2.3. For all $\Delta, \Gamma \in \Pi_{\text{fin}}$, and $t > 0$, the process $\nu(t, \Delta)$ is not D^Γ -differentiable, since the process $\nu(\cdot, \Delta)$ is not mean square differentiable. To avoid problems caused by the nondifferentiability of the initial noise we introduce the following notion.

Definition 2.2. A functional f is called almost surely differentiable with respect to a family $\mathcal{T}^\Gamma \equiv \{T_h^\Gamma, h \in H_0\}$ ($\Gamma \in \Pi_{\text{fin}}$ is fixed) if there exists a random element $\tilde{D}^\Gamma f$ in H such that

$$(2.1) \quad \frac{1}{\|\bar{\varepsilon}\|} \left[f \circ T_{\varepsilon_1 h_1 + \dots + \varepsilon_n h_n}^\Gamma - f - (\tilde{D}^\Gamma f, \varepsilon_1 h_1 + \dots + \varepsilon_n h_n) \right] \rightarrow 0, \quad \|\bar{\varepsilon}\| \rightarrow 0,$$

almost surely for all $h_1, \dots, h_n \in H_0$, where $\bar{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)$. The element $\tilde{D}^\Gamma f$ is called the almost sure derivative of f with respect to $\mathcal{T}^\Gamma$.

Example 2.3'. For all $\Delta, \Gamma \in \Pi_{\text{fin}}$ and $t > 0$, the process $\nu(t, \Delta)$ is almost surely differentiable with respect to $\mathcal{T}^\Gamma$ and $\tilde{D}^\Gamma \nu(t, \Delta) = 0$.

3. THE REGULARITY OF THE DISTRIBUTION OF DIFFERENTIABLE FUNCTIONALS

First we prove a general result on the regularity of differentiable functionals and then apply it to a particular case of solutions of stochastic differential equations with jumps.

Theorem 3.1. *Assume that for some $\Gamma \in \Pi_{\text{fin}}$ such that $\Pi(\Gamma) > 0$, all the functionals $f_1, \dots, f_m$ are either Γ -stochastic differentiable or surely differentiable with respect to $\mathcal{T}^\Gamma$, $h_1, \dots, h_m \in H$. Put $\sigma = \{(\widehat{D}^\Gamma f_k, h_j)\}_{k,j=1}^m$ where $\widehat{D}^\Gamma = D^\Gamma$ in the case of the Γ -differentiability or $\widehat{D}^\Gamma = \widetilde{D}^\Gamma$ in the case of the almost surely differentiability. Put $N_{f,\bar{h}}^\Gamma = \{\omega : \text{the matrix } \sigma(\omega) \text{ is nondegenerate}\}$. Then*

$$\mathbf{P} \big|_{N_{f,\bar{h}}^\Gamma} \ll \lambda^m.$$

The main idea of the proof is to apply an analog of the Davydov stratification method (see [12] for more details concerning this method). Unfortunately the straightforward application of the corresponding results, say Theorem 1 of [13], is not possible, since the differential structure constructed in Section 2 is nonflat as shown in the following example.

Example 3.1. Let $f = \tau_n^\Gamma$ and $h, g \in C(\mathbb{R}^+) \cap L_2(\mathbb{R}^+)$ be such that

$$h(t) \int_0^t g(s) ds \neq g(t) \int_0^t g(s) ds, \quad t > 0.$$

Then

$$(D^\Gamma(D^\Gamma f, h)_H, g)_H = h(\tau_n^\Gamma) \int_0^{\tau_n^\Gamma} g(s) ds \neq g(\tau_n^\Gamma) \int_0^{\tau_n^\Gamma} h(s) ds = (D^\Gamma(D^\Gamma f, g)_H, h)_H$$

almost surely. This means, in particular, that the family of transformations $\{T_h^\Gamma, h \in H_0\}$ is not commuting. Thus it cannot be represented as an additive group of transformations indexed with elements of H .

In what follows we make use of the result proved in [9] that the flat differentiation can be substituted for the ordinary differentiation in the case of the Poisson process with the help of some nonlinear transformation of the space. The precise statement is as follows.

We assume that

$$(\Omega, \mathcal{F}, \mathbf{P}) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, P_1 \times P_2),$$

where $\mathcal{F}_1 = \sigma\{\nu(\cdot, \Delta), \Delta \subset \Gamma\}$ and $\mathcal{F}_2 = \sigma\{\nu(\cdot, \Delta), \Delta \cap \Gamma = \emptyset\}$. If Ω_1 and Ω_2 are corresponding spaces of trajectories, then P_1 and P_2 are corresponding distributions of Poisson point intensity measures $\Pi|_\Gamma$ and $\Pi|_{\mathbb{R}^d \setminus \Gamma}$, respectively.

Define the mappings $\widetilde{\Psi}: \Omega_1 \rightarrow [\mathbb{R} \times \mathbb{R}^d]^\infty$ and $\Psi: \Omega \rightarrow [\mathbb{R} \times \mathbb{R}^d]^\infty \times \Omega_2$ by

$$\left[\widetilde{\Psi}(\omega_1) \right]_k = (\ln(\tau_k^\Gamma(\omega_1) - \tau_{k-1}^\Gamma(\omega_1)), p(\tau_k)), \quad k \geq 1, \quad \omega_1 \in \Omega_1, \quad \tau_0^\Gamma \equiv 0,$$

and

$$\Psi(\omega) = (\widetilde{\Psi}(\omega_1), \omega_2), \quad \omega = (\omega_1, \omega_2) \in \Omega,$$

respectively (τ_k^Γ is defined in Example 2.1). Denote by γ^∞ the infinite product of the measures

$$\gamma_1 \times \frac{\Pi|_\Gamma}{\Pi(\Gamma)},$$

where γ_1 is the measure on $\mathbb{R}$ with the density

$$p_1(x) = \frac{d\gamma_1}{d\lambda^1}(x) = \Pi(\Gamma) \exp[x - \Pi(\Gamma)e^x], \quad x \in \mathbb{R}.$$

It is easy to see that the image of the measure $\mathbf{P}$ under the transformation Ψ is equal to the product $\gamma^\infty \times P_2$. To make the notation shorter, let $\mathcal{X} = [\mathbb{R} \times \mathbb{R}^d]^\infty \times \Omega_2$ and $\varkappa = \gamma^\infty \times P_2$. Since the mapping $\Psi: \Omega \rightarrow \mathcal{X}$ maps $\mathbf{P}$ to $\varkappa$, it generates the mapping

$$L_0(\mathcal{X}, \varkappa) \rightarrow L_0(\Omega, \mathbf{P})$$

that is isometric in any space L_p ; the latter mapping is denoted by the same symbol Ψ .

Further, we introduce the linear shift Q_z in $\mathcal{X}$ by

$$Q_z \cdot (\{(x_k, p_k), k \geq 1\}, \omega_2) = (\{(x_k + z_k, p_k), k \geq 1\}, \omega_2)$$

for $z = \{z_k, k \geq 1\} \in \ell_2$.

