On differentiability of stochastic flow for a multidimensional SDE with discontinuous drift

Olga V. Aryasova*   Andrey Yu. Pilipenko†

Abstract

We consider a d-dimensional SDE with an identity diffusion matrix and a drift vector being a vector function of bounded variation. We give a representation for the derivative of the solution with respect to the initial data.

Keywords: Stochastic flow; Continuous additive functional; Differentiability with respect to initial data.
AMS MSC 2010: 60J65; 60H10.
Submitted to ECP on June 20, 2013, final version accepted on July 11, 2014.
Supersedes arXiv:1306.4816v3.

Introduction

Consider an SDE of the form

$$\begin{cases} d\varphi_t(x) = a(\varphi_t(x))dt + dw_t, \\ \varphi_0(x) = x, \end{cases} \tag{0.1}$$

where $x \in \mathbb{R}^d$, $d \geq 1$, $(w_t)_{t\geq 0}$ is a d-dimensional Wiener process, $a = (a^1, \ldots, a^d)$ is a bounded measurable mapping from $\mathbb{R}^d$ to $\mathbb{R}^d$.

According to [23] there exists a unique strong solution to equation (0.1).

It is well known that if a is continuously differentiable and its derivative is bounded, then equation (0.1) generates a flow of diffeomorphisms. It turns out that this condition can be essentially reduced [12], and a flow of diffeomorphisms exists in the case of possible unbounded Hölder continuous drift vector a. Recently the case of discontinuous drift was studied in [10, 11, 19, 20] and the weak differentiability of the solution to (0.1) was proved under rather weak assumptions on the drift. The authors of [10] consider a drift vector belonging to $L_q(0, T; L_p(\mathbb{R}^d))$ for some p, q such that

$$p \geq 2, \ q > 2, \ \frac{d}{p} + \frac{2}{q} < 1.$$

They establish the existence of the Gâteaux derivative in $L_2(\Omega \times [0, T]; \mathbb{R}^d)$. In [20] it is proved that for a bounded measurable drift vector a the solution belongs to the space

*Institute of Geophysics, National Academy of Sciences of Ukraine, Ukraine.
  E-mail: oaryasova@gmail.com
†Institute of Mathematics, National Academy of Sciences of Ukraine, Ukraine.
  E-mail: pilipenko.ay@yandex.ua

$L^2(\Omega; W^{1,p}(U))$ for each $t \in \mathbb{R}^d, p > 1$, and any open and bounded $U \in \mathbb{R}^d$. The Malliavin calculus is used in [19, 20].

The aim of our paper is to find a natural representation of the derivative $\nabla_x \varphi_t(x)$ if a is discontinuous. We suppose that for $1 \leq i \leq d$, a^i is a function of bounded variation on $\mathbb{R}^d$, i.e. for each $1 \leq j \leq d$, the generalized derivative $\mu^{ij} = \frac{\partial a^i}{\partial x_j}$ is a signed measure on $\mathbb{R}^d$. Let $\mu^{ij,+}, \mu^{ij,-}$ be measures from the Hahn-Jordan decomposition $\mu^{ij} = \mu^{ij,+} - \mu^{ij,-}$. Denote $|\mu^{ij}| = \mu^{ij,+} + \mu^{ij,-}$. Assume that for all $1 \leq i, j \leq d$, $|\mu^{ij}|$ is a measure of Kato's class, i.e.

$$\lim_{t \downarrow 0} \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_0^t \frac{1}{(2\pi s)^{d/2}} \exp \left\{ -\frac{\|y - x\|^2}{2s} \right\} ds \right) |\mu^{ij}|(dy) = 0.$$

The condition we impose on the drift is more restrictive than that of [10, 20], but it allows us to obtain a representation for the derivative in terms of intrinsic parameters of the initial equation (see Theorem 2.3). Our methods are different from those used in the papers cited above. We show that the derivative $Y_t(x)$ in x is a solution of the following integral equation

$$Y_t(x) = E + \int_0^t dA_s(\varphi(x)) Y_s(x),$$

where $A_t(\varphi(x))$ is a continuous additive functional of the process $(\varphi_t(x))_{t \geq 0}$, which is equal to $\int_0^t \nabla a(\varphi_s(x)) ds$ if a is differentiable, E is the d-dimensional identity matrix. This representation is a natural generalization of the expression for the derivative in the smooth case.

In the one-dimensional case (see [3, 4]) the derivative was represented via the local time of the process. It is well known that the solution of (0.1) does not have a local time at a point in the multidimensional situation. We use continuous additive functionals for the representation of the derivative. This method can be considered as a generalization of the local time approach to the multidimensional case.

Our method can be used in the case of non-constant diffusion and.

The paper is organized as follows. In Section 1 we collect some definitions and statements concerning continuous additive functionals. The main result of the paper is formulated in Section 2 (see Theorem 2.3). For the proof we approximate equation (0.1) by equations with smooth coefficients. The definitions and properties of approximating equations are given in Sections 3, 4. We prove Theorem 2.3 in Section 5.

1   Preliminaries: W-functionals

In this section we collect some facts about continuous additive functionals which will be used in the sequel. Further information can be found in [6]; [8], Ch. 6–8; [13], Ch. II, §6.

Let $(\xi_t, \mathcal{M}_t, \mathbb{P}_x)$ be a homogeneous Markov process with a phase space $\mathbb{R}^d$ (see notations in [8]). Assume that $\xi_t, \ t \geq 0$, has continuous trajectories and the infinite life-time. Denote $\mathcal{N}_t = \sigma\{\xi_s : 0 \leq s \leq t\}$.

Definition 1.1. *A random function $A_t, t \geq 0$, adapted to the filtration $\{\mathcal{N}_t\}$ is called a continuous additive functional of the process $(\xi_t)_{t \geq 0}$ if it is*

- *non-negative;*
- *continuous in t;*
- *homogeneous additive, i.e. for all $t \geq 0, \ s > 0, \ x \in \mathbb{R}^d$,*

$$A_{t+s} = A_s + \theta_s A_t \ \ \mathbb{P}_x - almost \ surely, \tag{1.1}$$

where θ is the shift operator.

If additionally for each $t \geq 0$,

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}_x A_t < \infty,$$

then A_t, $t \geq 0$, is called a W-functional.

Remark 1.2. *It follows from Definition 1.1 that a W-functional is non-decreasing as a function of t, and for all $x \in \mathbb{R}^d$*

$$\mathbb{P}_x\{A_0 = 0\} = 1.$$

Definition 1.3. *The function*

$$f_t(x) = \mathbb{E}_x A_t, \ t \geq 0, \ x \in \mathbb{R}^d,$$

is called the characteristic of a W-functional A_t.

Proposition 1.4 (See [8], Theorem 6.3)**.** *A W-functional is defined by its characteristic uniquely up to equivalence.*

The following theorem states the connection between the convergence of functionals and the convergence of their characteristics.

Theorem 1.5 (See [8], Theorem 6.4)**.** *Let $A_{n,t}$, $n \geq 1$, be W-functionals of the process $(\xi_t)_{t \geq 0}$ and $f_{n,t}(x) = \mathbb{E}_x A_{n,t}$ be their characteristics. Suppose that for each $t > 0$, a function $f_t(x)$ satisfies the condition*

$$\lim_{n \to \infty} \sup_{0 \leq u \leq t} \sup_{x \in \mathbb{R}^d} |f_{n,u}(x) - f_u(x)| = 0. \tag{1.2}$$

Then $f_t(x)$ is the characteristic of a W-functional A_t. Moreover,

$$A_t = \operatorname*{l.i.m.}_{n \to \infty} A_{n,t},$$

where $\mathrm{l.i.m.}$ denotes the convergence in mean square (for any initial distribution of ξ_0).

