ANNALES DE L'I. H. P., SECTION B

A. N. BORODIN
M. I. FREIDLIN

Fast oscillating random perturbations of dynamical systems with conservation laws

Annales de l'I. H. P., section B, tome 31, nº 3 (1995), p. 485-525
<http://www.numdam.org/item?id=AIHPB_1995__31_3_485_0>

Ann. Inst. Henri Poincaré,

Vol. 31, n° 3, 1995, p. 485-525. *Probabilités et Statistiques*

Fast oscillating random perturbations of dynamical systems with conservation laws

by

A. N. BORODIN (*) and M. I. FREIDLIN (**)
University of Maryland

ABSTRACT. – We consider fast oscillating random perturbations of dynamical systems with first integrals. We prove that if the dynamical system is ergodic in the subset of the phase space where the first integrals are constants, then the evolution of the first integrals in a proper time scale is described by a diffusion process.

Key words: Averaging principle, random perturbations of dynamical systems, conservation laws, diffusion approximation.

RÉSUMÉ. – On considère des systèmes dynamiques avec des intégrales premières perturbés par des perturbations aléatoires à oscillation rapide. On montre que si le système dynamique est ergodique sur le sous-ensemble de l'espace de phase dans lequel les premières intégrales sont constantes alors l'évolution de ces intégrales dans une escale de temps approprié est décrite par un processus de diffusion.

A.M.S. Classification 1980 : 60 H XX, 58 G 32.

(*) Permanently at Mathematics Institute of Russian Acad. of Sci. at St. Petersburg. The work was done while visiting the University of Maryland at College Park.

(**) Work partially supported by ARO Grant DAAL03-92-G-0219.

1. INTRODUCTION

Consider the equation

$$\dot{\tilde{X}}^{\varepsilon}(t) \equiv \frac{d\tilde{X}^{\varepsilon}(t)}{dt} = b(\tilde{X}^{\varepsilon}(t), \zeta_{t/\varepsilon}), \qquad \tilde{X}^{\varepsilon}_{(0)} = x \in R^r. \tag{1.1}$$

Here ζ_t is a stationary stochastic process, ε is a small positive parameter, and vector field $b(x, z)$ suppose to be smooth enough. Denote $b(x) = Eb(x, \zeta_t)$. One can prove that under some mild assumptions concerning the ergodicity of ζ_t the averaging principle is hold. The processes $\tilde{X}^{\varepsilon}(t)$ converge in probability as $\varepsilon \downarrow 0$ uniformly on any finite time interval $[0, T]$ to the solution of the averaged equation

$$\dot{X}(t) = b(X(t)), \; X(0) = x \in R^r. \tag{1.2}$$

This means that for any $\delta, T > 0$

$$\lim_{\varepsilon \downarrow 0} P\{ \max_{0 \le t \le T} |\tilde{X}^{\varepsilon}(t) - X(t)| > \delta \} = 0$$

(*see*, for example, [9], [5], Ch. 7), and we can look on (1.1) as a result of small (in the mean sense) random perturbations of the dynamical system (1.2). Moreover, one can prove that the processes $\varepsilon^{-1/2}(\tilde{X}^{\varepsilon}(t) - X(t))$ converge weakly, in the space of continuous functions C_{0T}, to a mean zero Gaussian Markov process. The last statement holds if we make certain assumptions about the mixing properties of the noise ζ_t (*see* [8], [5], § 7.1).

Let system (1.2) have a conservation law $H(x) : H(X(t)) = H(x)$, $t \ge 0$; and let $H(x)$ be a smooth function with compact connected level sets. Since $\tilde{X}^{\varepsilon}(t) \to X(t)$ as $\varepsilon \downarrow 0$, $H(\tilde{X}^{\varepsilon}(t)) \to H(X(t)) = H(x)$ for any $t > 0$ independent of ε. To observe the evolution of $H(\tilde{X}^{\varepsilon}(t))$ let us rescale the time: we denote $X^{\varepsilon}(t) = \tilde{X}^{\varepsilon}(t/\varepsilon)$. It is clear, that $X^{\varepsilon}(t)$ is the solution of the problem.

$$\dot{X}^{\varepsilon}(t) = \frac{1}{\varepsilon} b(X^{\varepsilon}(t), \zeta_{t/\varepsilon^2}), \qquad X^{\varepsilon}(0) = x \in R^r.$$

Now we have:

$$\begin{aligned} H(X^{\varepsilon}(t)) - H(x) &= \frac{1}{\varepsilon} \int_0^t (\nabla H(X^{\varepsilon}(t)), b(X^{\varepsilon}(s), \zeta_{s/\varepsilon^2}))\, ds \\ &= \frac{1}{\varepsilon} \int_0^t (\nabla H(X^{\varepsilon}(t)), b(X^{\varepsilon}(s), \zeta_{s/\varepsilon^2})) \\ &\quad - b(X^{\varepsilon}(s))\, ds. \end{aligned}$$

We used in the last equality the fact that $(\nabla H(x), b(x)) \equiv 0$, $x \in R^r$, which holds since $H(x)$ is a first integral for system (1.2). Taking into account that $E[b(x, \zeta_s) - b(x)] \equiv 0$, one can note that

$$\frac{1}{\varepsilon} \int_0^t (\nabla H(x), b(x, \zeta_{s/\varepsilon^2}) - b(x))\, ds \tag{1.4}$$

converges to a Gaussian variable as $\varepsilon \downarrow 0$, if the process ζ_s has good enough mixing properties. Of course, the characteristics of the limiting Gaussian distribution depend on the point x.

Taking into account that the rates of changing $X^\varepsilon(t)$ and ζ_{s/ε^2} have different order, and that $\tilde{X}^\varepsilon(t)$ converges weakly to $X(t)$ as $\varepsilon \downarrow 0$, one can expect that if the dynamical system $X(t)$ has some ergodic properties on the level sets $\{x : H(x) = y\}$, the characteristics for the limit of $dH(X^\varepsilon(t))$ as $\varepsilon \downarrow 0$ depend only on $H(X^\varepsilon(t))$. This means that the limiting process for $H(X^\varepsilon(t))$ as $\varepsilon \downarrow 0$ will be a diffusion process

$$dY(t) = \sigma(Y(t))\, dW_t + B(Y_t)\, dt, \qquad Y_0 = H(x), \tag{1.5}$$

where W_t is a standard Wiener process. Thus, the convergence to a Markov diffusion process is a result of averaging and a Gaussian approximation due to mixing properties of ζ_t and of ergodicity of the non-perturbed system on the level sets.

The formulation and the proof of the rigorous results concerning this convergence is the goal of this paper.

In the next section we introduce the conditions, formulate the main results and consider some examples. We consider in Section 2 the two-dimensional case. We prove these results in Section 3.

The last section contains some remarks and generalizations. In particular, we formulate a result for systems in R^r, $r \geq 2$, with $l \geq 1$ conservation laws.

Perturbations of dynamical systems which are not ergodic on the level sets are considered shortly in Section 4 as well. Roughly speaking, if the dynamical system is not ergodic on the intersection of the level sets of all first integrals under consideration, then the limiting process will not be Markovian one. One should extend the phase space to obtain a limiting process with Markov property. Because of non-ergodicity on the level sets the limiting process, for example, can have something like a hysteresis effect.

If the level sets of a first integral $H(x)$ have several connected components then the limiting process should be considered on a graph defined by the function $H(x)$. Only then the limiting process will be Markovian. Here the situation is similar to one considered in [FW 2, 3] for the white-noise-type perturbations of dynamical systems. We plan to consider those questions in the case of fast oscillating noise elsewhere.

2. MAIN RESULTS. EXAMPLES

Consider the following equation in the plane

$$\begin{cases} \dot{X}^\varepsilon(t) \equiv \dfrac{dX^\varepsilon(t)}{dt} = \dfrac{1}{\varepsilon}\, b(X^\varepsilon(t), \zeta_{t/\varepsilon^2}), \\ X^\varepsilon(0) = x \in R^2, \quad \varepsilon > 0. \end{cases} \tag{2.1}$$

Here ζ_t is a stationary process with values in R^m. We assume that the trajectories ζ_t have at most finite number of simple discontinuities on each finite time interval with probability 1. Equation (2.1) is fulfilled at the points t where ζ_{t/ε^2} is continuous. The vector field $b(x, z)$ supposed to be Lipschitz continuous and $|b(x, z)|$ grows not too fast. Then there exists an unique continuous for all $t \geq 0$ with probability 1 solution of (2.1). Denote

$$b(x) = Eb(x, \zeta_t)$$

and consider the non-perturbed equation

$$\dot{X}(t) = b(X(t)), \qquad X(0) = x. \tag{2.2}$$

We introduce the following conditions.

1. Assume that there exists a real valued three times differentiable function $H(x)$, $x \in R^2$ which is the first integral for the non-perturbed system (2.2), *i.e.* $H(X(t)) = H(x)$ for any starting point $x \in R^2$ and $t \geq 0$.

Suppose that the set $C(y) = \{x \in R^2 : H(x) = y\}$ is a closed connected curve in the plane without intersections for any y in the range of values of the function H. It means that H has only one extremum point, which is an equilibrium point of the field $b(x)$. We can assume without loss of generality that this point is the origin 0, and that $H(0) = 0$, $H(x) > 0$ for $x \neq 0$. The trajectory $X(t)$ performs periodic motion along the curve $C(y)$, $y = H(X(0))$, with some period $T(y)$. Assume that $T(y) < C(1+y^2)$. Note, that if $H(x)$ is a first integral for systems (2.2) and $f(y)$ is a real function, then $f(H(x))$ is also a first integral.

2. Let

$$(2.3)\quad \begin{cases} 0 \le H(X) < C, & \left| \dfrac{\partial H(x)}{\partial x_k} \right| < C, \quad \left| \dfrac{\partial^2 H(x)}{\partial x_k\, \partial X_l} \right| < C, \\ \left| \dfrac{\partial^3 H(x)}{\partial x_k\, \partial x_l\, \partial x_j} \right| < C, & k,\ l,\ j = 1,\ 2; \end{cases}$$

here and in the sequal C denotes a constant, not necessarily always the same.

3. Let $b(x, y)$ be a twice differentiable function with respect to x, and let for some $p > 8$, $k, l = 1, 2$

$$(2.4)\quad \begin{cases} E \sup\limits_x |b(x, \zeta_t)|^p < C, & E \sup\limits_x \left| \frac{\partial}{\partial x_k} b(x, \zeta_t) \right|^p < C, \\ E \sup\limits_x \left| \dfrac{\partial^2 b(x, \zeta_t)}{\partial x_k\, \partial x_l} \right|^p < C \end{cases}$$

Later we will weaken assumptions (2.3) and (2.4) to allow $H(x)$ and $|b(x, z)|$ to grow as $|x| \to \infty$.

4. Denote by $\mathcal{N}_s^t$ the σ-field generated by the process ζ_v when $-\infty \le s \le v \le t \le +\infty$. Suppose that the family of σ-fields $\{\mathcal{N}_s^t\}$ satisfies the absolute regularity mixing condition (Kolmogrov's condition):

$$\beta(\tau) = \operatorname*{Var}_{B \in \mathcal{N}_{-\infty}^0 \times \mathcal{N}_\tau^\infty} (P_{0,\tau}(B) - P_0 \times P_\tau(B)) \to 0 \quad \text{as } \tau \to \infty,$$

where for sets $A_1 \times A_2$, $A_1 \in \mathcal{N}_{-\infty}^0$, $A_2 \in \mathcal{N}_\tau^0$, the measures are defined by the relations

$$P_{0,\tau}(A_1 \times A_2) = P(A_1 A_2), \qquad P_0 \times P_\tau(A_1 \times A_2) = P(A_1)\, P(A_2).$$

In the special case, when $b(x, z)$ has the form

$$(2.5)\qquad b(x, z) = \sum_{k=1}^n u_k(x)\, v_k(x), \quad n < \infty,$$

it is sufficient to assume that the family $\{\mathcal{N}_s^t\}$ satisfies the strong mixing condition (Rosenblatt's condition):

$$\alpha(\tau) = \sup_{A_1 \in \mathcal{N}_{-\infty}^0,\, A_2 \in \mathcal{N}_\tau^\infty} |P(A_1 A_2) - P(A_1)\, P(A_2)| \to 0 \text{ as } \tau \to \infty.$$

It is known that the absolute regularity mixing condition is stronger than the Rosenblatt condition [10]. Some sufficient conditions for these mixing properties and bounds for the coefficients $\alpha(\tau)$ and $\beta(\tau)$ one can find in [10].

We assume that the mixing coefficients are such that

$$\int_0^\infty \tau^3 \beta^{1-\frac{8}{p}}(\tau)\, \alpha t < \infty, \qquad \int_0^\infty \tau^3 \alpha^{1-\frac{8}{p}}(\tau)\, \alpha t < \infty, \tag{2.6}$$

$$\beta^{1-\frac{8}{p}}(\tau) \le C \min(1, \tau^{-4}), \qquad \alpha^{1-\frac{8}{p}}(\tau) \le C \min(1, \tau^{-4}). \tag{2.7}$$

The conditions for $\alpha(\tau)$ are assumed if $b(x, z)$ satisfies (2.5).

5. Put

$$\begin{aligned} g(x, z) &= b(x, z) - b(x), \\ F(x, z) &= (\nabla H(x), g(x, z)), \\ D(x, s) &= EF(x, \zeta_s) F(x, \zeta_0), \\ Q(x, s) &= E(\nabla F(x, \zeta_s), g(x, \zeta_0)), \end{aligned}$$

where ∇ denotes the gradient in x, and the notation $(\cdot, \cdot)$ is used for the scalar product. Denote

$$D(x) = 2 \int_0^\infty D(x, s)\, ds, \qquad Q(x) = 2 \int_0^\infty Q(x, s)\, ds,$$

We will verify that under our assumptions these integrals are finite and define bounded Lipschitz continuous functions $D(x)$ and $Q(x)$.

