The Annals of Probability
1984, Vol. 12, No. 2, 458–479

A MARTINGALE APPROACH TO THE LAW OF LARGE NUMBERS FOR WEAKLY INTERACTING STOCHASTIC PROCESSES[1]

BY KARL OELSCHLÄGER

Universität Heidelberg

It is shown that certain measure-valued stochastic processes describing the time evolution of systems of weakly interacting particles converge in the limit, when the particle number goes to infinity, to a deterministic nonlinear process.

1. Introduction. We study two species of interacting processes, namely diffusion processes and jump processes. In each situation we consider a system of N particles moving in $\mathbb{R}^n$. We assume that the characteristics of the dynamics of one particle of the system (i.e. drift vector and diffusion matrix, resp. jump intensities) depend both on the actual position of the particle and moreover on the state of the whole N-particle system, i.e. on the positions of the other particles too. This interaction among the particles is "weak", which means, that as the particle number N goes to infinity, the range of the interaction remains fixed, whereas its strength is rescaled by $1/N$. Having in mind this limiting behaviour, it seems convenient to describe a particle in $x \in \mathbb{R}^n$ by $(1/N)\delta_x$ (i.e. the Dirac measure with mass $1/N$ concentrated at x). Therefore any state of the whole system of N particles can be identified with a probability measure of the form $(1/N) \sum_{i=1}^N \delta_{x_i}$, where x_i is the position of particle i.

In the diffusion case the dynamics of the N-particle system can be given either by a system of N stochastic differential equations

$$dX_t^i = b(X_t^i, (1/N) \textstyle\sum_{j=1}^N \delta_{X_t^j})\, dt + \sigma(X_t^i, (1/N) \sum_{j=1}^N \delta_{X_t^j})\, dW_t^i, \quad i = 1, \cdots, N$$

(X_t^i is the position of particle i at time t, W_t^i, $i = 1, \cdots, N$ are independent $\mathbb{R}^n$-valued Brownian motions) or by the infinitesimal generator A^N, which is defined in the following way: Let $f(\mu) = g(\langle \mu, h \rangle)$

$$\left(g \in C_b^2(\mathbb{R}),\ h \in C_b^2(\mathbb{R}^n),\ \mu = \frac{1}{N} \textstyle\sum_{i=1}^N \delta_{x_i},\ \langle \mu, h \rangle = \displaystyle\int h(x)\mu(dx)\right)$$

and $a(x, \mu) = \sigma(x, \mu)\sigma^T(x, \mu)$. Then we have

$$(A^N f)(\mu) = g'(\langle \mu, h \rangle)\{\langle \mu, b(\cdot, \mu) \cdot \nabla h \rangle + \tfrac{1}{2}\langle \mu, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, \mu)\partial_{ij} h \rangle\}$$
$$+ \tfrac{1}{2} g''(\langle \mu, h \rangle)(1/N)\langle \mu, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, \mu)\partial_i h \partial_j h \rangle.$$

In the jump process case, one may describe the time evolution by the jump

Received September 1982; revised July 1983.

[1] This work has been supported by the Deutsche Forschungsgemeinschaft.

AMS 1980 *subject classifications*. Primary 60K35; secondary 60F05.

Key words and phrases. Interacting stochastic processes, law of large numbers, martingales.

intensities $A(x_i, y; (1/N) \sum_{j=1}^N \delta_{x_j})\gamma_{x_i}(dy)$ for a jump of particle i from x_i to $x_i + y$ (if the state of the system of all N particles is $\mu = (1/N) \sum_{j=1}^N \delta_{x_j}$), where γ_x is a probability measure on $\mathbb{R}^n$ for all x, or more conveniently by the infinitesimal generator

$$(B^N f)(\mu) = N\left\langle \mu, \int A(\cdot, y; \mu)\left(f\left(\mu + \frac{1}{N}\delta_{\cdot+y} - \frac{1}{N}\delta_{\cdot}\right) - f(\mu)\right)\gamma_{\cdot}(dy)\right\rangle,$$

(which is defined for real valued bounded measurable functions f on the space of probability measures on $\mathbb{R}^n$ of the form $(1/N) \sum_{i=1}^N \delta_{x_i}$.)

Our aim is to show that under suitable smoothness conditions on b and a (resp. A and γ) the convergence of the initial distributions $\mathscr{L}(X_0^N)$ of the N-particle process X^N implies the convergence of the processes $X^N = (X_t^N)_{t\le T}$ to a deterministic probability-measure-valued process $X^* = (X_t^*)_{t\le T}$ (as $N \to \infty$), where T is a fixed positive number. Similar results have been derived in different contexts (see below), but it seems worthwhile to present a method, based on martingales, which works for this sort of limit problem (sometimes called the Vlassov-Limit; cf. [8]) for different species of interacting processes in essentially the same way.

In [9], the author applied these methods to processes describing certain models of populations of interacting individuals, which are characterized by their biological age. Furthermore he proved central limit thoerems for those population processes and the jump processes considered in the present paper.

2. Some notation.

$C_b(M)$	space of bounded continuous functions on the topological space M
$C_b^k(E)$	space of all functions in $C_b(E)$ with k bounded continuous derivatives (E Euclidean space)
$C_c^k(E)$	space of all functions in $C_b^k(E)$ with compact support
$\mathscr{L}(X)$	distribution of the random variable (process) X
$\langle \mu, f \rangle$	$= \int f(x)\cdot \mu(dx)$
$\mathscr{B}_1$	$= \left\{ f: \mathbb{R}^n \to \mathbb{R} : \sup_{x\in\mathbb{R}^n} \lvert f(x) \rvert \le 1, \sup_{x,y\in\mathbb{R}^n} \frac{\lvert f(x) - f(y) \rvert}{\lvert x - y \rvert} \le 1 \right\}$
$\mathscr{M}$	space of probability measures on $\mathbb{R}^n$
$(\mathscr{M}, \lVert\cdot\rVert_0)$	$\mathscr{M}$ equipped with the strong topology (defined by the total variation $\lVert\cdot\rVert_0$)

$(\mathcal{M}, \|\cdot\|_1)$	$\mathcal{M}$ equipped with the metric $\|\mu - v\|_1 = \sup_{f\in\mathscr{A}_1} \langle \mu - v, f\rangle$ (It should be noted that $(\mathcal{M}, \|\cdot\|_1)$ is a complete metric space, which is topologically equivalent to $\mathscr{P}(\mathbb{R}^n)$; cf. [4])
$\mathscr{P}(M)$	space of all probability measures on the topological space M ($\mathscr{P}(M)$ is equipped with the weak topology)
$\Omega_i = \mathscr{D}([0, T], (\mathcal{M}, \|\cdot\|_i))$ $(i = 0, 1)$	space of all right-continuous functions from $[0, T]$ to $(\mathcal{M}, \|\cdot\|_i)$, which have left limits at each point of their domain. These spaces are equipped with the Skorokhod topology
$Y = (Y_t)_{t\le T}$	typical element of Ω_i $(i = 0, 1)$
$\mathscr{F}_t$	σ-algebra generated by $\{Y_s : s \le t\}$
$x \wedge y = \min\{x, y\}$	
$x \vee y = \max\{x, y\}$	
$\mathcal{S}^n$	space of all positive definite, symmetric matrices on R^n
$\kappa(x) = \|x\|$, $\psi(x) = \|x\|^2$	for $x \in \mathbb{R}^n$
$c_1, c_2, \cdots$	positive constants $< \infty$.

3. Results. Before formulating the main results, we shall make the following assumptions on the functions $b:\mathbb{R}^n \times \mathcal{M} \to \mathbb{R}^n$, $a:\mathbb{R}^n \times \mathcal{M} \to \mathcal{S}^n$, $A:\mathbb{R}^n \times \mathbb{R}^n \times \mathcal{M} \to R^+$ and $\gamma:\mathbb{R}^n \to \mathcal{M}$:

(A1) $b:\mathbb{R}^n \times (\mathcal{M}, \|\cdot\|_1) \to \mathbb{R}^n$ is Lipschitz continuous and satisfies $|b(x, \mu)| \le c_1(1 + |x| + \langle \mu, \kappa\rangle)$.

(A2) $a:\mathbb{R}^n \times (\mathcal{M}, \|\cdot\|_1) \to \mathcal{S}^n$ has a representation $a = \sigma \cdot \sigma^T$, where each component of the matrix-valued function $(x, \mu) \to \sigma(x, \mu)$ is bounded and Lipschitz continuous from $\mathbb{R}^n \times (\mathcal{M}, \|\cdot\|_1)$ to $\mathbb{R}$.

(B1) For each pair $\{x, x'\} \subset R^n$ there exists a probability measure $\rho_{x,x'}$ on $\mathbb{R}^{2n}$, such that $\rho_{x,x'}$ has the marginal distributions γ_x and $\gamma_{x'}$, and

i) $$\int (A(x, y; \mu) \wedge A(x', y'; \mu))\,|y - y'|\,\rho_{x,x'}(dy, dy') \le |x - x'|\,c_2(1 + \langle \mu, \kappa\rangle)$$

ii) $$\left(\int |A(x, y; \mu) - A(x', y'; \mu)|^2\,\rho_{x,x'}(dy, dy')\right)^{1/2} \le (1 + \langle \mu, \kappa\rangle)c_6\,|x - x'|.$$

(B2) $$\int |y|^2\,\gamma_x(dy) \le c_3.$$

(B3) $$0 \le A(x, y; \mu) \le c_4(1 + |x| + \langle \mu, \kappa \rangle).$$

(B4) $$|A(x, y; \mu) - A(x, y; \nu)| \le c_5(1 + |x|) \| \mu - \nu \|_1.$$

Typically these assumptions are satisfied, if

$$\sigma(x, \mu) = \sigma(x, \langle \mu, f_1(x, \cdot) \rangle, \cdots, \langle \mu, f_r(x, \cdot) \rangle),$$

$$b(x, \mu) = \mathbf{b}(x, \langle \mu, g_1(x, \cdot) \rangle, \cdots, \langle \mu, g_p(x, \cdot) \rangle),$$

$$A(x, y; \mu) = \mathbf{A}(x, y; \langle \mu, h_1(x, y, \cdot) \rangle, \cdots, \langle \mu, h_q(x, y, \cdot) \rangle) \quad \text{and}$$

$$\gamma_x = \gamma \quad \text{(independent of } x\text{)},$$

where the functions σ, $\mathbf{b}$, $\mathbf{A}$, $f_1, \cdots, f_r, g_1, \cdots, g_p, h_1, \cdots, h_q$ are Lipschitz continuous and "not too rapidly" growing at infinity. Our results are summarized in the following two theorems:

