Theory of Stochastic Processes
Vol. 12 (28), no. 1–2, 2006, pp. 127–141

UDC 519.21

A. YU. PILIPENKO

SUPPORT THEOREM ON STOCHASTIC FLOWS WITH INTERACTION

We prove an analogue of the Stroock–Varadhan theorem for stochastic flows describing a motion of interacting particles in a random media. A version of the Itô lemma for functions on a measure-valued process is obtained.

Introduction

Let us consider a flow of interacting particles in a random media. Denote by $x_t(u)$ a position at an instant of time t of a particle starting from a point $u \in \mathbb{R}^d$. Let μ_t be a distribution of the mass of particles at the moment t. Suppose that each particle does not change its mass in time. So μ_t is the image of the measure μ_0 under the mapping x_t, i.e. $\mu_t = \mu_0 \circ x_t^{-1}$.

Assume that the motion of each particle depends not only on its position at the current time moment but also on the distribution of the total mass of particles and satisfies the following system of stochastic equations:

$$
\begin{cases} dx_t(u) = a(x_t(u), \mu_t)dt + \sum_{k=1}^m b_k(x_t(u), \mu_t)dw_k(t), \\ x_0(u) = u, \ u \in \mathbb{R}^d, \\ \mu_t = \mu_0 \circ x_t^{-1}, \ t \in [0, T], \end{cases} \tag{0.1}
$$

where $a, b_k : \mathbb{R}^d \times \mathcal{P} \to \mathbb{R}^d$, $\mu_0 \in \mathcal{P}$; $\{w_k(t)\}_{k=\overline{1,m}}$ are independent one-dimensional Wiener processes. Here, $\mathcal{P}$ is the space of probability measures on $\mathbb{R}^d$ with a topology of weak convergence.

Systems of stochastic equations of the type (0.1) for flows of interacting particles were introduced by A.A. Dorogovtsev and P. Kotelenez [1]. Measure-valued processes which are solutions of (0.1) were initially obtained in [2] as some weak limits of the finite number of systems of interacting particles. However, there was not considered any flow describing the individual behavior of particles for the limit process.

The theorem on uniqueness and existence can be proved under some natural conditions on the coefficients of (0.1) [3]. Moreover, if a, b_k are smooth enough, then a process $x_t(u)$ possesses a modification which is differentiable with respect to u (the corresponding number of times), and its derivatives $\frac{\partial^k x_t(u)}{\partial u^k}$ are continuous in (t, u). So, we can consider the process $x_t(\cdot), t \geq 0$ as a continuous stochastic process with values in $C^k(\mathbb{R}^d, \mathbb{R}^d)$. Here, a space $C^k(\mathbb{R}^d, \mathbb{R}^d)$ of k times continuously differentiable functions is provided by the topology of uniform convergence of functions and their derivatives on compact sets. On the other hand, $x = x_t(u)$ can be considered as a random element in the space $\mathcal{C}_k = C([0, T]; C^k(\mathbb{R}^d, \mathbb{R}^d))$. The aim of this article is a characterization of the support of a stochastic flow $x_t(u)$.

2000 *AMS Mathematics Subject Classification.* Primary 60K35, 60G57.

Key words and phrases. Support of stochastic flow, systems of interacting particles.

The corresponding result for the usual stochastic differential equation

$$d\xi_t = a(\xi_t)dt + \sum_{k=1}^{m} b_k(\xi_t) \circ dw_k(t)$$

is given by the well-known Stroock—Varadhan theorem [4]. It asserts that, under some smoothness assumptions on the coefficients of Eq. (0.1), the support of $\xi(t)$ distribution is equal to the closure in $C([0,T],\mathbb{R}^d)$ of the set

$$\{x^\psi, \psi \text{ is piecewise smooth}\},$$

where x^ψ is a solution of the non-random equation

$$dx^\psi(t) = a(x^\psi(t))dt + \sum_{k=1}^{m} b_k(x^\psi(t))d\psi(t).$$

A similar result for the stochastic flow $x_t(u)$ generated by s.d.e. depending on the initial condition

$$\text{(0.2)} \qquad \begin{cases} dx_t(u) = a(x_t(u))dt + \sum_k b_k(x_t(u)) \circ dw_k(t), \\ x_0(u) = u. \end{cases}$$

was obtained by H.Kunita [5].

Stochastic flows generated by the equation

$$\text{(0.3)} \quad dx_t(u) = \int_{\mathbb{R}^d} \alpha(x_t(u), x_t(v))\mu(dv)dt + \sum_{k=1}^{m} \left(\int_{\mathbb{R}^d} \beta_k(x_t(u), x_t(v))\mu(dv) \right) \circ dw_k(t)$$

with interaction were considered in [6, 7].

Note that Eq. (0.3) is a partial case of Eq. (0.1) with $a(u,\nu) = \int_{\mathbb{R}^d} \alpha(u,v)\nu(dv)$, $b_k(u,\nu) = \int_{\mathbb{R}^d} \beta_k(u,v)\nu(dv)$.

Usually, it is convenient to formulate a Stroock—Varadhan-type theorem for stochastic equations written in the Stratonovich form. Moreover, it is always supposed that the coefficients of equations are differentiable more than one time. The coefficients of (0.1) depends on the measure-valued process μ_t. So we need some version of the Itô formula for functions depending on a measure-valued process. We prove the corresponding result in §2.

The rest of the proof of the Stroock—Varadhan theorem for a stochastic flow $x_t(u)$ is quite standard.

We estimate the "small-balls" probability in §3 and prove that, for each $\varepsilon > 0$ and a piecewise smooth function ψ, the conditional probability

$$P\left(\sup_{|u|<R} \sup_{t\in[0,T]} |\nabla(x_t(u) - x_t^\psi(u))| / \sup_{t\in[0;T]} |w(t) - \psi(t)| \right)$$

converges to zero as $\delta \to 0+$. This gives us the inclusion

$$\{x^\psi, \psi \text{ is piecewise smooth}\} \subset \operatorname{supp} P_x.$$

The support is a closed set, so it contains the closure of $\{x^\psi, \psi \textit{ is piecewise smooth}\}$,

Section 4 is devoted to the proof of the approximation theorem. Particularly, we prove that if a sequence of processes $\{\xi_k^\varepsilon(t),\ \varepsilon > 0\}$ is such that $w_k^\varepsilon(t) = \int_0^t \xi_k^\varepsilon(s)ds$ converges as $\varepsilon \to 0$ to a Wiener process $w_k(t)$ in some sense, then a solution of Eq. (0.1) with $w^\varepsilon(t)$ instead of $w(t)$ converges to the solution of (0.1), but may be with corrected coefficients.

It is worth to mention one principal difference between Eqs. (0.1) and (0.2).

While studying Eq. (0.2), we can consider different starting points independently. Then the differentiability with respect to u is just a differentiability with respect to a parameter. However, all the equations in (0.1) are coupled, and (0.1) should be considered as an infinite (even uncountable) system of stochastic equations or as a stochastic equation in the functional space.

1. Preliminaries

Let $\mathcal{P}$ be a set of probability measures on $\mathbb{R}^d$ with a topology of weak convergence.