Let

$$\sigma^2(y) = \left(\int_{C(y)} \frac{dl}{|b(x)|} \right)^{-1} \int_{C(y)} \frac{D(x)\, dl}{|b(x)|},$$
$$B(y) = \left(\int_{C(y)} \frac{dl}{|b(x)|} \right)^{-1} \int_{C(y)} \frac{Q(x)\, dl}{|b(x)|},$$

where dl is the length element on $C(y)$, be Lipschitz continuous in y. Note that $\int_{C(y)} |b(x)|^{-1}\, dl = T(y)$ is the period of the rotation along the curve $C(y)$. It follows from our assumptions that $\sigma^2(y)$ and $|B(y)|$ are bounded.

THEOREM 2.1. – *Let* $Y^\varepsilon(t) = H(X^\varepsilon(t))$, *where* $X^\varepsilon(t)$ *is the solution of problem* (2.1), *and* $H(x)$, $x \in R^2$, *is the first integral for non-perturbed equation* (2.2). *Let conditions* 1-5 *hold. Then for any* $T < \infty$ *the processes*

$Y^{\varepsilon}(t)$ converges weakly in C_{0T} as $\varepsilon \downarrow 0$ to the diffusion process $Y(t)$ determined by the stochastic differential equation

$$dY(t) = \sigma(Y(t))\, dW_t + B(Y(t))\, dt, \qquad Y(0) = H(x), \tag{2.8}$$

where W_t is the standard Wiener process.

Now we will modify the conditions on the system (2.1) to include the case of unbound functions $H(x)$ and $|b(x, z)|$. Instead of conditions 2-4 we introduce the following ones.

2′. Let for some constant C and $k, l, j \in \{1, 2\}$

$$\begin{cases} |H(x)| < C(|x|^2 + 1), & \left|\dfrac{\partial H(x)}{\partial x_k}\right| < C(|x| + 1), \\ \left|\dfrac{\partial^2 H(x)}{\partial x_k\, \partial x_l}\right| < C, & \left|\dfrac{\partial^3 H(x)}{\partial x_k\, \partial x_l\, \partial x_j}\right| < C; \end{cases} \tag{2.9}$$

and for some μ, $1 < \mu \leq 2$,

$$|x|^{\mu} < C(|H(x)| + 1) \tag{2.10}$$

3′. Let $g(x, z) = b(x, z) - b(x)$. Suppose that for some constant C and some positive function $q(z)$

$$\begin{cases} |b(x)| < C(|x| + 1), & \left|\dfrac{\partial b(x)}{\partial x_k}\right| < C, \\ \left|\dfrac{\partial^2 b(x)}{\partial x_k\, \partial x_l}\right| < C, & k, l \in \{1, 2\} \end{cases} \tag{2.11}$$

$$|g(x, z)| < Cq(z), \qquad \left|\frac{\partial g(x, z)}{\partial x_k}\right| \leq \frac{C}{1 + |x|}\, q(z). \tag{2.12}$$

Assume that

$$E\,|q(\zeta_t)|^p < C,$$

for some $p > \dfrac{4(\mu + 1)}{(\mu - 1)}$.

4′. Let condition 4 holds with (2.6), (2.7) replaced by

$$\int_0^{\infty} \tau^3 \, \beta^{((\mu-1)/\mu+1)-4/p} (\tau) \, dt < \infty \tag{2.13}$$

$$\beta^{((\mu-1)/\mu+1)-4/p} (\tau) < C \, \min (1, \, \tau^{-4}) \tag{2.14}$$

and the same condition for the coefficient $\alpha(\tau)$.

THEOREM 2.2. – *Let $Y^{\varepsilon}(t) = H(X^{\varepsilon}(t))$, where $X^{\varepsilon}(t)$ is the solution of problem* (2.1), *and $H(x)$, $x \in R^2$, is the first integral for the non-perturbed equation* (2.2). *Suppose that conditions* 1, 2′-4′, 5 *hold. Then the process $Y^{\varepsilon}(t)$ converges as $\varepsilon \downarrow 0$ weakly in C_{0T} for any $0 \le T < \infty$ to the diffusion process $Y(t)$ determined by stochastic differential equation* (2.8).

We give the proofs of Theorems 2.1 and 2.2 in the next section. Now let us consider some examples.

Let $H(x)$, $x \in R^2$, be a smooth function satisfying condition 2′. Assume that the origin 0 is the only point where $\nabla H(x) = 0$ and let it be the minimum point and $H(0) = 0$. Let $\Lambda(x)$, $x \in R^2$, is a $2 \in 2$ matrix such that the function $g(x, z) = \Lambda(x) z$, $z \in R^2$, satisfies condition (2.12). Let ζ_t be a two dimensional stationary process such that $E\zeta_t = 0$ and let conditions 3′, 4′ and 5 be fulfilled for $g(x, \zeta_t) = \Lambda(x) \zeta_t$.

Consider the equation

$$\dot{\tilde{X}}_t^{\varepsilon} \equiv \overline{\nabla} H(\tilde{X}_t^{\varepsilon}) + \Lambda(\tilde{X}_t^{\varepsilon}) \zeta_{t/\varepsilon}, \qquad \tilde{X}_0^{\varepsilon} = x, \tag{2.15}$$

where $\overline{\nabla} H(x_1, x_2) = \left(\frac{\partial H(x)}{\partial x^2}, -\frac{\partial H(x)}{\partial x^1} \right)$ is the Hamiltonian vector field corresponding to $H(x)$. One can look on (2.15) as on result of small in the mean sense perturbations of the Hamiltonian system

$$\dot{X}_t = \overline{\nabla} H(X_t), \qquad X_0 = x. \tag{2.16}$$

It follows from the averaging principle (*see*, for example [8], [5]) that $\tilde{X}_t^{\varepsilon}$ converges weakly in C_{0T}, $0 < T < \infty$, to X_t as $\varepsilon \downarrow 0$. Thus $H(\tilde{X}_t^{\varepsilon}) \to H(X_t) = H(x)$ for small ε. But on larger, growing together with ε^{-1} time intervals $H(\tilde{X}_t^{\varepsilon})$ will change. One can apply here the Theorem 2.2. Here

$$\begin{cases} g(x, z) = \Lambda(x) z, \\ F(x, z) = (\nabla H(x), \Lambda(x) z) = (\Lambda^T(x) \nabla H(x), z), \\ D(x, s) = E\Lambda^T(x) \nabla H(x), \zeta_s) (\Lambda^T(x) \nabla H(x), \zeta_0) \\ Q(x, s) = E(\nabla(\nabla H(x), \Lambda(x) \zeta_s), \Lambda(x) \zeta_0). \end{cases} \tag{2.17}$$

Let $K(\tau) = (K^{ij}(\tau))$ be the correlation matrix for the stationary process $\zeta_t = (\zeta_t^1, \zeta_t^2)$; $K^{ij}(t) = E\zeta_\tau^i \dot{\zeta}_0^j$. Denote $\bar{K} = \int_0^\infty K(\tau)\, dt$. The finiteness of this integral follows from the assumption 5 if one takes into account that $E\zeta_t \equiv 0$.

One can derive from (2.17), that

$$D(x) = 2\int_0^\infty D(x,\, s)\, ds = (\Lambda(x)\, \bar{K}\, \Lambda^T(x)\, \nabla H(x),\, \nabla H(x)).$$

Since $K(\tau)$ is a positively defined function, $\bar{K}$, and thus $\Lambda\bar{K}\Lambda^T$, are also positively defined. Thus one can introduce $\sigma(x)$ such that

$$\sigma^2(y) = \left(\int_{C(y)} \frac{dl}{|\nabla H(x)|}\right)^{-1} \times \int_{C(y)} \frac{(\Lambda(x)\, \bar{K}\, \Lambda^T(x)\, \nabla H(x),\, \nabla H(x))}{|\nabla H(x)|}\, dl. \tag{2.18}$$

To write down the drift coefficient for the limiting process, we need some notations. For any smooth vector field $e(x)$ in R^2 denote $\nabla e(x)$ the matrix $(e_{ij}(x))$, $e_{ij}(x) = \dfrac{\partial e_j(x)}{\partial x_i}$. Then simple calculations show that

$$Q(x) = \int_0^\infty Q(x,\, s)\, ds = \mathrm{tr}\,(\Lambda^T(x)\, \nabla\, (\Lambda^T(x)\, \nabla H(x)\, \bar{K}).$$

and we have the following expression for the drift

$$B(y) = \left(\int_{C(y)} \frac{dl}{|\nabla H(x)|}\right)^{-1} \times \int_{C(y)} \frac{\mathrm{tr}\,(\Lambda^T(x)\, \nabla\, (\,\Lambda^T(x)\, \nabla H(x)\, \bar{K})\, dl}{|\nabla H(x)|}\, dl. \tag{2.19}$$

Let, for example, $\Lambda(x)$ be the unit matrix. Then

$$D(x) = (\bar{K}\nabla H(x),\, \nabla H(x)), \qquad Q(x) = \mathrm{tr}\hat{H}(x)\, \bar{K}$$

where $\hat{H}(x)$ is the Hessian matrix for $H(x)$: $\hat{H}_{ij}(x) = \dfrac{\partial^2 H(x)}{\partial x^i\, \partial x^j}$. Formulas (2.18) and (2.19) give the diffusion coefficient and the drift for the limiting process.

Let us consider the harmonic oscillator: $H(x) = |x|^2$ and assume that the components of the noise ζ_t are independent and have the same correlation function, $\bar{K}^{11} = \bar{K}^{22} = K = \int_0^\infty K(\tau)\, dt$. Then $C(y) = \{x \in R^2 : |x| = \sqrt{y}\}$, $|\nabla H(x)| = 2\sqrt{H(x)}$, $D(x) = 4K|x|^2$, $Q(x) = 4K$, $\int_{C(y)} |\nabla H(x)|^{-1}\, dl = \pi$, $\sigma^2(y) = 4Ky$, $B(y) = 4K$. If, for example, we are interested in the expectation of the time $\tau_{H_1}^\varepsilon$ when the energy reaches the level H_1 starting from a point x with the energy $H(x) < H_1$, we should solve the problem

$$2Ky\frac{d^2 u(y)}{dy^2} + 4K\frac{du}{dy} = -1, \qquad 0 < y < H_1, \tag{2.20}$$

$u(H_1) = 0$, $u(y)$ is bounded for $y \in [0, H_1]$. Then $\lim_{\varepsilon \downarrow 0} \varepsilon E_x\, \tau_{H_1}^\varepsilon = u(H(x))$. Problem (2.20) can be solved explicitly. Note that the diffusion coefficient is degenerate at the critical point $y = 0$ and the drift is positive, so that the point $y = 0$ is inaccessible for the limiting process and no boundary conditions should be added at this point.

After this paper was written (Technical Report TR92-25, 1992, University of Maryland), an article [2] appeared where a similar problem is considered. But the assumptions, methods, and part of results are different. In particular, we make substantially less restrictive assumptions on the mixing properties of the noise, and allow some growth of the right side of the equations. This allows to consider some interesting examples. We shortly consider also the case when the averaged dynamical system is not ergodic on the energy level, and the limiting process should be considered on a graph (*see* Section 4).

Finally, we would like to mention that if the fast process ζ_t is a diffusion process, our results imply some new results concerning second order partial differential equations with a small parameter. If ζ_t is a Markov process with finite phase space then Theorems 2.1 and 2.2 allow us to consider a small parameter problem for some systems of partial differential equations (compare [4], Ch. 4,6).

3. PROOF OF THEOREMS 2.1 AND 2.2

There are two main parts in the proof of these theorems. The first one concerns the averaging with respect to the fast oscillating noise. The second part deals with the averaging provided by the ergodic properties of the non-perturbed dynamical system. The rates of change of the noise and of the motion of the dynamical system have different order as $\varepsilon \downarrow 0$. It allows, in a sense, to make those two averating successively.

We need some auxilliary results from [1] (Corollaries 1 and 1′).

LEMMA 3.1. – *Suppose that the family of σ–fields $\mathcal{N}_s^t$, $0 \le s \le t < \infty$, satisfies absolute regularity mixing condition. Let $G(x, \omega)$ be $\mathcal{N}_t^\infty$-measurable random variable for each $x \in R^2$, $E \sup_x |G(x, \omega)|^\gamma < \infty$ for some $\gamma > 1$, and the random variable $\zeta(\omega)$ to be $\mathcal{N}_0^s$-measurable. Set $g(x) = EG(x, \omega)$. Then for $s < t$ and any set $A \in \mathcal{N}_0^t$ with $P(A) \neq 0$*

$$(3.1)\quad |E\{G(\zeta(\omega), \omega) \mid A\} - E\{g(\zeta(\omega)) \mid A\}| < \frac{C}{p(A)} \beta^{1-1/\gamma}(t-s),$$

$$(3.2)\qquad |EG(\zeta(\omega), \omega) - Eg(\zeta(\omega))| < \beta^{1-1/\gamma}(t-s).$$

LEMMA 3.1′. – *Let*

$$G(x, w) = \sum_{k=1}^n U_k(x) V_k(\omega), \qquad n < \infty.$$

Suppose that $|U_k(x)| < C$, $E|V_k(\omega)|^\gamma < C$ for some $\gamma > 1$. Then Lemma 2.1 holds when the condition of absolute regularity of the family $\mathcal{N}_s^t$, $0 \le s \le t < \infty$, is replaced by the strong mixing condition and correspondingly the coefficient $\beta(\tau)$ in (3.1), (3.2) is replaced by $\alpha(\tau)$.