THEOREM 1 (diffusion processes). *Assume that the drift vector and the diffusion matrix satisfy* (A1) *and* (A2). *Furthermore suppose, that*

$$\sup_{N \in \mathbb{N}} E_{\mathscr{L}(X^N)}[\langle Y_0, \psi \rangle] = \sup_{N \in \mathbb{N}} E[\langle X_0^N, \psi \rangle] < \infty$$

$$\mathscr{L}(X_0^N) \to_{N \to \infty} \delta_{X_0^*} \quad \textit{weakly in} \quad \mathscr{P}(\mathscr{M}, \| \cdot \|_1).$$

($\delta_{X_0^*}$ is the Dirac measure concentrated in $X_0^* \in \mathscr{M}$). Then we have the convergence

$$\lim_{N \to \infty} \mathscr{L}(X^N) = \delta_{X^*} \quad \textit{in} \quad \mathscr{P}(\Omega_1),$$

where the deterministic process $X^* = (X_t^*)_{t \le T}$ satisfies $X_t^* = \mathscr{L}(\zeta_t)$ for the unique solution $\zeta = (\zeta_t)_{t \le T}$ of

$$d\zeta_t = b(\zeta_t, \mathscr{L}(\zeta_t))dt + \sigma(\zeta_t, \mathscr{L}(\zeta_t))d\tilde{W}_t, \quad \mathscr{L}(\zeta_0) = X_0^*.$$

($\tilde{W} = (\tilde{W}_t)_{t \le 0}$ n-dimensional Brownian motion).

REMARKS. Since all the processes X^N have continuous trajectories, and since the deterministic limit process X^* is continuous too, it can be concluded from the above theorem that the convergence $\mathscr{L}(X^N) \to \delta_{X^*}$ takes place also in $\mathscr{P}(C([0, T], (\mathscr{M}, \| \cdot \|_1)))$, where the space $C([0, T], (\mathscr{M}, \| \cdot \|_1))$ of continuous functions from $[0, T]$ to $(\mathscr{M}, \| \cdot \|_1)$ is equipped with the usual supremum-norm topology. Nevertheless, for showing the analogy to Theorem 2, we formulate and prove Theorem 1 using the space $\mathscr{D}([0, T], (\mathscr{M}, \| \cdot \|_1))$.

Similar problems about interacting diffusion processes have been studied by McKean [5], who proved propagation of chaos for such processes. Moreover Marchioro and Pulvirenti [6] considered (among many other things) the same question as in Theorem 1 for systems, in which interaction occurs only through the drift.

THEOREM 2 (jump processes). *Suppose the validity of* (B1) *to* (B4). *Moreover* *assume for the initial distributions* $\mathscr{L}(X_0^N)$ *of the N-particle processes* X^N:

$$\lim_{N\to\infty}\mathscr{L}(X_0^N) = \delta_{X_0^*} \quad \textit{weakly in} \quad \mathscr{P}(\mathscr{M}, \|\cdot\|_1)$$

$$\sup_{N\in\mathbb{N}} E_{\mathscr{L}(X^N)}[\langle Y_0, \kappa\rangle] = \sup_{N\in\mathbb{N}} E[\langle X_0^N, \kappa\rangle] < \infty.$$

Then the sequence $\{\mathscr{L}(X^N)\}_{N\in\mathbb{N}}$ *of the distributions of the N-particle processes* X^N *converges weakly in* $\mathscr{P}(\mathscr{D}([0, T], (\mathscr{M}, \|\cdot\|_1)))$ *to the Dirac measure* δ_{X^*}, *concentrated at the solution of the integral equation*

$$\langle X_t^*, f\rangle = \langle X_0^*, f\rangle + \int_0^t \left\langle X_s^*, \int A(\cdot, y; X_s^*)(f(\cdot + y) - f(\cdot))\gamma\cdot(dy)\right\rangle ds, \quad f\in C_c^2(\mathbb{R}^n).$$

REMARKS. Results similar to Theorem 2 have been announced by Skorokhod [7].

Since the proof of Theorem 1 is essentially the same as the proof of Theorem 2, we shall omit a verification of the last theorem, except the proof of the uniqueness of the limit dynamics.

4. Proof. To prove Theorem 1, we proceed in small steps, with each step formulated as a lemma.

LEMMA 1. *For each function* $\phi_{f,g}\in C_b(\mathscr{M}, \|\cdot\|_i)$ $(i = 0, 1)$

$$\phi_{f,g}(\mu) = f(\langle\mu, g\rangle), \quad g\in C_c^2(\mathbb{R}^n), \quad f\in C_b^2(\mathbb{R})$$

and each $N\in\mathbb{N}$ *the process*

$$E_{f,g}^N = (E_{f,g}^N(t))_{t\le T}, \quad E_{f,g}^N(t)(Y) = \phi_{f,g}(Y_t) - \int_0^t A^N\phi_{f,g}(Y_s)\,ds$$

is a martingale relative to $(\Omega_i, \mathscr{F}_T, (\mathscr{F}_t)_{t\le T}, \mathscr{L}(X^N))$. (This lemma is a consequence of Dynkin's Formula; cf. [1]).

LEMMA 2.

a) $\mathscr{P}_{\mathscr{L}(X^N)}[\|Y_t\|_0 = 1\ \forall t\le T] = 1$.

b) *There exist positive constants* c_7, c_8, c_9, c_{10}, *such that the process*

$$t\to\langle Y_t, \psi\rangle\, e^{-c_7 t} - c_8 t \quad (\text{respect. } t\to\langle Y_t, \psi\rangle e^{c_9 t} + c_{10}t)$$

is a super- (respect. sub-) martingale with respect to

$$(\Omega_0, \mathscr{F}_T, (\mathscr{F}_t)_{t\le T}, \mathscr{L}(X^N)).$$

PROOF. a) is trivial. For the proof of b) we apply Lemma 1 to the functions $f(x) = x$ and $g(x) = \psi(x)$. Since Lemma 1 is not formulated for those functions, a correct application of that lemma includes the use of the stopping-times $\tau_r =$

$\inf\{t \geq 0 : \langle Y_t, \psi\rangle > r\} \wedge T$ in the following way:

Let $\psi_r \in C_c^2(\mathbb{R}^n)$, such that $\psi_r(x) = \psi(x)$ if $|x| \leq \sqrt{rN} + 1$, and $\eta_r \in C_b^2(\mathbb{R})$, such that $\eta_r(x) = x$ if $|x| \leq r + 1$. Then we have $\langle Y_t, \psi\rangle = \eta_r(\langle Y_t, \psi_r\rangle)$ for all $t \leq \tau_r(Y)$ $\mathscr{L}(X^N)$—a.s. Thus, by Lemma 1, for all sufficiently large $r = r(Y)$

$$\begin{aligned}
&E_{\mathscr{L}(X^N)}[\langle Y_{(t\wedge\tau_r(Y))\vee s}, \psi\rangle \mid \mathscr{F}_s] \\
&\quad = E_{\mathscr{L}(X^N)}[\eta_r(\langle Y_{(t\wedge\tau_r(Y))\vee s}, \psi_r\rangle) \mid \mathscr{F}_s] \\
&\quad = \eta_r(\langle Y_s, \psi_r\rangle) + E_{\mathscr{L}(X^N)}\left[\int_s^{(t\wedge\tau_r(Y))\vee s} du \left\{\langle Y_u, b(\cdot, Y_u)\cdot\nabla\psi\rangle \right.\right. \\
&\qquad\qquad \left.\left. + \frac{1}{2}\langle Y_u, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_u)\partial_{ij}\psi\rangle\right\} \,\middle|\, \mathscr{F}_s\right] \\
(2.1)\qquad &\quad \leq \eta_r(\langle Y_s, \psi_r\rangle) + E_{\mathscr{L}(X^N)}\left[\int_s^{(t\wedge\tau_r(Y))\vee s} du\{\langle Y_u, (1 + \kappa + \langle Y_u, \kappa\rangle)\kappa\rangle \right. \\
&\qquad\qquad \left. + \|Y_u\|_0 n^2\}c_{11} \mid \mathscr{F}_s\right] \\
&\quad \leq \langle Y_s, \psi\rangle + E_{\mathscr{L}(X^N)}\left[\int_s^t du\{\langle Y_{(u\wedge\tau_r(Y))\vee s}, \psi\rangle\}c_{12} \mid \mathscr{F}_s\right] + c_{13}(t-s).
\end{aligned}$$

(2.1) and Gronwall's Lemma yield

$$E_{\mathscr{L}(X^N)}[\langle Y_{(t\wedge\tau_r(Y))\vee s}, \psi\rangle \mid \mathscr{F}_s] \leq \langle Y_s, \psi\rangle e^{c_{12}(t-s)} + c_{13}(t-s)e^{c_{12}(t-s)},$$

and so by Fatou's Lemma,

$$\begin{aligned}
E_{\mathscr{L}(X^N)}[\langle Y_t, \psi\rangle \mid \mathscr{F}_s] &= E_{\mathscr{L}(X^N)}\,[\liminf_{r\to\infty}\langle Y_{(t\wedge\tau_r(Y))\vee s}, \psi\rangle \mid \mathscr{F}_s] \\
&\leq \langle Y_s, \psi\rangle e^{c_{12}(t-s)} + c_{13}(t-s)e^{c_{12}t}.
\end{aligned}$$

This proves the first part of Lemma 2b). In a similar way as (2.1) we can prove

$$\begin{aligned}
&E_{\mathscr{L}(X^N)}[\langle Y_{(t\wedge\tau_r(Y))\vee s}, \psi\rangle \mid \mathscr{F}_s] \\
&\qquad \geq \langle Y_s, \psi\rangle - E_{\mathscr{L}(X^N)}\left[\int_s^t du\langle Y_{(u\wedge\tau_r(Y))\vee s}, \psi\rangle c_{14} \,\middle|\, \mathscr{F}_s\right] - c_{15}(t-s).
\end{aligned}$$

Using Lemma 1 and the supermartingale property of $t \to \langle Y_t, \psi\rangle e^{-c_9 t} - c_{10}t$, we can prove uniform integrability (in $r > 0$) for both sides of the last inequality. Therefore

$$E_{\mathscr{L}(X^N)}[\langle Y_t, \psi\rangle \mid \mathscr{F}_s] \geq \langle Y_s, \psi\rangle - E_{\mathscr{L}(X^N)}\left[\int_s^t du\langle Y_u, \psi\rangle c_{14} \,\middle|\, \mathscr{F}_s\right] - c_{15}(t-s).$$

Another application of Gronwall's Lemma finishes the proof of Lemma 2.