Definition 1.1. A pair (x_t, μ_t) is said to be a solution to Eq. (0.1) if

1) mapping $x = x_t(u, \omega) : [0, T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is measurable in a triple of its arguments;

2) processes $x_t(u), \mu_t$ are adapted to a filtration generated by the Wiener processes $w_k(t), k = \overline{1, m}$;

3) for each $u \in \mathbb{R}^d$ with probability one, the equality

$$x_t(u) = u + \int_0^t a(x_s(u), \mu_s)ds + \sum_{k=1}^m \int_0^t b_k(x_s(u), \mu_s)dw_k(s), \ t \in [0, T],$$

where $\mu_t = \mu_0 \circ x_t^{-1}$ is the image of the initial measure μ_0 under the random mapping x_t, is satisfied.

Let us introduce the Wasserstein metric on the space $\mathcal{P}$:

$$\gamma(\mu_1, \mu_2) := \inf_{\varkappa \in Q(\mu_1, \mu_2)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |u - v| \wedge 1 \varkappa(du, dv),$$

where $Q(\mu_1, \mu_2)$ is a set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ_1 and μ_2. It is well known that the metric space $(\mathcal{P}, \gamma)$ is complete and separable [8], and the convergence in metric γ is equivalent to the weak convergence of measures.

Let us formulate the theorem of existence and uniqueness for the solution of Eq. (0.1).

Theorem 1.1 [3]**.** *Assume that*
$\exists L > 0 \ \forall u_1, u_2 \in \mathbb{R}^d \ \forall \mu_1, \mu_2 \in \mathcal{P}$:

$$\begin{aligned} |a(u_1, \mu_1) - a(u_2, \mu_2)| + \sum_{k=1}^m |b_k(u_1, \mu_1) - b_k(u_2, \mu_2)| \le \\ \le L(|u_1 - u_2| + \gamma(\mu_1, \mu_2)). \end{aligned} \tag{1.1}$$

Then there exists a unique solution for Eq. (0.1).

Further, we always assume that the conditions of Theorem 1.1 are satisfied.

Put $\widetilde{a}(x, t) := a(x, \mu_t(\omega)), \ \widetilde{b}_k(x, t) := b_k(x, \mu_t(\omega))$. Observe that if x_t, μ_t satisfy (0.1) then x_t satisfies the following Itô equation:

$$\begin{cases} dx_t(u) = \widetilde{a}(x_t(u), t)dt + \sum_{k=1}^m \widetilde{b}_k(x_t(u), t)dw_k(t), t \ge 0, \\ x_0(u) = u. \end{cases} \tag{1.2}$$

The functions $\widetilde{a}, \widetilde{b}_k$ are continuous and satisfy the Lipschitz condition in x. So, the flow $x_t(u)$ has a modification which is continuous in (t, u) (cf.[5]). Moreover, if the coefficients a, b_k are $(n+1)$ times continuously differentiable with respect to u and have bounded derivatives, then the partial derivatives $\frac{\partial^j x_t(u)}{\partial u^j}, j = \overline{1, n}$ exist and are continuous in (t, u).

Therefore, we can consider the process $x_t(\cdot)$ as a continuous process with values in $C^n(\mathbb{R}^d, \mathbb{R}^d)$ or as a random element in $\mathcal{C}_n = C([0, T], C^n(\mathbb{R}^d, \mathbb{R}^d))$, where the space of functions $C^n(\mathbb{R}^d, \mathbb{R}^d)$ differentiable n-times has a topology of uniform convergence of the functions and their derivatives on compact sets.

Usually, to study the support of a solution for some stochastic equation, one needs a stronger condition of smoothness than the Lipschitz condition. That's why we need the following definition of a derivative for functions depending on a measure-valued argument [9].

Let $\mathfrak{M}$ be a space of finite measures on $\mathbb{R}^d$ with a topology of weak convergence, and let F be a continuous function from $\mathfrak{M}$ to $\mathbb{R}$.

Definition 1.2. Assume that, for each $\mu \in \mathfrak{M},\ u \in \mathbb{R}^d$, there exists the limit

$$\frac{\delta F(\mu)}{\delta\mu}(u) := \lim_{\varepsilon\to 0+} \frac{F(\mu + \varepsilon\delta_u) - F(\mu)}{\varepsilon} \tag{1.3}$$

which is continuous in (μ, u), where δ_u is a unit measure concentrated at the point $u \in \mathbb{R}^d$.

Then function F is said to be continuous differentiable, and $\frac{\delta F}{\delta\mu}$ is its derivative.

Higher order derivatives $\frac{\delta^n F}{\delta\mu^n}(u_1, \ldots, u_n)$ are defined iteratively.

Remark 1. Let μ_t be a process from (0.1). Then the full mass of μ_t is not changing in time, $\mu_t(\mathbb{R}^d) = \mu_0(\mathbb{R}^d) = 1$. So, in general, in order to consider Eq. (0.1), it is natural to know the coefficients in the space $\mathbb{R}^d \times \mathcal{P}$ only, but not in the larger space $\mathbb{R}^d \times \mathfrak{M}$. Note also that if $\mu \in \mathcal{P}$, then $\mu + \varepsilon\delta_u \notin \mathcal{P}$ as $\varepsilon \neq 0$. We can replace the definition of continuous differentiability (1.3) remaining in a class $\mathcal{P}$ by using Newton—Leibnitz formula:

$\exists \frac{\delta F}{\delta\mu} \in C(\mathcal{P} \times \mathbb{R}^d)\ \forall \mu_1, \mu_2 \in \mathcal{P}:$

$$F(\mu_2) - F(\mu_1) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta F(\tau\mu_2 + (1-\tau)\mu_1)}{\delta\mu}(u)(\mu_2 - \mu_1)(du)d\tau. \tag{1.4}$$

The proof of (1.4) can be done for discrete measures at first. Then it is not difficult to extend (1.4) for an arbitrary measure by continuity (recall that $\frac{\delta F}{\delta\mu}$ is continuous in (μ, u)). However, formula (1.4) seems to be less natural than (1.3). Moreover, if the function F is defined on $\mathcal{P}$ and satisfies (1.4) for all $\mu_1, \mu_2 \in \mathcal{P}$, then its extension $\widetilde{F}(\mu) := F\left(\frac{\mu}{\mu(\mathbb{R}^d)}\right)$ to the space of all finite measures $\mathfrak{M}$ is continuous differentiable in the sense of Definition 1.2.

The following statement will be useful in our investigations.

Lemma 1.1. *Assume that $F : \mathcal{P} \to \mathbb{R}$ satisfies (1.4), where the function $\frac{\delta F}{\delta\mu}$ is bounded and satisfies the Lipschitz condition in u:*

$$\exists L_1\ \forall \mu \in \mathcal{P}\ \forall u \in \mathbb{R}^d : \quad \left|\frac{\delta F(\mu)}{\delta\mu}(u)\right| \le L_1;$$

$$\exists L_2\ \forall \mu \in \mathcal{P}\ \forall u_1, u_2 \in \mathbb{R}^d : \quad \left|\frac{\delta F(\mu)}{\delta\mu}(u_1) - \frac{\delta F(\mu)}{\delta\mu}(u_2)\right| \le L_2|u_1 - u_2|.$$

Let $\nu \in \mathcal{P}$ be a probability measure. Then there exists a constant $L = L(F)$ such that, for all measurable functions $f_1, f_2 : \mathbb{R}^d \to \mathbb{R}^d$, the following inequality holds:

$$|F(\nu_2) - F(\nu_1)| \le L \int_{\mathbb{R}^d} |f_1(u) - f_2(u)| \wedge 1 \nu(du), \tag{1.5}$$

where $\nu_i = \nu \circ f_i^{-1},\ i = 1, 2.$

Proof. Put $\mu_\tau = \tau\nu_2 + (1-\tau)\nu_1$. Then, due to (1.4), we have

$$|F(\nu_2) - F(\nu_1)| \le \int_0^1 \left| \int_{\mathbb{R}^d} \frac{\delta F(\mu_\tau)}{\delta\mu}(u)(\nu_2(du) - \nu_1(du)) \right| d\tau =$$

$$= \int_0^1 \left| \int_{\mathbb{R}^d} \left(\frac{\delta F(\mu_\tau)}{\delta \mu}(f_2(u)) - \frac{\delta F(\mu_\tau)}{\delta \mu}(f_1(u)) \right) \nu(du) \right| d\tau \le$$

$$\le \int_0^1 \int_{\mathbb{R}^d} (L_2 |f_2(u) - f_1(u)|) \wedge 2L_1 \nu(du) d\tau \le$$

$$\le (L_2 + 2L_1) \int_{\mathbb{R}^d} |f_2(u) - f_1(u)| \wedge 1 \nu(du).$$

Lemma 1.1 is proved.