These lemmas allows us to use conditions (2.6), (2.7) for the proof of Theorem 2.1. Taking into account that $H(x)$ is the first integral we have that $(\nabla H(x), b(x)) = 0$ for $x \in R$. Due to this fact the increments of the process $Y^\varepsilon(t) = H(X^\varepsilon(t))$ can be represented as follows.

$$(3.3)\quad \begin{aligned} Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1) &= \frac{1}{\varepsilon}\int_{\tau_1}^{\tau_2} (\nabla H(X^\varepsilon(s)), b(X^\varepsilon(s), \zeta_{s/\varepsilon^2}))\, ds \\ &= \frac{1}{\varepsilon}\int_{\tau_1}^{\tau_2} (\nabla H(X^\varepsilon(s)), g(X^\varepsilon(s), \zeta_{s/\varepsilon^2}))\, ds \\ &= \frac{1}{\varepsilon}\int_{\tau_1}^{\tau_2} F(X^\varepsilon(s), \zeta_{s/\varepsilon^2})\, ds. \end{aligned}$$

Denote by $[a]$ the greatest integer not exceeding a. Let $\delta = \varepsilon^{3/2+1/(p-2)}$. Obviously $\varepsilon^{5/3} \le \delta \le \varepsilon^{3/2}$, since $p > 8$. We devide any time interval (τ_1, τ_2) in subintervals (v_k, v_{k+1}), $v_k = k\,\delta$, where k is an integer number. Let also $v_{[\tau_1/\delta]} = \tau_1$, $v_{[\tau_2/\delta]+1} = \tau_2$, $v_{[\tau_1/\delta]-1} = \tau_1 - \min\{\tau_2 - \tau_1, \delta\}$. This agreement in fact implies that the length of the first and last subintervals can be less then δ.

To prove Theorems 2.1 and 2.2 we need many technical results. To clarify the main ideas of the proof of these results we present first some rough calculations.

Let A be an arbitrary subset of $N_0^{\tau_1/\varepsilon^2}$ with $P(A) \neq 0$. Note that the process $X^\varepsilon(t)$, $0 \leq t \leq \tau$, is measurable with respect to $N_0^{\tau_1/\varepsilon^2}$. Using (3.3) and the fact that as $\delta/\varepsilon^2 \to \infty$, $\delta \to 0$ the process $X^\varepsilon(v_{k-1})$ weakly depends on ζ_{s/ε^2}, $s \in [v_k, v_{k+1})$, we can for any $\tau_1 < \tau_1$ obtain

$$E\{Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1) \mid A\} = \frac{1}{\varepsilon} \int_{\tau_1}^{\tau_2} E\{F(X^\varepsilon(s), \zeta_{s/\varepsilon^2}) \mid A\} \, ds$$

$$\approx \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \frac{1}{\varepsilon} \int_{v_k}^{v_{k+1}} E\{F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}) \mid A\} \, ds$$

$$+ \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \frac{1}{\varepsilon^2} \int_{v_k}^{v_{k+1}} ds \int_{v_k}^{s} dt \; E\{(\nabla F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}),$$

$$g(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})) \mid A\}.$$

The first term here is negligible. Using stationarity of the process ζ_t and the definition of functions $Q(x, s)$, $Q(x)$ the second term can be rewritten in a form which implies

$$E\{Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1) \mid A\}$$

$$\approx \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} E\left\{\frac{1}{\varepsilon} \int_0^\delta ds \int_{v_k}^{s} dt \; Q\left(X^\varepsilon(v_{k-1}), \frac{s-t}{\varepsilon^2}\right) \middle| A\right\}$$

$$\approx \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \delta E\{Q(X^\varepsilon(v_{k-1})) \mid A\}.$$

At this stage we carry out the first averaging with respect to the fast oscillating noise. Next one must use the closeness of the non-perturbed solution $X(t)$ with $\tilde{X}^\varepsilon(t) = X^\varepsilon(\varepsilon t)$ and the fact that the solution $X_x(t)$ of the equation (2.2) starting at the point x possess the ergodic properties.

If we choose $\Delta \to 0$ such that $\Delta/\delta \to \infty$, $\Delta/\varepsilon \to \infty$, then we will have

$$
\begin{aligned}
&E\{Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1) \mid A\} \\
&\approx \sum_{l=[\tau_1/\delta]}^{[\tau_2/\delta]} \delta \sum_{k=[t_l/\delta]}^{[t_{l+1}/\delta]} E\{(Q(X^\varepsilon(v_k)) \mid A\} \\
&\approx \sum_{l=[\tau_1/\delta]}^{[\tau_2/\delta]} \Delta E\left\{\frac{\varepsilon}{\Delta}\int_0^{\Delta/\varepsilon} Q(X_{x^\varepsilon(t_l)}(s))\,ds \mid A\right\} \\
&\approx \sum_{l=[\tau_1/\delta]}^{[\tau_2/\delta]} \Delta E\{B(H(X^\varepsilon(t_l)) \mid A\} \\
&\approx E\left\{\int_{\tau_1}^{\tau_2} B(H(X^\varepsilon(t))\,dt \mid A\right\} = E\left\{\int_{\tau_1}^{\tau_2} B(Y^\varepsilon(t))\,dt \mid A\right\}.
\end{aligned}
$$

Analogously one can obtain

$$
\begin{aligned}
&E\{(Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1))^2 \mid A\} \\
&\approx E\left\{\left(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \frac{1}{\varepsilon}\int_{v_k}^{v_{k+1}} F(X^\varepsilon(v_{k-1}, \zeta_{s/\varepsilon^2})\,ds\right)^2 \Bigg| A\right\} \\
&\approx \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} E\left\{\frac{1}{\varepsilon^2}\left(\int_{v_k}^{v_{k+1}} F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2})\,ds\right)^2 \Bigg| A\right\} \\
&= \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \frac{2}{\varepsilon^2}\int_{v_k}^{v_{k+1}} ds \\
&\quad \times \int_{v_k}^{s} dt\, E\{F(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})\, F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}) \mid A\} \\
&= \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \frac{2}{\varepsilon^2}\int_{v_k}^{v_{k+1}} ds \int_{v_k}^{s} dt\, E\left\{D\left(X^\varepsilon(v_{k-1}), \frac{s-t}{\varepsilon^2}\right) \Bigg| A\right\} \\
&\approx \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \delta E\{D(X^\varepsilon(v_{k-1})) \mid A\} \\
&\approx \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \Delta E\left\{\frac{\varepsilon}{\Delta}\int_0^{\Delta/\varepsilon} D(X_{x^\varepsilon(t_l)}(s))\,ds \mid A\right\}
\end{aligned}
$$

$$\approx \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \Delta E\{\sigma^2(H(X^\varepsilon(t_l)) \mid A\}$$

$$\approx E\left\{\int_{\tau_1}^{\tau_2} \sigma^2(Y^\varepsilon(t))\, dt \mid A\right\}.$$

Suppose that we can prove the existence of the weak limit

$$Y^\varepsilon(t) \to Y(t), \quad \text{as } \varepsilon \to 0,$$

where $Y(t)$ is some limiting process. Then for the process $Y(t)$ the following relations for the conditional first and second moments will be hold:

$$E\{Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1) \mid A\} = E\left\{\int_{\tau_1}^{\tau_2} B(Y(t))\, dt \mid A\right\},$$

$$E\{(Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1))^2 \mid A\} = E\left\{\int_{\tau_1}^{\tau_2} \sigma^2(Y(t))\, dt \mid A\right\},$$

where A is some arbitrary subset from the σ-field generated by the process $Y(s)$ up to time τ_1. These two equalities will imply that $Y(t)$ is the diffusion process with drift coefficient $B(x)$ and diffusion coefficient $\sigma^2(x)$.

LEMMA 3.2. – *The increments of the process $Y^\varepsilon(t)$ can be written in the form*

$$(3.4) \quad Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1) = \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \{\xi_k + \eta_k + \varphi_k\} + \theta_\varepsilon(\tau_2) - \theta_\varepsilon(\tau_1),$$

where

$$\xi_k = \frac{1}{\varepsilon} \int_{v_k}^{v_{k+1}} F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2})\, ds,$$

$$\eta_k = \frac{1}{\varepsilon^2} \int_{v_k}^{v_{k+1}} ds \int_{v_k}^{s} dt\, (\nabla F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}), g(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})),$$

$$\varphi_k = \frac{1}{\varepsilon^2} \int_{v_k}^{v_{k+1}} ds \int_{v_{k-1}}^{v_k} dt\, (\nabla F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}), g(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})),$$

and $\theta_\varepsilon(\tau)$ is some process for which

$$(3.5) \qquad E(\theta_\varepsilon(\tau_2) - \theta_\varepsilon(\tau_1))^2 = (\tau_2 - \tau_1)^{3/2}\, 0(1).$$

(Here and in the sequel the letter θ with some argument or index is used for the random variables which are not essential for the limiting behavior of the process $Y^\varepsilon(\tau)$. The symbol $o(1)$ stands for a quantity which tends to zero as $\varepsilon \to 0$ uniformly with respect to any parameters and $0(1)$ is uniformly bounded quantity.)

Proof. – Substituting in the relation

$$Y^\varepsilon(v_{k+1}) - Y^\varepsilon(v_k) = \frac{1}{\varepsilon}\int_{v_k}^{v_k} F(X^\varepsilon(s), \zeta_{s/\varepsilon^2})\, ds,$$

the Taylor expansion of the function $F(x, \zeta_{s/\varepsilon^2})$ with the remainder term of the second order we obtain:

$$\begin{aligned}(3.6)\qquad & Y^\varepsilon(v_{k+1}) - Y^\varepsilon(v_k)\\ &= \xi_k + \frac{1}{\varepsilon}\int_{v_k}^{v_{k+1}} (\nabla F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2})),\\ &(X^\varepsilon(s) - X^\varepsilon(v_{k-1}))\, ds + \theta_{1,k}\end{aligned}$$

where

$$\begin{aligned}\theta_{1,k} = \frac{1}{2\varepsilon}\int_{v_k}^{v_{k+1}} & (X^\varepsilon(s) - X^\varepsilon(v_{k-1}))^T \nabla(\nabla F(\tilde{X}_s^\varepsilon, \zeta_{s/\varepsilon^2}))^T\\ & \times (X^\varepsilon(s) - X^\varepsilon(v_{k-1}))\, ds.\end{aligned}$$

Here the symbol T is used for transposition and $\tilde{X}_s^\varepsilon$ is some intermediate point. Substituting the Taylor expansion of the vector function $b(x, \zeta_{t/\varepsilon^2})$ with remainder term of the first order in the relation

$$(3.7)\qquad X^\varepsilon(s) - X^\varepsilon(v_{k-1}) = \frac{1}{\varepsilon}\int_{v_{k-1}}^{s} b(X^\varepsilon(t), \zeta_{t/\varepsilon^2})\, dt,$$

we have

$$\begin{aligned}X^\varepsilon(s) - X^\varepsilon(v_{k-1}) &= \frac{1}{\varepsilon}\int_{v_{k-1}}^{s} b(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})\, dt\\ &+\frac{1}{\varepsilon}\int_{v_{k-1}}^{s} (\nabla b^T(\tilde{X}_t^\varepsilon, \zeta_{t/\varepsilon^2}))^T (X^\varepsilon(t) - X^\varepsilon(v_{k-1}))\, dt.\end{aligned}$$

Using this relation and (2.13), we find that

$$(3.8)\qquad Y^\varepsilon(v_{k+1}) - Y^\varepsilon(v_k) = \xi_k + \eta_k + \varphi_k + \theta_{1,k} + \theta_{2,k} + \theta_{3,k},$$

where

$$\theta_{2,k} = \frac{1}{\varepsilon^2} \int_{v_k}^{v_{k+1}} ds \int_{v_k}^{s} dt \, (\nabla F (X^{\varepsilon}_{k-1}, \zeta_{s/\varepsilon^2}),$$
$$(\nabla b^T (\tilde{X}^{\varepsilon}_t, \zeta_{t/\varepsilon^2}))^T (X^{\varepsilon}(t) - X^{\varepsilon}(v_{k-1}))),$$
$$\theta_{3,k} = \frac{1}{\varepsilon^2} \int_{v_k}^{v_{k+1}} ds \, (s - v_{k-1}) \, (\nabla F (X^{\varepsilon}(v_{k-1}), \zeta_{s/\varepsilon^2}), \, b(X^{\varepsilon}(v_{k-1}))).$$

Sometimes to simplify the notations we denote $X^{\varepsilon}(v_{k-1}) = X^{\varepsilon}_{k-1}$.