Proposition 1.6 (See [8], Lemma 6.1′)**.** *If for any $t \geq 0$ the sequence of non-negative additive functionals $\{A_{n,t} : n \geq 1\}$ of the Markov process $(\xi_t)_{t \geq 0}$ converges in probability to a continuous functional A_t, then the convergence in probability is uniform, i.e.*

$$\forall \, T > 0 \ \sup_{t \in [0,T]} |A_{n,t} - A_t| \to 0, \ n \to \infty, \ \text{in probability.}$$

Let h be a non-negative bounded measurable function on $\mathbb{R}^d$, let the process $(\xi_t)_{t \geq 0}$ has a transition probability density $p_t(x, y)$. Then

$$A_t := \int_0^t h(\xi_s) ds$$

is a W-functional of the process $(\xi_t)_{t \geq 0}$ and its characteristic is equal to

$$f_t(x) = \int_{\mathbb{R}^d} \left(\int_0^t p_s(x, y) ds \right) h(y) dy = \int_{\mathbb{R}^d} k_t(x, y) h(y) dy,$$

where

$$k_t(x, y) = \int_0^t p_s(x, y) ds.$$

Let a measure ν be such that $\int_{\mathbb{R}^d} k_t(x, y) \nu(dy)$ is a function continuous in (t, x). If we can choose a sequence of non-negative bounded continuous functions $\{h_n : n \geq 1\}$ such that for each $T > 0$,

$$\lim_{n \to \infty} \sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} \left| \int_{\mathbb{R}^d} k_t(x, y) h_n(y) dy - \int_{\mathbb{R}^d} k_t(x, y) \nu(dy) \right| = 0,$$

then by Theorem 1.5 there exists a W-functional corresponding to the measure ν with its characteristic being equal to $\int_{\mathbb{R}^d} k_t(x, y)\nu(dy)$. Formally we will denote this functional by $\int_0^t \frac{d\nu}{dy}(\xi_s)ds$.

A sufficient condition for the existence of a W-functional corresponding to a given measure is stated in the following theorem.

Theorem 1.7 (See [8], Theorem 6.6). *Let the condition*

$$\lim_{t \downarrow 0} \sup_{x \in \mathbb{R}^d} f_t(x) = \lim_{t \downarrow 0} \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} k_t(x, y)\nu(dy) = 0 \tag{1.3}$$

hold. Then $f_t(x)$ is the characteristic of a W-functional A_t^ν. Moreover,

$$A_t^\nu = \underset{h \to 0}{\text{l.i.m.}} \int_0^t \frac{f_h(\xi_u)}{h} du,$$

and the sequence of characteristics of functionals $\int_0^t \frac{f_h(\xi_u)}{h} du$ converges to $f_t(x)$ in sense of the relation (1.2).

Let us return to the SDE (0.1). Let $(\varphi_t)_{t \geq 0}$ be a solution of equation (0.1) with bounded measurable a. The transition probability density $p_t^\varphi(y, z)$ of the process $(\varphi_t)_{t \geq 0}$ satisfies the Gaussian estimates (see [2])

$$\frac{K_1}{t^{d/2}} \exp\left\{-k_1 \frac{\|y - z\|^2}{t}\right\} \leq p_t^\varphi(y, z) \leq \frac{K_2}{t^{d/2}} \exp\left\{-k_2 \frac{\|y - z\|^2}{t}\right\} \tag{1.4}$$

valid in every domain of the form $t \in [0, T], y \in \mathbb{R}^d, z \in \mathbb{R}^d$, where $T > 0$, K_1, k_1, K_2, k_2 are positive constants that depend only on d, T, and $\|a\|_\infty$.

Denote by $k_t^w(x, y)$ the kernel $k_t(x, y)$ built on the transition density of the Wiener process, i.e.

$$k_t^w(x, y) = \int_0^t \frac{1}{(2\pi s)^{d/2}} \exp\left\{-\frac{\|y - x\|^2}{2s}\right\} ds. \tag{1.5}$$

It is easily to see ([8], Ch. 8, §1) that for all $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, $x \neq y$, $k_t^w(x, y) = \widetilde{k}_t(\|x - y\|)$, where

$$\widetilde{k}_t(r) = \frac{1}{2\pi^{d/2}} r^{2-d} \int_{r^2/2t}^\infty s^{d/2-2} e^{-s} ds, \ r > 0. \tag{1.6}$$

Therefore, the kernel $k_t^w(x, y)$ has a singularity if $x = y$ (for $d > 1$) and the integral

$$f_t(x) = \int_{\mathbb{R}^d} k_t^w(x, y)\nu(dy)$$

is not well defined for all measures.

Definition 1.8 (see [16]). *A measure ν is a measure of Kato's class if*

$$\lim_{t \downarrow 0} \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} k_t^w(x, y)\nu(dy) = 0. \tag{1.7}$$

It follows from (1.4) that a measure ν satisfies the condition (1.3) if and only if it belongs to Kato's class.

Remark 1.9. *A measure ν satisfies the condition (1.7) if and only if*

$$\sup_{x \in \mathbb{R}} \int_{|x-y| \leq 1} \nu(dy) < \infty, \quad \textit{when } d = 1;$$

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \mathbb{R}^2} \int_{|x-y| \leq \varepsilon} \ln \frac{1}{|x - y|} \nu(dy) = 0, \quad \textit{when } d = 2;$$

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \mathbb{R}^d} \int_{|x-y| \leq \varepsilon} |x - y|^{2-d} \nu(dy) = 0, \quad \textit{when } d \geq 3.$$

The proof is a slight modification of that for the case of $\nu(dx) = f(x)dx$ given in [1], Theorem 4.5 (see also [22], Exercise 1 on p. 12). Here f is a non-negative Borel measurable function. We use the representation (1.6) in the proof.

Example 1.10. *Let $d = 1$. For each $y \in \mathbb{R}^d$, the measure $\nu = \delta_y$ belongs to Kato's class and corresponds to the W-functional*

$$L_t(y) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{[y-\varepsilon, y+\varepsilon]}(w_s)\, ds,$$

which is called the local time of a Wiener process at the point y. Assume that ν is a measure of Kato's class. This means now that $\sup_{x \in \mathbb{R}} \nu([x, x+1]) < \infty$. Then (see [21], Ch. X, §2) the corresponding W-functional can be represented in the form

$$A_t^\nu = \int_{\mathbb{R}} L_t(y)\nu(dy).$$

Remark 1.11. *If $d \geq 2$, then δ_y does not belong to Kato's class. This is in consistency with the well-known fact that the local time at a point for a multidimensional Wiener process does not exist.*

Example 1.12. *If $\nu(dy) = f(y)dy$, where f is a non-negative bounded function, then ν is a measure of Kato's class and $A_t^\nu = \int_0^t f(\xi_s)ds$.*

Example 1.13. *Let $S \subset \mathbb{R}^d$ be a compact $(d-1)$-dimensional C^1-manifold. Denote by σ_S the surface measure. Then for any non-negative bounded function f, the measure $\nu(dy) = f(y)\sigma_S(dy)$ belongs to Kato's class.*

Example 1.14. *Let $d \geq 2$. Assume that a measure ν is such that*

$$\exists k, \gamma > 0 \; \forall x \in \mathbb{R}^d \; \forall \rho \in (0,1]: \quad \nu(B(x,\rho)) \leq k\rho^{d-2+\gamma}.$$

Then (c.f. [5], §2)

$$\exists c = c(d, \gamma) \; \forall x \in \mathbb{R}^d \; \forall \rho \in (0,1]: \quad \int_{B(x,\rho)} |x - y|^{2-d} \nu(dy) \leq ck\rho^\gamma.$$

This inequality together with Remark 1.9 yields that ν is a measure of Kato's class. In particular, the Hausdorff measure on the Sierpinski carpet in $\mathbb{R}^2$ is such a measure (see [5], Example 2.2).

We will need the uniform estimates on the moments of a W-functional.

Proposition 1.15 ([13], Ch. II, §6, Lemma 3)**.** *For all $m \geq 1, t > 0$,*

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}_x(A_t)^m \leq m! \left(\sup_{x \in \mathbb{R}^d} f_t(x) \right)^m. \tag{1.8}$$

Making use of this proposition one can easily obtain the following modification of Khas'minskii's Lemma (see [14] or [22], Ch.1 Lemma 2.1).

Lemma 1.16. *Let the W-function f_t satisfies the condition (1.3). Let A_t be the corresponding W-functional. Then for all $p > 0$, $t \geq 0$, there exists a constant C depending on p, t, and $\|f_t\|_\infty$ such that for all $x \in \mathbb{R}^d$,*

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}_x \exp\{pA_t\} \leq C. \tag{1.9}$$

By the definition of a W-functional, A_t^ν is measurable w.r.t. the σ-algebra generated by the Markov process. Since we have assumed that all the processes are continuous and have the infinite life-times, we may assume that $A_t^\nu = A_t^\nu(\cdot)$ is a measurable function defined on $C([0,\infty), \mathbb{R}^d)$ that depends only on behavior of functions on $[0,t]$ (if there is no misunderstanding, sometimes we will consider A_t^ν as a function on $C([0,t], \mathbb{R}^d)$).