LEMMA 3. *There exists a constant c_{16}, such that for all $0 < \Delta$, $t \leq T - \Delta$ and $N \in \mathbb{N}$*

$$E_{\mathscr{L}(X^N)}[\|Y_{t+\Delta} - Y_t\|_1 \mid \mathscr{F}_t] \leq \sqrt{\Delta}(\langle Y_t, \psi\rangle + 1)c_{16}.$$

PROOF. We choose an enumeration of the N particles. By $X_i(t)$ we denote the position of particle i at time t. Then we obtain

$$\begin{aligned}
&E_{\mathscr{L}(X^N)}[\| Y_t - Y_{t+\Delta} \|_1 \mid \mathscr{F}_t] \\
&\le \frac{1}{N} \textstyle\sum_{i=1}^N E_{\mathscr{L}(X^N)}[\,| X_i(t+\Delta) - X_i(t) |\mid \mathscr{F}_t] \\
&\le \frac{1}{N} \textstyle\sum_{i=1}^N (E_{\mathscr{L}(X^N)}[\,| X_i(t+\Delta) - X_i(t) |^2 \mid \mathscr{F}_t])^{1/2} \\
&= \frac{1}{N} \textstyle\sum_{i=1}^N \left(E_{\mathscr{L}(X^N)}\left[\int_t^{t+\Delta} ds\{2b(X_i(s), Y_s) \cdot (X_i(s) - X_i(t)) \right.\right. \\
&\qquad\qquad \left.\left. + \textstyle\sum_{k=1}^n a_{kk}(X_i(s), Y_s)\} \,\middle|\, \mathscr{F}_t \right]\right)^{1/2}
\end{aligned} \tag{3.1}$$

(by Dynkin's Formula)

$$\begin{aligned}
&\le \frac{1}{N} \textstyle\sum_{i=1}^N \left(E_{\mathscr{L}(X^N)}\left[\int_t^{t+\Delta} ds\{| X_i(s) |^2 + | X_i(t) |\, | X_i(s) | \right.\right. \\
&\qquad \left.\left. + (\langle Y_s, \kappa \rangle + 1)(| X_i(t) | + | X_i(s) | + 1)\} \,\middle|\, \mathscr{F}_t \right]\right)^{1/2} c_{17} \\
&\le \frac{1}{N} \textstyle\sum_{i=1}^N \left(E_{\mathscr{L}(X^N)}\left[\int_t^{t+\Delta} ds\{| X_i(t) |^2 + | X_i(s) |^2 \right.\right. \\
&\qquad \left.\left. + \langle Y_s, \psi \rangle + 1\} \,\middle|\, \mathscr{F}_t \right]\right)^{1/2} c_{18}.
\end{aligned}$$

Using similar methods as in the proof of Lemma 2, we can show that for $s \ge t$

$$E_{\mathscr{L}(X^N)}[\,| X_i(s) |^2 \mid \mathscr{F}_t] \le (| X_i(t) |^2 + \langle Y_t, \psi \rangle + 1) c_{19}.$$

Using this and Lemma 2, we obtain from (3.1)

$$\begin{aligned}
E_{\mathscr{L}(X^N)}[\| Y_{t+\Delta} - Y_t \|_1 \mid \mathscr{F}_t] &\le (1/N) \textstyle\sum_{i=1}^N (\Delta(| X_i(t) |^2 + \langle Y_t, \psi \rangle + 1))^{1/2} c_{20} \\
&\le (1/N) \textstyle\sum_{i=1}^N \{| X_i(t) |^2 + \langle Y_t, \psi \rangle + 1\} c_{20} \Delta^{1/2}.
\end{aligned}$$

Since $(1/N) \sum_{i=1}^N | X_i(t) |^2 = \langle Y_t, \psi \rangle$, the proof of Lemma 3 is finished.

LEMMA 4. *For each $\Delta > 0$ there exists a random variable $\gamma(\Delta) \ge 0$, such that*

$$E_{\mathscr{L}(X^N)}[\| Y_{s+\Delta} - Y_s \|_1 \mid \mathscr{F}_s] \le E_{\mathscr{L}(X^N)}[\gamma(\Delta) \mid \mathscr{F}_s]$$

$\mathscr{L}(X^N)$ a.s. for all $N \in \mathbb{N}$ and $s \in [0, T - \Delta]$. Furthermore we have

$$\lim_{\Delta \to 0} \lim_{N \to \infty} E_{\mathscr{L}(X^N)}[\gamma(\Delta)] = 0.$$

PROOF. From the foregoing two lemmata we obtain

$$E_{\mathscr{L}(X^N)}[\| Y_{s+\Delta} - Y_s \|_1 \,|\, \mathscr{F}_s] \le \sqrt{\Delta}\langle Y_s, \psi + 1\rangle c_{16}$$
$$\le \sqrt{\Delta} E_{\mathscr{L}(X^N)}[\langle Y_T, \psi + 1\rangle \,|\, \mathscr{F}_s] c_{21}.$$

Thus Lemma 4 is valid for

$$\gamma(\Delta) = \langle Y_T, \psi + 1\rangle \sqrt{\Delta} c_{21}.$$

LEMMA 5. *For all $\varepsilon > 0$ there exists a compact set $\mathscr{K}_\varepsilon$ in $(\mathscr{M}, \|\cdot\|_1)$, such that* $\inf_{N\in\mathbb{N}} P_{\mathscr{L}(X^N)}[Y_t \in \mathscr{K}_\varepsilon\ \forall t \le T] \ge 1 - \varepsilon$.

PROOF. By using the semimartingales of Lemma 2, we obtain for each set

$$B_\lambda^c = \{x \in \mathbb{R}^n : |x| > \lambda\}, \quad \lambda > 0 \quad \text{and} \quad \varepsilon > 0$$

$$P_{\mathscr{L}(X^N)}[\sup_{t\le T} \| Y_t \,|\, B_\lambda^c \|_0 > \varepsilon] \le P_{\mathscr{L}(X^N)}[\sup_{t\le T} \langle Y_t, \psi\rangle > \lambda^2 \varepsilon]$$
$$\le P_{\mathscr{L}(X^N)}[\sup_{t\le T}\{\langle Y_t, \psi\rangle e^{c_7 t} + c_8 t\} > \lambda^2\varepsilon]$$
$$\le (1/\lambda^2\varepsilon) E_{\mathscr{L}(X^N)}[\langle Y_T, \psi\rangle e^{c_7 T} + c_8 T]$$
(by Doob's Inequality)
$$\le (1/\lambda^2\varepsilon) E_{\mathscr{L}(X^N)}[(\langle Y_0, \psi\rangle + c_{10} T) e^{(c_9 + c_7)T} + c_8 T].$$

For fixed $\varepsilon > 0$ this expression is less than a given $\delta > 0$, if λ is sufficiently large.

Combining this result, Lemma 2a and Prohorov's Theorem (cf. [1]), we obtain Lemma 5.

By a theorem of Kurtz ([2], page 10, Theorem 2.7), Lemma 4 and Lemma 5 imply:

LEMMA 6. *The sequence of distributions $\{\mathscr{L}(X^N)\}_{N\in\mathbb{N}}$ is relatively compact in the space $\mathscr{P}(\Omega_1)$ of all probability measures on Ω_1, where $\mathscr{P}(\Omega_1)$ is equipped with the weak topology.*

Our next aim is to characterize the possible limit points $\mathscr{L}$ of the sequence $(\mathscr{L}(X^N)_{N\in\mathbb{N}})$.

LEMMA 7. *For each limit point $\mathscr{L}$, $\Theta \in \mathbb{R}$ and $f \in C_c^2(\mathbb{R}^n)$ the process*

$$\tilde{E}_\theta^f = (\tilde{E}_\theta^f(t))_{t\le T},$$

$$\tilde{E}_\theta^f(t)(Y) = \exp\left(i\Theta\left(\langle Y_t, f\rangle - \int_0^t ds \left\{\langle Y_s, b(\cdot, Y_s)\cdot \nabla f\rangle + \frac{1}{2}\langle Y_s, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_s)\partial_{ij} f\rangle\right\}\right)\right)$$

is a martingale relative to $(\Omega_1, \mathscr{F}_T, (\mathscr{F}_t)_{t\le T}, \mathscr{L})$.

PROOF. Let us first give a heuristic derivation of the result of Lemma 7. When making a formal limit in Lemma 1 we presume the martingale property of the process

$$t \to f(\langle Y_t, g\rangle) - \int_0^t f'(\langle Y_s, g\rangle)(A^\infty g)(Y_s)\, ds$$

relative to $(\Omega_1, \mathscr{F}_T, (\mathscr{F}_t)_{t\le T}, \mathscr{L})$ for all $f \in C_b^2(\mathbb{R})$, $g \in C_c^2(\mathbb{R}^n)$, *where*

$$(A^\infty g)(\mu) = \langle \mu, b(\cdot, \mu) \cdot \nabla g\rangle + ½\langle \mu, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, \mu)\partial_{ij} g\rangle.$$

Using Theorem 4.2.1 of [3] about the equivalence of certain martingales, we infer the martingale property of

$$t \to \exp\left(i\Theta\left(\langle Y_t, g\rangle - \int_0^t (A^\infty g)(Y_s)\, ds\right)\right)$$

for all $\Theta \in \mathbb{R}$.