Due to the condition (2.4) it follows from (3.7) that for $s \in [v_k, v_{k+1}]$

$$E \, | X^{\varepsilon}(s) - X^{\varepsilon}(v_{k-1}) |^6 \le C (\delta/\varepsilon)^6. \tag{3.9}$$

Conditions (2.3), (2.4) imply

$$E \, | \nabla (\nabla F (\tilde{X}^{\varepsilon}_s, \zeta_{s/\varepsilon^2}))^T |^6 \le C, \qquad E \, | (\nabla b^T (\tilde{X}^{\varepsilon}_s, \zeta_{t/\varepsilon^2}))^T |^6 \le C.$$

Then, using Hölder's inequality, we have

$$E\theta^2_{l,k} \le C (\delta/\varepsilon)^6, \qquad l = 1, 2,$$

and thus

$$E \left(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \theta_{l,k} \right)^2 \le C (\tau_2 - \tau_1)^2 \, \delta^4 \, \varepsilon^{-6} \tag{3.10}$$
$$= C (\tau_2 - \tau_1)^2 \, \varepsilon^{4/(p-2)}, \qquad l = 1, 2.$$

To estimate the moments of the variable $\theta_{3,k}$ note that in view of (2.3) and (2.5)

$$\begin{cases} Eg(x, \zeta_s) = 0, & E \dfrac{\partial}{\partial x_l} g(x, \zeta_s) = 0, \quad l = 1, 2. \\ EF(x, \zeta_s) = 0, & E\nabla F(x, \zeta_s) = 0. \end{cases} \tag{3.11}$$

Denote

$$U(x, \zeta) = (\nabla F(x, \zeta), \, b(x)).$$

Using this notation, we can write

$$\theta_{3,k} = \frac{1}{\varepsilon^2} \int_{v_k}^{v_{k+1}} ds \, (s - v_{k-1}) \, U(X^{\varepsilon}(v_{k-1}), \zeta_{s/\varepsilon^2}) \, ds.$$

Since the function $g(x, \zeta_{s/\varepsilon^2})$ is $\mathcal{N}^{s/\varepsilon^2}_{s/\varepsilon^2}$-measurable and the variable $X^{\varepsilon}(v_{k-1})$ is $\mathcal{N}^{v_{k-1}/\varepsilon^2}_0$-measurable, applying Lemma 3.1 with $\gamma = p/2$ and

σ-fields $\mathcal{N}^{\infty}_{s/\varepsilon^2}$, $\mathcal{N}_0^{t/\varepsilon^2}$, $v_k \le t \le s \le v_{k+1}$, and taking into account (2.4), (2.5), (3.11), we have

$$\begin{aligned} &| EU(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})\, U(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}) | \\ &\le C\beta^{(p-2)/p}((s-t)/\varepsilon^2). \end{aligned}$$

Then

$$\begin{aligned} E\theta^2_{3,k} &\le C\frac{\delta^2}{\varepsilon^4}\int_{v_k}^{v_{k+1}} ds \int_{v_k}^{s} dt\, \beta^{(p-2)/p}((s-t)/\varepsilon^2) \\ &\le C\delta^2 \int_0^{\delta/\varepsilon^2} ds \int_0^s dt \min\{1, (s-t)^4\} \le C\delta^3\varepsilon^{-2}. \end{aligned}$$

Here we use condition (2.7). Let $l + 2 \le k$ and $s \in \{v_k, v_{k+1}\}$, $t \in [v_l, v_{l+1}]$. Applying Lemma 3.1 with $\gamma = p/2$ and σ-fields $\mathcal{N}^{\infty}_{s/\varepsilon^2}$, $\mathcal{N}_0^{v_{k-1}}$, we obtain:

$$\begin{aligned} &| EU(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})\, U(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}) | \\ &\le C\beta^{(p-2)/p}((s-v_{k-1})/\varepsilon^2) \le C(\varepsilon^2/\delta)^4, \end{aligned}$$

and thus

$$| E\theta_{3,l}\,\theta_{3,k} | \le C\frac{\delta^4}{\varepsilon^4}\left(\frac{\varepsilon^2}{\delta}\right)^4 = C\varepsilon^4$$

Now we have

$$\begin{aligned} (3.12)\quad & E\left(\sum_{K=[\tau_1/\delta]}^{[\tau_2/\delta]} \theta_{3,k}\right)^2 \\ &\le 3\left(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} E\theta^2_{3,k} + \sum_{l=[\tau_1/\delta]}^{[\tau_2/\delta]-2}\ \sum_{k=l+2}^{[\tau_2/\delta]} E\theta_{3,l}\,\theta_{3,k}\right) \\ &\le C\left((\tau_2-\tau_1)\frac{\delta^2}{\varepsilon^2} + (\tau_2-\tau_1)^2\cdot\frac{\varepsilon^4}{\delta^2}\right) \le C(\tau_2-\tau_1)^{3/2}\,.\,\varepsilon^{1/4}. \end{aligned}$$

Here we used the agreement $v_{[t_1/\delta]} = \tau_1$, $v_{[t_2/\delta]} = \tau_2$, which implies that if $\tau_2 - \tau_1 < \delta$, then for $E\theta^2_{3,k}$, $k = [\tau_1/\delta]$, we have the bound

$$\begin{aligned} (3.13)\quad E\theta^2_{3,k} &\le C(\tau_2-\tau_1)^3\varepsilon^{-2} \le C(\tau_2-\tau_1)^{3/2}\delta^{3/2}\varepsilon^{-2} \\ &\le C(\tau_2-\tau_1)^{3/2}\varepsilon^{1/4}. \end{aligned}$$

Let

$$\theta_\varepsilon(\tau) = \sum_{k=0}^{[\tau/\delta]} (\theta_{1,k} + \theta_{2,k} + \theta_{3,k}).$$

Adding up (3.8) we obtain (3.4). The estimate (3.5) is the consequence of the bounds (3.10), (3.12). □

In the following lemmas we assume that $A \in \mathcal{N}_0^{\tau_1/\varepsilon^2}$ and $P(A) \neq 0$.

LEMMA 3.3. – *The following estimates hold for the expectations of* ξ_k, η_k, φ_k:

$$(3.14)\quad \begin{cases} E\{\xi_k \mid A\} = \delta\, \dfrac{o(1)}{P(A)}, \\ E\{\eta_k \mid A\} - \delta E\{Q(X^\varepsilon(v_{k-1})) \mid A\} = \delta\, \dfrac{o(1)}{P(A)}, \\ E\{\varphi_k \mid A\} = \delta\, \dfrac{o(1)}{P(A)}. \end{cases}$$

Proof. – Since $A \in \mathcal{N}_0^{v_{k-1}/\varepsilon^2}$, the variable $X^\varepsilon(v_{k-1})$ is $\mathcal{N}_0^{v_{k-1}/\varepsilon^2}$-measurable and for $s \in [v_k, v_{k+1}]$ the process ζ_{s/ε^2} is $\mathcal{N}_{v_k/\varepsilon^2}^{\infty}$-measurable, we obtain from (3.1), with $\gamma = p$, (2.4), (2.5) and (3.11):

$$|E\{F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}) \mid A\}| \leq \frac{C}{P(A)}\, \beta^{(p-1)/p}(\delta/\varepsilon^2) \leq \frac{C\mathcal{E}^{4/3}}{P(A)}.$$

Integrating this relation with respect to s from v_k to v_{k+1}, we obtain the first of the estimates (3.14). Applying Lemma 3.1 with $\gamma = p/2$ and the σ-fields $\mathcal{N}_{t/\varepsilon^2}^{\infty}$, $\mathcal{N}_0^{s/\varepsilon^2}$ we have for $v_{k-1} \leq t \leq v_k \leq s \leq v_{k+1}$ that

$$\begin{aligned} |E\{(\nabla F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}), g(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})) \mid A\}| \\ \leq \frac{C}{P(A)}\, \beta^{(p-2)/p}((s-t)/\varepsilon^2). \end{aligned}$$

Since

$$\begin{aligned} \int_{v_k}^{v_{k+1}} ds \int_{v_{k-1}}^{v_k} dt\, \beta^{(p-2)/p}((s-t)/\varepsilon^2) \\ \leq C\varepsilon^4 \int_{\delta/\varepsilon^2}^{2\delta/\varepsilon^2} ds \int_0^{\delta/\varepsilon^2} dt\, \min\{1, (s-t)^{-4}\} \leq C\varepsilon^4, \end{aligned}$$

we obtain the third estimate in (3.14). When $v_k \leq t \leq s \leq v_{k+1}$ using

Lemma 3.1 with σ-fields $\mathcal{N}_0^{v_{k-1}/\varepsilon^2}$, $\mathcal{N}_{v_k/\varepsilon^2}^{\infty}$, we have

$$\begin{aligned}
&| E\{ (\nabla F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}), g(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2})) \mid A\} \\
&\qquad - E\{ Q(X^\varepsilon(v_{k-1}), (s-t)/\varepsilon^2) \mid A\} | \\
&\qquad \leq \frac{C}{P(A)} \beta^{(p-2)/p}(\delta/\varepsilon^2) \mid \leq \frac{C\varepsilon^8}{P(A)\,\delta^4},
\end{aligned}$$

and thus

$$\begin{aligned}
(3.15)\qquad & | E\{ \eta_k \mid A\} \\
& -E\left\{ \frac{1}{\varepsilon^2} \int_0^\delta ds \int_0^s dt\, Q\left(X^\varepsilon(v_{k-1}), \frac{s-t}{\varepsilon^2} \right) \mid A \right\} \mid \leq \frac{c\,\delta}{P(A)} \cdot \varepsilon.
\end{aligned}$$

From (3.2) and (3.11) one can deduce that

$$(3.16)\qquad | Q(x, u) | \leq c\beta^{(p-2)/p}(u).$$

Thus $Q(x, u)$ is an integrable function with respect to u and

$$\begin{aligned}
(3.17)\qquad & \left| \varepsilon^2 \int_0^{\delta/\varepsilon^2} ds \int_0^s Q(x, s-t)\, dt - \delta \int_0^\infty Q(x, u)\, du \right| \\
& \qquad \leq \varepsilon^2 \int_0^{\delta/\varepsilon^2} ds \int_s^\infty \beta^{(p-2)/p}(u)\, du \leq C\varepsilon^2.
\end{aligned}$$

Thus the second estimate in (3.14) is also proved. □

Remark 3.1. – Using (3.2) and taking into account both relations 3.11, one can obtain the following bound for the derivative of the function $Q(x, u)$:

$$\left| \frac{\partial}{\partial x} Q(x, u) \right| \leq C\beta^{(p-2)/p}(u).$$

This bound and (3.16) imply that the function $Q(x) = \int_0^\infty Q(x, u)\, du$ is bounded and satisfies the Lipschitz condition

$$(3.18)\qquad | Q(y) - Q(x) | \leq L\, | y - x |.$$

As a consequence of Lemmas 3.2, 3.3, we have the following result.

LEMMA 3.4. – *As* $\varepsilon \to 0$

$$E\{Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1) \mid A\} = E\left\{\delta \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} Q(X^\varepsilon(v_k)) \mid A\right\} + \frac{o(1)}{P(A)}. \tag{3.19}$$

We note only that since $|Q(x)| \le C$ *then*

$$\left| \delta \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} Q(X^\varepsilon(v_{k-1})) - \delta \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} Q(X^\varepsilon(v_k)) \right| \le 2\delta C. \quad \square \tag{3.20}$$

LEMMA 3.5. – *As* $\varepsilon \to 0$ *the following relations hold*

$$\begin{cases} E\{\xi_k^2 \mid A\} - \delta D(X^\varepsilon(v_{k-1})) \mid A\} = \dfrac{\delta}{P(A)}\, o(1), \\ E\{\eta_k^4 \mid A\} = \dfrac{\delta^4}{P(A)}\, 0(1), \qquad E\{\varphi_k^4 \mid A\} = \dfrac{\delta^4}{P(A)}\, 0(1). \end{cases} \tag{3.21}$$

Proof. – Let $v_k \le t \le s \le v_{k+1}$. Using Lemma 3.1 with $\gamma = p/2$ and the σ-fields $\mathcal{N}_0^{v_{k-1}/\varepsilon^2}$, $\mathcal{N}_{v_k/\varepsilon^2}^{\infty}$, we obtain

$$\begin{aligned} &|E\{F(X^\varepsilon(v_{k-1}), \zeta_{t/\varepsilon^2}) - F(X^\varepsilon(v_{k-1}), \zeta_{s/\varepsilon^2}) \mid A\} \\ &\quad - E\{D(X^\varepsilon(v_{k-1}), (s-t)/\varepsilon^2) \mid A\}| \\ &\le \frac{C}{P(A)} \beta^{(p-2)/p}(\delta/\varepsilon^2) \le \frac{C\varepsilon^8}{P(A)\,\delta^4}, \end{aligned}$$

Now calculations analogous to (3.15)-(3.17) yield the first of the relations in (3.21). To prove the other relations it is sufficient to estimate $E\eta_k^4$ and $E\varphi_k^4$. Let $s_i \in [v_k, v_{k+1}]$, $t_i \in [v_k, s_i]$, $i = 1, 2, 3, 4$,

$$f(x, t_1, \ldots, t_4, s_1, \ldots, s_4) = E\prod_{i=1}^{4} (\nabla F(x, \zeta_{s_i/\varepsilon^2}), g(x, \zeta_{t_i/\varepsilon^2})).$$

Applying Lemma 3.1 with $\gamma = p/8$ and σ-fields $\mathcal{N}_{v_k/\varepsilon^2}^{\infty}$, $\mathcal{N}_0^{v_{k-1}/\varepsilon^2}$, we obtain

$$\begin{aligned} &\left| E\prod_{i=1}^{4} (\nabla F(X^\varepsilon(v_{k-1}), \zeta_{s_i/\varepsilon^2}), g(X^\varepsilon(v_{k-1}), \zeta_{t_i/\varepsilon^2})) \right. \\ &\quad \left. - Ef(X^\varepsilon(v_{k-1}), t_1, \ldots, t_4, s_1, \ldots, s_4) \right| \\ &\le C\beta^{(p-8)/p}(\delta/\varepsilon^2) \le C(\varepsilon^2/\delta)^4. \end{aligned}$$

Integrating this relation with respect to s_i, t_i from the region $\{ v_k \leq t_i \leq s_i \leq v_{k+1} \}$ and taking into account the bound (*see* [8], Lemma 2.1),

$$\underset{v_k \leq t_i \leq s_i \leq v_{k+1}}{\int, \ldots, \int} f(x, t_1, \ldots, t_4, s_1, \ldots, s_4)\, dt_1, \ldots, ds_4 \leq C\,\delta^4\,\varepsilon^8,$$

we conclude that $E\eta_k^4 \leq C\,\delta^4$. Note that the method of the proof of this bound does not permit to improve the conditions (2.6), (2.7). It is a little more difficult to obtain the estimate for $E\varphi_k^4$. The detailed proof of an analogous estimate is given in ([1], Lemma 5), and we refer the reader to this paper. □

LEMMA 3.6. – *The following equality holds as* $\varepsilon \downarrow 0$:

$$E\left\{\left(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \xi_k\right)^2 \,\middle|\, A\right\} \tag{3.22}$$
$$= \delta E\left\{\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} D(X^\varepsilon(v_k)) \,\middle|\, A\right\} + \frac{(\tau_2 - \tau_1)}{P(A)}\, o(1)$$
$$+ \min\{(\tau_2 - \tau_1),\, \varepsilon^{3/2}\}\, 0(1).$$