Let $(\varphi_t(x))_{t\geq 0}$ be a solution of (0.1) defined on a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. By $\mathbb{P}_x$ denote the distribution of the process $(\varphi_t(x))_{t\geq 0}$. In Dynkin's notation [8] $((\varphi_t(x))_{t\geq 0}, \mathcal{F}_t, \mathbb{P})$ is called a Markov family of random functions (the measures $\mathbb{P}_x$ are measures on the space of continuous functions, the measure $\mathbb{P}$ is a probability on $(\Omega, \mathcal{F})$). The composition $A_t^\nu(\varphi_\cdot(x)), t \geq 0$, is an additive functional of $(\varphi_t(x))_{t\geq 0}$ corresponding to the measure ν. Note that $A_t^\nu(\varphi_\cdot(x))$ is defined on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ for any $x \in \mathbb{R}^d$.

If the measure ν belongs to Kato's class, then the corresponding additive functionals of $(\varphi_t)_{t\geq 0}$ and the Wiener process are well defined. Denote the corresponding measurable mappings by $A_t^{\nu,\varphi}$ and $A_t^{\nu,w}$. By the Girsanov theorem, for each $x \in \mathbb{R}^d$, the distributions of the processes $(\varphi_t(x))_{t\geq 0}$ and $(x + w_t)_{t\geq 0}$ are equivalent. The question naturally arises whether the mappings $A_t^{\nu,\varphi}$ and $A_t^{\nu,w}$ are the same. The answer is positive and it is formulated in the next Lemma.

Lemma 1.17. *Let ν be a measure of Kato's class. Then for any $x \in \mathbb{R}^d$,*

$$A_t^{\nu,w}(\varphi_\cdot(x)) = A_t^{\nu,\varphi}(\varphi_\cdot(x)) \ \ \mathbb{P} - \text{almost surely.}$$

Proof. For $x \in \mathbb{R}^d$, denote by $(w_t(x))_{t\geq 0}$ the process $(x + w_t)_{t\geq 0}$. According to Theorem 1.7,

$$A_t^{\nu,w}(w_\cdot(x)) = \underset{h\downarrow 0}{\text{l.i.m.}} \int_0^t \frac{f_h^w(w_s(x))}{h} ds.$$

Then by the Girsanov theorem,

$$A_t^{\nu,w}(\varphi_\cdot(x)) = \underset{h\downarrow 0}{\mathbb{P}\text{–lim}} \int_0^t \frac{f_h^w(\varphi_s(x))}{h} ds, \tag{1.10}$$

where $\mathbb{P}$–lim means the limit in probability.

It remains to show that the characteristics of $\int_0^t \frac{f_h^w(\varphi_s(x))}{h} ds$ converge uniformly to $\int_{\mathbb{R}^d} k_t^\varphi(x,y)\nu(dy)$ as $h \downarrow 0$ (see Theorem 1.5). This proof is routine and technical, so we postpone it to Appendix. $\square$

2 The main result

Let a be a bounded measurable function of bounded variation. Denote by ∇a the matrix $\left(\frac{\partial a^i}{\partial x_j}\right)_{1\leq i,j\leq d}$ and for $1 \leq i,j \leq d$, by μ^{ij} the signed measure $\frac{\partial a^i}{\partial x_j}$. Let $\mu^{ij} = \mu^{ij,+} - \mu^{ij,-}$ be the Hahn-Jordan decomposition of μ^{ij}. Further on we suppose that for all $1 \leq i,j \leq d$, the measure $|\mu^{ij}| = \mu^{ij,+} + \mu^{ij,-}$ belongs to Kato's class.

By Theorem 1.7, there exist W-functionals $A_t^{\mu^{ij,\pm},w}$ (we will denote the corresponding mappings by $A_t^{ij,\pm}(\cdot)$) with their characteristics defined according to the formula

$$f_t^{ij,\pm}(x) = \int_{\mathbb{R}^d} k_t^w(x,y)\mu^{ij,\pm}(dy).$$

Denote $A_t^{ij} = A_t^{ij,+} - A_t^{ij,-}, \ A_t = (A_t^{ij})_{1\leq i,j\leq d}$.

Remark 2.1. *Assume that the measure μ^{ij} can be represented in the form $\mu^{ij} = \tilde{\mu}^{ij,+} - \tilde{\mu}^{ij,-}$, where $\tilde{\mu}^{ij,+}, \tilde{\mu}^{ij,-}$ are from Kato's class and are not necessarily orthogonal. Then $A^{\mu^{ij,+}} - A^{\mu^{ij,-}} = A^{\tilde{\mu}^{ij,+}} - A^{\tilde{\mu}^{ij,-}}$.*

Remark 2.2. *Recall that the mappings $A_t^{ij,+}, A_t^{ij,-}$ are continuous and monotonous in t. So the function $t \to A_t^{ij}$ is a continuous function of bounded variation on $[0, T]$ almost surely.*

The main result on differentiability of a flow generated by equation (0.1) with respect to the initial conditions is given in the following theorem.

Theorem 2.3. *Let $a : \mathbb{R}^d \to \mathbb{R}^d$ be such that for all $1 \le i \le d$, a^i is a function of bounded variation on $\mathbb{R}^d$. Put $\mu^{ij} = \frac{\partial a^i}{\partial x_j}$, $1 \le i, j \le d$. Assume that the measures $|\mu^{ij}|, 1 \le i, j \le d$, belong to Kato's class. Let $Y_t(x)$, $t \ge 0$, be a solution to the integral equation*

$$Y_t(x) = E + \int_0^t dA_s(\varphi(x))Y_s(x), \tag{2.1}$$

where E is the $d \times d$-identity matrix, the integral on the right-hand side of (2.1) is the Lebesgue-Stieltjes integral with respect to the continuous function of bounded variation $t \to A_t(\varphi(x))$.

Then $Y_t(x)$ is the derivative of $\varphi_t(x)$ in L_p-sense: for all $p > 0$, $x \in \mathbb{R}^d$, $h \in \mathbb{R}^d$, $t \ge 0$,

$$\mathbb{E} \left\| \frac{\varphi_t(x + \varepsilon h) - \varphi_t(x)}{\varepsilon} - Y_t(x)h \right\|^p \to 0, \ \varepsilon \to 0, \tag{2.2}$$

where $\| \cdot \|$ is a norm in the space $\mathbb{R}^d$. Moreover,

$$P \left\{ \forall t \ge 0 : \varphi_t(\cdot) \in W_{p,loc}^1(\mathbb{R}^d, \mathbb{R}^d), \nabla \varphi_t(x) = Y_t(x) \text{ for } \lambda\text{-a.a. } x \right\} = 1,$$

where λ is the Lebesgue measure on $\mathbb{R}^d$.

Remark 2.4. *The differentiability was proved in [10, 20]. We give a representation for the derivative. Note that the Sobolev derivative is defined up to the Lebesgue null set.*

Remark 2.5. *Consider the non-homogeneous SDE*

$$\begin{cases} d\varphi_t(x) = a(t, \varphi_t(x))dt + dw_t, \\ \varphi_0(x) = x. \end{cases}$$

Similarly to the arguments given in Section 1 a theory of non-homogeneous additive functionals of non-homogeneous Markov processes can be constructed. All the formulations and proofs can be literally rewritten with natural necessary modifications. Unfortunately, there are no corresponding references, therefore we did not carry out the corresponding reasonings.

Consider examples of functions a for which $|\mu^{ij}|, 1 \le i, j \le d$, are measures of Kato's class.

Example 2.6. *Let for all $1 \le i \le d$, a^i be a Lipschitz function. By Rademacher's theorem [9] the Frechét derivatives $\mu^{ij} = \frac{\partial a^i}{\partial x_j}$ exist almost surely w.r.t. the Lebesgue measure. It is easy to verify that they are bounded and the Frechét derivative coincides with the derivative considered in the generalized sense. Then $|\mu^{ij}|$ belongs to Kato's class.*

Let now $h \in C^1(\mathbb{R}^d, \mathbb{R}^d), D$ be a bounded domain in $\mathbb{R}^d$ with C^1 boundary ∂D. Put $a(x) = h(x) \mathbb{1}_{x \in D}$. It follows from Example 1.13 that for all $1 \le i, j \le d, |\mu^{ij}|$ is a measure of Kato's class because (cf. [24])

$$\mu^{ij}(dx) = \frac{\partial a^i}{\partial x_j}(x) \mathbb{1}_{x \in D} dx + h^i(x) \cos(n_j(x)) \sigma_{\partial D}(dx),$$

where $n(x) = (n_1(x), \dots, n_d(x))$ is the outward unit normal vector at the point $x \in \partial D$.