Remark 2. While studying carefully the proof of Theorem 1.1 (see [3]), one can notice that it is sufficient to demand a condition of type (1.5) instead of (1.1). So, the existence of the bounded derivatives $\frac{\partial a(x,\mu)}{\partial x}$, $\frac{\partial}{\partial u}\frac{\delta a(x,\mu)}{\delta \mu}(u)$, $\frac{\partial b_k(x,\mu)}{\partial x}$, and $\frac{\partial}{\partial u}\frac{\delta b_k(x,\mu)}{\delta \mu}(u)$ is enough for Theorem 1.1. This condition will always be supposed in §2, §3. However, it seems to the author that the formulation of Theorem 1.1 is methodologically more useful than a similar theorem with the condition

$$\exists L > 0 \ \forall x, u \in \mathbb{R}^d \ \forall \mu :$$

$$\left| \frac{\partial a(x,\mu)}{\partial x} \right| + \left| \frac{\partial}{\partial u} \frac{\delta a(x,\mu)}{\delta \mu}(u) \right| + \sum_{k=1}^m \left(\left| \frac{\partial b_k(x,\mu)}{\partial x} \right| + \left| \frac{\partial}{\partial u} \frac{\delta b_k(x,\mu)}{\delta \mu}(u) \right| \right) \le L$$

instead of (1.1).

2. Itô formula for functions of measure-valued argument

In this section, we prove a version of the Itô formula for a process $F(\xi_t, \mu_t)$, where μ_t is defined in (0.1), and the process ξ_t has a stochastic differential

$$d\xi_t = \alpha_t dt + \sum_{j=1}^m \beta_t^j dw_j(t). \tag{2.1}$$

The main result is as follows:

Theorem 2.1. *Assume that a function $F : \mathbb{R}^k \times \mathfrak{M} \to \mathbb{R}$ is such that the following derivatives exist, bounded, and continuous in all arguments:*

$$\frac{\partial F(x,\mu)}{\partial x}, \ \frac{\partial^2 F(x,\mu)}{\partial x^2}, \ \frac{\delta F(x,\mu)}{\delta \mu}(u), \ \frac{\delta^2 F(x,\mu)}{\delta \mu^2}(u,v),$$

$$\frac{\partial}{\partial u}\frac{\delta F(x,\mu)}{\delta \mu}(u), \ \frac{\partial^2}{\partial u^2}\frac{\delta F(x,\mu)}{\delta \mu}(u), \ \frac{\partial^2}{\partial x \partial u}\frac{\delta F(x,\mu)}{\delta \mu}(u), \ \frac{\partial^2}{\partial u \partial v}\frac{\delta^2 F(x,\mu)}{\delta \mu^2}(u,v).$$

Suppose that a process $\xi_t = (\xi_t^1, \dots, \xi_t^k)$ has the stochastic differential (2.1), and the functions a, b_k satisfy the assumptions of Theorem 1.1 and are bounded.

Let us introduce the following notations:

$$\langle f, \mu \rangle = \int f d\mu, \ \sigma_\xi^2(t) = \sum_{l=1}^m \beta_t^l \, (\beta_t^l)^*,$$

$$\sigma_{\xi\mu}^2(t,x,\mu) = \sum_{l=1}^m \beta_t^l b_l^*(x,\mu), \ \sigma_{\mu\mu}^2(x,y,\mu) = \sum_{l=1}^m b_l(x,\mu) b_l^*(y,\mu),$$

$$\sigma_\mu^2(x,\mu) = \sigma_{\mu\mu}^2(x,x,\mu),$$

where * is the transposition operator of a matrix.

Set

$$L_1(t)F(x,\mu) = F_x'(x,\mu)\alpha_t + \frac{1}{2}\mathrm{sp}(F_{xx}''(x,\mu)\sigma_{\xi,\xi}^2(t)),$$

$$L_2F(x,\mu) = \int_{\mathbb{R}^d} \left(\frac{\partial}{\partial u} \frac{\delta F(x,\mu)}{\delta \mu}(u) a(u,\mu) + \frac{1}{2}\mathrm{sp}\frac{\partial^2}{\partial u^2}\frac{\delta F(x,\mu)}{\delta\mu}(u)\sigma^2_{\mu}(u,\mu)\right)\mu(du)+$$

$$+\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{sp}\frac{\partial^2}{\partial u \partial v}\frac{\delta^2 F(x,\mu)}{\delta \mu^2}(u,v)\sigma^2_{\mu\mu}(u,v,\mu)\mu(du)\mu(dv). \tag{2.2}$$

Then

$$dF(\xi_t,\mu_t) = L_1(t)F(\xi_t,\mu_t)dt + F_1'(\xi_t,\mu_t)\sum_{l=1}^m \beta_t^l dw_l(t)+$$

$$+L_2F(\xi_t,\mu_t)dt + \sum_{l=1}^m \int_{\mathbb{R}^d}\frac{\partial}{\partial u}\frac{\delta F(\xi_t,\mu_t)}{\delta\mu}(u) b_l(u,\mu_t)\mu_t(du)dw_l(t)+$$

$$+\frac{1}{2}\mathrm{sp}\int_{\mathbb{R}^d}\frac{\partial^2}{\partial\xi\partial u}\frac{\delta F(\xi_t,\mu_t)}{\delta\mu}(u)\sigma^2_{\xi\mu}(t,u,\mu_t)\mu(du)dt. \tag{2.3}$$

Proof. Let us verify formula (2.3) for a function $F(x,\mu) = F(\mu)$ depending on the second argument only. The general case can be proved similarly but with additional routine calculations.

Let

$$\mu^n = \sum_{j=1}^n c_{j,n}\delta_{u_{j,n}},\ n\in\mathbb{N},\ c_{j,n}\ge 0,\ \sum_j c_{j,n} = 1$$

be a sequence of discrete measures which converges weakly to μ_0 as $n\to\infty$.

Denote, by $x_t^n(u)$, a solution of Eq. (0.1) with initial measure μ^n instead of μ_0, $\mu_t^n := \mu^n \circ (x_t^n)^{-1}$. The plan of the proof is to obtain the Itô formula for μ_t^n at first, and then to pass to a limit as $n\to\infty$.

Remark 3. The Itô formula for superprocesses with interaction was obtained in [10]. There, a function $F(\mu)$ was approximated by polynomials in μ, and the proof is a hard technical work. It is worth mentioning that the process μ_t is neither a superprocess nor the Fleming—Viot process. This fact can be easily checked if we write down and compare their generators [3,9].