An exact proof of Lemma 7 can be given as follows: Let

$$\eta_{\theta,f,t}(Y) = \exp\left(-i\Theta \int_0^t du \left\{\langle Y_u, b(\cdot, Y_u) \cdot \nabla f\rangle + \frac{1}{2}\langle Y_u, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_u)\partial_{ij} f\rangle\right\}\right).$$

Moreover we define the process $\tilde{E}_\theta^{f,N} = (\tilde{E}_\theta^{f,N}(t))_{t\le T}$ by

$$(*) \qquad \tilde{E}_\theta^{f,N}(t) = E_{g_\theta,f}^N(t)\eta_{\theta,f,t} - \int_0^t E_{g_\theta,f}^N(s)\eta_{\theta,f,ds},$$

where $g_\theta(x) = \exp(i\Theta x)$, $\eta_{\theta,f,ds} = ((d/ds)\eta_{\theta,f,s})\, ds$ and $E_{g_\theta,f}^N$ is as in Lemma 1. Then we obtain

$$\begin{aligned}
&\tilde{E}_\Theta^{f,N}(t)(Y) \\
&= \tilde{E}_\Theta^f(t)(Y) - \eta_{\Theta,f,t}(Y) \int_0^t ds \left\{ i\Theta \exp(i\Theta\langle Y_s, f\rangle)(A^\infty f)(Y_s) \right. \\
&\qquad \left. - \frac{\Theta^2}{2N} \exp(i\Theta\langle Y_s, f\rangle)\langle Y_s, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_s)\partial_i f \partial_j f\rangle \right\} \\
(7.1)\quad &- \int_0^t \left\{ \exp(i\Theta\langle Y_s, f\rangle) - \int_0^s du \left\{ i\Theta \exp(i\Theta\langle Y_u, f\rangle)(A^\infty f)(Y_u) \right.\right. \\
&\qquad \left.\left. - \frac{\Theta^2}{2N} \exp(i\Theta\langle Y_u, f\rangle)\langle Y_u, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_u)\partial_i f \partial_j f\rangle \right\}\right\} \\
&\qquad \cdot (-i\Theta)(A^\infty f)(Y_s) \cdot \eta_{\Theta,f,s}(Y)\, ds.
\end{aligned}$$

Using integration by parts we see that the third term in (7.1) equals

$$
\begin{aligned}
&+\int_0^t ds\{i\Theta \exp(i\Theta\langle Y_s, f\rangle)(A^\infty f)(Y_s)\eta_{\Theta,f,s}(Y)\}\\
&-\int_0^t ds\Big\{i\Theta \exp(i\Theta\langle Y_s, f\rangle)(A^\infty f)(Y_s)\\
&\qquad -\frac{\Theta^2}{2N}\exp(i\Theta\langle Y_s, f\rangle)\langle Y_s, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_s)\partial_i f\partial_j f\rangle\Big\}\eta_{\Theta,f,s}(Y)\\
&+\eta_{\Theta,f,t}(Y)\int_0^t ds\Big\{i\Theta \exp(i\Theta\langle Y_s, f\rangle)(A^\infty f)(Y_s)\\
&\qquad -\frac{\Theta^2}{2N}\exp(i\Theta\langle Y_s, f\rangle)\langle Y_s, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_s)\partial_i f\partial_j f\rangle\Big\}.
\end{aligned}
\tag{7.2}
$$

(7.1) and (7.2) yield

$$
\begin{aligned}
\tilde{E}_\theta^{f,N}(t)(Y) &= \tilde{E}_\theta^f(t)(Y)\\
&\quad+\int_0^t ds\tilde{E}_\theta^f(s)(Y)\frac{\Theta^2}{2N}\langle Y_s, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_s)(\partial_i f)(\partial_j f)\rangle.
\end{aligned}
\tag{7.3}
$$

Using representation (*) of $\tilde{E}_\theta^{f,N}$, Lemma 1, Lemma 2 and Theorem 1.2.8 in [3], we infer the martingale property of $\tilde{E}_\theta^{f,N}$ relative to $(\Omega_1, \mathscr{F}_T, (\mathscr{F}_t)_{t\le T}, \mathscr{L}(X^N))$. Since $|\tilde{E}_\Theta^f(s)| \le 1$, we can conclude from Lemma 2 and (7.3)

$$
\lim_{N\to\infty} E_{\mathscr{L}(X^N)}[|\tilde{E}_\theta^{f,N}(t) - \tilde{E}_\theta^f(t)|] = 0, \quad \forall t \le T. \tag{7.4}
$$

The following steps are necessary since we are working in $\mathscr{D}([0, T], (\mathscr{M}, \|\cdot\|_1))$ and not in a space of continuous functions. This especially implies that for fixed $f \in C_c^2(\mathbb{R}^n)$ *and* $t \le T$, the function $Y \to \langle Y_t, f\rangle$ is not continuous.

We define for fixed $h \in C_b([0, T])$, $N \in \mathbb{N}$, $\Theta \in \mathbb{R}$, $f \in C_c^2(\mathbb{R}^n)$ and $\phi \in C_b(\Omega_1)$ the realvalued functions

$$
Z_{\theta,f,h}^{N,\phi}(Y) = \int_0^T h(t)\tilde{E}_\theta^{f,N}(t)(Y)\phi(Y)\,dt
$$

and

$$
Z_{\theta,f,h}^{\phi}(Y) = \int_0^T h(t)\tilde{E}_\Theta^{f}(t)(Y)\phi(Y)\,dt \quad \text{on} \quad \Omega_1.
$$

Obviously $Z_{\theta,f,h}^\phi \in C_b(\Omega_1)$, and so we have for each subsequence $(\mathscr{L}(X^{N'}))_{N'\in I\subseteq\mathbb{N}}$, which converges to $\mathscr{L}$ in $\mathscr{P}(\Omega_1)$

$$
\lim_{N'\to\infty} E_{\mathscr{L}(X^{N'})}[Z_{\theta,f,h}^\phi] = E_{\mathscr{L}}[Z_{\theta,f,h}^\phi]
$$

and by (7.4)

$$
\lim_{N'\to\infty} E_{\mathscr{L}(X^{N'})}[Z_{\theta,f,h}^{N',\phi}] = E_{\mathscr{L}}[Z_{\theta,f,h}^\phi]. \tag{7.5}
$$

This is valid for each $h \in C_b([0, T])$. Moreover, Lemma 3 implies the equicontinuity of the functions

$$\{t \to E_{\mathscr{L}(X^N)}[\tilde{E}_\theta^{f,N}(t)\phi]\}_{N\in\mathbb{N}},$$

Therefore (7.5) yields

$$\lim_{N'\to\infty} E_{\mathscr{L}(X^{N'})}[\tilde{E}_\theta^{f,N'}(t)\phi] = E_{\mathscr{L}}[\tilde{E}_\theta^f(t)\phi], \quad \forall t \le T.$$

Now let ϕ be $\mathscr{F}_s$-measurable. Then for $t \ge s$

$$E_{\mathscr{L}}[\tilde{E}_\theta^f(t)\phi] = \lim_{N'\to\infty} E_{\mathscr{L}(X^{N'})}[\tilde{E}_\theta^{f,N'}(t)\phi] = \lim_{N'\to\infty} E_{\mathscr{L}(X^{N'})}[\tilde{E}_\Theta^{f,N'}(s)\phi]$$

(by the martingale property of $\tilde{E}_\theta^{f,N'}$ with respect to $(\Omega_1, \mathscr{F}_T, (\mathscr{F}_t)_{t\le T}, \mathscr{L}(X^{N'}))$

$$= E_{\mathscr{L}}[\tilde{E}_\theta^f(s)\phi]).$$

This completes the proof of Lemma 7.

LEMMA 8. *The process* $\eta = (\eta_t)_{t\le T}$ *defined by*

$$\eta_t(f, Y) = \langle Y_t, f\rangle - \int_0^t ds \left\{ \langle Y_s, b(\cdot, Y_s)\cdot\nabla f\rangle + \frac{1}{2}\langle Y_s, \textstyle\sum_{i,j=1}^n a_{ij}(\cdot, Y_s)\partial_{ij} f\rangle \right\};$$
$$t \le T, \quad f \in C_c^2(\mathbb{R}^n)$$

satisfies

$$\eta_t(f, Y) = \eta_0(f, Y) = \langle X_0^*, f\rangle, \quad \forall t \le T, \quad \forall f \in C_c^2(\mathbb{R}^n) \quad \mathscr{L}\text{—a.s.}$$

PROOF. Using the foregoing lemma, we may compute the characteristic function of the random variable $\eta_t(f, \cdot)$ (t, f fixed)

$$E_{\mathscr{L}}[\exp(i\theta\eta_t(f, Y))] = E_{\mathscr{L}}[\exp(i\theta\eta_0(f, Y))]$$
$$= \exp(i\theta\langle X_0^*, f\rangle), \quad \text{i.e.} \quad \eta_t(f, Y) = \langle X_0^*, f\rangle \quad \mathscr{L}\text{—a.s.}$$

Consequently we have:

$$(8.1) \qquad \eta_t(f, Y) = \langle X_0^*, f\rangle \quad \forall t \in \mathscr{T} \quad \text{and} \quad \forall f \in \mathscr{D} \quad \mathscr{L}\text{—a.s.},$$

where $\mathscr{T}$ (resp. $\mathscr{D}$) is any countable subset of $[0, T]$ (resp. $C_c^2(\mathbb{R}^n)$).

Since $C_c^2(\mathbb{R}^n)$ is a separable space if equipped with the norm defined by the supremum of the function and its first two derivatives and since $t \to \eta_t(f, \cdot)$ is right-continuous, which is a consequence of the right-continuity of $t \to Y_t$, (8.1) implies the conclusion of Lemma 8.

At this stage the following remark should be made: From Lemma 8, we know that $\mathscr{L}$ is concentrated on that subset M of Ω_1, which consists of solutions $Y = (Y_t)_{t\le T}$ of

$$(8.2) \quad \langle Y_t, f\rangle = \langle X_0^*, f\rangle + \int_0^t \langle Y_s, L(Y_s)f\rangle\, ds, \quad 0 \le t \le T, \quad f \in C_c^2(\mathbb{R}^n),$$

with

$$(L(\mu)f)(x) = \sum_{i=1}^n b_i(x, \mu)\partial_i f(x) + \tfrac{1}{2} \sum_{i,j=1}^n a_{ij}(x, \mu)\partial_{ij} f(x).$$

If we would impose in addition the assumption that the partial derivatives of $b_k(x, \mu)$ and $\sigma_{k,\ell}(x, \mu)$ of first and second order with respect to x exist, are continuous and uniformly bounded in (x, μ), then we could argue as in Lemma 10 below, show uniqueness of the solution of $(Y_t)_{t\le T}$ of (8.2), and $Y_t = \mathscr{L}(\zeta_t)$ for the process ζ defined in (9.2), and the proof of Theorem 1 would be finished. If the above differentiability assumptions are not satisfied, we can introduce smooth modifications $b^\varepsilon(x, \mu)$ and $\sigma^\varepsilon(x, \mu)$ of $b(x, \mu)$ and $\sigma(x, \mu)$, and processes $X^{N,\varepsilon}$, controlled by those coefficients, and finally make the limit $\varepsilon \to 0$. This is done in the following lemma.