Condition (1.7) is also satisfied by the measure generated by a being a linear combination of the form

$$h_0(x) + \sum_{k=1}^{m} h_k(x) \mathbb{1}_{x \in D_k}, \tag{2.3}$$

where $h_0 \in \mathrm{Lip}(\mathbb{R}^d, \mathbb{R}^d)$, $h_k \in C^1(\mathbb{R}^d, \mathbb{R}^d)$, $1 \leq k \leq d$, D_k is a bounded domain in $\mathbb{R}^d$ with C^1 boundary.

Further examples of a can be obtained as the limits of sequences of the functions of form (2.3).

In one-dimensional case all the functions of bounded variation generate measures belonginig to Kato's class (see Example 1.10).

See also Example 1.14 showing that if $|\mu^{ij}|$ are "Hausdorff-type" measures with a parameter greater than $(d-1)$, then a satisfies assumptions of the Theorem.

The idea of the proof of Theorem 2.3 is to approximate the solution of equation (0.1) by solutions of SDEs with smooth coefficients. The definition and properties of approximating equations are given in Sections 3, 4. The proof of the Theorem itself is presented in Section 5.

3 Approximation by SDEs with smooth coefficients

For $n \geq 1$, let $g_n \in C_0^\infty(\mathbb{R}^d)$ be a non-negative function such that $\int_{\mathbb{R}^d} g_n(z)dz = 1$, and $g_n(x) = 0$, $|x| \geq 1/n$. Put

$$a_n(x) = (g_n * a)(x) = \int_{\mathbb{R}^d} g_n(x-y)a(y)dy, \ x \in \mathbb{R}^d, \ \ n \geq 1, \tag{3.1}$$

where the function a satisfies the assumptions of Theorem 2.3. Note that

$$\sup_n \|a_n\|_\infty \leq \|a\|_\infty, \tag{3.2}$$

and $a_n \to a$, $n \to \infty$, in $L_{1,loc}(\mathbb{R}^d)$. Passing to subsequences we may assume without loss of generality that $a_n(x) \to a(x)$, $n \to \infty$, for almost all x w.r.t. the Lebesgue measure.

Consider the SDE

$$\begin{cases} d\varphi_{n,t}(x) = a_n(\varphi_{n,t}(x))dt + dw_t, \\ \varphi_{n,0}(x) = x, \ x \in \mathbb{R}^d. \end{cases} \tag{3.3}$$

Put $\nabla a_n = \left(\frac{\partial a_n^i}{\partial x_j} \right)_{1 \leq i,j \leq d}$. Denote by $Y_{n,t}(x)$ the matrix of derivatives of $\varphi_{n,t}(x)$ in x, i.e., $Y_{n,t}^{ij}(x) = \frac{\partial \varphi_{n,t}^i(x)}{\partial x_j}$. Then $Y_{n,t}(x)$ satisfies the equation

$$Y_{n,t}(x) = E + \int_0^t \nabla a_n(\varphi_{n,s}(x)) Y_{n,s}(x) ds, \tag{3.4}$$

where E is the d-dimensional identity matrix.

Lemma 3.1. *For each $p \geq 1$,*

1) for all $t \geq 0$ and any compact set $U \in \mathbb{R}^d$,

$$\sup_{x \in U, \ n \geq 1} (\mathbb{E}(\|\varphi_{n,t}(x)\|^p + \|\varphi_t(x)\|^p)) < \infty;$$

2) for all $x \in \mathbb{R}^d$, $T \geq 0$,

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} \|\varphi_{n,t}(x) - \varphi_t(x)\|^p \right) \to 0 \text{ as } n \to \infty,$$

where $\| \cdot \|$ is a norm in the space $\mathbb{R}^d$.

Proof. Statement 1) follows from the uniform boundedness of the coefficients and the finiteness of the moments of a Wiener process; 2) is proved in [18], Theorem 3.4. □

For $1 \leq i, j \leq d$, put $\mu_n^{ij} = \frac{\partial a_n^i}{\partial x_j}$. By the properties of convolution of a generalized function (see [24], Ch. 2, §7),

$$\nabla a_n = \nabla a * g_n.$$

For each $n \geq 1$, $1 \leq i, j \leq d$, put $\mu_n^{ij,\pm} = \mu^{ij,\pm} * g_n$ and $\mu_n^{ij} = \mu_n^{ij,+} - \mu_n^{ij,-}$ (c.f. Remark 2.1). Then, according to Theorem 1.7, there exist W-functionals $A_{n,t}^{ij,\pm}$ of a Wiener process on $\mathbb{R}^d$ which correspond to the measures $\mu_n^{ij,\pm}$ and have characteristics of the form

$$f_{n,t}^{ij,\pm}(x) = \int_{\mathbb{R}^d} k_t^w(x,y) \mu_n^{ij,\pm}(dy), \ 1 \leq i, j \leq d. \tag{3.5}$$

The functional $A_{n,t}^{ij} = A_{n,t}^{ij,+} - A_{n,t}^{ij,-}$ is given by the formula

$$A_{n,t}^{ij} = \int_0^t \frac{\partial a_n^i}{\partial x_j}(w_u) du \tag{3.6}$$

(see Example 2).

Lemma 3.2. *For each $T > 0$, $x \in \mathbb{R}^d$, $\varepsilon > 0$, $1 \leq i, j \leq d$,*

$$\mathbb{P}_{w(x)} \left\{ \sup_{0 \leq t \leq T} \left| A_{n,t}^{ij,\pm} - A_t^{ij,\pm} \right| > \varepsilon \right\} \to 0, \ n \to \infty,$$

where $\mathbb{P}_{w(x)}$ is the distribution of the process $(x + w_t)_{t \geq 0}$.

The following simple proposition used for the proof of Lemma 3.2 is easily checked.

Proposition 3.3. *Let ν, ν_n, $n \geq 1$, be from the Kato class, f, f_n, $n \geq 1$, be the characteristics of the corresponding W-functionals of a Wiener process, and the representation $\nu_n = g_n * \nu$ hold true. Then the relation $f_{n,t} = g_n * f_t$ is fulfilled.*

Proof of Lemma 3.2. To prove the convergence of functionals in mean square it is sufficient to show that for each $T > 0$, $1 \leq i, j \leq d$,

$$\lim_{n \to \infty} \sup_{0 \leq t \leq T} \sup_{x \in \mathbb{R}^d} |f_{n,t}^{ij,\pm}(x) - f_t^{ij,\pm}(x)| = 0 \tag{3.7}$$

(see Theorem 1.5). Then the uniform convergence in probability follows from Proposition 1.6.

For each $0 < \delta < t$,

$$\sup_{x \in \mathbb{R}^d} \left| f_{n,t}^{ij,\pm}(x) - f_t^{ij,\pm}(x) \right| = \sup_{x \in \mathbb{R}^d} \left| \int_{\mathbb{R}^d} k_t^w(x,y) \left(\mu_n^{ij,\pm}(dy) - \mu^{ij,\pm}(dy) \right) \right| = I^\pm + II^\pm,$$

where

$$I^\pm = \sup_{x \in \mathbb{R}^d} \left| \int_{\mathbb{R}^d} \left(\mu_n^{ij,\pm}(dy) - \mu^{ij,\pm}(dy) \right) \int_0^\delta \frac{1}{(2\pi s)^{d/2}} \exp \left\{ -\frac{\|y - x\|^2}{2s} \right\} ds \right|, \tag{3.8}$$

$$II^\pm = \sup_{x \in \mathbb{R}^d} \left| \int_{\mathbb{R}^d} \left(\mu_n^{ij,\pm}(dy) - \mu^{ij,\pm}(dy) \right) \int_\delta^t \frac{1}{(2\pi s)^{d/2}} \exp \left\{ -\frac{\|y - x\|^2}{2s} \right\} ds \right|.$$