Observe that the measure μ_t^n is also discrete, $\mu_t^n = \sum_{j=1}^n c_{j,n}\delta_{x_t^n(u_{j,n})}$. To simplify notations, we will write c_j, u_j instead of $c_{j,n}, u_{j,n}$.

The processes $x_t^n(u_k), k=\overline{1,n}$ satisfy a finite system of stochastic equations

$$dx_t^n(u_k) = a(x_t^n(u_k), \sum_{j=1}^n c_j\delta_{x_t^n(u_j)})dt+$$

$$+\sum_{l=1}^m b_l(x_t^n(u_k), \sum_{j=1}^n c_j\delta_{x_t^n(u_j)})dw_l(t), k=\overline{1,n}.$$

Consider a function $f:\mathbb{R}^{nd}\to\mathbb{R}$ defined as follows:

$$f(v_1,\dots,v_n) = F(\sum_{j=1}^n c_j\delta_{v_j}).$$

Then $F(\mu_t^n) = f(x_t^n(u_1),\dots,x_t^n(u_n))$. Let us check that f is twice continuously differentiable.

Assume for simplicity that $n=2$ and calculate $\langle\frac{\partial f}{\partial v_1}, e\rangle_{\mathbb{R}^d}$, where $e\in\mathbb{R}^d$ (recall that $v_1\in\mathbb{R}^d$).

Set $\nu = c_1\delta_{v_1} + c_2\delta_{v_2}$, $\nu_\tau = (1-\tau)(c_1\delta_{v_1} + c_2\delta_{v_2}) + \tau(c_1\delta_{v_1+\varepsilon e} + c_2\delta_{v_2})$. Formula (1.4) implies

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{f(v_1+\varepsilon e, v_2) - f(v_1,v_2)}{\varepsilon} \\
&\quad = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^1\Bigg(\int_{\mathbb{R}^d}\frac{\delta F(\nu_\tau)}{\delta\mu}(u)(c_1\delta_{v_1+\varepsilon e} + c_2\delta_{v_2})(du) \\
&\qquad\qquad\qquad - \int_{\mathbb{R}^d}\frac{\delta F(\nu_\tau)}{\delta\mu}(u)(c_1\delta_{v_1} + c_2\delta_{v_2})(du)\Bigg) \\
&\quad = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^1 c_1\left(\frac{\delta F(\nu_\tau)}{\delta\mu}(v_1+\varepsilon e) - \frac{\delta F(\nu_\tau)}{\delta\mu}(v_1)\right)d\tau \\
&\quad = c_1\left(\frac{\partial}{\partial v_1}\frac{\delta F(\nu)}{\delta\mu}, e\right)_{\mathbb{R}^d}.
\end{aligned}$$

Here, while grounding the passage to the limit, we use the continuous differentiability of $\frac{\delta F(\mu)}{\delta\mu}(u)$ with respect to the parameter u and the boundedness of its derivative.

So, we have verified that

$$\frac{\delta f}{\partial v_j} = c_j\nabla\frac{\delta F(\nu)}{\delta\mu}(v_j), \tag{2.4}$$

where ∇ is the derivative with respect to the argument v_j.

Analogously,

$$\frac{\partial^2 f}{\partial v_i\partial v_j} = c_ic_j\nabla_1\nabla_2\frac{\delta^2 F(\nu)}{\delta\mu^2}(v_i,v_j) + c_j\nabla^2\frac{\delta F(\nu)}{\delta\mu}(v_j)\delta_{ij}, \tag{2.5}$$

where δ_{ij} is the Kronecker symbol.

Let us apply the usual Itô lemma to the process

$$\eta_t^n = f(x_t^n(u_1),\dots,x_t^n(u_n)) = F(\sum_{j=1}^n c_j\delta_{x_t^n(u_j)}) = F(\mu_t^n)$$

and use (2.4), (2.5):

$$\begin{aligned}
d\eta_t^n &= \sum_i c_i\nabla\frac{\delta F(\mu_t^n)}{\delta\mu}(x_t^n(u_i))\left(a(x_t^n(u_i),\mu_t^n)dt + \sum_l b_l(x_t^n(u_i),\mu_t^n)dw_l(t)\right) + \\
&+\frac{1}{2}\sum_{i,j}c_ic_j\mathrm{sp}\left(\nabla_1\nabla_2\frac{\delta^2 F(\mu_t^n)}{\delta\mu^2}(x_t^n(u_i),x_t^n(u_j))\sigma^2_{\mu\mu}(x_t^n(u_i),x_t^n(u_j),\mu_t^n)\right)dt+ \\
&+\frac{1}{2}\sum_i c_i\mathrm{sp}\left(\nabla^2\frac{\delta F(\mu_t^n)}{\delta\mu}(x_t^n(u_i))\sigma^2_\mu(x_t^n(u_i),\mu_t^n)\right)dt.
\end{aligned} \tag{2.6}$$

Observe that, for each $g = g(u), h = h(u,v)$, we have

$$\sum_i c_ig(x_t^n(u_i)) = \int_{\mathbb{R}^d} g(u)\mu_t^n(du),$$

$$\sum_{i,j}c_ic_jh(x_t^n(u_i),x_t^n(u_j)) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}h(u,v)\mu_t^n(du)\mu_t^n(dv).$$

So, the expression in (2.6) is equal to

$$d\eta^n_t = L_2F(\mu^n_t)dt + \sum_{l=1}^m \int_{\mathbb{R}^d} \frac{\partial}{\partial u}\frac{\delta F(\mu^n_t)}{\delta \mu}(u) b_l(u, \mu^n_t)\mu^n_t(du)dw_l(t), \tag{2.7}$$

where the operator L_2 is defined in (2.2).

The initial distributions μ^n are weakly convergent to μ_0. Therefore [3], for each $t \geq 0$, we have the convergence μ^n_t to μ_t in probability. Taking a subsequence, if it is needed, it can be assumed without loss of generality that, for almost all $\omega \in \Omega$ and almost all $t \geq 0$,

$$\mu^n_t(\omega) \Rightarrow \mu_t(\omega),\ n \to \infty.$$

All functions in the differentiated expression in (2.7) are bounded and continuous. The proof that we can pass to a limit under the integral sign in (2.7) follows from the next lemma.

Lemma 2.1. *Let $\{\nu^n, n \geq 1\} \subset \mathcal{P}$ be a non-random sequence of probability measures which converges weakly to ν_0. Assume that the function $g : \mathbb{R}^d \times \mathcal{P} \to \mathbb{R}$ is continuous and bounded.*

Then

$$\int_{\mathbb{R}^d} g(u, \nu_n)\nu_n(du) \to \int_{\mathbb{R}^d} g(u, \nu_0)\nu_0(du),\ n \to \infty.$$

Proof.

$$\left|\int_{\mathbb{R}^d} g(u, \nu_n)\nu_n(du) - \int_{\mathbb{R}^d} g(u, \nu_0)\nu_0(du)\right| \leq$$

$$\leq \int_{\mathbb{R}^d} |g(u, \nu_n) - g(u, \nu_0)|\nu_n(du) + \left|\int_{\mathbb{R}^d} g(u, \nu_0)\nu_n(du) - \int_{\mathbb{R}^d} g(u, \nu_0)\nu_0(du)\right|.$$

The second item converges to zero due to the weak convergence $\nu_n \Rightarrow \nu_0,\ n \to \infty$. To estimate the first item, observe that the set $\{\nu_n, n \geq 0\}$ is a weak compact. By the Prokhorov theorem, for each $\varepsilon > 0$, there exists a compact $K_\varepsilon \subset \mathbb{R}^d$ such that $\nu_n(\mathbb{R}^d \setminus K_\varepsilon) < \varepsilon$ for all $n \geq 0$.