Proof. – The left-hand side of (3.22) can be represented in the form

$$\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} E\{\xi_k^2 \,|\, A\} + 2 \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]-1} E\{\xi_k\, \xi_{k+1} \,|\, A\}$$
$$+2 \sum_{l=[\tau_1/\delta]}^{[\tau_2/\delta]-2} \sum_{k=l+2}^{[\tau_2/\delta]} E\{\xi_k\, \xi_l \,|\, A\}.$$

Using (3.1), for $\gamma = p/2$, we obtain

$$|E\{\xi_k\, \xi_{k+1} \,|\, A\}|$$
$$\leq \frac{C}{P(A)\,\varepsilon^2} \int_{v_k}^{v_{k+1}} ds \int_{v_{k+1}}^{v_{k+2}} dt\, \beta^{(p-2)/p}((t-s)/\varepsilon^2) \leq \frac{C\varepsilon^2}{P(A)}$$

and for $l + 2 \leq k$

$$|E\{\xi_l\, \xi_k \,|\, A\}| \leq \frac{C\,\delta^2}{P(A)\,\varepsilon^2}\, \beta^{(p-2)/p}(\delta/\varepsilon^2) \leq \frac{C\,\delta^2}{P(A)\,\varepsilon^2}$$
$$\leq \left(\frac{\varepsilon^2}{\delta}\right)^{4(p-2)/(p-8)} = \frac{C\,\delta^2}{P(A)}\, \varepsilon^{8/(p-8)}.$$

Here we used the special choice of the parameter $\delta\left(\delta=\frac{3}{2}+\frac{1}{p-2}\right)$ and the condition (2.7). These estimates together with the first of the relations (3.21) imply (3.22). It is necessary only to take into account the estimate for the function D similar to (3.20), what explains the appearance of the last term in the right-hand side of (3.22). □

LEMMA 3.7. – *The following inequality holds as* $\varepsilon < 0$:

$$\begin{aligned} \Big| & E\{(Y^\varepsilon(\tau_2)-Y^\varepsilon(\tau_1))^2 \mid A\} \\ & - \delta E\left\{ \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} D(X^\varepsilon(v_k)) \,\Big|\, A\right\}\Big| \\ & \le CP^{-1}(A)(|\tau_2-\tau_1|^{3/2}+o(1)). \end{aligned} \tag{3.23}$$

Proof. – Due to (3.4) and Hölder's inequality

$$\begin{aligned} & \left| E\left\{ (Y^\varepsilon(\tau_2)-Y^\varepsilon(\tau_1))^2 - \Big(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \xi_k\Big) \,\Big|\, A\right\}\right| \\ & \le C\left\{ E^{1/2}\left\{\Big(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \xi_k\Big)^2 \,\Big|\, A\right\}\left(E^{1/2}\left\{\Big(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} (\eta_k+\varphi_k)\Big)^2 \,\Big|\, A\right\}\right.\right. \\ & \left. + E^{1/2}\{(\theta_\varepsilon(\tau_2)-\theta_\varepsilon(\tau_1))^2 \mid A\}\right) \\ & \left. + E\left\{\Big(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} (\eta_k+\varphi_k)\Big)^2 \,\Big|\, A\right\} + E\{(\theta_\varepsilon(\tau_2)-\theta_\varepsilon(\tau_1))^2 \mid A\}\right\}. \end{aligned}$$

The function $D(x)$ is bounded for the same reason as it holds for the function $Q(x)$ (*see* Remark 2.1). Therefore, we have from (3.22):

$$E\left\{\Big(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \xi_k\Big)^2 \,\Big|\, A\right\} \le \frac{C}{P(A)}(\tau_2-\tau_1).$$

By virtue of (3.21)

$$E\left\{\Big(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} (\eta_k+\varphi_k)\Big)^2 \,\Big|\, A\right\} \le \frac{C}{P(A)}(\tau_2-\tau_1)^2.$$

Now (3.23) is an immediate consequence of (3.5) and (3.22).

COROLLARY 3.1. – *For* τ, $\tau_l \in [0, T]$, $l = 1, 2$, *the following relations hold:*

$$E\,|\,Y^{\varepsilon}(\tau)\,|^2 \leq C, \qquad E\,|\,Y^{\varepsilon}(\tau_2) - Y^{\varepsilon}(\tau_1)\,|^2 \leq C\,|\,\tau_2 - \tau_1\,| + o(1).$$

It follows from this corollary (*see* [5], remark on p. 13) that for any sequence $\varepsilon_n \to 0$ there is a subsequence $\varepsilon_{n_m} \to 0$ such that on some probability space it is possible to define random processes $Y_0^{\varepsilon_{nm}}(t)$, which have the same finite-dimensional distributions as $Y^{\varepsilon_{nm}}(t)$ and $Y_0^{\varepsilon_{nm}}(t) \to Y(t)$ as $m \to \infty$ in probability, where $Y(t)$ is some stochastically continuous separable random process.

The end of the proof of Theorem 2.1 will consist of the following steps. We shall prove that the process $Y(t)$ is continuous with probability one. Then it will be shown that the finite-dimensional distributions of the process $Y(t)$ obtained in this way are independent of the choice of the subsequence ε_{n_m}. Moreover, it will be verified that $Y(t)$ is a diffusion process determined by the stochastic differential equation (2.8). Finally, we establish the weak compactness of the family $\{Y^{\varepsilon}(t),\ t \in [0, T]\}_{\varepsilon>0}$ of random process. All these statement together imply that the processes $Y^{\varepsilon}(t)$, $t \in [0, T]$, converges weakly as $\varepsilon \to 0$ to the process $Y(t)$ determined by (2.8).

LEMMA 3.8. – *The trajectories of the process* $Y(t)$ *are continuous with probability one and for* $0 \leq \tau_1 \leq \tau_2 \leq T$

$$E\,(Y(\tau_2) - Y(\tau_1))^4 < C\,(\tau_2 - \tau_1)^2. \tag{3.24}$$

Proof. – As is known from the Kolomogorov theorem (*see*, for example [12]), (3.24) implies the first statement of the Lemma 3.8. To prove (3.24) let

$$Y_*^{\varepsilon}(t) = \sum_{k=0}^{[t/\delta]} \{\,\xi_k + \eta_k + \varphi_k\,\}.$$

It follows from (3.4) that

$$Y^{\varepsilon}(t) = Y_*^{\varepsilon}(t) + \theta_{\varepsilon}(t), \tag{3.25}$$

and hence the limiting behavior of the process $Y_*^{\varepsilon}(t)$ and $Y^{\varepsilon}(t)$ is the same. Therefore to prove (3.24) it is sufficient to verify that

$$E\,(Y_*^{\varepsilon}(t_2) - Y_*^{\varepsilon}(t_1))^4 \leq C\,(\tau_2 - \tau_1)^2. \tag{3.26}$$

Taking into account the convergence $Y_0^\varepsilon(t) \to Y(t)$ and Foutu's lemma, (3.25) and (3.26) imply (3.24).

Obviously

$$E(Y_*^\varepsilon(t_2) - Y_*^\varepsilon(t_1))^4$$
$$\le CE\left\{\left(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \xi_k\right)^4 + \left(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} (\eta_k + \varphi_k)\right)^4\right\}.$$

Due to (3.21)

$$E\left(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} (\eta_k + \varphi_k)\right)^4$$
$$\le C\left(\frac{\tau_2 - \tau_1}{\delta}\right)^3 \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} (E\eta_k^4 + E\varphi_k^4) \le C(\tau_2 - \tau_1)^4.$$

The proof of the bound

$$E\left(\sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} \xi_k\right)^4 \le C(\tau_2 - \tau_1)^2$$

require more detailed consideration. Similar calculations in a more general case were made in [8] (Lemma 3.2) or in [1] (Lemma 8). For our case one should put there $\phi^0(x, s) = 0$, $\delta(T) = 0$. We thus omit the proof of this bound. The mentioned estimates imply (3.26). The Lemma is proved. □

The main estimates connected with the averaging due to the mixing properties of the stationary process ζ_t have been obtained. Now we pass to the second averaging connected with the asymptotic behavior of the solutions of the non-perturbed equation (2.2) when $t \to \infty$.

Let $\tilde{X}^\varepsilon(t) = X^\varepsilon(\varepsilon t)$. It is obvious that the process $\tilde{X}^\varepsilon(t)$ satisfies the equation

$$\frac{d\tilde{X}^\varepsilon(t)}{dt} = b(\tilde{X}^\varepsilon(t), \zeta_{t/\varepsilon}), \qquad \tilde{X}^\varepsilon(0) = X_0. \tag{3.27}$$

It turns out (*see*, for example [8], Theorem 1.1 or [5], Ch. 7) that the process $\tilde{X}^\varepsilon(t)$ converges as $\varepsilon \to 0$ to the solution $X(t)$ of the non-perturbed equation (2.2). The following result gives an L_1 estimate for this convergence

LEMMA 3.9. – *For any* $t \ge 0$

$$E\,|\,\tilde{X}^\varepsilon(t) - X(t)\,| \le C\sqrt{\varepsilon t(1+t)}\,\exp(Lt), \tag{3.28}$$

where L and C are some constants.

Proof. – It follows from (3.27) and (2.2) that

$$\tilde{X}^{\varepsilon}(t) - X(t) = \int_0^t (b(\tilde{X}^{\varepsilon}(s)) - b(X(s))\, ds + \int_0^t g(\tilde{X}^{\varepsilon}(s), \zeta_{s/\varepsilon})\, ds.$$

In view of the definition of the function $b(x)$ and (2.5)

$$|b(y) - b(x)| \leq L\,|y - x|$$

for some constant L. Then

$$E\,|\tilde{X}^{\varepsilon}(t) - X(t)| = L \int_0^t E\,|\tilde{X}^{\varepsilon}(s) - X(s)|\, ds + E\left|\int_0^t g(\tilde{X}^{\varepsilon}(s), \zeta_{s/\varepsilon})\, ds\right|.$$

By Gronwell's lemma we have

$$E\,|\tilde{X}^{\varepsilon}(t) - X(t)| \leq \sup_{0 \leq v \leq t} E\left|\int_0^v g(\tilde{X}^{\varepsilon}(s), \zeta_{s/\varepsilon})\, ds\right| \exp(Lt),$$

and we must verify only that

$$E\left(\int_0^t g(\tilde{X}^{\varepsilon}(s), \zeta_{s/\varepsilon})\, ds\right)^2 \leq C\varepsilon\,(t + t^2). \tag{3.29}$$

Using Taylor expansion we can write

$$\int_{\sqrt{\varepsilon}}^t g(\tilde{X}^{\varepsilon}(s), \zeta_{s/\varepsilon})\, ds = \int_{\sqrt{\varepsilon}}^t g(\tilde{X}^{\varepsilon}(s - \sqrt{\varepsilon}), \zeta_{s/\varepsilon})\, ds$$
$$+ \int_{\sqrt{\varepsilon}}^t (\nabla g^T(\tilde{X}_s^{\varepsilon}, \zeta_{s/\varepsilon}))^T (\tilde{X}^{\varepsilon}(s) - \tilde{X}^{\varepsilon}(s - \sqrt{\varepsilon}))\, ds = I_1 + I_2,$$

where symbol T is used for the transposition and $\tilde{X}_s^{\varepsilon}$ is some intermediate point.

Because of (2.5) and (3.27)

$$E\,|I_2|^2 \leq C\varepsilon t^2.$$

Applying (3.2) and taking into account (3.11) we have

$$|E(g(\tilde{X}^{\varepsilon}(s - \sqrt{\varepsilon}), \zeta_{s/\varepsilon}), g(\tilde{X}^{\varepsilon}(v - \sqrt{\varepsilon}), \zeta_{v/\varepsilon}))|$$
$$\leq C\beta^{(p-2)/2}\left(\min\left\{\sqrt{\varepsilon}, \frac{|s - v|}{\varepsilon}\right\}\right),$$

and therefore

$$E\,|\,I_1\,|^2 \le C \int_{\sqrt{\varepsilon}}^{t} \int_{\sqrt{\varepsilon}}^{t} \beta^{(p-2)/2} \left(\min\left\{\sqrt{\varepsilon},\, \frac{|\,s-v\,|}{\varepsilon}\right\}\right)$$
$$\times\, ds\, dv \le C\,(\varepsilon^2\, t^2 + \varepsilon t).$$

Thus the estimate (3.29) and hence Lemma 3.9 is proved. □

Let us consider the sums

$$S_1^\varepsilon = \delta \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} Q\,(X^\varepsilon\,(V_k)), \qquad S_2^\varepsilon = \delta \sum_{k=[\tau_1/\delta]}^{[\tau_2/\delta]} D\,(X^\varepsilon\,(V_k))$$

which are contained in the relations (3.19) and (3.22) correspondingly. These sums are similar, so we will consider only the first one. Let Δ satisfy the conditions: $\Delta = \dfrac{1}{3\,L}\,\varepsilon \ln(1/\varepsilon)$, where L is the constant from (3.28); Δ/δ is an integer number. Let $t_l = l\,\Delta$. Without loss of generality one can assume that $t_{[\tau_1/\Delta]} = \tau_1$, $t_{[\tau_2/\Delta]+1} = \tau_2$. For arbitrary δ and any variable v put $[v]_\delta = [v/\delta]\,\delta$. Then the sum S_1^ε can be rewritten as follows

$$\begin{aligned}
S_1^\varepsilon &= \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \delta \sum_{k=[t_l/\delta]}^{[t_{l+1}/\delta]} \cdot\, Q\,(X^\varepsilon\,(V_k)) = \sum_{k=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \int_{t_l}^{t_{l+1}} Q\,(X^\varepsilon\,([v]_\delta))\,dv \\
&= \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \varepsilon \int_{t_l/\varepsilon}^{t_{l+1}/\varepsilon} Q\,(X^\varepsilon\,(\varepsilon\,[u]_{\delta/\varepsilon}))\,du \\
&= \sum_{k=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \varepsilon \int_{t_l/\varepsilon}^{t_{l+1}/\varepsilon} Q\,(\tilde{X}^\varepsilon\,([u]_{\delta/\varepsilon}))\,du \\
&= \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \varepsilon \int_{0}^{\Delta/\varepsilon} Q\,(\tilde{X}_{l,\,X^\varepsilon(t_l)}^\varepsilon\,([s]_{\delta/\varepsilon}))\,ds,
\end{aligned}$$

where

$$\tilde{X}_{l,\,x}^\varepsilon\,(s) = x + \int_0^s b\,(\tilde{X}_{l,\,x}^\varepsilon\,(u),\, \zeta_{(t_l+u)/\varepsilon})\,du.$$