We have

$$I^{\pm} \leq \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |\mu_n^{ij}|(dy) \int_0^{\delta} \frac{1}{(2\pi s)^{d/2}} \exp\left\{-\frac{\|y-x\|^2}{2s}\right\} ds +$$

$$\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |\mu^{ij}|(dy) \int_0^{\delta} \frac{1}{(2\pi s)^{d/2}} \exp\left\{-\frac{\|y-x\|^2}{2s}\right\} ds = I_1 + I_2.$$

Because of the condition (1.7), for each $\varepsilon > 0$, we can choose δ so small that I_2 is less then $\varepsilon/4$. To obtain the same estimate for I_1, note that by the associative, distributive and commutative properties of convolution (see [24], Ch. II, §7),

$$I_1 = \sup_{x \in \mathbb{R}^d} (|\mu_n| * k_{\delta})(x) \leq ((|\mu| * g_n) * k_{\delta})(x) = \sup_{x \in \mathbb{R}^d} (|\mu| * (g_n * k_{\delta}))(x) =$$

$$\sup_{x \in \mathbb{R}^d} (|\mu| * (k_{\delta} * g_n))(x) = \sup_{x \in \mathbb{R}^d} ((|\mu| * k_{\delta}) * g_n)(x) \leq \sup_{x \in \mathbb{R}^d} (|\mu| * k_{\delta})(x) = I_2 < \varepsilon/4.$$

We get $I^{\pm} < \varepsilon/2$.

Consider $II^{\pm}$. The functions

$$q_{\delta,t}^{ij,\pm}(x) := \int_{\mathbb{R}^d} \mu^{ij,\pm}(dy) \int_{\delta}^t \frac{1}{(2\pi s)^{d/2}} \exp\left\{-\frac{\|x-y\|^2}{2s}\right\} ds$$

are equicontinuous in x for $t \in [\delta, T]$. We have

$$\sup_{\delta < t < T} II^{\pm} = \sup_{\delta < t < T} \sup_{x \in \mathbb{R}^d} |(q_{\delta,t}^{ij,\pm} * g_n)(x) - q_{\delta,t}^{ij,\pm}(x)| \to 0, \ n \to \infty.$$

Then there exists n_0 such that for all $n > n_0$, $\sup_{\delta < t < T} II^{\pm} < \varepsilon/2$. $\qquad\square$

Lemma 3.4. *For each $T > 0$, $x \in \mathbb{R}^d$, $\varepsilon > 0$, $1 \leq i, j \leq d$,*

$$\mathbb{P}\left\{\sup_{0 \leq t \leq T} \left|A_{n,t}^{ij,\pm}(\varphi_n(x)) - A_t^{ij,\pm}(\varphi(x))\right| > \varepsilon\right\} \to 0, \ n \to \infty.$$

For the proof we make use of the following proposition

Proposition 3.5. *Let X, Y be complete separable metric spaces, $(\Omega, \mathcal{F}, P)$ be a probability space. Let measurable mappings $\xi_n : \Omega \to X$, $h_n : X \to Y$, $n \geq 0$, be such that*

1) $\xi_n \to \xi_0$, $n \to \infty$, in probability;

2) $h_n \to h_0$, $n \to \infty$, in measure ν, where ν is a probability measure on X;

3) for all $n \geq 1$ the distribution P_{ξ_n} of ξ_n is absolutely continuous w.r.t. the measure ν;

4) the sequence of densities $\{\frac{dP_{\xi_n}}{d\nu} : n \geq 1\}$ is uniformly integrable w.r.t. the measure ν.

Then $h_n(\xi_n) \to h_0(\xi_0)$, $n \to \infty$, in probability.

The proof can be found, for example, in [7], Corollary 9.9.11 or [15], Lemma 2.

Proof of Lemma 3.4. Fix $T > 0$ and $x \in \mathbb{R}^d$. Since φ_t, $\varphi_{n,t}$, $n \geq 1$, are measurable functions of a Wiener process, we may assume without loss of generality that $\Omega = C([0,T], \mathbb{R}^d)$, $\mathcal{F} = \sigma\{w_t : 0 \leq t \leq T\}$, $P = \mathbb{P}$ is the Wiener measure, and put $\xi_n = (\varphi_{n,t}(x))_{0 \leq t \leq T}$, $\xi_0 = (\varphi_t(x))_{0 \leq t \leq T}$, $\nu = \mathbb{P}_{w(x)}$ is the distribution of the process $(w_t(x))_{0 \leq t \leq T}$, $X = C([0,T], \mathbb{R}^d)$, $Y = C([0,T])$, $h_n^{\pm} = A_{n,t}^{ij,\pm}(\cdot)$, $h_0^{\pm} = A_t^{ij,\pm}(\cdot)$. Then

$\{\xi_n : n \geq 0\}$ is a sequence of random elements in the space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values on $C([0,T], \mathbb{R}^d)$. Lemma 3.1 entails the convergence $\xi_n \to \xi_0$, $n \to \infty$, in probability $\mathbb{P}$ uniformly in $t \in [0,T]$. This implies the first assertion of Proposition 3.5.

According to Lemma 3.2, $A_{n,t}^{ij} \to A_t^{ij}$ as $n \to \infty$, in probability measure $\mathbb{P}_{w(x)}$ uniformly in $t \in [0,T]$. This means that $h_n^{\pm} \to h^{\pm}$, $n \to \infty$, as elements of $C([0,T])$ in measure $\mathbb{P}_{w(x)}$. So the second assertion of Proposition 3.5 is justified. The absolute continuity of the distribution of $(\varphi_{n,t}(x))_{0 \leq t \leq T}$ w.r.t. the measure $\mathbb{P}_{w(x)}$ follows from Girsanov's theorem. The density is defined by the formula

$$\beta_n = \frac{d\mathbb{P}_{\varphi_n(x)}}{d\mathbb{P}_{w(x)}} = \exp\left\{ \int_0^T (a_n(w_s(x)), dw_s(x)) - \frac{1}{2} \int_0^T \|a_n(w_s(x))\|^2 ds \right\}.$$

As

$$\mathbb{E} \exp\left\{ \frac{1}{2} \int_0^T \|a_n(w_s(x))\|^2 ds \right\} \leq \exp\left\{ \frac{T}{2} \sup_{y \in \mathbb{R}^d} \|a(y)\|^2 \right\} < \infty,$$

where $\| \cdot \|$ is a norm in $\mathbb{R}^d$, we have that for each $p > 1$,

$$\mathbb{E} \exp\left\{ p \int_0^T (a_n(w_s(x)), dw_s(x)) - \frac{p^2}{2} \int_0^T \|a_n(w_s(x))\|^2 ds \right\} = 1$$

(cf. [17], Theorem 6.1). The uniform integrability of the family $\{ \frac{d\mathbb{P}_{\varphi_n(x)}}{d\mathbb{P}_{w(x)}} : n \geq 1 \}$ follows from the estimate

$$\mathbb{E} \exp\left\{ p \left(\int_0^T (a_n(w_s(x)), dw_s(x)) - \frac{1}{2} \int_0^T \|a_n(w_s(x))\|^2 ds \right) \right\} =$$

$$\mathbb{E} \exp\left\{ p \int_0^T (a_n(w_s(x)), dw_s(x)) - \frac{p^2}{2} \int_0^T \|a_n(w_s(x))\|^2 ds \right\} \times$$

$$\exp\left\{ \frac{1}{2}(p^2 - p) \int_0^T \|a_n(w_s(x))\|^2 ds \right\} \leq$$

$$\exp\left\{ (p^2 - p)\|a_n\|_\infty^2 T \right\} \mathbb{E} \exp\left\{ p \int_0^T (a_n(w_s(x)), dw_s(x)) - \frac{p^2}{2} \int_0^T \|a_n(w_s(x))\|^2 ds \right\} =$$

$$\exp\left\{ (p^2 - p)\|a_n\|_\infty^2 T \right\} \leq \exp\left\{ (p^2 - p)\|a\|_\infty^2 T \right\}$$

valid for $p > 1$. Thus all the assertions of Proposition 3.5 are fulfilled and we have

$$\sup_{0 \leq t \leq T} \left| A_{n,t}^{ij,\pm}(\varphi_n(x)) - A_t^{ij,\pm}(\varphi(x)) \right| \to 0, \ n \to \infty,$$

in probability $\mathbb{P}$. The Lemma is proved. $\qquad\square$

4 Convergence of the derivatives of solutions

Recall that $Y_t(x), Y_{n,t}(x)$, $t \geq 0$, $x \in \mathbb{R}^d$, are the solutions of equations (2.1), (3.4), respectively. In this section we show the convergence of the sequence $\{Y_{n,t}(x) : n \geq 1\}$ in probability uniformly in t. This together with Lemma 3.1 allow us to prove Theorem 2.3.