Since the logarithmic derivative of the measure γ_1 is square integrable, the measure $\varkappa$ is differentiable with respect to the group of shifts $\mathcal{Q} = \{Q_z, z \in \ell_2\}$, and its logarithmic derivative has the weak moment of the second order. Thus one can apply the construction of the preceding section to the measure $\varkappa$ and then introduce the stochastic derivative D^γ with respect to the family $\mathcal{Q}$. We also introduce the almost sure derivative $\tilde{D}^\gamma$ with respect to $\mathcal{Q}$ similarly to Definition 2.2 but, instead of the set H_0 , we use the set ℓ_0 of the sequences whose members become zero starting with some number.

Finally we define for all $\omega \in \Omega$, the unbounded operators $T_\omega, S_\omega: H \rightarrow \ell_2$ by

$$(T_\omega h)_k = -\frac{1}{\tau_k^\Gamma(\omega) - \tau_{k-1}^\Gamma(\omega)} \left(h, \mathbf{1}_{[\tau_{k-1}^\Gamma(\omega), \tau_k^\Gamma(\omega))} \right)_H,$$

$$(S_\omega h)_k = - \left(h, \mathbf{1}_{[\tau_{k-1}^\Gamma(\omega), \tau_k^\Gamma(\omega))} \right)_H,$$

$k \geq 1$. The following result describes the relationship between derivatives on the spaces $(\Omega, \mathbf{P})$ and $(\mathcal{X}, \varkappa)$.

Lemma 3.1. 1) Let $F \in L_2(\mathcal{X}, \varkappa)$ have the stochastic derivative $D^\gamma F$ with respect to $\mathcal{Q}$ and $[T.]^* \{D^\gamma F(\Psi \cdot)\} \in L_2(\Omega, \mathbf{P}, H)$. Then $f = \Psi F$ has the stochastic derivative with respect to $\mathcal{T}^\Gamma$, and this derivative is equal to $D^\Gamma f(\cdot) = [T.]^* \{D^\gamma F(\Psi \cdot)\}$.

2) Let $f \in L_2(\Omega, \mathbf{P})$ have the stochastic derivative $D^\Gamma f$ with respect to $\mathcal{T}^\Gamma$ and

$$\Psi^{-1}[S_\omega D^\Gamma f] \in L_2(\mathcal{X}, \varkappa, \ell_2).$$

Then $F = \Psi^{-1}f$ has the stochastic derivative with respect to $\mathcal{Q}$, and this derivative is equal to $D^\gamma F = \Psi^{-1}[S_\omega D^\Gamma f]$.

3) Let $f \in L_0(\Omega, \mathbf{P})$ be almost surely differentiable with respect to $\mathcal{T}^\Gamma$ and

$$\tilde{D}^\Gamma f \in \text{Dom}(S_\omega)$$

for all $\omega \in \Omega$. Then $F = \Psi^{-1}f$ has the almost sure derivative with respect to $\mathcal{Q}$, and this derivative is equal to $\tilde{D}^\gamma F = \Psi^{-1}[S_\omega \tilde{D}^\Gamma f]$.

Proof. Assertions 1) and 2) are similar to Lemma 4.1 in [9] and can be proved in the same way with the help of Lemmas 2.1 and 2.2 of [9]. Now we turn to the proof of assertion 3). Assume that $z^1, \dots, z^m \in \ell_0$ are fixed and N is such that $z_n^j = 0$ for all $n > N, j = 1, \dots, m$. Fix $\bar{\alpha} = \{\alpha_j\}, \bar{\beta} = \{\beta_j\}$, and c such that

$$0 < \alpha_1 < \beta_1 < \alpha_2 < \dots < \beta_N < c < \infty.$$

We show that

$$(3.1) \quad \frac{1}{\|\bar{\varepsilon}\|} \left[F \circ Q_{\varepsilon_1 z_1 + \dots + \varepsilon_m z_m}(x) - F(x) - (\tilde{D}^\gamma F(x), \varepsilon_1 z_1 + \dots + \varepsilon_m z_m) \right] \rightarrow 0,$$

$$\|\bar{\varepsilon}\| \rightarrow 0,$$

for $\varkappa$ -almost all x of the set

$$\mathcal{X}_{\bar{\alpha}, \bar{\beta}, c} = \Psi \{ \omega: \tau_j^\Gamma(\omega) \in (\alpha_j, \beta_j), j \leq N, \tau_{N+1}^\Gamma > c \}.$$

Let $\delta = \min_{j=1, \dots, N+1} \{\alpha_j - \beta_{j-1}\}$, where $\beta_0 \equiv 0$ and $\alpha_{N+1} \equiv c$. Consider functions $h_1, \dots, h_N$ of $C^\infty(\mathbb{R}^+)$ such that H_j assumes the value 1 on the interval $[\alpha_j - \frac{\delta}{4}, \beta_j + \frac{\delta}{4}]$ and the value 0 on $\mathbb{R}^+ \setminus [\alpha_j - \frac{\delta}{3}, \beta_j + \frac{\delta}{3}]$, where $H_j(\cdot) = \int_0^\cdot h_j(s) ds$, $j = 1, \dots, N$. Then

$$T_{\zeta_1 h_1 + \dots + \zeta_N h_N}^\Gamma(\Psi^{-1}x) = \Psi^{-1}(Q_{\zeta_1 e_1 + (\zeta_2 - \zeta_1)e_2 + \dots + (\zeta_N - \zeta_{N-1})e_N}x), \quad x \in \mathcal{X}_{\bar{\alpha}, \bar{\beta}, c},$$

for some constant $\zeta = \zeta(\bar{\alpha}, \bar{\beta}, c) > 0$ and all vectors $\bar{\eta} = (\eta_1, \dots, \eta_N)$ such that $\|\bar{\eta}\| < \zeta$, where $e_j = (0, \dots, 1_j, 0, \dots)$. This together with Definition 2.2 implies (3.1) for almost all $x \in \mathcal{X}_{\bar{\alpha}, \bar{\beta}, c}$. Choosing all well-ordered families of rational numbers as $\bar{\alpha}$, $\bar{\beta}$, and c , we obtain the assertion desired. The lemma is proved. $\square$

Proof of Theorem 3.1. According to Lemma 3.1, we may assume in what follows that the initial functional is defined on the product space equipped with the linear differentiation. Thus we can apply the standard technique of the stratification method. Now the result for stochastic differentiable functionals follows from Theorem 1 of [13] if the moment conditions of Lemma 3.1 hold. These conditions can be checked with the help of the corresponding localization arguments (we omit the details).

Theorem 3.1 in the case of almost surely differentiable functionals can be obtained with the help of the stratification method and by using the following result, which is a particular case of Theorem 3.1.16 in [14].