LEMMA 9. *For any limit point $\mathscr{L}$ of the sequence $\{\mathscr{L}(X^N)\}_{N\in\mathbb{N}}$ we have*

(9.1) $$Y_t = \mathscr{L}(\zeta_t) \quad \forall t \le T \quad \mathscr{L}\text{—a.s.},$$

where $\mathscr{L}(\zeta_t)$ is the distribution of the unique solution of

(9.2) $$d\zeta_t = b(\zeta_t, \mathscr{L}(\zeta_t))dt + \sigma(\zeta_t, \mathscr{L}(\zeta_t))d\tilde{W}_t,$$

(9.3) $$\mathscr{L}(\zeta_0) = X_0^*.$$

($\tilde{W} = (\tilde{W}_t)_{t\ge 0}$ n-dimensional Brownian motion).

PROOF. We have for any $Y \in M$ and any $\tilde{\psi}_r \in C_c^2(\mathbb{R}^n)$ with $\tilde{\psi}_r(x) = \psi(x) = x^2$ for $|x| \le r$, $|\partial_i \tilde{\psi}_r(x)| \le 2|x|$ and $|\partial_{ij}\tilde{\psi}_r(x)| \le 2$ $(i, j = 1, 2, \cdots, n;\ x \in \mathbb{R}^n)$

$$\begin{aligned}
&\langle Y_t, \tilde{\psi}_r\rangle \\
&= \langle X_0^*, \tilde{\psi}_r\rangle + \int_0^t \langle Y_s, L(Y_s)\tilde{\psi}_r\rangle\, ds \\
&\le \langle X_0^*, \tilde{\psi}_r\rangle + \int_0^t \langle Y_s, (1 + \kappa + \langle Y_s, \kappa\rangle)2\kappa + 1\rangle c_{11}\, ds \quad \text{(by (A1) and (A2))} \\
&\le \langle X_0^*, \psi\rangle + c_{12}\int_0^t (\langle Y_s, \psi\rangle + 1)\, ds.
\end{aligned}$$

Fatou's Lemma yields

$$\langle Y_t, \psi\rangle \le \liminf_{r\to\infty}\langle Y_t, \tilde{\psi}_r\rangle \le \langle X_0^*, \psi\rangle + c_{12}\int_0^t (\langle Y_s, \psi\rangle + 1)\, ds.$$

Therefore we obtain by Gronwall's Lemma

(9.4) $$\sup_{t\le T}\langle Y_t, \psi\rangle \le (\langle X_0^*, \psi\rangle + c_{12}T)\exp(c_{12}T) := c_{13} < \infty$$

uniformly in $Y \in M$.

From (9.4) we obtain for any $Y \in M$ and $f \in C_c^2(\mathbb{R}^n)$

$$
\begin{aligned}
&|\langle Y_t, f\rangle - \langle Y_s, f\rangle| \\
&= \left| \int_s^t \langle Y_u, L(Y_u)f\rangle \, du \right| \\
&\le c_{14} \int_s^t \langle Y_u, (1 + \kappa + \langle Y_u, \kappa\rangle)\sup_{i=1,\dots,n,x\in \mathbb{R}^n} |\partial_i f(x)| \\
&\qquad + \sup_{i,j=1,\dots,n,x\in\mathbb{R}^n} |\partial_{ij} f(x)|\rangle \, du \qquad (9.5)\\
&\le c(f) \int_s^t (\langle Y_u, \psi\rangle + 1) \, du
\end{aligned}
$$

(the constant $c(f)$ depends on the first and second partial derivatives of f)

$$\le |t - s|(c_{13} + 1)c(f) \quad \text{(by (9.4))},$$

i.e. the functions $t \to \langle Y_t, f\rangle$ ($Y \in M, f \in C_c^2(\mathbb{R}^n)$) are Lipschitz continuous. This implies that for each $t \in [0, T]$ the left and right limits $Y_{t^-} = \lim_{s\nearrow t} Y_s$ (resp. $Y_{t^+} = \lim_{s\searrow t} Y_s$) coincide with Y_t. We have obtained:

(9.6) For any $Y \in M$ the function $t \to Y_t$ is a continuous function from $[0, T]$ to $(\mathcal{M}, \|\cdot\|_1)$.

Let us define now the functions

$$(x, \mu) \to \int_{\mathbb{R}^n} b(x - y, \mu) g_\varepsilon(y)\, dy =: b^\varepsilon(x, \mu)$$

and

$$(x, \mu) \to \int_{\mathbb{R}^n} \sigma(x - y, \mu) g_\varepsilon(y)\, dy =: \sigma^\varepsilon(x, \mu) \qquad (9.7)$$

with $g_\varepsilon(y) = (2\pi\varepsilon)^{-n/2}\exp(-(y^2/2\varepsilon))$.

Using b^ε and σ^ε as drift vector (resp. diffusion matrix) we may define N-particle-processes $X^{N,\varepsilon}$ in exactly the same way as we have introduced the processes X^N. When we assume $X_0^{N,\varepsilon} = X_0^N$, we obtain all the results proved so far for the sequence $\{X^N\}_{N\in\mathbb{N}}$ for the sequence $\{X^{N,\varepsilon}\}_{N\in\mathbb{N}}$ too. In particular we have the tightness of the sequence $\{\mathcal{L}(X^{N,\varepsilon})\}_{N\in\mathbb{N}}$ and

$$
\begin{aligned}
&\eta_t^\varepsilon(f, Y) := \langle Y_t, f\rangle - \int_0^t \langle Y_u, L^\varepsilon(Y_u)f\rangle \, du = \langle X_0^*, f\rangle \\
&\qquad\qquad \forall t \in T, \quad \forall f \in C_c^2(\mathbb{R}^n) \quad \mathcal{L}^\varepsilon\text{—a.s.}
\end{aligned} \qquad (9.8)
$$

for all limit points $\mathcal{L}^\varepsilon$ of $\{\mathcal{L}(X^{N,\varepsilon})\}_{N\in\mathbb{N}}$. Here we set

$$(L^\varepsilon(\mu)f)(x) = \sum_{i=1}^n b_i^\varepsilon(x, \mu)\partial_i f(x) + \tfrac{1}{2} \sum_{i,j=1}^n a_{ij}^\varepsilon(x, \mu)\partial_{ij} f(x) \qquad (9.9)$$

with $a^\varepsilon(x, \mu) = \sigma^\varepsilon(x, \mu) \cdot \sigma^{\varepsilon T}(x, \mu)$.

Using Lemma 10 below, we obtain

$$\mathscr{L}^\varepsilon = \delta_{X^{*,\varepsilon}}, \tag{9.10}$$

where $X^{*,\varepsilon} = (X_t^{*,\varepsilon})_{t\le T}$, $X_t^{*,\varepsilon} = \mathscr{L}(\zeta_t^\varepsilon)$ for the uniquely existing nonlinear diffusion process ζ^ε, which solves the stochastic differential equation

$$d\zeta_t^\varepsilon = b^\varepsilon(\zeta_t^\varepsilon, \mathscr{L}(\zeta_t^\varepsilon))\,dt + \sigma^\varepsilon(\zeta_t^\varepsilon, \mathscr{L}(\zeta_t^\varepsilon))\,dW_t, \quad \mathscr{L}(\zeta_0^\varepsilon) = X_0^*. \tag{9.11}$$

Let us assume now that particles i in both processes X^N and $X^{N,\varepsilon}$ are driven by the same Brownian motion W_t^i, i.e.

$$\begin{aligned} dX_i^N(t) &= b(X_i^N(t), (1/N)\textstyle\sum_{j=1}^N \delta_{X_j^N(t)})\,dt \\ &\quad + \sigma(X_i^N(t), (1/N)\textstyle\sum_{j=1}^N \delta_{X_j^N(t)})\,dW_t^i \end{aligned} \tag{9.12}$$

and

$$\begin{aligned} dX_i^{N,\varepsilon}(t) &= b^\varepsilon(X_i^{N,\varepsilon}(t), (1/N)\textstyle\sum_{j=1}^N \delta_{X_j^{N,\varepsilon}(t)})\,dt \\ &\quad + \sigma^\varepsilon(X_i^{N,\varepsilon}(t), (1/N)\textstyle\sum_{j=1}^N \delta_{X_i^{N,\varepsilon}(t)})\,dW_t^i, \quad (i = 1, 2, \cdots, n). \end{aligned} \tag{9.13}$$

This assumption is allowed because by the Lipschitz continuity of the functions $(x, \mu) \to b(x, \mu)$, $(x, \mu) \to \sigma(x, \mu)$, $(x, \mu) \to b^\varepsilon(x, \mu)$ and $(x, \mu) \to \sigma^\varepsilon(x, \mu)$ both systems of equations (9.12) and (9.13) have a unique strong solution. Moreover we assume

$$X_i^{N,\varepsilon}(0) = X_i^N(0), \quad i = 1, 2, \cdots, n \quad \text{a.s.}$$