The function g is continuous on a compact $K_\varepsilon \times \{\nu_n; n \geq 0\} \subset \mathbb{R}^d \times \mathcal{P}$ and, hence, it is uniformly continuous. Thus,

$$\sup_{u \in K_\varepsilon} |g(u, \nu_n) - g(u, \nu_0)| \to 0,\ n \to \infty.$$

Therefore,

$$\overline{\lim_{n\to\infty}} \int_{\mathbb{R}^d} |g(u, \nu_n) - g(u, \nu_0)|\nu_n(du) \leq$$

$$\leq \overline{\lim_{n\to\infty}} \int_{K_\varepsilon} |g(u, \nu_n) - g(u, \nu_0)|\nu_n(du) + \overline{\lim_{n\to\infty}} \int_{\mathbb{R}^d\setminus K_\varepsilon} (|g(u, \nu_n)| + |g(u, \nu_0)|)\nu_n(du) \leq$$

$$\leq 2 \sup_{u\in\mathbb{R}^d} \sup_{\nu\in\mathcal{P}} |g(u, \nu)|\varepsilon.$$

The number $\varepsilon > 0$ is arbitrary. Thus, Lemma 2.1 and also Theorem 2.1 are proved.

3. Lower estimate for support of $x_t(u)$

Let $\psi = (\psi_1, \dots, \psi_m) : [0, T] \to \mathbb{R}^m$ be a piecewise smooth function. Denote, by x^ψ, a solution of the following (deterministic) equation

$$\begin{cases} dx^\psi_t(u) = \widetilde{a}(x^\psi_t(u), \mu^\psi_t)dt + \sum_{k=1}^m b_k(x^\psi_t(u), \mu^\psi_t)\psi_k(t),\ u \in \mathbb{R}^d, \\ \mu^\psi_t = \mu_0 \circ (x^\psi_t)^{-1},\ t \in [0, T], \\ x^\psi_0(u) = u, \end{cases} \tag{3.1}$$

where

$$\widetilde{a}(x,\mu) = a(x,\mu) - \frac{1}{2}\sum_{k=1}^{m}\left(\frac{\partial b_k(x,\mu)}{\partial x}b_k(x,\mu) + \int_{\mathbb{R}^d}\frac{\partial}{\partial v}\frac{\delta b_k(x,\mu)}{\delta\mu}(v)b_k(v,\mu)\mu(dv)\right) \tag{3.2}$$

is a corrected coefficient.

The aim of this section is to show that the support supp$x$ of the flow $x_t(u)$ contains the set

$$S = \{x^{\psi}:\ \psi \text{ is piecewise smooth}\}.$$

Note that if we replace the integral with respect to $\psi_k(t)$ by the Stratonovich integral w.r.t. $w_k(t)$, then we obtain Eq. (0.1).

Theorem 3.1. *Assume that the coefficients a, b_k are such that the derivatives*

$$\frac{\partial^{\alpha}}{\partial\bar{u}^{\alpha}}\frac{\partial^{\beta}}{\partial x^{\beta}}\frac{\delta^{j}a(x,\mu)}{\delta\mu^{j}}(\bar{u}) \quad \textit{and} \quad \frac{\partial^{\alpha}}{\partial\bar{u}^{\alpha}}\frac{\partial^{\beta}}{\partial x^{\beta}}\frac{\delta^{j}b_k(x,\mu)}{\delta\mu^{j}}(\bar{u})$$

are bounded and continuous in all their arguments, where $j = \overline{0, n+3}$, $\bar{u} = (u_1, \dots, u_j)$, $\alpha = (\alpha_1, \dots, \alpha_j), \alpha_i \geq 0$, $\beta = (\beta_1, \dots, \beta_d)$, $\beta_i \geq 0$, $\alpha_1 + \dots + \alpha_j + \beta_1 + \dots + \beta_d \leq n+3$. Then the support of x_t considered as a random element in $C([0,T], C^n(\mathbb{R}^d, \mathbb{R}^d))$ contains the set S.

Proof. We restrict ourselves only to the case $n = 0$ and $\psi \equiv 0$. The case of arbitrary ψ can be considered with the use of the Girsanov theorem (see the reasoning of Theorem [11]). The reasoning for arbitrary n is similar to the case $n = 0$ (see [6] for details).

Lemma 3.1. *Assume that the conditions of Theorem 3.1 are satisfied. Then, for each ball $U \subset \mathbb{R}^d$ and each $\varepsilon > 0, \delta > 0$, we have the following convergence of conditional probabilities:*

$$\lim_{\delta\to 0+} P\left(\sup_{\substack{u\in U\\ t\in[0,T]}}\left|\int_0^t b_k(x_s(u),\mu_s)\circ dw_k(s)\right| > \varepsilon / \|w\| < \delta\right) = 0,$$

where $|\cdot|$ is a norm in $\mathbb{R}^d$, $\|w\| = \sup_{t\in[0,T]}\max_{i=\overline{1,m}}|w_i(t)|$ is a norm in $C([0,T],\mathbb{R}^m)$.

Proof. Due to the Sobolev embedding theorems [12], it is enough to verify that

$$\lim_{\delta\to 0+} P\left(\sup_{t\in[0,T]}\|\int_0^t b_k(x_s(\cdot),\mu_s)\circ dw_k(s)\|_{W_p^1(U)} \geq \varepsilon / \|w\| < \delta\right) = 0,$$

where $p > d$, $\|f\|_{W_p^1(U)} = \left(\int_U(|f|^p + |\nabla f|^p)dx\right)^{1/p}$. The proof of the corresponding fact can be done similarly [7]. We need to use the Itô formula (2.3), boundedness of a, b_k and their derivatives, and some estimates of the moments for $x_t(u)$ and $\frac{\partial x_t(u)}{\partial u}$ (to estimate the moments, one can apply results in [5], §4.5, 4.6, to (1.2)).

Let us show now that $x^0 \in$ supp$x$. To verify this, it is sufficient to check that, for every ball $U \subset \mathbb{R}^d$ and $\varepsilon > 0$,

$$P\left(\sup_{t\in[0,T]}\sup_{u\in U}|x_t(u) - x_t^0(u)| < \varepsilon\right) > 0.$$

Denote $\sum_{k=1}^m \int_0^t b_k(x_s(u),\mu_s)\circ dw_k(s)$ by $o_t(u)$. Then

$$x_t(u) = u + \int_0^t \widetilde{a}(x_s(u),\mu_s)ds + o_t(u),$$
$$x_t^0(u) = u + \int_0^t \widetilde{a}(x_s^0(u),\mu_s^0)ds,$$

where $\widetilde{a}$ is defined in (3.2), $x_t^0(u)$ is a solution of (3.1) with $\psi\equiv 0$, $\mu_t^0 = \mu_0\circ(x_t^0)^{-1}$.

Let us use inequality (1.5) to estimate the difference $|x_t(u)-x_t^0(u)|, u\in U:$

$$|x_t(u)-x_t^0(u)| \le L\int_0^t \left(|x_s(u)-x_s^0(u)| + \int_{\mathbb{R}^d} |x_s(v)-x_s^0(v)|\wedge 1\mu_0(dv)\right)ds + |o_t(u)| \le$$

$$\le 2L\int_0^t (\sup_{v\in U}|x_s(v)-x_s^0(v)| + \mu(\mathbb{R}^d\setminus U))ds + \sup_{\substack{s\in[0,t]\\ v\in U}} |o_s(v)|,$$

where $L>0$ is a constant.