Lemma 3.2. *Let $F: \mathbb{R}^m \rightarrow \mathbb{R}^n$ and $G: \mathbb{R}^m \rightarrow \mathbb{R}^m \times \mathbb{R}^n$ be such that*

$$\frac{1}{\|a\|} \|F(x+a) - F(x) - (G(x), a)_{\mathbb{R}^m}\| \rightarrow 0, \quad \|a\| \rightarrow 0,$$

for λ^m -almost all $x \in \mathbb{R}^m$. Then, for all $\varepsilon > 0$, there exists $F_\varepsilon \in C^1(\mathbb{R}^m, \mathbb{R}^n)$ such that

$$\lambda^m(\{x: F(x) \neq F_\varepsilon(x)\} \cup \{x: G(x) \neq \nabla F_\varepsilon(x)\}) < \varepsilon.$$

Theorem 3.1 is proved. $\square$

4. THE REGULARITY OF DISTRIBUTIONS OF STOCHASTIC DIFFERENTIAL EQUATIONS

Consider a stochastic differential equation in $\mathbb{R}^m$, $m \geq 1$,

$$(4.1) \quad x(t) = x_0 + \int_0^t a(s, x(s)) ds + \iint_{[0,t] \times \mathbb{R}^d} c(s, x(s-), u) \tilde{\nu}(ds, du), \quad t \in \mathbb{R}^+.$$

We assume that the coefficients $a: \mathbb{R}^+ \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ and $c: \mathbb{R}^+ \times \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ are measurable functions that are differentiable with respect to the variables (s, x) and locally bounded together with their derivatives. We also assume that there exists a constant $K < +\infty$ such that

$$\int_{\mathbb{R}^d} \|c(s, x, u) - c(s, y, u)\|^2 \Pi(du) \leq K \|x - y\|^2, \quad x, y \in \mathbb{R}^m, s \in \mathbb{R}^+.$$

Thus equation (4.1) has a unique strong solution whose trajectories have limits on the left and are right continuous. We further assume that

$$(4.2) \quad \sup_{\mathbb{R}^d} \sup_{s \in [0, T], \|x\| \leq R} [\|c(s, x, u)\|_{\mathbb{R}^m} + \|\nabla_x c(s, x, u)\|_{\mathbb{R}^m \times \mathbb{R}^m}] \Pi(du) < +\infty,$$

$$T, R < +\infty,$$

and put $\tilde{a}(s, x) = a(s, x) - \int_{\mathbb{R}^d} c(s, x, u) \Pi(du)$ and

$$\Delta(s, x, u) = [\tilde{a}(s, x + c(s, x, u)) - \tilde{a}(s, x)] - [\nabla_x c(s, x, u)] \tilde{a}(s, x) - c'_s(s, x, u),$$

where ∇_x is the vector derivative with respect to the variable x .

Lemma 4.1. *For all $t \in \mathbb{R}^+$ and $\Gamma \in \Pi_{\text{fin}}$, each component of the vector $x(t)$ has the almost sure derivative with respect to $\mathcal{T}^\Gamma$, and, for all $h \in H$, the process*

$$Y_h^\Gamma(\cdot) = \left((\tilde{D}^\Gamma x_1(\cdot), h), \dots, (\tilde{D}^\Gamma x_m(\cdot), h) \right)$$

is such that

$$(4.3) \quad \begin{aligned} Y_h^\Gamma(t) = & \iint_{[0,t] \times \Gamma} \Delta(s, x(s-), u) H(s) \nu(ds, du) + \int_0^t [\nabla_x \tilde{a}](s, x(s)) Y_h^\Gamma(s) ds \\ & + \iint_{[0,t] \times \mathbb{R}^d} [\nabla_x c](s, x(s-), u) Y_h^\Gamma(s-) \nu(ds, du), \quad t \geq 0, \end{aligned}$$

where $H(\cdot) = \int_0^\cdot h(s) ds$.

Proof. Let $t > 0$ be fixed. Since $\Pi(\Gamma) < +\infty$, $\Omega = \bigcup_{k \geq 0} \Omega_k$ almost surely, where $\Omega_k \equiv \{\nu(t, \Gamma) = k\}$. To prove Lemma 4.1 it is sufficient to prove that (2.1) holds almost surely on every set Ω_k , $k \geq 0$. The case of $k = 0$ is trivial; we consider the case of $k = 1$ in more detail (the reasoning for general $k > 1$ is the same).

By τ we denote the moment of the unique jump of ν that happens before the time t and belongs to the set Γ . For $h \in H_0$, we introduce the process z_h by

$$z_h(r) = x_0 + \int_0^r \tilde{a}(s, z_h(s)) ds + \iint_{[0,r] \times (\mathbb{R}^d \setminus \Gamma)} c(s, z_h(s-), u) \nu(ds, du)$$

if $r < T_{-h}\tau$ and

$$(4.4) \quad \begin{aligned} z_h(r) = & z_h(T_{-h}\tau-) + c(T_{-h}\tau-, z_h(T_{-h}\tau-), p(\tau)) + \int_{T_{-h}\tau}^r \tilde{a}(z_h(s)) ds \\ & + \iint_{[T_{-h}\tau-, r] \times (\mathbb{R}^d \setminus \Gamma)} c(s, z_h(s-), u) \nu(ds, du) \end{aligned}$$

if $r > T_{-h}\tau$. One can check that $z_h|_{[0,t]} \equiv T_h^\Gamma x|_{[0,t]}$ on the set

$$\Omega_1 \cap T_h^\Gamma(\Omega_1) \supset \left\{ \tau < t - \sqrt{t\|h\|_H}, \tilde{\tau} > t + \sqrt{t\|h\|_H} \right\}$$

where $\tilde{\tau}$ is the moment of the next jump belonging to the set Γ . Thus it remains to prove that z_h is almost surely Frechét differentiable in h and the corresponding derivative satisfies equality (4.3). Put $h = \varepsilon h_0$. Considering the limit behavior, as $\varepsilon \rightarrow 0$, of terms on the right-hand side of (4.4) we have

$$\frac{d}{d\varepsilon} T_{-\varepsilon h_0} v|_{\varepsilon=0} = -H_0(v) = -\int_0^v h_0(s) ds.$$

This implies that

$$\frac{d}{d\varepsilon} \int_0^{T_{-\varepsilon h_0}\tau} \tilde{a}(s, z_{\varepsilon h_0}(s)) ds|_{\varepsilon=0} = -H_0(\tau) \tilde{a}(\tau, x(\tau-)).$$

Without loss of generality one can assume that

$$\|x(s)\| < R, \quad \|z_h(s)\| < R$$

for some fixed R , all $s \leq t$, and all sufficiently small $\|h\|$. Put

$$\eta_v = \iint_{[0,v] \times (\mathbb{R}^d \setminus \Gamma)} \sup_{s \in [0,T], \|x\| \leq R} [\|c(s, x, u)\|_{\mathbb{R}^m} + \|\nabla_x c(s, x, u)\|_{\mathbb{R}^m \times \mathbb{R}^m}] \nu(ds, du).$$

According to condition (4.2) η . almost surely is a singular function of bounded variation. Applying the Lebesgue theorem we obtain that almost all trajectories η . have zero derivative for λ^1 -almost all points of the interval $[0, t]$. Since τ and η . are independent and the distribution of τ is absolutely continuous, this implies that

$$\frac{d}{d\varepsilon} \iint_{[0, T-\varepsilon h_0 \tau) \times (\mathbb{R}^d \setminus \Gamma)} c(s, z_{\varepsilon h_0}(s), u) \nu(ds, du) \Big|_{\varepsilon=0} = 0$$

almost surely and therefore

$$\begin{aligned} \frac{d}{d\varepsilon} z_{\varepsilon h_0}(T-\varepsilon h_0 \tau-) \Big|_{\varepsilon=0} &= -H_0(\tau) \tilde{a}(\tau, x(\tau-)), \\ \frac{d}{d\varepsilon} c(T-\varepsilon h_0 \tau-, z_{\varepsilon h_0}(T-\varepsilon h_0 \tau-), p(\tau)) \Big|_{\varepsilon=0} &= -H_0(\tau) [c'_s + \nabla_x c \cdot \tilde{a}](\tau, x(\tau-), p(\tau)). \end{aligned}$$