Then Ito's Formula yields

$$\begin{aligned} &E[(X_i^N(t) - X_i^{N,\varepsilon}(t))^2] \\ &= E\Bigg[\int_0^t \Bigg\{ 2(X_i^N(s) - X_i^{N,\varepsilon}(s)) \\ &\qquad \cdot \Bigg(b\Big(X_i^N(s), \frac{1}{N}\textstyle\sum_{j=1}^N \delta_{X_j^N(s)}\Big) - b^\varepsilon\Big(X_i^{N,\varepsilon}(s), \frac{1}{N}\sum_{j=1}^N \delta_{X_j^{N,\varepsilon}(s)}\Big)\Bigg)\Bigg\}\,ds\Bigg] \\ &\quad + E\Bigg[\int_0^t \textstyle\sum_{k=1}^n \tilde{a}_{kk}^\varepsilon\Big(X_i^N(s), X_i^{N,\varepsilon}(s), \frac{1}{N}\sum_{j=1}^N \delta_{X_j^N(s)}, \frac{1}{N}\sum_{j=1}^N \delta_{X_j^{N,\varepsilon}(s)}\Big)\,ds\Bigg] \end{aligned} \tag{9.14}$$

with $\tilde{a}^\varepsilon(x, y, \mu, \nu) = (\sigma(x, \mu) - \sigma^\varepsilon(y, \nu))(\sigma(x, \mu) - \sigma^\varepsilon(y, \nu))^T$. From (A1) we obtain

$$\begin{aligned} |b^\varepsilon(x, \mu) - b(x, \mu)| &= \left|\int_{\mathbb{R}^n} (b(y, \mu) - b(x, \mu))g_\varepsilon(x - y)\,dy\right| \\ &\le \int_{\mathbb{R}^n} |b(y, \mu) - b(x, \mu)|\,g_\varepsilon(x - y)\,dy \\ &\le c_{14}\int_{\mathbb{R}^n} |x - y|\,g_\varepsilon(x - y)\,dy| \le c_{14}\sqrt{\varepsilon}. \end{aligned} \tag{9.15}$$

Similarly we have

$$|\sigma_{ij}^{\varepsilon}(x, \mu) - \sigma_{ij}(x, \mu)| \le c_{15}\sqrt{\varepsilon}. \tag{9.16}$$

Using (9.14), (9.15), (9.16), (A1) and (A2) we obtain

$$\begin{aligned} &E[(X_i^N(t) - X_i^{N,\varepsilon}(t))^2] \\ &\le c_{16}\Bigg\{E\Bigg[\int_0^t |X_i^N(s) - X_i^{N,\varepsilon}(s)| \\ &\qquad \Bigg\{|X_i^N(s) - X_i^{N,\varepsilon}(s)| + \Bigg\| \frac{1}{N}\sum_{j=1}^N \delta_{X_j^N(s)} - \frac{1}{N}\sum_{j=1}^N \delta_{X_j^{N,\varepsilon}(s)} \Bigg\|_1 + \sqrt{\varepsilon}\Bigg\}\, ds\Bigg] \\ &\qquad + E\Bigg[\int_0^t \Bigg\{|X_i^N(s) - X_i^{N,\varepsilon}(s)|^2 \\ &\qquad\qquad + \Bigg\| \frac{1}{N}\sum_{j=1}^N \delta_{X_j^N(s)} - \frac{1}{N}\sum_{j=1}^N \delta_{X_j^{N,\varepsilon}(s)} \Bigg\|_1^2 + \varepsilon\Bigg\}\, ds\Bigg]\Bigg\}. \end{aligned} \tag{9.17}$$

Summing over $i = 1, \cdots, N$, noting that

$$\begin{aligned} &\| (1/N) \textstyle\sum_{j=1}^N \delta_{X_j^N(s)} - (1/N) \sum_{j=1}^N \delta_{X_j^{N,\varepsilon}(s)} \|_1 \\ &\qquad = \sup_{f\in\mathscr{A}_1} |(1/N) \textstyle\sum_{j=1}^N \{f(X_j^N(s)) - f(X_j^{N,\varepsilon}(s))\}| \\ &\qquad \le (1/N) \textstyle\sum_{j=1}^N |X_j^N(s) - X_j^{N,\varepsilon}(s)| \\ &\qquad \le ((1/N) \textstyle\sum_{j=1}^N |X_j^N(s) - X_j^{N,\varepsilon}(s)|^2)^{1/2}, \end{aligned} \tag{9.18}$$

and using the inequality $|uv| \le u^2 + v^2$, we get from (9.17)

$$\begin{aligned} &E\Bigg[\frac{1}{N}\sum_{i=1}^N |X_i^N(t) - X_i^{N,\varepsilon}(t)|^2\Bigg] \\ &\qquad \le c_{17} \int_0^t \Bigg\{E\Bigg[\frac{1}{N}\sum_{i=1}^N |X_i^N(s) - X_i^{N,\varepsilon}(s)|^2\Bigg] + \varepsilon\Bigg\}\, ds, \end{aligned}$$

and therefore by Gronwall's Lemma

$$E[(1/N) \textstyle\sum_{i=1}^N |X_i^N(t) - X_i^{N,\varepsilon}(t)|^2] \le (c_{17}\varepsilon T)\exp(c_{17}T). \tag{9.19}$$

Furthermore we have by (9.12) and (9.13)

$$
\begin{aligned}
&E\left[\sup_{t\le T}\frac{1}{N}\sum_{i=1}^N |X_i^N(t)-X_i^{N,\varepsilon}(t)|^2\right]\\
&= E\left[\sup_{t\le T}\frac{1}{N}\sum_{i=1}^N \left|\int_0^t\left(b\left(X_i^N(s),\frac{1}{N}\sum_{j=1}^N\delta_{X_j^N(s)}\right)\right.\right.\right.\\
&\qquad \left.-\,b^\varepsilon\left(X_i^{N,\varepsilon}(s),\frac{1}{N}\sum_{j=1}^N\delta_{X_j^{N,\varepsilon}(s)}\right)\right)ds\\
&\qquad +\int_0^t\left(\sigma\left(X_i^N(s),\frac{1}{N}\sum_{j=1}^N\delta_{X_j^N(s)}\right)\right.\\
&\qquad \left.\left.\left.-\,\sigma^\varepsilon\left(X_i^{N,\varepsilon}(s),\frac{1}{N}\sum_{j=1}^N\delta_{X_j^{N,\varepsilon}(s)}\right)\right)dW_s^i\right|^2\right]\\
&\le c_{18}\left\{E\left[\frac{1}{N}\sum_{i=1}^N\int_0^T\left|b\left(X_i^N(s),\frac{1}{N}\sum_{j=1}^N\delta_{X_j^N(s)}\right)\right.\right.\right.\\
&\qquad \left.\left.-\,b^\varepsilon\left(X_i^{N,\varepsilon}(s),\frac{1}{N}\sum_{j=1}^N\delta_{X_j^{N,\varepsilon}(s)}\right)\right|^2ds\right]\\
&\qquad + E\left[\sup_{t\le T}\frac{1}{N}\sum_{i=1}^N\left(\int_0^t\left(\sigma\left(X_i^N(s),\frac{1}{N}\sum_{j=1}^N\delta_{X_j^N(s)}\right)\right.\right.\right.\\
&\qquad\qquad \left.\left.\left.\left.-\,\sigma^\varepsilon\left(X_i^{N,\varepsilon}(s),\frac{1}{N}\sum_{j=1}^N\delta_{X_j^{N,\varepsilon}(s)}\right)\right)dW_s^i\right)^2\right]\right\} \qquad (9.20)\\
&\le c_{19}\left\{E\left[\frac{1}{N}\sum_{i=1}^N\int_0^T\left\{\left|X_i^N(s)-X_i^{N,\varepsilon}(s)\right|^2\right.\right.\right.\\
&\qquad\qquad \left.\left.+\frac{1}{N}\sum_{j=1}^N|X_j^N(s)-X_j^{N,\varepsilon}(s)|^2+\varepsilon\right\}ds\right]\\
&\qquad + E\left[\frac{1}{N}\sum_{i=1}^N\int_0^T\sum_{k=1}^n\tilde{a}_{kk}^\varepsilon\left(X_i^N(s),X_i^{N,\varepsilon}(s),\right.\right.\\
&\qquad\qquad \left.\left.\left.\frac{1}{N}\sum_{j=1}^N\delta_{X_j^N(s)},\frac{1}{N}\sum_{j=1}^N\delta_{X_j^{N,\varepsilon}(s)}\right)ds\right]\right\}
\end{aligned}
$$

(by (A1), (9.15), (9.18) and Doob's Inequality)

$$
\le c_{20}\frac{1}{N}\sum_{i=1}^N\int_0^T\{E[|X_i^N(s)-X_i^{N,\varepsilon}(s)|^2]+\varepsilon\}\,ds
$$

(by (A2), (9.16) and (9.18))

$$
\le c_{21}\varepsilon \quad \text{(by (9.19))}.
$$

Consequently by (9.18) and (9.20):

$$
\begin{aligned}
E[\sup_{t\le T}\| X_t^N - X_t^{N,\varepsilon}\|_1] &\le E[\sup_{t\le T}(1/N)\textstyle\sum_{j=1}^N |X_j^N(t) - X_j^{N,\varepsilon}(t)|] \\
&\le (E[\sup_{t\le T}(1/N)\textstyle\sum_{j=1}^N |X_j^N(t) - X_j^{N,\varepsilon}(t)|^2])^{1/2} \qquad (9.21)\\
&\le \sqrt{c_{21}}\sqrt{\varepsilon} \quad \text{(uniformly in } N\in\mathbb{N}).
\end{aligned}
$$

In the same way as uniqueness of the solution of (9.11) (cf. Lemma 10) we may show uniqueness of the solution of (9.2), (9.3). Moreover, by similar arguments that lead to (9.21), we may prove that

$$(9.22)\qquad \lim_{\varepsilon\to 0}\mathscr{L}(\zeta^\varepsilon) = \mathscr{L}(\zeta) \quad \text{in} \quad \mathscr{P}(\mathscr{D}([0,T],\mathbb{R}^n)).$$

Using (9.10), (9.21) and (9.22) we finish the proof of Lemma 9.

LEMMA 10. *Let L^ε be defined as in* (9.9). *Then the system of integral equations*

$$(10.1)\qquad \langle Y_t, f\rangle = \int_0^t \langle Y_s, L^\varepsilon(Y_s)f\rangle\, ds + \langle X_0^*, f\rangle, \quad 0\le t\le T, \quad f\in C_c^2(\mathbb{R}^n)$$

has a unique solution in $\mathscr{D}([0,T],(\mathscr{M},\|\cdot\|_1)) = \Omega_1$. Moreover this solution equals $\mathscr{L}(\zeta_t^\varepsilon)$ for all $t\le T$, where $\zeta^\varepsilon = (\zeta_t^\varepsilon)_{t\le T}$ is the unique solution of

$$(10.2)\qquad d\zeta_t^\varepsilon = b^\varepsilon(\zeta_t^\varepsilon, \mathscr{L}(\zeta_t^\varepsilon))\,dt + \sigma^\varepsilon(\zeta_t^\varepsilon, \mathscr{L}(\zeta_t^\varepsilon))\,d\tilde W_t,$$

$$(10.3)\qquad \mathscr{L}(\zeta_0^\varepsilon) = X_0^*$$

($\tilde W_t$ n-dimensional Brownian motion).