Let

$$X_x\,(s) = x + \int_0^s b\,(X_x\,(u))\,du.$$

If one looks over the proof of Lemma 3.9, one can see that even for the random starting point $X^\varepsilon(t_l)$ the estimate (3.18) holds:

$$E \mid \tilde{X}^\varepsilon_{l,\, X^\varepsilon(t_l)}(s) - X_{X^\varepsilon(t_l)}(s) \mid \leq C \sqrt{\varepsilon s (1+s)} \exp(Ls).$$

Let

$$\tilde{S}^\varepsilon_1 = \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \varepsilon \int_0^{\Delta/\varepsilon} Q(X_{x^\varepsilon(t_l)}(s))\, ds.$$

Since the function $Q(x)$ satisfies Lipschitz condition (3.18) we have

$$\begin{aligned} (3.30)\quad & E \mid S^\varepsilon_1 - \tilde{S}^\varepsilon_1 \mid \\ & \leq L \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \varepsilon \int_0^{\Delta/\varepsilon} E \mid \tilde{X}^\varepsilon_{l,\, X^\varepsilon(t_l)}([s]_{\delta/\varepsilon}) - X_{x^\varepsilon(t_l)}(s) \mid ds \\ & \leq LC(\tau_2 - \tau_1)(\sqrt{\varepsilon} \exp(L\Delta/\varepsilon) + \delta/\varepsilon) = LC(\tau_2 - \tau_1)\, \varepsilon^{1/6}. \end{aligned}$$

LEMMA 3.10. – *The following relation holds:*

$$(3.31)\qquad \begin{cases} \dfrac{1}{V} \displaystyle\int_0^V Q(X_x(s))\, ds - B(H(x) = 0(1) \dfrac{1 + H^2(x)}{V}, \\ V \to \infty. \end{cases}$$

Proof. – Let $h = H(x)$ for some fixed point x. The trajectory $X(t)$ of the equation (2.2) starting at the point x moves periodically along the curve $C(h) = \{y : H(y) = h\}$ with some period $T(h)$. It is easy to verify that

$$\begin{aligned} (3.32)\quad & \left| \frac{1}{V} \int_0^V Q(X_x(s))\, ds - \frac{1}{T(h)} \int_0^{T(h)} Q(X_x(s))\, ds \right| \\ & \leq 2 \sup_y Q(y) \frac{T(h)}{V} \leq \frac{C}{V}(1 + h^2) \end{aligned}$$

since we assumed that $T(h) < C(1 + h^2)$. One obtains after the change of variable $y = X_x(s)$:

$$\begin{aligned} (3.33)\qquad & \frac{1}{T(h)} \int_0^{T(h)} Q(X_x(s))\, ds \\ & = \int_{C(h)} \frac{Q(y)}{|b(y)|} dl \Big/ \int_{C(h)} \frac{dl}{|b(y)|} = B(h). \end{aligned}$$

The integrals along the curve $C(h)$ exist since the function $Q(y)$ is bounded and

$$\int_{C(h)} \frac{dl}{|b(y)|} = T(h) < \infty.$$

Relations (3.32), (3.33) imply (3.31). □

In view of (3.31) and Corollary 3.1 the sum $\tilde{S}_1^\varepsilon$ can be represented as follows

$$\begin{aligned}
\tilde{S}_1^\varepsilon &= \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \Delta \frac{\varepsilon}{\Delta} \int_0^{\Delta/\varepsilon} Q(X_{X^\varepsilon(t_l)}(s))\, ds \\
&= \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \Delta B(H(X^\varepsilon(t_l)) + 0(1) \sum_{l=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \Delta \frac{\varepsilon}{\Delta} (1 + H^2(X^\varepsilon(t_l))) \\
&= \int_{\tau_1}^{\tau_2} B(Y^\varepsilon([t]_\Delta)\, dt + 0(1) \frac{\varepsilon}{\Delta} \int_{\tau_1}^{\tau_2} (1 + |Y^\varepsilon([t]_\Delta|^2)\, dt \\
&= \int_{\tau_1}^{\tau_2} B(Y^\varepsilon([t]_\Delta)\, dt + (\tau_2 - \tau_1)\, o(1).
\end{aligned}$$

In the last equality we use the estimate

$$E\,|Y^\varepsilon(t)|^2 \le C$$

from Corollary 3.11.

Now we have from (3.19), (3.30) and this representation:

$$\begin{aligned}
&E\{Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1) \mid A\} \\
&\quad = E\left\{\int_{\tau_1}^{\tau_2} B(Y^\varepsilon[t]_\Delta))\, dt \,\middle|\, A\right\} + \frac{o(1)}{P(A)}.
\end{aligned} \tag{3.34}$$

In a similar way (3.23) can be transformed into the relation

$$\begin{aligned}
&E\{(Y^\varepsilon(\tau_2) - Y^\varepsilon(\tau_1))^2 \mid A\} \\
&\quad = E\left\{\int_{\tau_1}^{\tau_2} \sigma^2(Y^\varepsilon([t]_\Delta))\, dt \,\middle|\, A\right\} \\
&\quad \le CP^{-1}(A)\,((\tau_2 - \tau_1)^{3/2} + o(1)).
\end{aligned} \tag{3.35}$$

The following lemma gives the main relations for the limiting process $Y(t)$.

LEMMA 3.11. – *Let $\mathcal{F}_\tau$ be a σ-field generated by the process $Y(t)$, when $0 \le t \le \tau$. Let $A \in \mathcal{F}_\tau$ and $P(A) \ne 0$. Then the following relations hold for the process $Y(\tau)$ for all τ_1, τ_2, $0 \le \tau_1 \le \tau_2 \le T$:*

$$E\{Y(\tau_2) - Y(\tau_1) \mid \mathcal{F}_{\tau_1}\} = E\left\{\int_{\tau_1}^{\tau_2} B(Y(t))\, dt \,\middle|\, \mathcal{F}_{\tau_1}\right\}, \tag{3.36}$$

$$E\{(Y(\tau_2) - Y(\tau_1))^2 \mid A\} = E\left\{\int_{\tau_1}^{\tau_2} \sigma^2(Y(t))\, dt \,\middle|\, A\right\} + P^{-1}(A) \mid \tau_2 - \tau_1 \mid^{3/2} 0(1). \tag{3.37}$$

Proof. – We need the following result (*see* for example, [8]): Let the random variables f_n converge to the variable f in probability as $n \to \infty$. Let

$$E \mid f_n \mid^\rho < c, \qquad E \mid f \mid^\rho < C, \qquad \rho > 1,$$
$$P\{(A_n \backslash A) \cup (A \backslash A_n)\} \to 0 \quad \text{as } n \to \infty.$$

Then

$$E\{f_n \mid A_n\} \to E\{f \mid A\}.$$

To prove Lemma 3.11, it is sufficient to consider only the following sets $A = \{\omega : (Y(t_1, \omega), \ldots, Y(t_k, \omega)) \in B_k\}$, where B_k is some Borel set in R^k, $t_i \in [0, \tau_1]$. Let $A_m = \{\omega : (Y_0^{\varepsilon_{nm}}(t_1, \omega), \ldots, Y_0^{\varepsilon_{nm}}(t_k, \omega)) \in B_k\}$. Using the result formulated above, (3.34), Corollary 3.1 and the assumption that $B(y)$ is uniformly continuous, we obtain:

$$\begin{aligned} E\{Y(\tau_2) - Y(\tau_1) \mid A\} &= \lim_{m\to\infty} E\{Y_0^{\varepsilon_{nm}}(t_2) - Y_0^{\varepsilon_{nm}}(t_1) \mid A_m\} \\ &= \lim_{m\to\infty} E\left\{\int_{\tau_1}^{\tau_2} B(Y_0^{\varepsilon_{nm}}[t]_\Delta)\, dt \,\middle|\, A_m\right\} \\ &= E\left\{\int_{\tau_1}^{\tau_2} B(Y(t))\, dt \,\middle|\, A\right\}, \end{aligned}$$

This proves (3.36). To prove (3.37) we cannot apply the result formulated above directly to the process $Y_0^{\varepsilon_{nm}}(\tau)$ because we have no estimates for the moments of this process of the order greater than two. But we do have such an estimate for the process $Y_*^\varepsilon(\tau)$; it is (3.26). We also have the bound (3.5) for the second moment of the difference $Y^{\varepsilon}(t) - Y_*^{\varepsilon_{nm}}(\tau_1)$. Then taking into account (3.35) we obtain

$$\begin{aligned} &E\{(Y(\tau_2) - Y(\tau_1))^2 \mid A\} \\ &= \lim_{m\to\infty} E\{(Y_*^{\varepsilon_{nm}}(t_2) - Y_*^{\varepsilon_{nm}}(t_1))^2 \mid A_m\} \\ &= \lim_{m\to\infty} E\left\{(Y_0^{\varepsilon_{nm}}(t_2) - Y_0^{\varepsilon_{nm}}(t_1))^2 \mid A_m\right\} \\ &= E\left\{\int_{\tau_1}^{\tau_2} \sigma^2(Y(t))\,dt \,\middle|\, A\right\} + \frac{|\tau_2 - \tau_1|^{3/2}}{P(A)}\, 0(1). \quad \square \end{aligned}$$

Let us complete now the proof of the theorem. Put

$$\tilde{Y}(\tau) = Y(\tau) - \int_0^{\tau} B(Y(t))\,dt.$$

By (3.36) the process $\tilde{Y}(\tau)$ is a martingale with respect to the family of σ-fields $\mathcal{F}_\tau$. By (3.37) and because of the fact that $|B(y)| \le \sup_x |Q(x)| < C$ we have:

$$\left| E\{(\tilde{Y}(\tau_2) - \tilde{Y}(\tau_1))^2 \mid A\} - E\left\{\int_{\tau_1}^{\tau_2} \sigma^2(Y(t))\,dt \,\middle|\, A\right\}\right| \le C P^{-1}(A)\,|\tau_2 - \tau_1|^{3/2}.$$

Let $t_l = l\,\Delta$, $t_{[t_1/\Delta]} = \tau_1$, $t_{[t_2/\Delta]} = \tau_2$. Then it follows from this estimate and the martingale property of the process $\tilde{Y}(\tau)$ that

$$\begin{aligned} &\left| E\{(\tilde{Y}(\tau_2) - \tilde{Y}(\tau_1))^2 \mid A\} - E\left\{\int_{\tau_1}^{\tau_2} \sigma^2(Y(t))\,dt \,\middle|\, A\right\}\right| \\ &= \left| \sum_{k=[\tau_1/\Delta]}^{[\tau_2/\Delta]} \left(E\{\tilde{Y}(\tau_{k+1}) - \tilde{Y}^{\varepsilon}(\tau_k)^2 \mid A\} \right.\right. \\ &\quad \left.\left. - E\left\{\int_{t_k}^{t_{k+1}} \sigma^2(Y(t))\,dt \,\middle|\, A\right\}\right)\right| \\ &\le C P^{-1}(A) \sum_{k=[\tau_1/\Delta]}^{[\tau_2/\Delta]} |t_{k+1} - t_k|^{3/2} \\ &\le C P^{-1}(A)\,\Delta^{1/2}\,|\tau_2 - \tau_1| \to 0 \quad \text{as } \Delta \to 0. \end{aligned}$$

Hence

$$|E\{(\tilde{Y}(\tau_2) - \tilde{Y}(\tau_1))^2 \mid \mathcal{F}_{\tau_1}\} = E\left\{\int_{\tau_1}^{\tau_2} \sigma^2(Y(t))\,dt \,\middle|\, \mathcal{F}_{\tau_1}\right\}.$$

Then by Theorem 5.3, Chapter IX of [3], it follows that there exists a Wiener process $\{w(t),\ 0 \le t \le T\}$ such that for each $\tau \in [0, T]$

$$\tilde{Y}(\tau) = \tilde{Y}(0) + \int_0^\tau \sigma(Y(t))\, dw(t)$$

with probability 1. Thus the finite-dimensional distributions of the process $Y(\tau)$ are uniquely defined and $Y(t)$ satisfies the equation

$$Y(\tau) = x_0 + \int_0^\tau B(Y(t))\, dt + \int_0^\tau \sigma(Y(t))\, dw(t).$$

The solution of this equation is determined only by the coefficients $B(y)$, $\sigma(y)$ and the starting point x_0 and independent of the choice of the sequence $\varepsilon_n \to 0$. Therefore, the finite-dimensional distributions of the process $Y^\varepsilon(t)$ converge to the finite-dimensional distributions of the process $Y(t)$ determined by (2.8).

The weak compactness of the family $\{Y^\varepsilon(t),\ t \in [0, T]\}$ can be established in the following way. According to (3.25) the process $Y^\varepsilon(t)$ is the sum of two processes $Y_*^\varepsilon(t)$ and $\theta_\varepsilon(t)$. Each of these sequences of processes satisfies the condition of the weak compactness, that is the conditions (3.26) and (3.5) correspondingly. From this we have the weak compactness of the processes $Y^\varepsilon(t)$, which together with the convergence of the finite-dimensional distributions implies (*see* [11], Theorem 2.1) the weak convergence of $Y^\varepsilon(t)$ to $Y(t)$ as $\varepsilon \to 0$. Theorem 2.1 is proved. □

Now we turn into the proof of Theorem 2.2. The general scheme of this proof is the same as for Theorem 2.1. Therefore we point out only the main differences. In view of conditions 2′, 3′ we need good bounds for the moments of the variable $|X^\varepsilon(t)|$. For this purpose we need a generalization of Lemma 3.1.