Analogously, differentiating the third and fourth terms in (4.4) with respect to the lower bound of integration we obtain almost surely $H_0(\tau) \tilde{a}(\tau, x(\tau-) + c(\tau, x(\tau-), p(\tau)))$ and 0, respectively. Differentiating formally (4.4) we get equation (4.3). In order to justify this formal differentiation we apply the former results and the following estimate:

$$\|x(v) - z_{\varepsilon h_0}(v)\| \leq \int_0^t \|x(s-) - z_{\varepsilon h_0}(s)\| d\eta_s + O(\varepsilon),$$

where $O(\varepsilon)/\varepsilon$ is almost surely and uniformly bounded for $v > \tau + \delta$ and $|\varepsilon| < \varepsilon_0$. Now we apply the Gronwall–Bellman lemma and obtain that the fraction

$$\frac{\|x(v) - z_{\varepsilon h_0}(v)\|}{\varepsilon}$$

is almost surely and uniformly bounded for the same v and ε . This allows one to differentiate (4.4) with respect to the parameter h .

Up to this moment we have considered a fixed direction $h_0 \in H$. Since

$$\frac{T-\varepsilon h v - v}{\varepsilon} \rightarrow - \int_0^v h(s) ds, \quad \varepsilon \rightarrow 0,$$

uniformly in $v \leq t$ and $h \in B_H(0, 1)$, the previous estimates hold uniformly in

$$h_0 \in B_H(0, 1),$$

where $B_H(0, 1)$ is the unit ball in H . This means that $z_h(t)$ is Frechét differentiable. The remaining part of the proof is similar to what we proved already; thus we omit the details. The lemma is proved. $\square$

In what follows we use the additional assumption

$$(4.5) \quad \sup_{s, x, u} \|\nabla_x c(s, x, u)\| < 1.$$

A somewhat weaker assumption is discussed in Remark 4.1 after the proof of Theorem 4.2. If condition (4.5) holds, then the solution of equation (4.3) can be written explicitly:

$$Y_h^\Gamma(t) = \mathcal{E}_0^t \iint_{[0, t] \times \Gamma} H(s) [\mathcal{E}_0^{s-}]^{-1} \tilde{\Delta}(s, x(s-), u) \nu(ds, du), \quad t \geq 0,$$

where

$$\tilde{\Delta}(s, x, u) = [I_{\mathbb{R}^m} + \nabla_x c(s, x, u)]^{-1} \Delta(s, x, u)$$

and $\{\mathcal{E}_r^t, 0 \leq r \leq t\}$ is a process assuming values in the space of $m \times m$ matrices satisfying the following stochastic differential equation:

$$\mathcal{E}_r^t = I_{\mathbb{R}^m} + \int_r^t \nabla_x a(s, x(s)) \mathcal{E}_r^s ds + \iint_{[r, t] \times \mathbb{R}^d} \nabla_x c(s, x(s-), u) \mathcal{E}_r^{s-} \nu(ds, du).$$

Our first result concerning the absolute continuity of the distribution of solution (4.1) follows from Theorem 2.1 and Lemma 3.1. By S_t we denote the linear span of the set of vectors

$$\{[\mathcal{E}_0^{\tau-}]^{-1}\tilde{\Delta}(\tau, x(\tau-), p(\tau)), \tau \leq t\},$$

where the moments τ are taken from the domain of the point process associated with the random point measure ν .

Theorem 4.1. *Let $\Omega_t = \{\omega: \dim S_t(\omega) = m\}$. Then*

$$\mathbf{P}|_{\Omega_t} \circ [x(t)]^{-1} \ll \lambda^m.$$

Proof. First we introduce a generalized almost sure derivative. Since this notion does not look natural in the general case, we did not mention it in Sections 1 and 2. The main problem can be explained as follows. Lemma 4.1 can be treated as a result concerning the differentiability of solutions of the initial equation with respect to a direction indexed with a function $h(s) \cdot \mathbf{I}_\Gamma(u)$. A usual approach in the Wiener case consists of two steps. The first step is to evaluate the gradient derivative that “collects all the properties” of the directional derivatives. After studying the set where the gradient derivative is nondegenerate one uses a general result about the regularity of distributions (similar to Theorem 3.1 in our case).

The method described above meets with some problems in our case because the gradient derivative is a random element in a space of rather complicated structure (see [9]), namely in the space of (random) kernels. To avoid these problems we follow another approach extending the space of directions of the differentiability.

Let $0 < \alpha_1 < \beta_1 < \alpha_2 < \dots < \beta_N < \infty$ be fixed. By $H_{\bar{\alpha}, \bar{\beta}}$ we denote the set of stochastic processes of the form

$$(4.6) \quad g(s, u) = \sum_{k=1}^N \varphi_k h_k(s) \mathbf{I}_{\Gamma_k}(u), \quad s \in \mathbb{R}^+, u \in \mathbb{R}^d,$$

where $\text{supp } H_k \subset (\alpha_k, \beta_k)$, φ_k is an $\mathcal{F}_{\alpha_k}$ -measurable random variable, and $\Gamma_k \in \Pi_{\text{fin}}$ for $k = 1, \dots, N$.

For all k , the transformation $T_{h_k}^{\Gamma_k}$ changes the point process associated with the measure ν only in the interval (α_k, β_k) . The values of ν on subsets of this interval are independent of $\mathcal{F}_{\alpha_k}$. Thus the transformation

$$T_{\varphi_k h_k \mathbf{I}_{\Gamma_k}} \equiv T_{\varphi_k h_k}^{\Gamma_k}$$

is well defined at every point. Given g of the form (4.6), we put

$$T_g = T_{\varphi_N h_N \mathbf{I}_{\Gamma_N}} \cdots T_{\varphi_1 h_1 \mathbf{I}_{\Gamma_1}}.$$

The following result contains the main properties of the family of transformations

$$\mathcal{T}^{\bar{\alpha}, \bar{\beta}} \equiv \{T_g, g \in H_{\bar{\alpha}, \bar{\beta}}\}.$$

Lemma 4.2. 1) *For all $g \in H_{\bar{\alpha}, \bar{\beta}}$, there exists the density $p_g = P \circ T_g^{-1}$ and, moreover,*

$$\frac{d}{d\varepsilon} p_{\varepsilon g} \Big|_{\varepsilon=0} = \iint_{\mathbb{R}^+ \times \mathbb{R}^d} g(s, u) \tilde{\nu}(ds, du).$$

2) *For all $t \in \mathbb{R}^+$, there exists the almost sure derivative of the solution $x(t)$ of equation (4.1) with respect to $\mathcal{T}^{\bar{\alpha}, \bar{\beta}}$; that is, there exists a linear mapping*

$$\tilde{D}x(t): H_{\bar{\alpha}, \bar{\beta}} \rightarrow L_2(\Omega, \mathbf{P})$$

such that

$$\frac{1}{\|\bar{\varepsilon}\|} \left[f \circ T_{\varepsilon_1 g_1 + \dots + \varepsilon_n g_n} - f - [\tilde{D}x(t)](\varepsilon_1 g_1 + \dots + \varepsilon_n g_n) \right] \rightarrow 0, \quad \|\bar{\varepsilon}\| \rightarrow 0,$$

almost surely for all $g_1, \dots, g_n \in H_{\bar{\alpha}, \bar{\beta}}$. Moreover, the mapping $\tilde{D}x(t)$ can be written explicitly:

$$[\tilde{D}x(t)](g) = \mathcal{E}_0^t \iint_{[0, t] \times \mathbb{R}^d} [jg](s, u) [\mathcal{E}_0^{s-}]^{-1} \tilde{\Delta}(s, x(s-), u) \nu(ds, du),$$

where

$$[jg](s, u) = \int_0^s g(r, u) dr.$$

3) Let the functionals $f_1, \dots, f_m$ be differentiable almost surely with respect to $\mathcal{T}^{\bar{\alpha}, \bar{\beta}}$,

$$g_1, \dots, g_m \in H_{\bar{\alpha}, \bar{\beta}}.$$

Denote by σ the corresponding matrix of derivatives. Let

$$N_{\bar{f}, \bar{g}} = \{\omega: \text{the matrix } \sigma(\omega) \text{ is nondegenerate}\}.$$

Then

$$\mathbf{P} \big|_{N_{\bar{f}, \bar{g}}} \circ \bar{f}^{-1} \ll \lambda^m.$$

The proof of assertions 1)–3) is absolutely analogous to the proofs of Lemma 2.1, Lemma 4.1, and Theorem 3.1 (details are omitted).

Now we come back to the proof of Theorem 4.1. For arbitrary $\bar{\alpha}$ and $\bar{\beta}$, denote by $L_{\bar{\alpha}, \bar{\beta}}$ the subspace of $L_2(\Omega \times \mathbb{R}^+ \times \mathbb{R}^d, \mathbf{P} \times \lambda^1 \times \Pi)$ generated by all processes of the form

$$\sum_{k=1}^N \varphi_k f_k(s) \mathbf{1}_{\Gamma_k}(u), \quad s \in \mathbb{R}^+, u \in \mathbb{R}^d,$$

where $\text{supp } f_k \subset [\alpha_k, \beta_k]$, φ_k is an $\mathcal{F}_{\alpha_k}$ -measurable random variable, and $\Gamma_k \in \Pi_{\text{fin}}$ for $k = 1, \dots, N$. For all rational vectors $\bar{\alpha}$ and $\bar{\beta}$, choose a countable subset $\mathcal{A}_{\bar{\alpha}, \bar{\beta}}$ in $H_{\bar{\alpha}, \bar{\beta}}$ such that the set $\{jg, g \in \mathcal{A}_{\bar{\alpha}, \bar{\beta}}\}$ is dense in $L_{\bar{\alpha}, \bar{\beta}}$.

Put

$$\Omega_{\bar{\alpha}, \bar{\beta}} = \left\{ \omega: \text{rank}\{\tilde{D}x(t)(\bar{g})\} < m \text{ for all } g_1, \dots, g_m \in \mathcal{A}_{\bar{\alpha}, \bar{\beta}} \right\}.$$

It follows from assertion 3) that

$$\mathbf{P} \big|_{\Omega \setminus \Omega_{\bar{\alpha}, \bar{\beta}}} \circ [x(t)]^{-1} \ll \lambda^m.$$

It remains to prove that

$$\bigcap_{(\bar{\alpha}, \bar{\beta})} \Omega_{\bar{\alpha}, \bar{\beta}} \subset \Omega \setminus \Omega_t$$

except for some set of probability zero.

The mapping $\tilde{D}x(t)$ defined in statement 2) of Lemma 4.2 is continuous in the space of all anticipating processes with respect to the distance in the space

$$L_2(\Omega \times \mathbb{R}^+ \times \mathbb{R}^d, \mathbf{P} \times \lambda^1 \times \Pi).$$

If condition (4.5) holds, then the matrix $\mathcal{E}_0^t$ is almost surely nondegenerate. Since the set $\bigcup_{(\bar{\alpha}, \bar{\beta})} \mathcal{A}_{\bar{\alpha}, \bar{\beta}}$ is dense in the space of all anticipating processes, the matrix

$$M_t(\psi) = \iint_{[0, t] \times \mathbb{R}^d} [\mathcal{E}_0^{s-}]^{-1} [\tilde{\Delta}(s, x(s-), u) \otimes \psi(s, u)] \nu(ds, du)$$

is degenerate almost surely on the set

$$\bigcap_{(\bar{\alpha}, \bar{\beta})} \Omega_{\bar{\alpha}, \bar{\beta}}$$

for any anticipating process $\psi: \Omega \times \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^m$. On the other hand, choosing $h \in L_2(\mathbb{R}^+)$, $r \in L_2(\mathbb{R}^d, \Pi)$, $h, r > 0$ and putting

$$\hat{\psi}(s, u) = \frac{h(s)r(u)}{\|[\mathcal{E}_0^{s-}]^{-1}\tilde{\Delta}(s, x(s-), u)\|^2} [\mathcal{E}_0^{s-}]^{-1}\tilde{\Delta}(s, x(s-), u),$$

we get the matrix

$$M_t(\hat{\psi}) = \iint_{[0, t] \times \mathbb{R}^d} \frac{h(s)r(u)}{\|[\mathcal{E}_0^{s-}]^{-1}\tilde{\Delta}(s, x(s-), u)\|^2} \left([\mathcal{E}_0^{s-}]^{-1}\tilde{\Delta}(s, x(s-), u)\right)^{\otimes 2} \nu(ds, du)$$

that almost surely is nondegenerate on Ω_t . The theorem is proved. $\square$

According to Theorem 4.1, the sufficient condition for the existence of the density of the distribution of $x(t)$ is that the random space S_t is almost surely equal to the whole space $\mathbb{R}^m$.

Our next aim is to obtain another sufficient condition expressed in terms of coefficients (this kind of condition is more convenient to check).

Theorem 4.2. *Let*

$$(4.7) \quad \Pi\left(u: (\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d} \neq 0\right) = +\infty$$

for all $x \in \mathbb{R}^m$, $s \in \mathbb{R}^+$, and $\bar{l} \in \mathbb{R}^m \setminus \{0\}$. Then

$$\mathbf{P} \circ [x(t)]^{-1} \ll \lambda^m$$

for all $t > 0$.

Proof. We need the following auxiliary result.

Lemma 4.3. *If condition (4.7) holds for any compact subset*

$$K \subset \mathbb{R}^m \times \mathbb{R}^+,$$

then

$$(4.8) \quad \lim_{n \rightarrow \infty} \inf_{(x, s) \in K, \bar{l} \neq 0} \Pi\left(u: \|u\| \geq \frac{1}{n}, (\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d} \neq 0\right) = +\infty.$$

First we derive Theorem 4.2 from (4.8); then we provide the proof of Lemma 4.3.

Without loss of generality, we assume that there exists a compact set $\tilde{K} \subset \mathbb{R}^m$ such that $x(s) \in \tilde{K}$ almost surely, $s \leq t$. It is seen from the proof of Theorem 4.1 that the absolute continuity of the distribution of $x(t)$ holds if there is a sequence of anticipating $\mathbb{R}^m$ -valued processes $\{\psi_n(s, u), s \in \mathbb{R}^+, u \in \mathbb{R}^d\}$, $m \geq 1$, such that

$$\mathbf{P}\left(\text{rank} \iint_{[0, t] \times \mathbb{R}^d} [\mathcal{E}_0^{s-}]^{-1}\tilde{\Delta}(s, x(s-), u) \otimes \psi_n(s, u) \nu(ds, du) = m\right) \rightarrow 1, \quad n \rightarrow \infty.$$

Let $\psi_n(s, u) = 0$ for $\|u\| < n^{-1}$, and $\psi_n(s, u) = [\mathcal{E}_0^{s-}]^{-1}\tilde{\Delta}(s, x(s-), u)$ otherwise. It is clear that the rank of the matrix $M_t(\psi_n)$ is a nondecreasing function in t . Thus it remains to prove that

$$(4.9) \quad \mathbf{P}\left(\text{rank} M_t(\psi_n) = \text{rank} M_s(\psi_n) / \text{rank} M_s(\psi_n) < m\right) \rightarrow 0, \quad n \rightarrow +\infty,$$

for all $s < t$. Put $\tau_k^n \equiv \tau_k^{u: \|u\| \geq 1/n}$ and $\zeta_n = \inf\{s: \text{rank } M_s(\psi_n) = m\}$. One can show that there exists an anticipating $\mathbb{R}^m$ -valued process $\{l_t, t \geq 0\}$ such that with probability 1,

- a) the process $\{l_t\}$ is constant on every interval of the form $(\tau_{k-1}^n, \tau_k^n]$, $k \geq 1$, and for all $j < k$ the vectors $l_{\tau_k^n}$ and $[\mathcal{E}_0^{\tau_j^n}]^{-1} \tilde{\Delta}(\tau_j^n, x(\tau_j^n -), p(\tau_j^n))$ are orthogonal;
- b) $l_t \neq 0 \Leftrightarrow t \leq \zeta_n$.

In other words, the process $\{l_t\}$ is an appropriate vector, being orthogonal to the linear span $S_n(t)$ of the vectors

$$\left\{ [\mathcal{E}_0^{\tau_k^n}]^{-1} \tilde{\Delta}(\tau_k^n, x(\tau_k^n -), p(\tau_k^n)), \tau_k^n < t \right\}.$$

Using the above notation we get

$$\begin{aligned} & \mathbb{P}(\text{rank } M_t(\psi_n) = \text{rank } M_s(\psi_n) / \text{rank } M_s(\psi_n) < m) \\ &= \mathbb{P}\left(\tilde{\Delta}(\tau_k^n, x(\tau_k^n -), p(\tau_k^n)) \perp \{[\mathcal{E}_0^{\tau_k^n}]^{-1}\}^* l_{\tau_k^n} \text{ for every } k: \tau_k \in (s, t] / l_s \neq 0\right) \\ &= \mathbb{E} \prod_{k: \tau_k^n \in (s, t]} \mathbb{P}\left(\tilde{\Delta}(\tau_k^n, x(\tau_k^n -), p(\tau_k^n)) \perp \{[\mathcal{E}_0^{\tau_k^n}]^{-1}\}^* l_{\tau_k^n} / l_{\tau_k^n} \neq 0\right). \end{aligned}$$

For all k , the distribution of the jump $p(\tau_k^n)$ is

$$\frac{\Pi|_{\{u: \|u\| \geq 1/n\}}}{\Pi(\{u: \|u\| \geq 1/n\})}$$

and does not depend on the σ -algebra $\mathcal{G}_k^n$ generated by the random measure ν on

$$\mathbb{R}^+ \times \left\{ u: \|u\| < \frac{1}{n} \right\},$$

by moments $\{\tau_j^n, j \leq k\}$, and by $\{p(\tau_j^n), j < k\}$. On the other hand, the random variables τ_k^n , $x(\tau_k^n -)$, $l_{\tau_k^n}$, and $\{[\mathcal{E}_0^{\tau_k^n}]^{-1}\}^* l_{\tau_k^n}$ are $\mathcal{G}_k^n$ -measurable. Putting

$$\begin{aligned} \gamma_n &= \inf_{(x,s) \in \bar{K} \times [0,t], \bar{l} \neq 0} \Pi\left(u: \|u\| > \frac{1}{n}, (\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d} \neq 0\right), \\ \lambda_n &= \Pi\left(\left\{u: \|u\| \geq \frac{1}{n}\right\}\right), \end{aligned}$$

we obtain the estimate

$$\mathbb{P}\left(\tilde{\Delta}(\tau_k^n, x(\tau_k^n -), p(\tau_k^n)) \perp \{[\mathcal{E}_0^{\tau_k^n}]^{-1}\}^* l_{\tau_k^n} / l_{\tau_k^n} \neq 0\right) \leq 1 - \frac{\gamma_n}{\lambda_n}, \quad k \geq 1.$$

The distribution of the number $N_{s,t}^n$ of k is such that $\tau_k^n \in (s, t]$ is Poisson with intensity $(t-s)\lambda_n$, whence

$$\begin{aligned} & \mathbb{P}\left(\text{rank } M_t(\psi_n) = \text{rank } M_s(\psi_n) / \text{rank } M_s(\psi_n) < m\right) \\ & \leq \mathbb{E} \left(1 - \frac{\gamma_n}{\lambda_n}\right)^{N_{s,t}^n} = \exp\{-(t-s)\gamma_n\} \rightarrow 0, \quad n \rightarrow \infty. \end{aligned}$$

The theorem is proved. □

Proof of Lemma 4.3. Without loss of generality, we assume that

$$\|\tilde{\Delta}(s, x, u)\| \leq 1, \quad (s, x) \in K, \quad u \in \mathbb{R}^d,$$

and that $\tilde{\Delta}$ is a Lipschitz function of the variables (s, x) whose Lipschitz constant L is nonrandom. By S we denote the unit sphere in $\mathbb{R}^m$. For an arbitrary k , we have

$$\begin{aligned}
 (4.10) \quad & \lim_{n \rightarrow \infty} \inf_{(x,s) \in K, \bar{l} \in S} \Pi \left(u: \|u\| \geq \frac{1}{n}, (\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d} \neq 0 \right) \\
 & \geq \lim_{n \rightarrow \infty} \inf_{(x,s) \in K, \bar{l} \in S} \Pi \left(u: \|u\| \geq \frac{1}{n}, |(\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d}| > \frac{1}{3k} \right) \\
 & \geq \lim_{n \rightarrow \infty} \min_{(x,s) \in K_k, \bar{l} \in S_k} \Pi \left(u: \|u\| \geq \frac{1}{n}, |(\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d}| > \frac{1}{k} \right) \\
 & = \min_{(x,s) \in K_k, \bar{l} \in S_k} \Pi \left(u: |(\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d}| > \frac{1}{k} \right) \\
 & \geq \inf_{(x,s) \in K, \bar{l} \in S} \Pi \left(u: |(\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d}| > \frac{1}{k} \right),
 \end{aligned}$$

where K_k and S_k are some finite $(3Lk)^{-1}$ - and $(3k)^{-1}$ -coverings of the compact sets K and S , respectively. Now the proof is completed by applying the following analog of the Dini theorem. $\square$

Lemma 4.4. *Let $\{f_k, k \geq 1\}$ be a sequence of functions defined on a compact space X . Let the functions assume values in $\mathbb{R} \cup \{+\infty\}$ and satisfy the following conditions:*

- a) *for all $z \in X$, the sequence $f_k(z)$ is increasing to $+\infty$;*
- b) *every function f_k is upper semicontinuous, that is,*

$$f_k(z) \leq \liminf_n f_k(z_n)$$

for an arbitrary sequence $z_n \rightarrow z$.

Then the sequence f_k tends to $+\infty$ uniformly on X .

Proof. The proof is simple: if the sequence f_k does not tend to $+\infty$ uniformly on X , then there exists a sequence of points $z_k \rightarrow z$, $k \rightarrow \infty$, such that

$$f_k(z_k) \leq C < +\infty, \quad k \geq 1.$$

Thus $f_k(z) \leq C$, $k \geq 1$, in view of the monotonicity and semicontinuity. The latter inequality does not hold by condition a).

Properties a) and b) are easy to check in the case of $X = K \times S$ and

$$f_k(s, x, l) = \Pi \left(u: |(\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d}| > \frac{1}{k} \right).$$

Thus Lemma 4.4 implies that

$$\inf_{(x,s) \in K, \bar{l} \in S} \Pi \left(u: |(\tilde{\Delta}(x, s, u), \bar{l})_{\mathbb{R}^d}| > \frac{1}{k} \right) \rightarrow +\infty.$$

This together with (4.10) completes the proof. $\square$

Remark 4.1. The measure Π of the set

$$\Delta_{T,R} \equiv \left\{ u: \sup_{s \leq T, \|x\| \leq R} \|\nabla_x c(s, x, u)\| \geq 1 \right\}$$

is finite for all $T, R < +\infty$ in view of (4.2). Put

$$\Omega_{T,R} = \{\nu([0, T] \times \Delta_{T,R}) = 0\} \cap \left\{ \sup_{s \leq T} \|x(s)\| \leq R \right\}.$$

If condition (4.5) does not hold, then we proceed in the same way as in the proof of Theorem 4.1 and obtain that

$$\mathbf{P}|_{\Omega_t \cap \Omega_{t,R}} \circ [x(t)]^{-1} \ll \lambda^m$$

for all $R < +\infty$. Let

$$\Omega_{z,t,R} = \{\nu([z,t] \times \Delta_{T,R}) = 0\} \cap \left\{ \sup_{s \leq t} \|x(s)\| \leq R \right\}, \quad \Omega_{z,t} = \{\omega : \dim S_{z,t}(\omega) = m\},$$

where $S_{z,t}$ is the linear span of the vectors

$$\left\{ [\mathcal{E}_0^{\tau-}]^{-1} \tilde{\Delta}(\tau, x(\tau-), p(\tau)), \tau \in [z, t] \right\}.$$

Since the solution of equation (4.1) is evolutionary, we get for all $R < +\infty$ that

$$\mathbf{P}|_{\Omega_{z,t} \cap \Omega_{z,t,R}} \circ [x(t)]^{-1} \ll \lambda^m.$$

Note that $\mathbf{P}(\Omega_{z,t,R}) \rightarrow 1$ as $R \rightarrow +\infty$ and $z \rightarrow t-$. On the other hand, it is shown in the proof of Theorem 4.2 that $\mathbf{P}(\Omega_{z,t}) = 1$, $z < t$, if condition (4.7) is satisfied. Thus Theorem 4.2 holds even if condition (4.5) is not assumed.

To conclude Section 4 we discuss condition (4.7) in more detail. It can be viewed as a natural condition that the vector function $\tilde{\Delta}$ is nondegenerate; namely, it says that the intensity is infinite for any proper subspace $L \subset \mathbb{R}^m$ and for $u \in \mathbb{R}^d$ such that $\tilde{\Delta}(\cdot, \cdot, u) \notin L$. This condition is equivalent to the assumption that a combination of certain differential and difference operators is nondegenerate. This interpretation is close to the well-known Hörmander condition for diffusion. Below we give a related statement for a homogeneous-in-time equation.

In what follows we need a family of operators $\{\Psi_u, u \in \mathbb{R}^d\}$, where any operator Ψ_u acts from $C^1(\mathbb{R}^m)$ to $C(\mathbb{R}^m)$ and the function $u \mapsto \Psi_u f(x)$ is measurable for all fixed $x \in \mathbb{R}^m$ and $f \in C^1(\mathbb{R}^m)$.

Definition 4.1. A family $\{\Psi_u\}$ is nondegenerate with respect to the measure Π if

$$\Pi(\{u : \Psi_u f(x) \neq 0\}) = +\infty$$

for all $x \in \mathbb{R}^m$ and for all $f \in C^1(\mathbb{R}^m)$ such that $\nabla f(x) \neq 0$.

The following result is a corollary of Theorem 4.2.

Theorem 4.3. Suppose a and c do not depend on s . By A and C_u , $u \in \mathbb{R}^d$, we denote the operators

$$Af(x) = (\nabla f(x), \tilde{a}(x))_{\mathbb{R}^m}^m, \quad C_u f(x) = f(x + c(x, u)) - f(x), \quad x \in \mathbb{R}^m, \quad f \in C^1(\mathbb{R}^m),$$

respectively, where I is the unit operator. If the family

$$\Psi_u = [A, C_u](I + C_u)^{-1} = A - (I + C_u)A(I + C_u)^{-1}, \quad u \in \mathbb{R}^d,$$

is nondegenerate with respect to the measure Π , then the distribution of the random variable $x(t)$ has the density with respect to λ^m for all $t > 0$.

Proof. Theorem 4.3 follows from the relation

$$\Psi_u f(x) = (\nabla f(x), \tilde{\Delta}(x, u))_{\mathbb{R}^m}, \quad x \in \mathbb{R}^m, \quad f \in C^1(\mathbb{R}^m), \quad u \in \mathbb{R}^d,$$

which can be checked directly. $\square$

Corollary 4.1. *Let a pseudo-differential operator L be defined by*

$$L = A + \int_{\mathbb{R}^d} C_u \Pi(du).$$

If all the assumptions of Theorem 4.3 hold, then the fundamental solution of the equation

$$u'_t = Lu$$

is an ordinary (not generalized) function.

5. SMOOTH DENSITY: A COUNTEREXAMPLE

A natural question arises on whether or not the density defined in Theorems 4.2 and 4.3 is regular (continuous or of the class C^∞)? We mentioned in Section 1 that the positive answer follows from some conditions posed on the regularity of the intensity measure Π . Thus it is interesting to clarify whether the approach based on the differentiation in time allows one to obtain any result concerning the smoothness of the density (recall that this approach does not require any condition posed on Π).

The standard idea (see [2, 12]) is to apply several times the integration-by-parts formula and finally use a variant of the Sobolev Lemma.

There are two problems when following this idea in our case. Being combined, these problems do not allow one to apply the integration-by-parts formula more than one time. The first of these problems is that the logarithmic derivative of the initial measure is not stochastically differentiable (see Example 2.3). In other words, this means that the initial measure is differentiable only once with respect to $\mathcal{T}^\Gamma$. Another problem is that the solution of a stochastic differential equation also is not stochastically differentiable.