Lemma 4.1. *1) For all $T \geq 0$, $x \in \mathbb{R}^d$, $p > 0$,*

$$\sup_{n \geq 1} \mathbb{E} \sup_{0 \leq t \leq T} \|Y_{n,t}(x)\|^p < \infty,$$

2) For all $T \geq 0$, $x \in \mathbb{R}^d, p > 0$,

$$\mathbb{E} \sup_{0 \leq t \leq T} \|Y_{n,t}(x) - Y_t(x)\|^p \to 0, \ n \to \infty,$$

where

$$\|Y\| = \max_{1 \leq i,j \leq d} |Y^{ij}|.$$

For the proof we need the following two propositions. The first one is a version of the Gronwall-Bellman inequality and can be obtained by a standard argument.

Proposition 4.2. *Let $x(t)$ be a continuous function on $[0, +\infty)$, $C(t)$ be a non-negative continuous function on $[0, +\infty)$, $K(t)$ be a non-negative, non-decreasing function, and $K(0) = 0$. If for all $0 \leq t \leq T$,*

$$x(t) \leq C(t) + \left| \int_0^t x(s) dK(s) \right|,$$

then

$$x(T) \leq \left(\sup_{0 \leq t \leq T} C(t) \right) \exp\{K(T)\}.$$

Proposition 4.3. *For all $t \geq 0$, $p > 0$, $1 \leq i,j \leq d$, there exists a constant C such that*

$$\sup_{x \in \mathbb{R}^d} \sup_{n \geq 1} \mathbb{E} \left(\exp \left\{ pA_{n,t}^{ij,\pm}(\varphi_n(x)) \right\} + \exp \left\{ pA_t^{ij,\pm}(\varphi(x)) \right\} \right) < C. \tag{4.1}$$

Proof. The statement of the Proposition follows from Lemma 1.16 and the inequalities (1.4), which allow us to obtain the estimates uniform in $n \geq 1$. $\qquad\square$

Proof of Lemma 4.1. For all $t > 0$, define the variation of A^{ij} on $[0, t]$ by

$$\operatorname{Var} A_t^{ij}(\varphi(x)) := A_t^{ij,+}(\varphi(x)) + A_t^{ij,-}(\varphi(x)),$$

and put

$$\operatorname{Var} A_t(\varphi(x)) := \Sigma_{1 \leq i,j \leq d} \operatorname{Var} A_t^{ij}(\varphi(x)).$$

The variations of $A_{n,t}(\varphi_n(x))$, $n \geq 1$, are defined similarly.

The proof of 1). We have

$$\|Y_{n,t}(x)\| \leq 1 + \left\| \int_0^t (dA_{n,s}(\varphi_n(x))) Y_{n,s}(x) \right\| \leq 1 + \int_0^t \|Y_{n,s}(x)\| \, d \left(\operatorname{Var} A_{n,s}(\varphi_n(x)) \right).$$

Making use of the Gronwall-Bellman lemma we get

$$\|Y_{n,t}(x)\| \leq \exp \left\{ \operatorname{Var} A_{n,t}(\varphi_n(x)) \right\} \leq \exp \left\{ \operatorname{Var} A_{n,T}(\varphi_n(x)) \right\}. \tag{4.2}$$

The statement 1) follows now from the estimate (4.2) and Proposition 4.3.

The proof of 2). We have

$$\|Y_{n,t}(x) - Y_t(x)\| \leq \left\| \int_0^t (dA_{n,s}(\varphi_n(x)) - dA_s(\varphi(x))) Y_s(x) \right\| +$$

$$\left\| \int_0^t dA_{n,s}(\varphi_n(x)) (Y_{n,s}(x) - Y_s(x)) \right\| \leq$$

$$\left\| \int_0^t (dA_{n,s}(\varphi_n(x)) - dA_s(\varphi(x))) Y_s(x) \right\| + \int_0^t \|Y_{n,s}(x) - Y_s(x)\| \, d \left(\operatorname{Var} A_{n,s}(\varphi(x)) \right).$$

By Proposition 4.2,

$$\|Y_{n,t}(x) - Y_t(x)\| \leq \sup_{0 \leq u \leq t} \left\| \int_0^u (dA_{n,s}(\varphi_n(x)) - dA_s(\varphi(x))) Y_s(x) \right\| \exp \left\{ \operatorname{Var} A_{n,t}(\varphi_n(x)) \right\}.$$

$$\tag{4.3}$$

To estimate the right-hand side of (4.3) we make use of the following Proposition.

Proposition 4.4. *Let $\{g_n : n \geq 1\}$ be a sequence of continuous monotonic functions on $[0,T]$, and $f \in C([0,T])$. Suppose that for each $t \in [0,T]$, $g_n(t) \to g(t)$, as $n \to \infty$. Then*

$$\sup_{t \in [0,T]} \left| \int_0^t f(s) dg_n(s) - \int_0^t f(s) dg(s) \right| \to 0, \ n \to \infty.$$

We get

$$\sup_{0 \leq u \leq t} \left\| \int_0^u \left(dA_{n,s}(\varphi_n(x)) - dA_s(\varphi(x)) \right) Y_s(x) \right\| \exp\{\mathrm{Var}\, A_{n,t}(\varphi_n(x))\} \leq$$

$$\sup_{0 \leq u \leq t} \left\| \int_0^u \left(dA_{n,s}^+(\varphi_n(x)) - dA_s^+(\varphi(x)) \right) Y_s(x) \right\| \exp\{\mathrm{Var}\, A_{n,t}(\varphi_n(x))\} +$$

$$\sup_{0 \leq u \leq t} \left\| \int_0^u \left(dA_{n,s}^-(\varphi_n(x)) - dA_s^-(\varphi(x)) \right) Y_s(x) \right\| \exp\{\mathrm{Var}\, A_{n,t}(\varphi_n(x))\}. \quad (4.4)$$

Consider the first summand in the right-hand side of (4.4). Put $g_n(s) = A_{n,s}^+(\varphi_n(x))$, $g(s) = A_s^+(\varphi(x))$, and $f(s) = Y_s(x)$. Then Lemma 3.4, Proposition 4.3, and Proposition 4.4 provide that

$$\sup_{0 \leq u \leq t} \left\| \int_0^u \left(dA_{n,s}^+(\varphi_n(x)) - dA_s^+(\varphi(x)) \right) Y_s(x) \right\| \exp\{\mathrm{Var}\, A_{n,t}(\varphi_n(x))\} \to 0 \text{ as } n \to \infty,$$

in probability. Similarly it is proved that the second summand in the right-hand side of (4.4) tends to 0 as $n \to \infty$.

This and statement 1) entail statement 2) of the Lemma. $\qquad\square$

5 The proof of Theorem 2.3

Proof. Define approximating equations by (3.3), where a_n, $n \geq 1$, are determined by (3.1). From Lemma 3.1 and the dominated convergence theorem we get the relation

$$\mathbb{E} \sup_{t \in [0,T]} \int_U |\varphi_{n,t}^i(x) - \varphi_t^i(x)|^p dx \to 0, \ n \to \infty,$$

valid for any bounded domain $U \subset \mathbb{R}^d$, $T > 0$, $p \geq 1$, and $1 \leq i \leq d$. So for each $1 \leq i \leq d$, there exists a subsequence $\{n_k^i : k \geq 1\}$ such that

$$\sup_{t \in [0,T]} \int_U |\varphi_{n_k^i,t}^i(x) - \varphi_t^i(x)|^p dx \to 0 \text{ a.s. as } k \to \infty.$$

Without loss of generality we can suppose that

$$\sup_{t \in [0,T]} \int_U |\varphi_{n,t}^i(x) - \varphi_t^i(x)|^p dx \to 0 \text{ a.s. as } n \to \infty. \quad (5.1)$$

Arguing similarly and taking into account Lemma 4.1 we arrive at the relation

$$\sup_{t \in [0,T]} \int_U |Y_{n,t}^{ij}(x) - Y_t^{ij}(x)|^p dx \to 0, \ n \to \infty, \text{ almost surely}, \quad (5.2)$$

which is fulfilled for all $1 \leq i,j \leq d$, $p \geq 0$.