PROOF. The existence of a solution of (10.1) follows from (9.8). Each limit point $\mathscr{L}^\varepsilon$ of the sequence $\{\mathscr{L}(X^{N,\varepsilon})\}_{N\in\mathbb{N}}$ is concentrated at solutions of (10.1). Now let $t\to \bar Y_t$ be any solution of (10.1). Using an estimate analogous to (9.4), we obtain the validity of (10.1) for all $f\in C_b^2(\mathbb{R}^n)$. Conditions (A1) and (A2), which hold for b^ε and σ^ε too, imply uniqueness of the solution $\xi^\varepsilon = (\xi_t^\varepsilon)_{t\le T}$ of

$$(10.4)\qquad d\xi_t^\varepsilon = b^\varepsilon(\xi_t^\varepsilon, \bar Y_t)\,dt + \sigma^\varepsilon(\xi_t^\varepsilon, \bar Y_t)\,dW_t, \quad \mathscr{L}(\xi_0^\varepsilon) = X_0^*.$$

Dynkin's Formula implies that $P_t = \mathscr{L}(\xi_t^\varepsilon)$ satisfies

$$(10.5)\qquad \langle P_t, f\rangle = \int_0^t \langle P_s, L^\varepsilon(\bar Y_s)f\rangle\, ds + \langle X_0^*, f\rangle, \quad 0\le t\le T, \quad f\in C_b^2(\mathbb{R}^n).$$

Our aim now is to show that $\mathscr{L}(\xi_t^\varepsilon)$ is the only solution in $\mathscr{D}([0,T],(\mathscr{M},\|\cdot\|_1))$ of the linear equation (10.5). Let $f(t,x)$ be any function in $C_b([0,T]\times\mathbb{R}^n)$ with one bounded continuous partial derivative with respect to t and bounded continuous partial derivatives of first and second order with respect to the spatial variable x. Then we obtain for any solution $(P_t)_{t\le T}\in\mathscr{D}([0,T],(\mathscr{M},\|\cdot\|_1))$ of

(10.5):

$$
\begin{aligned}
&\langle P_t, f(t, \cdot)\rangle - \langle P_0, f(0, \cdot)\rangle \\
&\quad = \langle P_t, f(t, \cdot)\rangle - \langle X_0^*, f(0, \cdot)\rangle \\
&\quad = \textstyle\sum_{k=0}^{\ell-1} \{\langle P_{(k+1)\delta}, f((k+1)\delta, \cdot)\rangle - \langle P_{k\delta}, f(k\delta, \cdot)\rangle\} \quad (\delta = t/\ell) \\
&\quad = \textstyle\sum_{k=0}^{\ell-1} \Big\{\langle P_{(k+1)\delta}, f((k+1)\delta, \cdot)\rangle \\
&\qquad - \Big\langle P_{k\delta}, f((k+1)\delta, \cdot) - \int_{k\delta}^{(k+1)\delta} \frac{\partial}{\partial s} f(s, \cdot)\, ds\Big\rangle\Big\} \\
&\quad = \textstyle\sum_{k=0}^{\ell-1} \Big\{\int_{k\delta}^{(k+1)\delta} \Big\langle P_s, \Big(L^\varepsilon(\bar{Y}_s) + \frac{\partial}{\partial s}\Big) f(s, \cdot)\Big\rangle\, ds \\
&\qquad - \int_{k\delta}^{(k+1)\delta} \Big\langle P_s - P_{k\delta}, \frac{\partial}{\partial s} f(s, \cdot)\Big\rangle\, ds \\
&\qquad + \int_{k\delta}^{(k+1)\delta} \langle P_s, L^\varepsilon(\bar{Y}_s)(f((k+1)\delta, \cdot) - f(s, \cdot))\rangle\, ds\Big\}.
\end{aligned}
$$

(Note that by (10.5) we have

$$\langle P_u, g\rangle - \langle P_v, g\rangle = \int_v^u \langle P_w, L^\varepsilon(\bar{Y}_w)g\rangle\, dw,$$

$$0 \le v \le u \le T,\ g \in C_b^2(\mathbb{R}^n).)$$

$$= \int_0^t \Big\langle P_s, \Big(L^\varepsilon(Y_s) + \frac{\partial}{\partial s}\Big) f(s, \cdot)\Big\rangle\, ds + r(\delta)$$

with $\lim_{\delta\to 0} r(\delta) = 0$, since $t \to P_t \in \mathscr{D}([0, T], (\mathscr{M}, \|\cdot\|_1))$. Consequently we have in the limit $\delta \to 0$

$$\text{(10.6)} \quad \langle P_t, f(t, \cdot)\rangle = \langle X_0^*, f(0, \cdot)\rangle + \int_0^t \Big\langle P_s, \Big(L^\varepsilon(\bar{Y}_s) + \frac{\partial}{\partial s}\Big) f(s, \cdot)\Big\rangle\, ds.$$

Let $\xi^{\varepsilon,x,t} = (\xi^{\varepsilon,x,t}(s))_{s\in[t,T]}$ be the solution of (10.4) with initial condition $\xi^{\varepsilon,x,t}(t) = x$. By (9.6), (A1), (A2) and the definition of b^ε and σ^ε the assumptions of Theorem 6.1 [10], page 124 are satisfied and we may conclude that the function

$$(t, x) \to E[f(\xi^{\varepsilon,x,t}(\tilde{T}))] =: f^{\tilde{T}}(t, x), \quad 0 \le t \le \tilde{T}, \quad f \in C_b^2(\mathbb{R}^n)$$

has continuous partial derivatives

$$\frac{\partial}{\partial t} f^{\tilde{T}}(t, x), \quad \frac{\partial}{\partial x_i} f^{\tilde{T}}(t, x), \quad \frac{\partial^2}{\partial x_i \partial x_j} f^{T}(t, x),$$

which satisfy Kolmogoroff's equation

$$\text{(10.7)} \qquad \Big(\frac{\partial}{\partial t} + L^\varepsilon(\bar{Y}_t)\Big) f^{\tilde{T}}(t, x) = 0, \quad f^{\tilde{T}}(\tilde{T}, x) = f(x).$$

By Theorem 5.3 and 5.4 [10], page 120–122, we know that the partial derivatives

$$\frac{\partial}{\partial x_i}\,\xi^{\varepsilon,x,t}(s),\quad \frac{\partial^2}{\partial x_i \partial x_j}\,\xi^{\varepsilon,x,t}(s)$$

exist in L^2 and satisfy stochastic differential equations obtained by applying formally $\partial/\partial x_i$ (resp. $\partial^2/\partial x_i \partial x_j$) to the equation for $\xi^{\varepsilon,x,t}(s)$. Writing down these equations, and using the fact that $x \to b^\varepsilon(x, \bar{Y}_t)$ and $x \to \sigma^\varepsilon(x, \bar{Y}_t)$ have uniformly bounded partial derivatives of first and second order, we see that $(\partial/\partial x_i)\xi^{\varepsilon,x,t}(s)$ and $(\partial^2/\partial x_i \partial x_j)\xi^{\varepsilon,x,t}(s)$ have moments of all orders which are bounded uniformly in s, t, x. Since $(\partial/\partial x_i) f^{\tilde{T}}(t, x) = E[\nabla f(\xi^{\varepsilon,x,t}(\tilde{T})) \cdot (\partial/\partial x_i)\xi^{\varepsilon,x,t}(\tilde{T})]$ (cf. (5.12), [10], page 122), we see that for $f \in C_0^2(\mathbb{R}^n)$ $(\partial/\partial x_i) f^{\tilde{T}}(t, x)$ is bounded uniformly in $\tilde{T}$, t and x.

Similarly $(\partial^2/\partial x_i \partial x_j) f^{\tilde{T}}(t, x)$ is bounded uniformly too. Therefore we may apply (10.6) to the function $(t, x) \to f^{\tilde{T}}(t, x)$:

$$\langle P_{\tilde{T}}, f^{\tilde{T}}(\tilde{T}, \cdot)\rangle = \langle P_{\tilde{T}}, f\rangle = \langle X_0^*, f^{\tilde{T}}(0, \cdot)\rangle = \langle \mathscr{L}(\xi_{\tilde{T}}^\varepsilon), f\rangle. \tag{10.8}$$

Thus $P_{\tilde{T}} = \mathscr{L}(\xi_{\tilde{T}}^\varepsilon) = \bar{Y}_{\tilde{T}}$ for all $\tilde{T} \le T$, since $t \to \bar{Y}_t$ solves (10.5) too.

We have shown so far that to any solution $(\bar{Y}_t)_{t \le T}$ of (10.1), there corresponds a solution of

$$d\zeta_t = b^\varepsilon(\zeta_t, \mathscr{L}(\zeta_t))\, dt + \sigma^\varepsilon(\zeta_t, \mathscr{L}(\zeta_t))\, dW_t, \quad \mathscr{L}(\zeta_0) = X_0^* \tag{10.9}$$

with $\bar{Y}_t = \mathscr{L}(\zeta_t)$.