LEMMA 3.12. – *Suppose that the family of σ-fields N_s^t, $0 \le s \le t < \infty$, satisfies absolute regularity mixing condition. Let $G(x, w)$ for each x be a N_t^∞-measurable random variable and $E \sup_x |G(x, w)|^\gamma < \infty$ for some $\gamma > 1$. Let the random variables $\zeta(\omega)$, $\eta(\omega)$ be N_0^s-measurable and $E|\eta(\omega)|^\sigma < \infty$ for some $\sigma > 1$. Denote $g(x) = EG(x, \omega)$. Then for $s < t$, $1/\sigma + 1/\mu < 1$ and any set $A \in N_0^s$ with $P(A) \ne 0$*

$$\begin{aligned} (3.38)\qquad & |E\{\eta(\omega)\, G(\zeta(\omega), \omega) \mid A\} - E\{\eta(\omega)\, g(\zeta(\omega)) \mid A\}| \\ & \le \frac{C}{P(A)}\, \beta^{1-1/\sigma-1/\gamma}(t-s)\, E^{1/\sigma}|\eta(\omega)|^\sigma, \end{aligned}$$

and

$$(3.39)\quad \begin{aligned} &|E\{\eta(\omega)\,G(\zeta(\omega),\,\omega)\,|\,A\} - E\{\eta(\omega)\,g(\zeta(\omega))\,|\,A\}| \\ &\quad\le C\beta^{1-1/\sigma-1/\gamma}\,(t-s)\,E^{1/\sigma}\,|\,\eta(\omega)\,|^{\sigma}, \end{aligned}$$

This result is the consequence of Lemma 1 from [1] and Hölder's inequality. □

First, we obtain some preliminary estimate for the moments of the variables $X^{\varepsilon}(t)$, $Y^{\varepsilon}(t)$. From (2.1), (2.11) and the definition of the function $g(x,\,y)$ we have

$$\begin{aligned} |\,X^{\varepsilon}(t)\,| \le |\,x_0\,| &+ \frac{1}{\varepsilon}\int_0^t C(|\,X^{\varepsilon}(s)\,|+1)\,ds \\ &+ \frac{1}{\varepsilon}\int_0^t |\,g(X^{\varepsilon}(s),\,\zeta_{s/\varepsilon^2})\,|\,ds, \end{aligned}$$

and by Gronwall's lemma

$$|\,X^{\varepsilon}(t)\,| \le \left(|\,x_0\,| + \varepsilon^{-1}\left(Ct + \int_0^t |\,g(X^{\varepsilon}(s),\,\zeta_{s/\varepsilon^2})\,|\,ds\right)\right)\,\exp\left(\frac{C}{\varepsilon}\,t\right).$$

Then applying the first bound in (2.12), we obtain

$$E\,|\,X^{\varepsilon}(t)\,|^{p} \le C\,(x_0 + t/\varepsilon)^{p}\,\exp\,(Cpt/\varepsilon)$$

and taking into account the first bound in (2.9)

$$E\,|\,Y^{\varepsilon}(t)\,|^{p/2} \le C\,(x_0 + t/\varepsilon)^{p}\,\exp\,(Cpt/\varepsilon)$$

We need these bounds only to be sure that the moments of the variables $X^{\varepsilon}(t)$ and $Y^{\varepsilon}(t)$ are finite.

LEMMA 3.13. – *For all* $t \in [0,\,T]$ *and* $q = 2\,[2\,(\mu+1)/\mu]\,\mu$

$$(3.40)\qquad E\,|\,X^{\varepsilon}(t)\,|^{q} < C.$$

Proof. – Let $y_0 = H(x_0)$. Then using (3.3) we can write

$$(3.41)\quad (Y^{\varepsilon}(\tau) - y_0)^{2k} = \frac{2k}{\varepsilon}\int_0^{\tau} (Y^{\varepsilon}(s) - y_0)^{2k-1}\,F(X^{\varepsilon}(s),\,\zeta_{s/\varepsilon^2})\,ds.$$

Let $k = [2(\mu+1)/\mu]$ be an integer number. It follows from (2.1) and (3.3) that for any fixed z

$$\begin{aligned}
&(Y^\varepsilon(s) - y_0)^{2k-1} F(X^\varepsilon(s), z) \\
&= \frac{1}{\varepsilon}\int_0^s \{(2k-1)(Y^\varepsilon(t) - y_0)^{2k-2} F(X^\varepsilon(t), z) F(X^\varepsilon(t), \zeta_{t/\varepsilon^2}) \\
&\quad + (Y^\varepsilon(t) - y_0)^{2k-1} (\nabla F(X^\varepsilon(t), z), b(X^\varepsilon(t), \zeta_{t/\varepsilon^2}))\}\, dt.
\end{aligned}$$

We obtain from the last equality and (3.41):

$$\begin{aligned}
(3.42)\quad &(Y^\varepsilon(\tau) - y_0)^{2k} \\
&= \frac{2k}{\varepsilon^2}\int_0^\tau ds \int_0^s dt\, \{(2k-1)(Y^\varepsilon(t) - y_0)^{2k-2} \\
&\quad \times F(X^\varepsilon(t), \zeta_{t/\varepsilon^2}) F(X^\varepsilon(t), \zeta_{t/\varepsilon^2}) \\
&\quad + (Y^\varepsilon(t) - y_0)^{2k-1} (\nabla F(X^\varepsilon(t), \zeta_{s/\varepsilon^2}), b(X^\varepsilon(t), \zeta_{t/\varepsilon^2}))\}.
\end{aligned}$$

By (2.9) and (2.12)

$$\begin{aligned}
&|F(X^\varepsilon(t), \zeta_{s/\varepsilon^2})| \le C(|X^\varepsilon(t)| + 1)|q(\zeta_{s/\varepsilon^2})|, \\
&(\nabla F(X^\varepsilon(t), \zeta_{s/\varepsilon^2})| < C|q(\zeta_{s/\varepsilon^2})|.
\end{aligned}$$

Applying Lemma 3.12 with $\sigma = k/(k-1+1/\mu)$, $\gamma = p/2$ and the σ-fields $N^\infty_{s/\varepsilon^2}$, N_0^{t/ε^2}, and taking into account (3.11), (2.9)-(2.12), we obtain

$$\begin{aligned}
(3.43)\quad &|E\{(2k-1)(Y^\varepsilon(t) - y_0)^{2k-2} F(X^\varepsilon(t), \zeta_{s/\varepsilon^2}) F(X^\varepsilon(t), \zeta_{t/\varepsilon^2}) \\
&\quad + (Y^\varepsilon(t) - y_0)^{2k-1} (\nabla F(X^\varepsilon(t), \zeta_{s/\varepsilon^2}), b(X^\varepsilon(t), \zeta_{t/\varepsilon^2}))\}| \\
&\le C\beta^{(\mu-1)/\mu k - 2/p}\left(\frac{s-t}{\varepsilon^2}\right) E^{1/\sigma}((Y^\varepsilon(t) - y_0)^{2k} + 1) \\
&\quad \times E^{2/p}|q(\zeta_{t/\varepsilon^2})\, q(\zeta_{s/\varepsilon^2})|^{p/2}.
\end{aligned}$$

To apply Lemma 3.12 it is convenient to rewrite the scalar products of the vectors in the coordinate form. Especially it concerns the function F. Integrating (3.43) with respect to s, t, $0 \le t \le s < \tau$, we have from (3.42)

$$
\begin{aligned}
&E\,(Y^{\varepsilon}(\tau) - y_0)^{2k} \\
&= \frac{C}{\varepsilon^2} \int_0^{\tau} ds \int_0^{s} dt\, \beta^{(\mu-1)/\mu k - 2p} \left(\frac{s-t}{\varepsilon^2}\right) (E\,(Y^{\varepsilon}(t) - y_0)^{2k} + 1) \\
&\le C \int_0^{\tau} dt\, (E\,(Y^{\varepsilon}(\tau) - y_0)^{2k} + 1) \int_0^{(\tau - t)/\varepsilon^2} du\, \beta^{(\mu-1)/\mu k - 2p}(u) \\
&\le C \int_0^{\tau} (E\,(Y^{\varepsilon}(t) - y_0)^{2k} + 1)\, dt.
\end{aligned}
$$

It follows by Gronwell's lemma that

$$E\,(Y^{\varepsilon}(\tau) - y_0)^{2k} \le C\tau \exp(c\tau).$$

This bound together with (2.10) implies

$$E\,(X^{\varepsilon}(\tau) - y_0)^{2k\mu} \le C\,(E\,|\,Y^{\varepsilon}(\tau)\,|^{2k} + 1) \le C.$$

The lemma is proved. □

In view of the restrictions on the growth of the functions $H(x)$, $b(x)$ and their derivatives, the estimate (3.40) allows to prove Theorem 2.2 analogously to the proof of Theorem 2.1. Note only that for the proof we must use Lemma 2.12 instead of Lemma 2.1.

4. REMARKS AND GENERALIZATIONS

1. If system (2.2) in the plane R^2 has a smooth integral $H(x)$, then the non-perturbed vector field $b(x)$ is orthogonal to $\nabla H(x)$ and directed along $\bar{\nabla} H(x) = \left(\frac{\partial H(x)}{\partial x^2}, -\frac{\partial H(x)}{\partial x^1}\right)$, if x is not a critical point. This means that $b(x) = \beta(x)\,\bar{\nabla} H(x)$, and we can rewrite system (2.2) in the form

$$\dot{x}_t = \beta(X_t)\,\bar{\nabla} H(X_t) \tag{4.1}$$

with a proper scalar function $\beta(x)$. If $\nabla H(x) \neq 0$ for $x \neq 0$ and $\beta(x)$ does not changes the sign we have, roughly speaking, the situation considered in Section 2. Consider now the case when $\beta(x)$ change the sign.

A typical example is shown in the Figure 1. The level sets (the trajectories of the non-perturbed system) form a family of loops around the equilibrium point 0 where $\nabla H(0) = 0$, $H(x) > 0$ and $\nabla H(x) \neq 0$ for $x \neq 0$. The function $\beta(x)$ is equal to zero on the curve $\Gamma = ABCEFDA$, positive outside the domain bounded by Γ and negative inside. Let

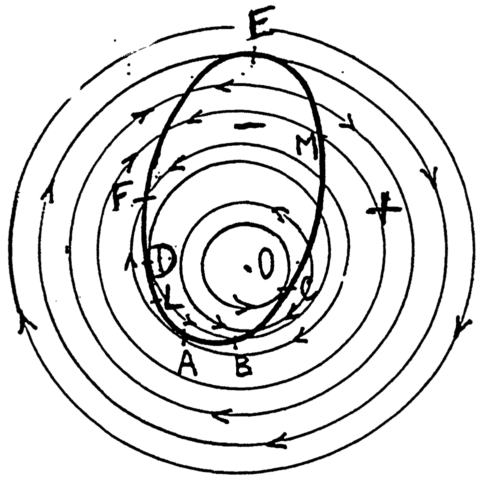

FIG. 1.

A, C, E, D be the points where Γ is tangent to a level set. Let $0 < H(C) < H(D) < H(A) \leq H(E) < \infty$.

In the case under consideration each level set $C(y) = \{x : H(x) = y\}$ with $H(C) \leq y \leq H(E)$ consist of more than one trajectory of the system (4.1), since at least one equilibrium point situated on such level sets. The dynamical system which is the restriction of our system to $C(y)$, $y \in (H(C), H(E))$ have many invariant measures. The extreme points of the cone of the invariant measures consist of δ-measures concentrated in each of equilibriums. Existence of such measures leads to new conservation laws, which are described by step-functions. Let $\tilde{H}(x)$ be equal to zero for x such that $H(x) \notin (H(C), H(E))$. If $H(x) \in (H(C), H(E))$, we put $\tilde{H}(x) = 1$ if x belongs to the domain of attraction of the points of the arc DFE, or $\tilde{H}(x) = -1$ if x is attracted to a point of the arc ABC. For the points of the arcs ALD and CME let us put $\tilde{H}(x) = 0$. It is easy to see that $\tilde{H}(x)$ is an integral for our system. Of course, any function of $\tilde{H}(x)$ and $H(x)$ will also be an integral.

Let us describe the limit of $Y_t^\varepsilon = H(X_t^\varepsilon)$ as $\varepsilon \downarrow 0$ in this case. Suppose we start at a point x with $H(x) > H(E)$. Then the limiting process Y_t will be as described in Section 2 until first time when $Y_t = H(E)$. Then the diffusion $\sigma_+^2(y)$ and drift $B_+(y)$ for $y \in [H(D), H(E)]$ will be equal to $D(X_{DE}(y))$ and $Q(X_{DE}(y))$, where $X_{DE}(y)$ is the point on the arc DFE with $H(X_{DE}(y)) = y$; the functions $D(x)$ and $Q(x)$ are defined in condition 5. This is so because the trajectories of the non-perturbed system are attracted to the points of the arc DFE. The exception is the unstable equilibrium points belonging to the arcs CME and ALD. But if we assume that the noise at these points is not degenerate $(D(x)) > D_0 > 0$ for the points of the arcs CME and ALD, the trajectories of the perturbed process will fast enough reach small neighborhoods of attractive points on the corresponding level sets.

If we started at x such that $H(x) > H(E)$, the limiting process Y_t will have the diffusion and drift coefficients $D(X_{DE}(y))$ and $Q(X_{DE}(y))$ up to the moment when the trajectory comes to the point $y = H(D)$.