In order to show that the problems mentioned above are principal (and not explained by the method, which can be inadequate in our case) we construct an example of an equation of the form (4.1) for which the conditions of Theorems 4.2 and 4.3 hold but the density of the distribution is discontinuous. Since the condition (3.7) posed on the coefficients of the equation holds at every point, one may conjecture that the smoothness of the density is not determined by properties of the coefficients of the equation; it also depends on the properties of the intensity measure.

Consider the equation

$$(5.1) \quad x(t) = x_0 + \int_0^t x(s) ds + \eta_t, \quad t \geq 0,$$

where

$$\eta_t = \sum_{k=1}^{\infty} 4^{-k} \eta_t^k,$$

and the $\{\eta_k\}$ are independent Poisson processes with the same intensity λ . Equation (5.1) is of the form (4.1) with $m = d = 1$, $\tilde{a}(x) = x$, $c(x, u) = u$, and $\Pi = \lambda \sum_k \delta_{1/4^k}$. Thus

$$\tilde{\Delta}(x, u) = u$$

for all $x \in \mathbb{R}$ and $\Pi(\tilde{\Delta} \neq 0) = \Pi(u \neq 0) = +\infty$. Therefore the solution $x(t)$ has the density p_t with respect to λ^1 by Theorem 4.2.

Put $x_0 = 0$. Then the density vanishes on the negative semiaxis. If

$$\limsup_{\varepsilon \rightarrow 0+} \frac{1}{\varepsilon^{k+1}} \mathbb{P}(x(t) < \varepsilon) = +\infty,$$

then $p_t \notin C^k(\mathbb{R})$. Since

$$x(t) = \int_0^t e^{t-s} \eta_s ds,$$

we have to prove that

$$\limsup_{\varepsilon \rightarrow 0+} \frac{1}{\varepsilon^{k+1}} \mathbf{P}(\eta_t < \varepsilon) = +\infty.$$

We have

$$\mathbf{P}(\eta_t < \varepsilon) \geq \prod_{k \geq 1} \mathbf{P}\left(\frac{1}{4^k} \eta_t^k < \frac{\varepsilon}{2^k}\right) = \prod_{k \geq 1} \mathbf{P}(\eta_t^k < \varepsilon \cdot 2^k).$$

For the sequence $\varepsilon_m = 2^{-m}$, we get

$$\mathbf{P}(\eta_t < \varepsilon_m) \geq [\mathbf{P}(\eta_t^1 = 0)]^m \cdot \theta = \theta \cdot [\varepsilon_m]^{\lambda t / \ln 2},$$

where

$$\theta = \prod_{k \geq 1} \mathbf{P}(\eta_t^k < 2^k) \in (0, 1).$$

Now $p_t \notin C^k(\mathbb{R})$ if $\lambda t / \ln 2 < k + 1$. In particular, if λt does not exceed $\ln 2$, then the density p_t is discontinuous.

The above construction may also serve as a counterexample for another method of proving the regularity of solutions of stochastic differential equations with jumps used by Picard [15]. The Picard method is based on the perturbation of the point measure by extending it with new points (details of this approach can be found in [15]). The Picard method requires that the Lévy measure is such that

$$\int_{\|u\| < \varepsilon} (\zeta, u)_{\mathbb{R}^d}^2 \Pi(du) \asymp \varepsilon^\beta, \quad \varepsilon \rightarrow 0+,$$

for some $\beta \in (0, 2)$ and for all $\zeta \in \mathbb{R}^d$ such that $\|\zeta\| = 1$, where $a_\varepsilon \asymp b_\varepsilon$ means that the fraction $a_\varepsilon/b_\varepsilon$ is far away from both 0 and $+\infty$. In the above example, $\beta = 2$. This means that the upper bound $\beta < 2$ is essential for the Picard method.

Acknowledgement. The author is grateful to M. A. Lifshits who turned the author's attention to the importance of an example where the density exists but is not regular. The author also thanks A. Yu. Pilipenko who conjectured that such an example can be constructed by considering equations similar to (5.1).

BIBLIOGRAPHY

1. J. M. Bismut, *Calcul des variations stochastiques et processus de sauts*, Z. Wahrsch. Verw. Gebiete **63** (1983), 147–235. MR0701527 (85a:60077)
2. K. Bichteler, J.-B. Gravereaux, and J. Jacod, *Malliavin Calculus for Processes with Jumps*, Gordon and Breach Science Publishers, New York–London–Paris–Tokyo, 1987. MR1008471 (90h:60056)
3. J. Norris, *Simplified Malliavin calculus*, Lecture Notes in Math. **1206** (1985), 101–130. MR0942019 (89f:60058)
4. T. Komatsu and A. Takeuchi, *On the smoothness of PDF of solutions to SDE of jump type*, Int. J. Differ. Equ. Appl. **2** (2001), no. 2, 141–197. MR1930241 (2003i:60096)
5. R. J. Elliott and A. H. Tsoi, *Integration by parts for Poisson processes*, J. Multivar. Anal. **44** (1993), no. 2, 179–190. MR1219202 (94c:60093)
6. A. M. Kulik, *Admissible transformations and Malliavin calculus for compound Poisson process*, Theory Stoch. Processes. **5(21)** (1999), no. 3–4, 120–126. MR2018406 (2004i:60080)
7. N. Privault, *Connections and curvature in the Riemannian geometry of configuration spaces*, J. Funct. Anal. **185** (2001), no. 2, 367–403. MR1856271 (2002g:58064)
8. N. Privault, *Hypothesis testing and Skorokhod stochastic integration*, J. Appl. Probab. **37** (2000), 560–574. MR1781013 (2002g:62028)
9. A. M. Kulik, *Some remarks on time-stretching differentiation for general Lévy processes*, Theory Stoch. Process. **7(23)** (2001), no. 3–4, 50–63.
10. A. M. Kulik, *Anticipating linear stochastic differential equations by Poisson random measures*, Theory Stoch. Process. **8(24)** (2002), no. 3–4, 233–242. MR2027395 (2005b:60160)

11. A. V. Skorokhod, *Random Processes with Independent Increments*, “Nauka”, Moscow, 1967; English transl., Kluwer Academic Publishers Group, Dordrecht, 1991. MR1155400 (93a:60114)
12. Yu. A. Davydov, M. A. Lifshits, and N. V. Smorodina, *Local Properties of Distributions of Stochastic Functionals*, “Nauka”, Moscow, 1995; English transl., American Mathematical Society, Providence, RI, 1998. MR1604537 (98k:60053b)
13. A. Yu. Pilipenko, *Properties of stochastic differential operators in non-Gaussian case*, Theory Stoch. Process. **2(18)** (1996), no. 3–4, 153–161.
14. H. Federer, *Geometric Measure Theory*, Springer-Verlag, New York, 1969. MR0257325 (41:1976)
15. J. Picard, *On the existence of smooth densities for jump processes*, Probab. Theory Related Fields **105** (1996), 481–511. MR1402654 (97h:60056)

INSTITUTE OF MATHEMATICS, NATIONAL ACADEMY OF SCIENCES OF UKRAINE, TERESHCHENKOVSKAYA
3, KIEV 01601, UKRAINE

E-mail address: kulik@imath.kiev.ua

Received 19/MAR/2004

Translated by OLEG KLESOV