So, by the Gronwall lemma, we get the estimate

$$\sup_{v\in U}|x_t(v)-x_t^0(v)| \le \left(\sup_{\substack{v\in U\\ s\in[0,t]}} |o_s(v)| + \mu(\mathbb{R}^d\setminus U)\right)e^{2LT}.$$

Let $\varepsilon>0$, and let $U\subset\mathbb{R}^d$ be fixed. Choose a ball $\widetilde{U}\supset U$ such that $\mu(\mathbb{R}^d\setminus\widetilde{U})e^{2LT} < \frac{\varepsilon}{2}$. Then we choose $\delta>0$ such that

$$P\left(\sup_{\substack{u\in\widetilde{U}\\ t\in[0,T]}} |o_t(u)| < \frac{e^{-2LT}\varepsilon}{2} \Big/ \|w\|<\delta\right) > 0.$$

Therefore,

$$P\left(\sup_{\substack{u\in U\\ t\in[0,T]}} |x_t(u)-x_t^0(u)| \le \varepsilon\right) \ge P\left(\sup_{\substack{u\in\widetilde{U}\\ t\in[0,T]}} |x_t(u)-x_t^0(u)| \le\varepsilon\right) \ge$$

$$\ge P\left((\sup_{\substack{u\in\widetilde{U}\\ t\in[0,T]}} |o_t(u)| + \mu(\mathbb{R}^d\setminus\widetilde{U}))e^{2LT} \le \varepsilon \ / \ \|w\|<\delta\right)\cdot P(\|w\|<\delta) > 0.$$

Moreover, it is easy to verify that

$$P\left(\sup_{\substack{u\in U\\ t\in[0,T]}} |x_t(u)-x_t^0(u)| \le \varepsilon \ / \ \|w\|<\delta\right) \to 0,\ \delta\to 0+.$$

Theorem 3.1 is proved.

4. Approximation theorem.

Let us consider the system

$$\begin{cases} \frac{d}{dt}x_t^\varepsilon(u) = a(x_t^\varepsilon(u), \mu_t^\varepsilon) + \sum_{k=1}^m b_k(x_t^\varepsilon(u), \mu_t^\varepsilon)\xi_k^\varepsilon(t), \\ x_0^\varepsilon(u) = u, \ u \in \mathbb{R}^d, \\ \mu_t^\varepsilon = \mu_0 \circ (x_t^\varepsilon)^{-1}, \ t \in [0, T], \end{cases}$$

where the stochastic processes $\xi_k^\varepsilon(t)$ are such that $w_k^\varepsilon(t) = \int_0^t \xi_k^\varepsilon(s)ds$ converges as $\varepsilon \to 0$ to a Wiener process $w_k(t)$ in some sense.

In this section, we give some sufficient conditions that ensure the convergence of $(x_t^\varepsilon, \mu_t^\varepsilon)$ to (x_t, μ_t), where (x_t, μ_t) is a solution of (0.1).

Theorem 4.1. *Assume that the functions $a, b_k, k = \overline{1, m}$ satisfy the conditions of Theorem 3.1 and*

A1. The collection of processes

$$K_\varepsilon(s) = \sum_{l,j=1}^N \left(\int_s^T |E\left(\xi_l^\varepsilon(r)/\mathcal{G}_s^\varepsilon\right)| dr \right) \left(1 + |\xi_j^\varepsilon(s)|\right), \quad \varepsilon > 0$$

is uniformly L_p-bounded for each $p \geq 1$:

$$\forall p \geq 1 \quad \sup_\varepsilon \sup_s EK_\varepsilon^p(s) = K_p < \infty$$

and is uniformly exponentially bounded in mean:

$$\forall \lambda > 0 \quad \sup_\varepsilon E \exp\{\lambda \int_0^T K_\varepsilon(s)\, ds\} < \infty,$$

where

$$\mathcal{G}_t^\varepsilon := \sigma(\xi_k^\varepsilon(z) : 0 \leq z \leq t, \ k = 1, \ldots, N).$$

A2. For each $s \in [0, T]$, the following convergence holds in L_2 :

$$\int_s^T |E\big(\xi_l^\varepsilon(z)/\mathcal{G}_s^\varepsilon\big)|\, dz \to 0, \ \varepsilon \to 0.$$

A3. There exist the deterministic bounded functions $\sigma^{lm}(t), 1 \leq l, m \leq N$ such that, for all $s < t$, we have the following convergence in L_1 :

$$E\bigg(\int_s^t \xi_l^\varepsilon(\tau)d\tau \int_s^\tau \xi_m^\varepsilon(v)dv/\mathcal{G}_s^\varepsilon \bigg) \underset{\varepsilon \to 0}{\longrightarrow} \int_s^t \sigma^{lm}(z)dz.$$

Then the distribution in the space

$$\mathcal{C}_n \times C([0; T], \mathbb{R}^m) \times C([0; T], \mathcal{P})$$

of the triple $(x_t^\varepsilon, w^\varepsilon(t), \mu_t^\varepsilon)$ converges weakly as $\varepsilon \to 0$ to a limit measure such that:

1) the distribution of the second coordinate $w(t) = (w_1(t), \ldots, w_m(t))$ is a Wiener process with the covariation matrix $\| \int_0^t (\sigma^{ij}(s) + \sigma^{ji}(s))ds \|$;

2) the processes $x_t(u), w(t)$ and μ_t are connected by the system

$$
(4.1)\quad \begin{cases} dx_t(u) = \Bigg[a(x_t(u),\mu_t) + \sum_{j,k=1}^m \Big(\frac{\partial b_k(x_t(u),\mu_t)}{\partial x} b_j(x_t(u),\mu_t) + \\ + \int_{\mathbb{R}^d} \frac{\partial}{\partial v}\frac{\delta b_k(x_t(u),\mu_t)}{\delta\mu}(v) b_j(v,\mu_t)\mu_t(dv)\Big)\Bigg]\sigma^{kj}(t)dt + \sum_{k=1}^m b_k(x_t(u),\mu_t)dw_k(t), \\ x_0(u) = u,\ u\in\mathbb{R}^d, \\ \mu_t = \mu_0\circ x_t^{-1},\ t\in[0,T]. \end{cases}
$$

If, in addition, the processes $w^\varepsilon(t)$ and $w(t)$ are given on the same probability space and $w^\varepsilon(t)\to w(t),\ \varepsilon\to 0$ in probability for all $t\in[0;T]$, then

$$
(4.2)\qquad \forall p\ge 1\ \forall R>0\quad E\sup_{|u|\le R}\sup_{t\in[0;T]}\sum_{k=0}^n |\nabla^k(x_t^\varepsilon(u)-x_t(u))|^p \to 0, \varepsilon\to 0.
$$

An example of the sequence $w^\varepsilon(t)$ with $\sigma^{ij} = \begin{cases} 0, & i\ne j \\ 0.5, & i=j\end{cases}$ is a polygonal approximation:

$$
\xi_\varepsilon^j(t) = \varepsilon^{-1}(w_j((k+1)\varepsilon) - w_j(k\varepsilon))
$$

if $t\in[k\varepsilon;(k+1)\varepsilon)$.

Theorem 4.1 and this example imply that the support of the $x_t(u)$ distribution in the space $\mathcal{C}_n$ is contained in a set

$$
\{x^\psi:\ \psi \text{ is a piecewise linear function, } x^\psi \text{ satisfies (3.1)}\}.
$$

It is easy to show that if $\{\psi_n\}$ is a sequence of piecewise smooth functions such that

$$
\sup_{t\in[0;T]}|\psi_n(t)-\psi_0(t)|\to 0,\ n\to\infty
$$

$$
\text{ess}\,\sup_{t\in[0;T]}|\psi_n'(t)-\psi_0'(t)|\to 0,\ n\to\infty,
$$

then we have the convergence $x^{\psi_n}\to x^{\psi_0},\ n\to\infty$ in $\mathcal{C}_n$.

So, the closure of

$$
\{x^\psi:\ \psi \text{ is piecewise linear function, } x^\psi \text{ satisfies (3.1)}\}.
$$

contains a set

$$
\{x^\psi:\ \psi \text{ is piecewise smooth function, } x^\psi \text{ satisfies (3.1)}\}.
$$

Therefore, we have obtained the following support theorem.

Theorem 4.2. *Assume that the conditions of Theorems 3.1, 4.1 are satisfied. Then the support of the $x_t(u)$ distribution in $\mathcal{C}_n$ is equal to the closure in $\mathcal{C}_n$ of the set*

$$
\{x^\psi:\ \psi \textit{ is a piecewise smooth function, } x^\psi \textit{ satisfies (3.1)}\}.
$$

Proof of the Theorem 4.1. We follow the ideas of [5] Ch.5, see also [7] for some details related to the equations with interaction. Note, at first, that, under the conditions of Theorem 4.1, the processes w^ε converge weakly to a Wiener process w with required covariation [5] §5.7.

Lemma 4.1.

$$\forall m \in \mathbb{N} \ \ \forall R>0 \ \ \exists C = C_{R,m} > 0 : \sup_{\varepsilon} E\|x_\varepsilon(\cdot,t)\|^{2m}_{W^k_{2m(B_R)}} \le C,$$
$$\forall t', t'' : \sup_{\varepsilon} E\|x_\varepsilon(\cdot,t') - x_\varepsilon(\cdot,t'')\|^{2m}_{W^k_{2m(B_R)}} \le C|t'-t''|^{2-\frac{1}{m}},$$

where B_R is a ball in $\mathbb{R}^d$ with the center at zero and with radius R.

The course of the proof is similar to [5] §5.2 (see also [7] for the equations with interaction), but we have to use Theorem 2.1 instead of the formula of the usual integration by parts.

It is well known that the Sobolev space $W^{n+1}_{p,loc}$ is compactly embedded into the space $C^n(\mathbb{R}^d, \mathbb{R}^d)$ if $p > d$. Thus, by Theorem 1.4.7 [5], we have the weak relative compactness of the family $x^\varepsilon, w^\varepsilon$ in $\mathcal{C}_n \times C([0;T], \mathbb{R}^m)$.

The next lemma implies that if a sequence of couples $(x^{\varepsilon_k}, w^{\varepsilon_k})$ converges weakly to some limit (x, w), then a sequence of triples $(x^{\varepsilon_k}_{\cdot}(\cdot), w^{\varepsilon_k}(\cdot), \mu^{\varepsilon_k}_{\cdot})$ converges weakly to $(x_{\cdot}(\cdot), w(\cdot), \mu_{\cdot})$, where $\mu_t = \mu_0 \circ (x_t)^{-1}$.

Lemma 4.2. *Assume that a sequence of random elements $y^n = y^n_t(u)$ converges weakly to $y^0 = y^0_t(u)$ in $C([0;T], C(\mathbb{R}^d, \mathbb{R}^d))$. Then, for each probability measure μ, we have the weak convergence of measure-valued processes $\mu^n_t = \mu \circ (y^n_t)^{-1} \to \mu^0_t = \mu \circ (y^0_t)^{-1}$ in the space $C([0;T], \mathcal{P})$.*

Proof. By the Skorokhod theorem [13], we can assume that all the elements $y^n, n \ge 0$ are given on the same probability space and with probability one:

$$\forall R > 0 : \sup_{t\in[0;T]} \sup_{|u|<R} |y^n_t(u) - y^0_t(u)| \to 0, \ n \to \infty.$$

Let ω be from the corresponding set of full probability. Denote, by $\widetilde{\varkappa}_t$, the image of the measure μ w.r.t. mapping (y^n_t, y^0_t). Then

$$\sup_{t\in[0;T]} \gamma(\mu^n_t, \mu^0_t) = \sup_{t\in[0;T]} \sup_{\varkappa \in Q(\mu^n_t;\mu^0_t)} \int\int |u-v| \wedge 1 \varkappa(du, dv) \le$$

$$\le \sup_{t\in[0;T]} \int\int |u-v| \wedge 1 \widetilde{\varkappa}_t(du, dv) = \sup_{t\in[0;T]} \int |y^n_t(u) - y^0_t(u)| \wedge 1 \mu(du). \tag{4.3}$$

Let $\varepsilon > 0$ be fixed. Choose $R > 0, \ n_0 \ge 1$ such that $\mu(u : \ |u| \ge R) < \varepsilon$ and

$$\forall n \ge n_0 : \sup_{t\in[0;T]} \sup_{|u|<R} |y^n_t(u) - y^0_t(u)| < \varepsilon.$$

Then the right-hand side of (4.3) is less than 2ε. Lemma 4.2 is proved.

If we verify that the limit triple (x_t, w_t, μ_t) satisfies Eq. (4.1), then, by uniqueness of the solution, we get the desired convergence $(x^\varepsilon_t, w^\varepsilon_t, \mu^\varepsilon_t) \to (x_t, w_t, \mu_t)$.

To check (4.1) for a limit, it is sufficient to verify that

$$M(u,t) = x(u,t) - u - \int_0^t \Bigg[a(x_z(u), \mu_z) + \sum_{j,k=1}^m \Big(\frac{\partial b_k(x_z(u), \mu_z)}{\partial x} b_j(x_z(u), \mu_z) +$$

$$+ \int_{\mathbb{R}^d} \frac{\partial}{\partial v} \frac{\delta b_k(x_z(u), \mu_t)}{\delta \mu}(v) b_j(v, \mu_z) \mu_z(dv) \Big) \sigma^{kj}(z) \Bigg] dz \tag{4.4}$$

is a continuous L_2-martingale with respect to $\mathcal{F}_t = \sigma\big(w(s), x_s(u), s \in [0,t], u \in \mathbb{R}^d\big), t \in [0,T]$, with the square characteristics

(4.5)
$$\langle M^{(i)}(u,t), M^{(j)}(v,t)\rangle = \sum_{k,l}\int_0^t \int b_k^{(i)}(x(u,s),\mu_s) b_l^{(j)}(x(v,s),\mu_s))(\sigma^{kl}(s)+\sigma^{lk}(s))ds$$

$$\langle M^{(i)}(u,t), w_j(t)\rangle = \sum_k \int_0^t \int b_k^{(i)}(x_s(u)\mu_s)(\sigma^{kj}(s)+\sigma^{jk}(s))ds. \tag{4.6}$$

The process $b_k(x_t^\varepsilon(u),\mu_t^\varepsilon)$ is differentiable with respect to t by Theorem 2.1 and

$$\frac{\partial}{\partial t}(b_k(x_t^\varepsilon(u),\mu_t^\varepsilon)) =$$

$$= \frac{\partial}{\partial x}(b_k(x_t^\varepsilon)(u),\mu_t^\varepsilon))\left(a(x_t^\varepsilon(u),\mu_t^\varepsilon) + \sum_{j=1}^m b_j(x_t^\varepsilon(u),\mu_t^\varepsilon)\xi_j^\varepsilon(t)\right) +$$

$$+ \int_{\mathbb{R}^d}\frac{\partial}{\partial v}\frac{\delta b_k(x_t^\varepsilon(u),\mu_t^\varepsilon)}{\delta\mu}(v)\left(a(v,\mu_t^\varepsilon) + \sum_{j=1}^m b_j(v,\mu_t^\varepsilon)\xi_j^\varepsilon(t)\right)\mu_t^\varepsilon(dv).$$

Then

$$x_t^\varepsilon(u) - x_s^\varepsilon(u) - \int_s^t a(x_z^\varepsilon(u),\mu_z^\varepsilon)dz = \sum_{k=1}^m \int_s^t b_k(x_z^\varepsilon(u),\mu_z^\varepsilon)\xi_k^\varepsilon(z)dz =$$

$$= \sum_{k=1}^m b_k(x_s^\varepsilon(u),\mu_s^\varepsilon)\int_s^t \xi_z^k dz + \sum_{k=1}^m \int_s^t\int_s^z \frac{\partial}{\partial r}\left(b_k(x_r^\varepsilon(u),\mu_r^\varepsilon)\right) dr\xi_k^\varepsilon(z)dz =$$

$$= \sum_{k=1}^m b_k(x_s^\varepsilon(u),\mu_s^\varepsilon)\int_s^t \xi_k^\varepsilon(z)dz +$$

$$+ \sum_{k=1}^m \int_s^t\int_s^z \left[\frac{\partial}{\partial x}(b_k(x_r^\varepsilon)(u),\mu_r))\left(a(x_r^\varepsilon(u),\mu_r^\varepsilon) + \sum_{j=1}^m b_j(x_r^\varepsilon(u),\mu_r^\varepsilon)\xi_j^\varepsilon(r)\right) +\right.$$

$$\left. + \int_{\mathbb{R}^d}\frac{\partial}{\partial v}\frac{\delta b_k(x_r^\varepsilon(u),\mu_r^\varepsilon)}{\delta\mu}(v)\left(a(v,\mu_r^\varepsilon) + \sum_{j=1}^m b_j(v,\mu_r^\varepsilon)\xi_j^\varepsilon(r)\right)\right]\mu_r^\varepsilon(dv)dr\xi_k^\varepsilon(z)dz.$$

Let $s \in [0;T]$, $s_j \in [0;s)$, $u_j \in \mathbb{R}^d$, $l \in \mathbb{N}$. Put

$$\Phi_\varepsilon = f(x_{s_1}^\varepsilon(u_1),\ldots,x_{s_l}^\varepsilon(u_l), w^\varepsilon(s_1),\ldots,w^\varepsilon(s_l)),$$

where f is some continuous bounded function,

$$\Phi = f(x_{s_1}(u_1),\ldots,x_{s_l}(u_l), w(s_1),\ldots,w(s_l)).$$

We recall that $\{\varepsilon_n\}$ is such that the sequence of triples $(x_\cdot^{\varepsilon_n}(\cdot), w^{\varepsilon_n}(\cdot), \mu_\cdot^{\varepsilon_n})$ converges weakly to some limit $(x_\cdot(\cdot), w(\cdot), \mu_\cdot)$

The following statement can be proved similarly to Lemma 5 [7]; see also the reasoning of Lemma 4.2.

Lemma 4.3. *Let $\alpha = \alpha(x, \mu)$ be a bounded continuous function. Then*

$$E\int_s^t \alpha(x_r^{\varepsilon_n}(u), \mu_r^{\varepsilon_n}) dr \Phi_{\varepsilon_n} \to E\int_s^t \alpha(x_r(u), \mu_r) dr\Phi, n \to \infty;$$

$$E\int_s^t \alpha(x_r^{\varepsilon_n}(u), \mu_r^{\varepsilon_n})\xi_k^{\varepsilon_n}(r) dr \Phi_{\varepsilon_n} \to 0, n \to \infty;$$

$$E\int_s^t\int_s^z \alpha(x_r^{\varepsilon_n}(u), \mu_r^{\varepsilon_n})\xi_k^{\varepsilon_n}(r)\xi_j^{\varepsilon_n}(z) dr dz \Phi_{\varepsilon_n} \to E\int_s^t \alpha(x_r(u), \mu_r)\sigma^{kj}(r) dr\Phi, n \to \infty.$$

As a corollary of Lemma 4.3, we have a fact that $M(u,t)$ is $\mathcal{F}_t-$martingale. The reasoning for (4.5), (4.6) is similar. That is, the limit triple satisfies (4.1). So we have proved the desired weak convergence.

The proof of (4.2) is quite classical and may be roughly formulated as follows. Let a sequence of solutions of SDEs given on the same probability space be weakly relatively compact, and let the uniqueness for the limit equation hold. Then the sequence converges strongly (cf. Theorem 5.2.8 [5]).

Bibliography

1. Dorogovtsev A.A. and Kotelenez P., *Smooth Stationary Solutions of Quasilinear Stochastic Partial Differential Equations: 1. Finite Mass*, Preprint No. 97-145, Cleveland (1997).
2. Kotelenez P., *A class of quasilinear stochastic partial differential equations of McKean-Vlasov type with mass conservation.*, Probab. Theory Related Fields **102** (1995), no. 2, 159–188.
3. Dorogovtsev A.A., Kotelenez P., *Stochastic flows with interaction and random measures.* (to appear)
4. Stroock D.W., Varadhan S.R.S., *On the support of diffusion processes with applications to the strong maximum principle.*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, pp. 333–359.
5. Kunita H., *Stochastic Flows and Stochastic Differential Equations*, vol. 24, Cambridge studies in advanced mathematics, 1990, pp. 346.
6. Pilipenko A. Yu., *Stroock-Varadhan theorem for flows generated by stochastic differential equations with interaction*, Ukrainian Math. J. **54** (2002), no. 2, 227–236.
7. Pilipenko A.Yu., *Approximation theorem for stochastic differential equations with interaction*, Random Oper. and Stoch. Equ. **11** (2003), no. 3,, 213–228.
8. Dudley R. M., *Real Analysis and Probability*, Pacific Grove, Calif.: Wodsworth & Brooks/Cole Pub. Co., 1989.
9. Dawson D.A., *Measure-Valued Markov processes. Ecole d'Ete de Probabilites de Saint-Flour XXI—1991, Lecture Notes in Math.*, vol. 1541, Springer, Berlin, 1993, pp. 1–260.
10. Jacka S., Tribe R., *Comparisons for measure valued processes with interactions.*, Ann. Probab. **31** (2003), no. 3, 1679–1712.
11. Ikeda N., Watanabe S., *Stochastic Differential Equations and Diffusion Processes*, North-Holland Mathematical Library, 24. North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1981.
12. Adams R.A., *Sobolev Spaces*, Academic Press, New York, 1975.
13. Ethier S.N, Kurtz T.G., *Markov Processes: Characterization and Convergence*, Wiley, New York, 1986.

Institute of Mathematics, National Academy of Sciences of Ukraine, 3, Tereshchenkivska Str., Kyiv 10601, Ukraine
E-mail: apilip@imath.kiev.ua