At the time when Y_t first hits $y = H(D)$ from the area of larger values the trajectory X_t^ε jumps to the point B and the diffusion and drift coefficients will be defined by the values of $D(x)$ and $Q(x)$ on the arc $ABC : \sigma_-^2(y) = D(X_{AC}(y), B_-(y) = Q(X_{AC}(y))$, where $x_{AC}(y)$ has the mining similar to $X_{DE}(y)$. After Y_t touches $y = H(A)$ the trajectory X_t^ε jumps to the point F and the coefficients again will be defined by the values of $D(x)$ and $Q(x)$ on the arc DE.

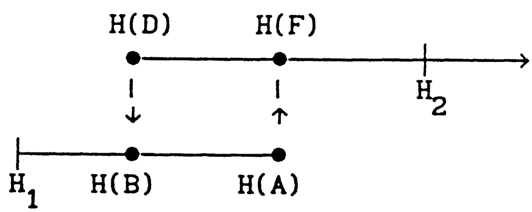

FIG. 2.

For $y < H(C)$ the diffusion process is defined as in Section 2. Thus the diffusion and drift coefficients for the limiting process on the interval $y \in (H(B), H(F))$ will be different depending on where from the trajectory entered the interval. The limiting process will have Markov property if considered not on the straight line but on the graph (*Fig.* 2). The coefficients will be $\sigma_+^2(y)$ and $B_+(y)$ on the upper line and $\sigma_-^2(y)$, $B_-(y)$ on the lower line. At the point $H(D)$ the trajectory of the limit process jumps instantaneously to $H(B)$ and continues the diffusion. At $H(A)$ the trajectory instantaneously jumps to $H(F)$. If, for example, we are interested in calculation of the expected value $u^\varepsilon(x)$ of the first time when $H(X_t^\varepsilon)$, $X_0^\varepsilon = x$, $H(x) \in (H_1, H_2)$, first exists the interval (H_1, H_2), we have the following non-standard boundary problem for

$u_{\pm}(H) = u_{\pm}(H(x)) = \lim_{\varepsilon \downarrow 0} u^{\varepsilon}(x)$, where the sign "+" corresponds to the case when x is attracted to DFE or if $H(x) > H(F)$, and the sign "−" corresponds to x attracted to DBC or if $H(x) < H(B)$:

$$(4.2) \quad \begin{cases} \frac{1}{2}\sigma_{+}^{2}(y)\, U_{+}''(y) + B_{+}(y)\, u_{+}'(y) = -1, & H(D) < y < H_2, \\ \frac{1}{2}\sigma_{-}^{2}(y)\, U_{-}''(y) + B_{-}(y)\, u_{-}'(y) = -1, & H_1 < y < H(A), \\ u_{+}(H(D)) = u_{-}(H(B)), \qquad u_{+}(H(F)) = u_{-}(H(A)), \\ \quad u_{+}(H_2) = 0, \qquad u_{-}(H_1) = 0. \end{cases}$$

It is not difficult to prove that the problem (4.2) has a unique solution, and it can be written down explicitly.

2. In the previous sections we considered the case when $H(x)$ had only one critical point, let us say a minimum. In general $H(x)$ can have many critical points and the situation becomes more complicated. Let $H(x)$ has two minima and one saddle point as on the Figure 3 *a*. Then corresponding non-perturbed trajectories behave as in Figure 3 *b* [we assume that $b(x)$ has no equilibrium points besides O_1, O_2, O_3]. It is easy to understand that the limit of $H(X_t^{\varepsilon})$ in this case will not be a Markov process: If, for example, we start at a point x such that $H(O_1) < H(O)_3 < H(x) < H(O_2)$, the behavior of the limiting process will be different for different connected components of the level set $\{z : H(z) = H(x)\}$. To have a Markov limiting process one should consider process on the graph Γ (*Fig.* 3 *c*) homeomorphic to the set of connected components of the level sets of the function $H(x)$. Here the situation is similar to the case of white noise perturbations of Hamiltonian systems considered in ([6], [7]). In that references one can find precise definition of the graph corresponding to $H(x)$, description of diffusion processes on a graph and the technique necessary to prove the convergence.

Let us consider for brevity the Hamiltonian case

$$\dot{X}_t^{\varepsilon} = \bar{\nabla} H(X_t^{\varepsilon}, \zeta_{t/\varepsilon}), \qquad X_0^{\varepsilon} = x \in R^2,$$
$$\bar{\nabla} H(x, y) = \left(\frac{\partial H(x, y)}{\partial x^2}, -\frac{\partial H(x, y)}{\partial x^1}\right).$$

Suppose that $H(x) = EH(x, \zeta_t)$ has the structure shown in Figure 3. The corresponding averaged system is

$$\dot{X}_t = \bar{\nabla} H(X_t), \qquad X_0 = x \in R^2,$$

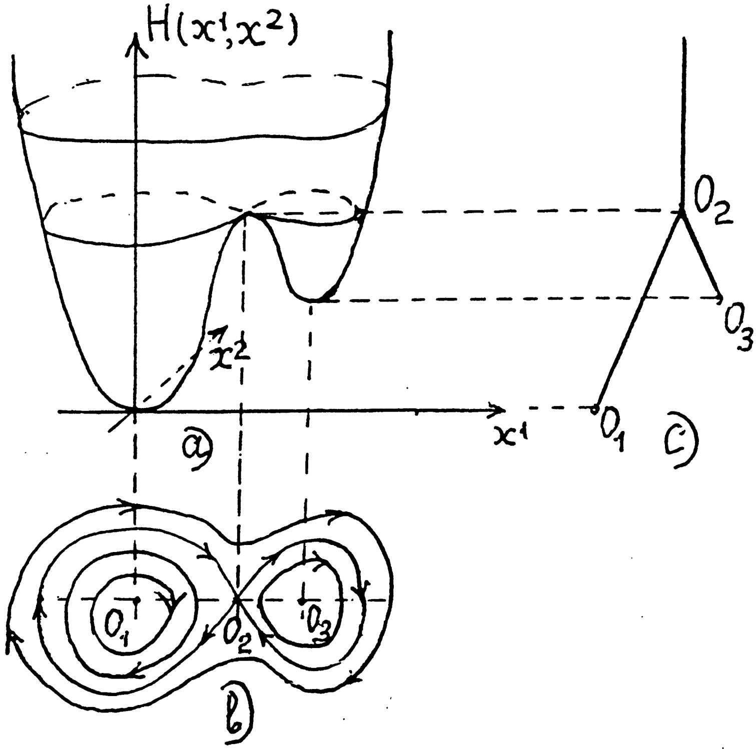

FIG. 3.

Denote $Y(x) : R^2 \to \Gamma$ the mapping such that $Y(x)$ is the point of Γ corresponding to the component of the level set of the Hamiltonian containing x. The calculation of the weak limit of the process $Y_t^\varepsilon = Y(X_{t/\varepsilon}^\varepsilon)$ on Γ as $\varepsilon \downarrow 0$ can be devided in the following steps: First, prove that the limiting process on the graph is a continuous Markov process with the Feller property. Each such process is defined by a family of differential operators corresponding to the segments of the graph and by the gluing conditions at the vertices ([6], [7]). Calculation of these characteristics is the next step. The differential operators defining the limiting process inside the segments (in our example there are 3 segments) are calculated exacly in the same way as in Section 2, if, of course, the mixing conditions and the restrictions on the growth are fulfilled. The vertices corresponding to the extrema of $H(x)$ (in our example, points O_1 and O_3) turn out to be inaccessible for the limiting process and thus no gluing conditions should be added at these points. The vertices corresponding to the seddle points (point O_2 in *Fig.* 3) are accessible, and one must add gluing conditions there. Each of these gluing conditions is determined by a finite number of constants (*see* [6], [7]). They can be calculated as follows. Denote $\mu(\cdot)$ the projection of the Lebesgue measure $\Lambda(\cdot)$ in R^2 on Γ given by the map $Y:\ \mu(\gamma)=\Lambda(Y^{-1}(\gamma))$, $\gamma\subset\Gamma$. Since the Lebesgue measure is invariant for Hamiltonian systems, the measure μ will be invariant for the limiting process on Γ. The constants in the gluing conditions should be choosen so that the measure μ would be invariant. We plan to consider this problem in detail elsewhere.

3. Now we describe the results for multidimensional dynamical systems with $l\geq 1$ conservation laws.

Consider a system

$$\dot{X}_t^\varepsilon=\frac{1}{\varepsilon}\,b(X_t^\varepsilon,\,\zeta_{t/\varepsilon^2}),\qquad X_0^\varepsilon=x\in R^\Gamma. \tag{4.3}$$

We assume that $b(x,z)$, $x\in R^\Gamma$, $z\in R^m$ is a smooth vector field with components bounded together with their first three derivatives. We suppose that the process ζ_t satisfies the absolute regularity mixing condition (*see* condition 4).

Denote $b(x)=Eb(x,\zeta_t)$ and consider the averaged system

$$\dot{X}_t^\varepsilon=b(X_t),\qquad X_0=x\in R^\Gamma. \tag{4.4}$$

Let $H_1(x), H_2(x), \ldots, H_l(x)$ be the first integrals of system (4.4) and assume that the functions $H_k(x)$, $k=1,\ldots,l$, are bounded together with their first three derivatives. Denote $D=\{y\in R^l:\ y=(H_1(x),\ldots,H_l(x))$ for some $x\in R^\Gamma\}$ and assume that D belongs to the closure of its interior (D) in R^l. Assume that the vectors $\nabla H_1(x), \nabla H_2(x), \ldots, \nabla H_1(x)$ are linearly independent if $(H_1(x),\ldots,H_l(x)\in(D)$.

Assume that for any $y=(y_1,\ldots,y_l)\in D$ the set $C(y)=\{x\in R^\Gamma:\ H_1(x)=y_1,\ldots,H_l(x)=y_l\}$ is compact and connected. Let a measure μ_y on $C(y)$ exists for any $y\in D$ such that $\mu_y(C(y))=1$ and for any continuous function $f(x)$ on $C(y)$

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T f(X_s)\,ds=\int_{C(y)} f(v)\,\mu_y(dv) \tag{4.5}$$

uniformly in the initial point $X_0=x\in C(y)$. Condition (4.5) replaces the periodicity condition.

Denote

$$g(x, z) = b(x, z) - b(x), \qquad F_j(x, z) = (\nabla H_j(x), g(x, z));$$
$$D(x, s) = (D_{ij}(x, s))_1^l \quad \text{where } D_{ij}(x, s) = E F_i(x, \zeta_s) F_j(x, \zeta_0),$$
$$D(x) = 2\int_0^\infty D(x, s)\, ds; \qquad Q_i(x, s) = \sum_{k=1}^{r} E \frac{\partial F_i(x, \zeta_s)}{\partial x_k} g_k(x, \zeta_0),$$
$$Q(x, s) = (Q_1(x, s), \ldots, Q_l(x, s)), \qquad Q(x) = \int_0^\infty Q(x, s)\, ds;$$
$$(A_{ij}(y))_1^l = \int_{C(y)} D(v)\, \mu_y(dv), \qquad B_i(y) = \int_{C(y)} Q_i(v)\, \mu_y(dv).$$

Then one can prove that the processes

$$H(X_t^\varepsilon) = (H_1(X_t^\varepsilon), \ldots, H_l(X_t^\varepsilon))$$

for any $T > 0$ converge weakly, in the space of continuous on $[0, T]$ functions with values in D, to the diffusion process in D governed by the operator.

$$L = \frac{1}{2} \sum_{i, j=1}^{l} A_{ij}(y) \frac{\partial^2}{\partial y_i\, \partial y_j} + \sum_{i=1}^{l} B_i(y) \frac{\partial}{\partial y_1}.$$

The points of the boundary ∂D of the domain D will be inaccessible for this process.

One can prove such result using the bounds given in Section 3. There are examples where the listed above conditions hold. But condition (4.5) turns out too restrictive if the dimension of the sets $C(y)$ bigger than 1. Apparently, the convergence can be proved if (4.5) is replaced by a weaker assumption for example, if (4.5) fulfilled for almost all $y \in D$, but this problem is still open.

In the more general case when the vectors $\nabla H_1(x), \ldots, \nabla H_l(x)$ are linearly dependent for some x, the limiting process should be considered on a complex consisting of l-dimensional pieces. Here the situation is similar to one discussed in Remark 2 of this section.

REFERENCES

[1] A. N. Borodin, Limit theorem for solutions of differential equations with random right-hand side, *Theory Prob. Appl.*, Vol. **22**, 3, 1977, pp. 482–497.

[2] R. Cogburn and J. A. Ellison, A stochastic theory of adiabatic invariance, *Comm. Math. Phys.,* Vol. **149**, 1992, pp. 97–126.
[3] J. L. Doob, *Stochastic Processes*, Wiley, New York, 1960.
[4] M. I. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton Univ. Press., 1985.
[5] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer-Verlag, 1984.
[6] M. I. Freidlin and A. D. Wentzell, Diffusion process on graphs and averaging principle, *The Annals of Probability*, Vol. **21**, 4, 1993, pp. 2215–2245.
[7] M. I. Freidlin and A. D. Wentzell, Random perturbation of Hamiltonian Systems, to appear in *The Memoirs AMS.*
[8] R. Z. Khasminskii, A limit thorem for solutions of differential equations with random right-hand side, *Theory Prob. Appl.,* Vol. **11**, 3, 1966, pp. 390–406.
[9] R. Z. Khasminskii, On processes defined by differential equations with a small parameter, *Theory Prob. Appl.,* Vol. **11**, 2, 1966, pp. 211–228.
[10] I. A. Ibragimov and Yu. A. Rozanov, *Gaussian Randum Processes*, Springer-Verlag, 1978.
[11] Yu V. Prokhorov, Convergence of random processes and limit theorems in probability theory, *Theory Prob. Appl.,* Vol. **1**, 2, 1956, pp. 157–214.
[12] A. D. Wentzell, *A course in the theory of stochastic processes*, McGraw-Hill Inc., 1981.

(Manuscript received June 1, 1993;
revised April 5, 1994.)