Since the Sobolev space is a Banach space, the relations (5.1), (5.2) mean that $Y_t(x)$ is the matrix of derivatives of the solution to (0.1).

Let us verify (2.2). We have for all $x, h \in \mathbb{R}^d, \alpha \in \mathbb{R}$,

$$\varphi_{n,t}(x + \alpha h) = \varphi_{n,t}(x) + \int_0^\alpha Y_{n,t}(x + uh) du.$$

On differentiability of stochastic flow

It follows from Lemmas 3.1 and 4.1 that

$$\varphi_t(x + \alpha h) = \varphi_t(x) + \int_0^\alpha Y_t(x + uh)du. \tag{5.3}$$

The following lemma implies the relation

$$\forall y_0 \in \mathbb{R}^d : \ Y_t(y) \to Y_t(y_0), \ y \to y_0, \tag{5.4}$$

in probability and hence in all L_p. This completes the proof of the Theorem, as (5.3) and (5.4) implies (2.2). □

Lemma 5.1. *Let ν be a measure of Kato's class. Then for any $t \geq 0$, $x_0 \in \mathbb{R}^d$, $\varepsilon > 0$,*

$$\mathbb{P}\left\{|A_t^\nu(\varphi(x)) - A_t^\nu(\varphi(x_0))| > \varepsilon\right\} \to 0 \text{ as } x \to x_0. \tag{5.5}$$

Proof. For $e \in \mathbb{R}^d$, denote by ν_e the shift of the measure ν by the vector e, i.e. for each $A \subset \mathbb{R}^d$,

$$\nu_e(A) = \nu(x : x - e \in A).$$

Then

$$A_t^\nu(\varphi(x)) = A_t^{\nu_{x-x_0}}(\varphi.(x) - x + x_0).$$

Note that for fixed x and x_0 the process $(\xi_t)_{t\geq 0} := (\varphi_t(x) - x + x_0)_{t\geq 0}$ can be considered as a Markov process starting from x_0, and its distribution is equivalent to the distribution $\mathbb{P}_{w(x_0)}$ of the Wiener process starting from x_0. Indeed,

$$\xi_t = x_0 + \int_0^t \tilde{a}(\xi_s)ds + w(t),$$

where $\tilde{a}(y) = a(y + x - x_0)$. Similarly to the proof of Lemma 5 it can be checked that the family of the Radon-Nikodym densities $\left\{\frac{d\mathbb{P}_{\varphi.(x)-x+x_0}}{d\mathbb{P}_{w(x_0)}}, \ x \in \mathbb{R}^d\right\}$ are uniformly integrable with respect to $P_{w(x_0)}$. By Proposition 3.1 and Lemma 3.5 to prove (5.5) it suffices to verify that

$$A_t^{\nu_{x-x_0}}(w(x_0)) \to A_t^\nu(w(x_0)), \ x \to x_0, \text{ in probability } \mathbb{P}. \tag{5.6}$$

By $\nu^{(R)}(dy) = \mathbb{1}_{|y|\leq R}\nu(dy)$ denote the restriction of the measure ν to the ball $\{y : |y| \leq R\}$. Put $f_t^w(y) = \mathbb{E}A_t^\nu(w(y))$, $f_{R,t}^w(y) = \mathbb{E}A_t^{\nu^{(R)}}(w(y))$. Then

$$\mathbb{E}A_t^{\nu_{x-x_0}}(w(y)) = f_t^w(y + x - x_0), \ \ \mathbb{E}A_t^{\nu_{x-x_0}^{(R)}}(w(y)) = f_{R,t}^w(y + x - x_0).$$

It is easy to see that the function $(s,y) \to f_{R,t}^w(y)$ is uniformly continuous in $(s,y) \in [0,t] \times \mathbb{R}^d$. So by Theorem 1.5 we have the convergence in probability

$$A_t^{\nu_{x-x_0}^{(R)}}(w(y)) \to A_t^{\nu^{(R)}}(w(y)), \ x \to x_0, \tag{5.7}$$

for any $y \in \mathbb{R}^d$.

It follows from [8], Theorem 8.4 that for any $R > 0$ and $y \in \mathbb{R}^d$ we have the equality $A_t^{\nu^{(R)}}(w(y)) = A_t^\nu(w(y))$ a.s. on the set $\{\sup_{s\in[0,t]} |y + w_s| < R\}$. This together with (5.7) entails (5.6). □

6 Appendix: The proof of Lemma 1.17

Note that $B_t^h(\varphi(x)) = \int_0^t \frac{f_h^w(\varphi_s(x))}{h} ds$ is a W-functional. Let us estimate its characteristic.

$$\mathbb{E} B_t^h(\varphi(x)) = \mathbb{E} \int_0^t \frac{f_h^w(\varphi_s(x))}{h} ds = \frac{1}{h} \int_0^h du \int_{\mathbb{R}^d} \left(\int_0^t ds \int_{\mathbb{R}^d} p_u^w(z,y) p_s^\varphi(x,z) dz \right) \nu(dy).$$

From the estimates (1.4) we obtain (see also the proof of Theorem 6.6. in [8])

$$\mathbb{E} B_t^h(\varphi(x)) \leq$$
$$\frac{1}{h} \int_0^h du \int_{\mathbb{R}^d} \left(\int_0^t ds \int_{\mathbb{R}^d} \frac{K}{u^{d/2}} \exp\left\{ -\frac{k\|y-z\|^2}{u} \right\} \frac{K}{s^{d/2}} \exp\left\{ -\frac{k\|z-x\|^2}{s} \right\} dz \right) \nu(dy) =$$
$$\widetilde{K} \frac{1}{h} \int_0^h du \int_{\mathbb{R}^d} \left(\int_0^t \frac{1}{(2\pi(u+s))^{d/2}} \exp\left\{ -\frac{k\|y-x\|^2}{u+s} \right\} ds \right) \nu(dy) =$$
$$\widetilde{K} \frac{1}{h} \int_0^h du \int_{\mathbb{R}^d} \left(\int_u^{t+u} \frac{1}{(2\pi s)^{d/2}} \exp\left\{ -\frac{k\|y-x\|^2}{s} \right\} ds \right) \nu(dy) =$$
$$\widehat{K} \frac{1}{h} \int_0^h du \int_{\mathbb{R}^d} \left(\int_{u/2k}^{(t+u)/2k} \frac{1}{(2\pi s)^{d/2}} \exp\left\{ -\frac{\|y-x\|^2}{2s} \right\} ds \right) \nu(dy) =$$
$$\widehat{K} \frac{1}{h} \int_0^h \left(f_{(t+u)/2k}^w(x) - f_{u/2k}^w(x) \right) du.$$

where $\widetilde{K} = K^2 \pi^2 (2/k)^{d/2}$, $\widehat{K} = 2K^2 k^{1-d} \pi^d$. Taking into account (1.1), we get

$$f_{(t+u)/2k}^w(x) - f_{u/2k}^w(x) = T_{u/2k}^w f_{t/2k}^w(x) \leq \|f_{t/2k}^w\|_\infty.$$

By Proposition 1.15,

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}_x \left(B_t^h(\varphi) \right)^2 \leq 2\widehat{K}^2 \left(\|f_{t/2k}^w\|_\infty \right)^2. \tag{6.1}$$

Therefore, the second moment of $B_t^h(\varphi)$ is bounded uniformly in h. This implies the uniform integrability and, consequently the convergence in L_1 holds in (1.10). Then the characteristic of the functional $A_t^{\nu,w}(\varphi(x))$ is equal to

$$\widetilde{f}_t(x) = \lim_{h \downarrow 0} \mathbb{E} \int_0^t \frac{f_h^w(\varphi_s(x))}{h} ds.$$