Therefore to finish the proof of Lemma 10 we have to show unique solubility of (10.9). Suppose $\zeta^1 = (\zeta_t^1)_{t \le T}$ and $\zeta^2 = (\zeta_t^2)_{t \le T}$ with $\zeta_0^1 = \zeta_0^2$ a.s. are solutions of (10.9). We can assume that both processes are "driven" by the same Brownian motion W_t, which implies the existence of a common distribution. Then we have:

$$\begin{aligned}
E[(\zeta_t^1 - \zeta_t^2)^2] &\le 2E\left[\left(\int_0^t (b^\varepsilon(\zeta_s^1, \mu_s^1) - b^\varepsilon(\zeta_s^2, \mu_s^2))\, ds\right)^2\right] \\
&\quad + 2E\left[\left(\int_0^t (\sigma^\varepsilon(\zeta_s^1, \mu_s^1) - \sigma^\varepsilon(\zeta_s^2, \mu_s^2))\, dW_s\right)^2\right] \\
&\qquad\qquad (\mu_s^i = \mathscr{L}(\zeta_s^i),\ i = 1, 2) \\
&\le \left(TE\left[\int_0^t |b^\varepsilon(\zeta_s^1, \mu_s^1) - b^\varepsilon(\zeta_s^2, \mu_s^2)|^2\, ds\right]\right. \\
&\quad \left. + E\left[\int_0^t \textstyle\sum_{i,j=1}^n |\sigma_{ij}^\varepsilon(\zeta_s^1, \mu_s^1) - \sigma_{ij}^\varepsilon(\zeta_s^2, \mu_s^2)|^2\, ds\right]\right) c_{22} \\
&\le \int_0^t ds \{E[|\zeta_s^1 - \zeta_s^2|^2] + \|\mu_s^1 - \mu_s^2\|_1^2\} c_{23}
\end{aligned}$$

(since the functions $x \to b^\varepsilon(x, \mu)$ and $x \to \sigma^\varepsilon(x, \mu)$ are smooth). Since

$$\begin{aligned}\| \mu_s^1 - \mu_s^2 \|_1^2 &= \left(\sup_{f\in\mathscr{A}_1} \int\int \mathscr{L}(\zeta_s^1, \zeta_s^2)(dx, dy)(f(x) - f(y))\right)^2 \\ &\le \left(\int\int \mathscr{L}(\zeta_2^1, \zeta_s^2)(dx, dy)\,|x - y|\right)^2 = (E[\,|\zeta_s^1 - \zeta_s^2|\,])^2 \\ &\le E[\,|\zeta_s^1 - \zeta_s^2|^2]\end{aligned}$$

we obtain

$$E[\,|\zeta_t^1 - \zeta_t^2|^2] \le \int_0^t dsE[\,|\zeta_s^1 - \zeta_s^2|^2]c_{24}.$$

Gronwall's Lemma shows: $\zeta_t^1 = \zeta_t^2$ almost surely (t fixed). This and the right-continuity of both processes imply that $\zeta_t^1 = \zeta_t^2$ for all $t \le T$ with probability 1.

Therefore the stochastic differential equation (10.9) and, moreover, the deterministic equation (10.1), has a unique solution.

As announced above, we now present the proof of the unique existence of the limit dynamics for Theorem 2.

LEMMA 11. *The system of integral equations*

$$\begin{aligned}\langle Y_t, f\rangle &= \int_0^t ds \left\langle Y_s, \int A(\cdot, y; Y_s)(f(\cdot + y) - f(\cdot))\gamma\cdot(dy)\right\rangle \\ &\quad + \langle X_0^*, f\rangle, \quad f \in C_c^2(\mathbb{R}^n), \quad t \le T\end{aligned} \tag{11.1}$$

has a unique solution $t \to Y_t$, such that Y_t is a probability measure on $\mathbb{R}^n$ for all $t \le T$.

PROOF. As in the proof of Lemma 10, we may assume that existence is already proved and as there we may conclude that uniqueness of (11.1) is related to uniqueness of a certain stochastic process. So we only need to show the uniqueness of a (nonlinear) jump process $Z = (Z_t)_{t\le T}$ on $\mathbb{R}^n$ with jump intensities $A(x, y; \mathscr{L}(Z_t))$ for a jump from x to $x + y$ at time t, and with initial distribution $\mathscr{L}(Z_0) = X_0^*$. Let Z^1 and Z^2 be two such processes. A coupling $\mathbf{Z} = (\mathbf{Z}_t)_{t\le T}$ of these processes is a process with state space $\mathbb{R}^n \times \mathbb{R}^n$, which consists of two components $\mathbf{Z}^1$ and $\mathbf{Z}^2$, which have the same distribution as Z^1 and Z^2 respectively.

We shall prove Lemma 11 by constructing a coupling $\mathbf{Z}$ with

$$E_{\mathscr{L}(\mathbf{Z})}[\,|\eta_t^1 - \eta_t^2|\,] = 0$$

for all $t \le T$. ($\eta = (\eta_t)_{t\le T}$, $\eta_t = (\eta_t^1, \eta_t^2)$ is a typical element of $\mathscr{D}([0, T], \mathbb{R}^{2n})$). $\mathbf{Z}$

is defined by its (time dependent) infinitesimal generator $\mathscr{A} = (\mathscr{A}_t)_{t\le T}$:

$$
\begin{aligned}
(\mathscr{A}_t f)(z_1, z_2)
&= \iint (f(z_1 + y_1, z_2 + y_2) - f(z_1, z_2)) \\
&\qquad \cdot (A(z_1, y_1; \kappa_t^1) \wedge A(z_2, y_2; \kappa_t^2))\rho_{z_1,z_2}(dy_1, dy_2) \\
&\quad + \iint (f(z_1 + y_1, z_2) - f(z_1, z_2)) \\
&\qquad \cdot ((A(z_1, y_1; \kappa_t^1) - A(z_2, y_2; \kappa_t^2)) \vee 0)\rho_{z_1,z_2}(dy_1, dy_2) \\
&\quad + \iint (f(z_1, z_2 + y_2) - f(z_1, z_2)) \\
&\qquad \cdot ((A(z_2, y_2; \kappa_t^2) - A(z_1, y_1; \kappa_t^1)) \vee 0)\rho_{z_1,z_2}(dy_1, dy_2), \\
&\qquad\qquad (\kappa_t^i = \mathscr{L}(\mathbf{Z}_t^i),\ i = 1, 2)
\end{aligned}
$$

and its initial distribution

$$P_{\mathscr{L}(\mathbf{Z})}[\eta_0^1 = \eta_0^2] = 1, \quad \mathscr{L}(\mathbf{Z}_0^1) = X_0^*.$$

It is easy to see that this definition indeed describes a coupling of the processes Z^1 and Z^2. The dynamics of $\mathbf{Z}$ was defined in such a way that both components jump with an intensity as large as possible in nearly the same direction.

Let $\mathbf{f}(z_1, z_2) = |z_1 - z_2|$. Dynkin's Formula implies

$$
\begin{aligned}
&E_{\mathscr{L}(\mathbf{Z})}[|\eta_t^1 - \eta_t^2|] \\
&= E_{\mathscr{L}(\mathbf{Z})}\left[\int_0^t \mathscr{A}_u \mathbf{f}(\eta_u)\, du\right] \\
&\le E_{\mathscr{L}(\mathbf{Z})}\Bigg[\int_0^t \Bigg\{\iint |y_1 - y_2|(A(\eta_u^1, y_1; \kappa_u^1) \wedge A(\eta_u^2, y_2; \kappa_u^2))\rho_{\eta_u^1,\eta_u^2}(dy_1, dy_2) \\
&\qquad + \iint (|y_1| + |y_2|)|A(\eta_u^1, y_1; \kappa_u^1) \\
&\qquad - A(\eta_u^2, y_2; \kappa_u^2)|\rho_{\eta_u^1,\eta_u^2}(dy_1, dy_2)\Bigg\}\, du\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&\le E_{\mathscr{L}(\mathbf{Z})}\Bigg[\int_0^t \Bigg\{\iint |y_1 - y_2|(A(\eta_u^1, y_1; \kappa_u^1) \wedge A(\eta_u^2, y_2; \kappa_u^1))\rho_{\eta_u^1,\eta_u^2}(dy_1, dy_2) \\
&\qquad + 2\iint (|y_1| + |y_2|)|A(\eta_u^2, y_2; \kappa_u^2) - A(\eta_u^2, y_2; \kappa_u^1)|\rho_{\eta_u^1,\eta_u^2}(dy_1, dy_2) \\
&\qquad + \iint (|y_1| + |y_2|)|A(\eta_u^1, y_1; \kappa_u^1) \\
&\qquad - A(\eta_u^2, y_2; \kappa_u^1)|\rho_{\eta_u^1,\eta_u^2}(dy_1, dy_2)\Bigg\}\, du\Bigg]
\end{aligned}
\tag{11.2}
$$

$$\le E_{\mathscr{L}(\mathbf{Z})}\left[\int_0^t du\{|\eta_u^1 - \eta_u^2|(1 + \langle\kappa_u^1, \kappa\rangle) + \|\kappa_u^1 - \kappa_u^2\|_1(1 + |\eta_u^2|)\}\right]c_{25}.$$

Since $\sup_{t\le T, i=1,2}\langle \kappa_t^i, \kappa\rangle < \infty$ and $\| \kappa_t^1 - \kappa_t^2 \|_1 \le E_{\mathscr{L}(\mathbf{Z})}[\,|\eta_t^1 - \eta_t^2|\,]$, $\forall t \le T$, Gronwall's Lemma and (11.2) imply

$$E_{\mathscr{L}(\mathbf{Z})}[\,|\eta_t^1 - \eta_t^2|\,] = 0.$$

REFERENCES

[1] WILLIAMS, D. (1979). *Diffusions, Markov Processes and Martingales.* Wiley, New York.

[2] KURTZ, T. G. (1981). Approximation of population processes. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM.

[3] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes.* Springer, Berlin.

[4] DUDLEY, R. M. (1966). Convergence of Baire measures. *Studia Math.* **27** 251–268.

[5] MCKEAN, H. P. (1967). Propagation of chaos for a class of nonlinear parabolic equations. *Lecture Series in Differential Equations* **7** 41–57, Catholic Univ.

[6] MARCHIORO, C. and PULVIRENTI, M. (1982). Vortex theory and mean field limit. Lectures given at Trento meeting on "Coulomb systems and fluid dynamics", Trento, Italy.

[7] SKOROKHOD, A. V. (1981). On randomly interacting particles. Report presented at sessions of the probability and statistics seminar at the T. G. Shevchenko Kiev State University 1979. *Theor. Probab. Appl.* **26** 204–206.

[8] SPOHN, H. (1980). Kinetic equations from Hamiltonian dynamics: The Markovian limit. *Rev. Mod. Phys.* **52** 569–616.

[9] OELSCHLÄGER, K. (1982). Grenzwertsätze für zwei Modelle von zufälligen Zeitentwicklungen für Systeme schwach miteinander wechselwirkender Teilchen. Dissertation, Universität Heidelberg.

[10] FRIEDMAN, A. (1975). *Stochastic Differential Equations and Applications,* Vol. 1. Academic, New York.

SONDERFORSCHUNGSBEREICH 123
UNIVERSITÄT HEIDELBERG
IM NEUENHEIMER FELD 293
D-6900 HEIDELBERG 1
WEST GERMANY