If we show that

$$\widetilde{f}_t(x) = \int_{\mathbb{R}^d} k_t^\varphi(x,y) \nu(dy), \tag{6.2}$$

then the statement of the Lemma follows from Proposition 1.4. We have, for each $0 < \delta < t$,

$$\left| \mathbb{E} \int_0^t \frac{f_h^w(\varphi_s(x))}{h} ds - \int_{\mathbb{R}^d} k_t^\varphi(x,y) \nu(dy) \right| \leq$$
$$\mathbb{E} \int_0^\delta \frac{f_h^w(\varphi_s(x))}{h} ds + \int_0^\delta \left(\int_{\mathbb{R}^d} p_s^\varphi(x,y) \nu(dy) \right) ds +$$
$$\left| \mathbb{E} \int_\delta^t \frac{f_h^w(\varphi_s(x))}{h} ds - \int_\delta^t \left(\int_{\mathbb{R}^d} p_s^\varphi(x,y) \nu(dy) \right) ds \right| = I + II + III.$$

Consider I. Arguing as in the proof of (6.1) we arrive at the inequality

$$I \leq \|f_{\delta/2k}^w\|_\infty.$$

Making use of (1.4) and changing the variables we get

$$II \leq 2K\pi^{d/2}(k)^{1-d/2} \int_0^{\delta/2k} ds \int_{\mathbb{R}^d} \frac{1}{(2\pi s)^{d/2}} \exp\left\{-\frac{\|y-x\|^2}{2s}\right\} \nu(dy)$$
$$\leq 2K\pi^{d/2}(k)^{1-d/2}\|f_{\delta/2k}^w\|_\infty.$$

For each $\varepsilon > 0$, the condition (1.7) allows us to choose δ so small that

$$I < \varepsilon/3, \ II < \varepsilon/3. \tag{6.3}$$

Further,

$$III = \left|\int_\delta^t ds \int_{\mathbb{R}^d} \nu(dy) \int_{\mathbb{R}^d} (p_s^\varphi(x,z) - p_s^\varphi(x,y))\left(\frac{1}{h}\int_0^h p_u^w(z,y)du\right) dz\right|.$$

The measure $\frac{1}{h}\left(\int_0^h p_u^w(z,y)du\right) dz$ converges weakly to the δ-measure at the point y. The function $p_s^\varphi(x,y)$ is equicontinuous in y for $s \in [\delta,t]$, $x \in \mathbb{R}^d$. So

$$\int_{\mathbb{R}^d} (p_s^\varphi(x,z) - p_s^\varphi(x,y))\left(\frac{1}{h}\int_0^h p_u^w(z,y)du\right) dz \to 0, \ h \downarrow 0,$$

uniformly in x and s. Besides, from (1.4)

$$\left|\int_{\mathbb{R}^d} (p_s^\varphi(x,z) - p_s^\varphi(x,y))\left(\frac{1}{h}\int_0^h p_u^w(z,y)du\right) dz\right| \leq$$
$$\int_{\mathbb{R}^d} \frac{K}{s^{d/2}}\left(\exp\left\{-\frac{k\|x-z\|^2}{s}\right\} + \exp\left\{-\frac{k\|x-y\|^2}{s}\right\}\right)\left(\frac{1}{h}\int_0^h p_u^w(z,y)du\right) dz \leq \frac{2K}{s^{d/2}}.$$

By the dominated convergence theorem,

$$III \to 0 \text{ as } h \downarrow 0. \tag{6.4}$$

Now the equality (6.2) follows from (6.3) and (6.4). The Lemma is proved.

References

[1] M. Aizenman and B. Simon, *Brownian motion and Harnack inequality for Schrödinger operators*, Communications on Pure and Applied Mathematics **35** (1982), no. 2, 209–273. MR-0644024

[2] D. G. Aronson, *Bounds for the fundamental solution of a parabolic equation*, Bull. Amer. Math. Soc. **73** (1967), 890–896. MR-0217444

[3] O. V. Aryasova and A. Yu. Pilipenko, *On properties of a flow generated by an SDE with discontinuous drift*, Electron. J. Probab. **17** (2012), no. 106, 1–20. MR-3015690

[4] S. Attanasio, *Stochastic flows of diffeomorphisms for one-dimensional SDE with discontinuous drift*, Electron. Commun. Probab. **15** (2010), no. 20, 213–226. MR-2653726

[5] R. F. Bass and Z.-Q. Chen, *Brownian motion with singular drift*, The Annals of Probability **31** (2003), no. 2, 791–817. MR-1964949

[6] R. M. Blumenthal and R. K. Getoor, *Markov Processes and Potential Theory. Reprint of the 1968 ed.*, Mineola, NY: Dover Publications. vi, 313 p., 2007 (English). MR-0264757

[7] V. I. Bogachev, *Measure theory*, vol. 2, Springer, Berlin, 2007. MR-2267655

[8] E. B. Dynkin, *Markov processes*, Fizmatlit, Moscow, 1963, [Translated from the Russian to the English by J. Fabius, V. Greenberg, A. Maitra, and G. Majone. Academic Press, New York; Springer, Berlin, 1965. vol. 1, xii + 365 pp.; vol. 2, viii + 274 pp.]. MR-0193671

[9] Federer H., *Geometric measure theory*, Springer-Verlag New York Inc. ed., Die Grundlehren der mathematischen Wissenschaften, vol. 153, New York, 1969. MR-0257325

[10] E. Fedrizzi and F. Flandoli, *Hölder flow and differentiability for SDEs with nonregular drift*, Stochastic Analysis and Applications **31** (2013), no. 4, 708–736. MR-3175794

[11] E. Fedrizzi and F. Flandoli, *Noise prevents singularities in linear transport equations*, Journal of Functional Analysis **264** (2013), no. 6, 1329 – 1354. MR-3017266

[12] F. Flandoli, M. Gubinelli, and E. Priola, *Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift*, Bulletin des Sciences Mathematiques **134** (2010), no. 4, 405 – 422. MR-2651899

[13] I. I. Gikhman and A. V. Skorokhod, *The theory of stochastic processes. ii*, Nauka, Moscow, 1973, [Translated from the Russian by S. Kotz. Corrected printing of the first edition. Berlin: Springer, 2004. viii, 441 p.]. MR-2058260

[14] R. Z. Khas'minskii, *On positive solutions of the equation $\mathfrak{A}u + vu = 0$*, Theory of Probability and Its Applications **4** (1959), no. 3, 309–318. MR-0123373

[15] A. M. Kulik and A. Yu. Pilipenko, *Nonlinear transformations of smooth measures on infinite-dimensional spaces*, Ukrainian Mathematical Journal **52** (2000), 1403–1431, 10.1023/A:1010380119199. MR-1816936

[16] K. Kuwae and M. Takahashi, *Kato class measures of symmetric Markov processes under heat kernel estimates*, Journal of Functional Analysis **250** (2007), no. 1, 86 – 113. MR-2345907

[17] R. S. Liptser and A. N. Shiryayev, *Statistics of random processes*, Nauka, Moscow, 1974, [Translated from the Russian to the English by A. B. Aries. Springer-Verlag, New York, 1977]. MR-0474486

[18] D. Luo, *Absolute continuity under flows generated by SDE with measurable drift coefficients*, Stochastic Processes and their Applications **121** (2011), no. 10, 2393 – 2415. MR-2822781

[19] T. Meyer-Brandis and F. Proske, *Construction of strong solutions of SDE's via Malliavin calculus*, Journal of Functional Analysis **258** (2010), no. 11, 3922 – 3953. MR-2606880

[20] S.E.A. Mohammed, T. Nilssen, and F. Proske, *Sobolev differentiable stochastic flows of SDE's with measurable drift and applications*, (2012), arXiv:1204.3867.

[21] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Springer-Verlag, Berlin, 1999. MR-1725357

[22] A.-S. Sznitman, *Brownian Motion, Obstacles and Random Media.*, Springer Monographs in Mathematics, Berlin: Springer, 1998 (English). MR-1717054

[23] A. Y. Veretennikov, *On strong solutions and explicit formulas for solutions of stochastic integral equations*, Math. USSR Sborn **39(3)** (1981), 387–403. MR-0568986

[24] V. S. Vladimirov, *The equation of mathematical phisics*, Nauka, Moscow, 1967, [Translated from the Russian to the English by A. Littlewood. Marcel Dekker, INC., New York, 1971.].

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS[1])

- Non profit, sponsored by IMS[2], BS[3], PKP[4]

- Purely electronic and secure (LOCKSS[5])

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]OJS: Open Journal Systems http://pkp.sfu.ca/ojs/
[2]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[3]BS: Bernoulli Society http://www.bernoulli-society.org/
[4]PK: Public Knowledge Project http://pkp.sfu.ca/
[5]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm