ASYMPTOTIC ANALYSIS OF TRANSPORT PROCESSES

BY GEORGE C. PAPANICOLAOU[1]

ABSTRACT. For a large class of processes called transport processes we study in detail a certain asymptotic limit, the diffusion limit. Transport processes arise in linear transport theory, learning theory, nonlinear oscillations in the presence of noise and other problems. We examine closely some examples and give complete proofs for the results stated here.

1. **Introduction.** The motion of a particle whose velocity undergoes jumps of random size at random times constitutes the prototype of a transport process. The general structure of transport processes however underlies a great variety of problems that arise in mathematical sciences and which may have no relation to the moving particle model. Our aim here is to study the properties of a large class of transport processes in a specific asymptotic limit, the diffusion limit. This limit corresponds roughly to the frequency of jumps becoming very large in the particle model. We give several examples to illustrate the scope of the asymptotic theory. We also give complete proofs for all results stated here. The contents are briefly as follows.

In §2 we outline the probabilistic construction of transport processes. The general theory of Markov processes [1], [2], [3] provides the necessary existence, uniqueness, and related information, so it is only sketched here. We present in detail, however, the connection of transport processes with linear transport theory [4]. In §3 we formulate the asymptotic problem that concerns us here by introducing a small parameter that corresponds to the mean free time between jumps. Diffusion approximations are well known in a variety of contexts [4], [5], [6] but for transport processes in the generality of our formulation many special features arise and the treatment of diverse problems is unified. The special features account for the versatility of the asymptotic theory. §4 illustrates this point. We consider a mathematical model of learning theory [7], [8], and show that some

---

asymptotic problems of interest in that theory are special cases within the framework of transport processes. Learning models do not, however, take full advantage of the available generality. We consider therefore a nonlinear oscillator in the presence of noise whose long-time statistical behavior is a problem that can also be formulated in the context of transport processes. These examples along with the diffusion or p_1 approximation of linear transport theory motivate the formulation we give here.

In §5 we discuss briefly the connection of the present problem with the operator formalism of random evolutions [9 (and references therein)], [10], [11] and related problems. §6 contains the statement of a theorem characterizing the asymptotic behavior of transport processes in the diffusion limit. The theorem is followed by three corollaries that provide related information. In §7 we collect some remarks that help explain the nature of the asymptotic theory and its present limitations. We refer here to previous works of Kramers [6], Gihman and Gihman and Skorohod [12], Khasminskii [13], [22], Kurtz [11], Norman [7] and others concerning similar problems. We also indicate how Smoluchovski's asymptotic theory [6, 16–18] fits into the present framework. Because of some technical difficulties we cannot give a general treatment to this problem at present.

In §8 we apply the diffusion approximation to some examples one of which has been studied by Stroock [14] and Baggett and Stroock [15]. §§9–16 contain the proof of the theorem and corollaries of §6. In §9 we give an outline of the approach we follow so we refer to that section for further information.

It is a pleasure to thank J. B. Keller, W. Kohler, H. P. McKean and B. White for several discussions on the problems considered here.

2. **Transport processes.** We shall consider a stochastic process $(X(\tau, \sigma, x, y), Y(\tau, \sigma, x, y))$ with values in $R^n \times R^m$, $0 \leq \sigma \leq \tau \leq T$, which is constructed as follows. Let $(\xi(\tau, \sigma, x, y), \eta(\tau, \sigma, x, y))$ denote the solution of the deterministic system of ordinary differential equations

$$
(2.1) \quad \frac{d\xi(\tau)}{d\tau} = F(\tau, \xi(\tau), \eta(\tau)), \qquad \frac{d\eta(\tau)}{d\tau} = H(\tau, \xi(\tau), \eta(\tau)), \quad \tau > \sigma,
$$

$$
\xi(\sigma, \sigma, x, y) = x, \qquad \eta(\sigma, \sigma, x, y) = y.
$$

We assume that the vector functions F and H are restricted appropriately (cf. §6) so that (2.1) has solutions for arbitrarily long times. When convenient we shall use the simplified notation $\xi(\tau) = \xi(\tau, \sigma, x, y)$, $\eta(\tau) = \eta(\tau, \sigma, x, y)$. Let $\tau_1 = \tau_1(\sigma, x, y)$ be a nonnegative random variable with exponential distribution

$$
(2.2) \quad P\{\tau_1 > t\} = \exp\left(-\int_\sigma^{\sigma+t} q(s, \xi(s), \eta(s))\, ds\right).
$$

Here $q(s, x, y)$ is a nonnegative bounded function on $[0, \infty) \times R^n \times R^m$ and, as the notation indicates, τ_1 depends on σ and the starting point (x, y) of $(\xi(\tau), \eta(\tau))$.

We now define

$$(2.3) \quad \begin{aligned} X(\tau, \sigma, x, y) &= \xi(\tau, \sigma, x, y), \\ Y(\tau, \sigma, x, y) &= \eta(\tau, \sigma, x, y), \quad \sigma \leq \tau < \sigma + \tau_1. \end{aligned}$$

At the instant $\sigma + \tau_1$ the Y process jumps by a random amount which is determined by a probability distribution. We assume that for each $s \geq 0$, $x \in R^n$, $y \in R^m$ and Borel set A in R^m a probability measure $\pi(s, x, y, A)$ is given so that

$$
\begin{aligned}
(2.4) \quad & P\{\tau_1 \leq t, \ Y(\sigma + \tau_1, \sigma, x, y) \in A\} \\
& = \int_\sigma^{\sigma+t} \pi(s, \xi(s), \eta(s), A) q(s, \xi(s), \eta(s)) \\
& \qquad\qquad \times \exp\left(-\int_\sigma^s q(\gamma, \xi(\gamma), \eta(\gamma))\, d\gamma\right) ds.
\end{aligned}
$$

Given the jump time τ_1 and $Y_1 = Y(\sigma + \tau_1, \sigma, x, y)$ then the process (X, Y) continues for an interval of length τ_2 along the trajectories of (2.1) with starting values at $\sigma + \tau_1$ equal to $X_1 = \xi(\sigma + \tau_1, \sigma, x, y)$ and Y_1. The distribution of τ_2 is

$$
\begin{aligned}
(2.5) \quad & P\{\tau_2 > t\} \\
& = \exp\left(\int_{\sigma+\tau_1}^{\sigma+\tau_1+t} q(s, \xi(s, \sigma + \tau_1, X_1, Y_1), \eta(s, \sigma + \tau_1, X_1, Y_1))\, ds\right).
\end{aligned}
$$

In $\sigma + \tau_1 \leq \tau < \sigma + \tau_1 + \tau_2$ the process (X, Y) is defined by

$$(2.6) \quad \begin{aligned} X(\tau, \sigma, x, y) &= \xi(\tau, \sigma + \tau_1, X_1, Y_1), \\ Y(\tau, \sigma, x, y) &= \eta(\tau, \sigma + \tau_1, X_1, Y_1). \end{aligned}$$

At the instant $\sigma + \tau_1 + \tau_2$ the Y process jumps again and the construction is now continued in the obvious manner. The result is the transport process, a pair of processes X and Y which are jointly Markov. The first component is continuous and the second undergoes jumps at random times. Aside from conditions on F and H that guarantee solutions of (2.1) for all times and aside from obvious measurability requirements on $q(s, x, y)$ and $\pi(s, x, y, A)$ as functions (s, x, y) we shall assume that

$$(2.7) \quad\quad\quad 0 \leq q(s, x, y) \leq M < \infty.$$

With this condition the process (X, Y) is well defined, that is, with probability one only finitely many jumps occur in every finite time interval.

Let us demonstrate this fact [3]. Let

$$\mathscr{F}_\sigma^\tau = \sigma\{(X(s), Y(s)); \ \sigma \leq s < \tau\}$$

i.e., $\mathscr{F}_\sigma^\tau$ is the σ-algebra generated by the process (X, Y) in the interval $[\sigma, \tau)$. A random variable $\tau^* \geq \sigma$ is called a Markov (or stopping) time if for all $\tau \geq \sigma$ the event $\{\tau^* \leq \tau\}$ belongs to $\mathscr{F}_\sigma^\tau$. The σ-algebra that contains the events $\{\tau^* \leq \tau\} \cap A$, $A \in \mathscr{F}_\sigma^\infty$, $\tau \geq \sigma$, is denoted by $\mathscr{F}_\sigma^{\tau^*}$.

For any $C > 0$ we have

$$P\{\tau_1 + \tau_2 + \cdots + \tau_n < t\}$$

$$\leq E\left\{\exp\left(C\left(t - \sum_{k=1}^n \tau_k\right)\right)\right\} = \exp(Ct)E\left\{\exp\left(-C\sum_{k=1}^n \tau_k\right)\right\}$$

$$= \exp(Ct)E\left\{\exp\left(-C\sum_{k=1}^{n-1} \tau_k\right)E\{\exp(-C\tau_n) \mid \mathscr{F}_\sigma^{\sigma_{n-1}}\}\right\}.$$

Here $\sigma_{n-1} = \sigma + \tau_1 + \cdots + \tau_{n-1}$ and it is a Markov time. From the construction of (X, Y) and with the notation $X(\sigma_{n-1}) = X_{n-1}$, $Y(\sigma_{n-1}) = Y_{n-1}$ it follows that

$$E\{\exp(-C\tau_n) \mid \mathscr{F}_\sigma^{\sigma_{n-1}}\}$$

$$= \int_0^\infty \exp(-Cs)q(\sigma_{n-1} + s, \xi(\sigma_{n-1} + s, \sigma_{n-1}, X_{n-1}, Y_{n-1}),$$

$$\eta(\sigma_{n-1} + s, \sigma_{n-1}, X_{n-1}, Y_{n-1}))$$

$$\times \exp\left[-\int_0^s q(\sigma_{n-1} + \gamma, \xi(\sigma_{n-1} + \gamma, \sigma_{n-1}, X_{n-1}, Y_{n-1}),\right.$$

$$\left.\eta(\sigma_{n-1} + \gamma, \sigma_{n-1}, X_{n-1}, Y_{n-1})) \, d\gamma\right] ds$$

$$\leq M/C.$$

Therefore,

$$P\{\tau_1 + \tau_2 + \cdots + \tau_n < t\} \leq \exp(Ct)(M/C)^n \to 0, \quad n \uparrow \infty,$$

if C is chosen so that $M/C = \frac{1}{2}$ say, and it is fixed along with $t \geq 0$ as $n \uparrow \infty$.

Let $f(x, y)$ be a bounded measurable function on $R^n \times R^m$ and set

(2.8) $u(\sigma, \tau, x, y) = E\{f(X(\tau, \sigma, x, y), Y(\tau, \sigma, x, y))\}.$

From the construction of (X, Y) and the law of total probability we obtain, by conditioning on the first jump time, the relation

$$u(\sigma, \tau, x, y) = f(\xi(\tau), \eta(\tau))\exp\left(-\int_\sigma^\tau q(s, \xi(s), \eta(s)) \, ds\right)$$

(2.9)

$$+ \int_\sigma^\tau\int u(s, \tau, \xi(s), z)\pi(s, \xi(s), \eta(s), dz)q(s, \xi(s), \eta(s))$$

$$\times \exp\left(-\int_\sigma^s q(\gamma, \xi(\gamma), \eta(\gamma)) \, d\gamma\right) ds.$$

This relation can be viewed as an integral equation for u and the construction of the Markov process (X, Y) may proceed in a more analytical manner by studying (2.9) [2]. When $f(x, y)$ is differentiable, (2.9) is equivalent to the integrodifferential equation

$$(2.10) \qquad \frac{\partial u(\sigma, \tau, x, y)}{\partial \sigma} + \mathscr{L}_\sigma u(\sigma, \tau, x, y) = 0, \qquad \sigma < \tau,$$

$$u(\tau, \tau, x, y) = f(x, y),$$

$$\mathscr{L}_\sigma g(x, y) \equiv F(\sigma, x, y)\frac{\partial g(x, y)}{\partial x} + H(\sigma, x, y)\frac{\partial g(x, y)}{\partial y} + Q_{\sigma,x}g(x, y),$$

$$Q_{\sigma,x}g(x, y) \equiv q(\sigma, x, y)\int g(x, z)\pi(\sigma, x, y, dz) - q(\sigma, x, y)g(x, y).$$

Equation (2.10) is a linear, conservative backward transport equation; it is the backward Kolmogorov equation for the process (X, Y) constructed above. We employ the notation $\partial/\partial x$ to denote the gradient operator and F times the gradient stands for the inner product of the vector function F with the gradient. If $P(\sigma, \tau, x, y, A, B)$ denotes the solution of (2.10) with $f(x, y) = \chi_{A \times B}(x, y)$, A and B are Borel sets on R^n and R^m respectively, then,

$$(2.11) \quad P(\sigma, \tau, x, y, A, B) = P\{X(\tau, \sigma, x, y) \in A, \ Y(\tau, \sigma, x, y) \in B\}.$$

Let $g(s, x, y)$ be a bounded measurable function such that

$$(\partial_s + \mathscr{L}_s)g(s, x, y)$$

is also bounded and measurable. Define $Z(\tau) = Z(\tau, \sigma, x, y; g)$ by

$$(2.12) \quad Z(\tau) = g(\tau, X(\tau), Y(\tau)) - \int_\sigma^\tau (\partial_s + \mathscr{L}_s)g(s, X(s), Y(s)) \, ds.$$

From the fact that the expectation (2.8) satisfies (2.10) it follows that

$$
\begin{aligned}
E\{Z(\tau)\} &= \iint P(\sigma, \tau, x, y, d\bar{x}, d\bar{y})g(\tau, \bar{x}, \bar{y}) \\
&\quad - \int_\sigma^\tau ds \iint P(\sigma, s, x, y, d\bar{x}, d\bar{y})(\partial_s + \mathscr{L}_s)g(s, \bar{x}, \bar{y}) \\
&= g(\sigma, x, y), \qquad \tau \geqq \sigma.
\end{aligned}
$$

(2.13)

Using (2.13), it is now easy to verify that $Z(\tau)$ is a martingale, i.e.,

$$(2.14) \qquad E\{Z(\tau) \mid \mathscr{F}_\sigma^t\} = Z(t), \qquad \sigma \leqq t \leqq \tau.$$

Let $\mathcal{D}$ be a bounded open set in R^n and let $\tau_{\mathcal{D}}$ be the first time the process $X(\tau)$ reaches the boundary $\partial\mathcal{D}$ starting from (x, y), $x \in \mathcal{D}$ at time σ. We recall that, by construction, $X(\tau)$ is continuous. Clearly $\tau_{\mathcal{D}}$ is a Markov time and so is $\tau_{\mathcal{D}} \wedge \tau = \min(\tau_{\mathcal{D}}, \tau)$. Thus, using (2.14) and the optional stopping theorem for martingales, we conclude that if $u(\sigma, \tau, x, y)$ is any bounded solution of

$$(2.15) \qquad \partial u/\partial\sigma + \mathcal{L}_\sigma u = 0, \qquad \sigma < \tau,$$

then, $u(s \wedge \tau_{\mathcal{D}}, \tau, X(s \wedge \tau_{\mathcal{D}}), Y(s \wedge \tau_{\mathcal{D}}))$, $\sigma \leq s \leq \tau$, is a martingale. We have therefore

$$(2.16) \qquad E\{u(\tau \wedge \tau_{\mathcal{D}}, \tau, X(\tau \wedge \tau_{\mathcal{D}}), Y(\tau \wedge \tau_{\mathcal{D}}))\} = u(\sigma, \tau, x, y).$$

This identity can be rewritten in the form

$$(2.17) \qquad \begin{aligned} u(\sigma, \tau, x, y) &= E\{u(\tau, \tau, X(\tau), Y(\tau)), \tau \leq \tau_{\mathcal{D}}\} \\ &\quad + E\{u(\tau_{\mathcal{D}}, \tau, X(\tau_{\mathcal{D}}), Y(\tau_{\mathcal{D}})), \tau > \tau_{\mathcal{D}}\}. \end{aligned}$$

Consider now the boundary value problem

$$\frac{\partial u(\sigma, \tau, x, y)}{\partial\sigma} + \mathcal{L}_\sigma u(\sigma, \tau, x, y) = 0, \qquad \sigma < \tau, \quad x \in \mathcal{D}, \quad y \in R^m,$$

$$(2.18) \quad u(\tau, \tau, x, y) = f(x, y), \qquad x \in \mathcal{D}, \quad y \in R^m,$$

$$u(\sigma, \tau, x, y) = h(\sigma, x, y),$$

$$x \in \partial\mathcal{D}, \quad y \in \{z \in R^m : F(\sigma, x, z) \cdot n(x) > 0\}.$$

Here f and h are bounded measurable functions and $n(x)$ denotes the unit outer normal to $\partial\mathcal{D}$ at the point $x \in \partial\mathcal{D}$. We shall denote the set of y values in the brackets above by $F(\sigma, x, y) \cdot n(x) > 0$ for notational convenience. From (2.15) and (2.17) it follows that the function $u(\sigma, \tau, x, y)$ defined by

$$(2.19) \qquad \begin{aligned} u(\sigma, \tau, x, y) &= E\{f(X(\tau), Y(\tau)), \tau \leq \tau_{\mathcal{D}}\} \\ &\quad + E\{h(\tau_{\mathcal{D}}, X(\tau_{\mathcal{D}}), Y(\tau_{\mathcal{D}})), \tau > \tau_{\mathcal{D}}\}, \end{aligned}$$

satisfies, in a generalized sense, the boundary value problem (2.18). In particular, the boundary values h are assumed at those points (x, y), $x \in \partial\mathcal{D}$, $F \cdot n > 0$, that are accessible from the interior and h is continuous. Throughout, we will consider the boundary value problem (2.18) within this generalized sense only, by accepting the representation (2.19) as its solution.

The renewal argument that led to (2.9) from (2.8) can be used here

again to show that (2.19) satisfies the integral equation,

$$u(\sigma, \tau, x, y) = \chi(\tau \leqq \bar{\tau}_{\mathscr{D}}) f(\xi(\tau), \eta(\tau)) \exp\left(-\int_{\sigma}^{\tau} q(s, \xi(s), \eta(s))\, ds\right)$$

$$+ \chi(\tau > \bar{\tau}_{\mathscr{D}}) h(\bar{\tau}_{\mathscr{D}}, \xi(\bar{\tau}_{\mathscr{D}}), \eta(\bar{\tau}_{\mathscr{D}}))$$

$$\times \exp\left(-\int_{\sigma}^{\bar{\tau}_{\mathscr{D}}} q(s, \xi(s), \eta(s))\, ds\right)$$

(2.20)

$$+ \iint_{\sigma}^{\tau \wedge \bar{\tau}_{\mathscr{D}}} u(s, \tau, \xi(s), z) \pi(s, \xi(s), \eta(s), dz) q(s, \xi(s), \eta(s)).$$

$$\times \exp\left(-\int_{\sigma}^{s} q(\gamma, \xi(\gamma), \eta(\gamma))\, d\gamma\right) ds.$$

Here $\bar{\tau}_{\mathscr{D}}$ is the first time $\xi(\tau) = \xi(\tau, \sigma, x, y)$ reaches the boundary $\partial \mathscr{D}$ starting from $x \in \mathscr{D}$, $y \in R^m$ at time σ and $\chi(\tau \leqq \bar{\tau}_{\mathscr{D}})$ is the characteristic function of the set $\{\tau \leqq \bar{\tau}_{\mathscr{D}}\}$ i.e., it is zero for those (σ, x, y) for which $\tau > \bar{\tau}$ and one otherwise. The integral equation (2.20) is equivalent to (2.18) in the generalized sense discussed above.

If in (2.19) $f(x, y) \equiv 0$, then, since $h(\sigma, x, y)$ is an arbitrary bounded measurable function, the solution of (2.20) (or (2.18)) yields the joint distribution of the exit time $\tau_{\mathscr{D}}$, the exit position $X(\tau_{\mathscr{D}})$ and $Y(\tau_{\mathscr{D}})$. If $h(\sigma, x, y) = 0$, then the solution of (2.20) (or (2.18)) yields the joint distribution of $(X(\tau), Y(\tau))$ conditional on not reaching the boundary $\partial \mathscr{D}$ in $[\sigma, \tau]$.

Let us assume that $\pi(\sigma, x, y, dz)$ has a density $\tilde{\pi}$ relative to Lebesgue measure on R^m,

$$\tilde{\pi}(\sigma, x, y, z) = \frac{\pi(\sigma, x, y, dz)}{dz}.$$

Let us also assume that F and G are smooth. We define the linear operator $\mathscr{L}_{\tau}^{*}$ on smooth functions as follows:

$$\mathscr{L}_{\tau}^{*} f(x, y) = -\frac{\partial}{\partial x}(F(\tau, x, y) f(x, y)) - \frac{\partial}{\partial y}(H(\tau, x, y) f(x, y))$$

(2.21)

$$+ \int q(\tau, x, z) \tilde{\pi}(\tau, x, z, y) f(x, z)\, dz - q(\tau, x, y) f(x, y).$$

Here $\partial(Ff)/\partial x$ denotes the divergence of the vector function Ff. Clearly, $\mathscr{L}_{\tau}^{*}$ is the formal adjoint of the operator $\mathscr{L}_{\tau}$ defined by (2.10). For smooth functions $f(s, x, y)$, $g(s, x, y)$ the following Green's identity holds.

$$\int_{\sigma}^{\tau} \int_{\mathscr{D}} \int [f(\partial_s + \mathscr{L}_s)g - g(-\partial_s + \mathscr{L}_s^{*})f]\, dy\, dx\, ds$$

(2.22)

$$= \int_{\mathscr{D}} \int fg\, dy\, dx\,\Big|_{\sigma}^{\tau} + \int_{\sigma}^{\tau} \int_{\partial\mathscr{D}} \int F \cdot nfg\, dy\, dS(x)\, ds.$$

Let $g(s, x, y)$ be a bounded smooth function that vanishes identically when s is near σ, when s is large, and when x is outside a compact subset of $\mathscr{D}$. For this g and with $\tau = \tau_\mathscr{D}$ in (2.12) we obtain

$$0 = E\left\{ \int_\sigma^{\tau_\mathscr{D}} (\partial_s + \mathscr{L}_s) g(s, X(s), Y(s)) \, ds \right\}$$

(2.23) $$= \int_\sigma^\infty E\{(\partial_s + \mathscr{L}_s) g(s, X(s), Y(s)), s \leqq \tau_\mathscr{D}\} \, ds$$

$$= \int_\sigma^\infty \int_\mathscr{D} \int p(\sigma, s, x, y; \bar{x}, \bar{y})(\partial_s + \mathscr{L}_s) g(s, \bar{x}, \bar{y}) \, d\bar{y} \, d\bar{x} \, ds.$$

Here $p(\sigma, \tau, x, y; \bar{x}, \bar{y})$ denotes the formal density of $(X(\tau), Y(\tau))$ under the condition that $X(\sigma) = x$, $Y(\sigma) = y$ and $\tau \leqq \tau_\mathscr{D}$. Under suitable conditions, since g is arbitrary, we may deduce from (2.23), by integration by parts, that p satisfies the forward Kolmogorov equation

$$(-\partial_\tau + \mathscr{L}_\tau^*) p(\sigma, \tau, x, y; \bar{x}, \bar{y}) = 0, \qquad \tau > \sigma, \quad \bar{x} \in \mathscr{D},$$

(2.24) $$p(\sigma, \sigma, x, y; \bar{x}, \bar{y}) = \delta(x - \bar{x}) \, \delta(y - \bar{y}),$$

$$p(\sigma, \tau, x, y; \bar{x}, \bar{y}) = 0, \qquad \bar{x} \in \partial\mathscr{D}, \quad F(\tau, \bar{x}, \bar{y}) \cdot n(\bar{x}) < 0.$$

The boundary condition on $\partial\mathscr{D}$ in (2.24) follows from the fact that $F(\tau_\mathscr{D}, X(\tau_\mathscr{D}), Y(\tau_\mathscr{D})) \cdot n(x(\tau_\mathscr{D})) \geqq 0$ i.e., $X(\tau)$ is moving towards the boundary from the interior at the instant before reaching the boundary. It follows immediately from (2.10) that for each fixed $(\tau, \bar{x}, \bar{y})$ the function $p(\sigma, \tau, x, y; \bar{x}, \bar{y})$ satisfies the boundary value problem (2.25) which is dual to (2.24).

$$(\partial_\sigma + \mathscr{L}_\sigma) p(\sigma, \tau, x, y; \bar{x}, \bar{y}) = 0, \qquad \sigma < \tau, \quad x \in \mathscr{D},$$

(2.25) $$p(\tau, \tau, x, y; \bar{x}, \bar{y}) = \delta(x - \bar{x}) \, \delta(y - \bar{y}),$$

$$p(\sigma, \tau, x, y; \bar{x}, \bar{y}) = 0, \qquad x \in \partial\mathscr{D}, \quad F(\tau, x, y) \cdot n(x) \geqq 0.$$

Clearly p is the density of the measure (2.11) relative to Lebesgue measure (assuming that it exists) under the condition $\tau \leqq \tau_\mathscr{D}$.

We shall employ the preceding analysis to express probabilistically quantities of interest in linear transport theory [4].

The basic boundary value problem in linear transport theory is the following. To find a function $\phi(\sigma, \tau, \bar{x}, \bar{y})$ satisfying

$$(-\partial_\tau + \mathscr{L}_\tau^*) \phi(\sigma, \tau, \bar{x}, \bar{y}) + A(\tau, \bar{x}, \bar{y}) \phi(\sigma, \tau, \bar{x}, \bar{y}) = -B(\tau, \bar{x}, \bar{y}),$$

$$\tau > \sigma, \quad \bar{x} \in \mathscr{D},$$

(2.26) $$\phi(\sigma, \sigma, \bar{x}, \bar{y}) = f(\bar{x}, \bar{y}),$$

$$-F(\tau, \bar{x}, \bar{y}) \cdot n(\bar{x}) \phi(\sigma, \tau, \bar{x}, \bar{y}) = g(\tau, \bar{x}, \bar{y}), \qquad \bar{x} \in \partial\mathscr{D},$$

$$F(\tau, \bar{x}, \bar{y}) \cdot n(\bar{x}) < 0$$

Here, A, B, f and g are bounded measurable functions. Usually, $\bar{x} \in R^3$, $\bar{y} \in R^3$ represent the position and velocity respectively of a particle of unit mass undergoing collisions. In this case $F(\tau, \bar{x}, \bar{y}) \equiv \bar{y}$ and $H(\tau, \bar{x}, \bar{y}) \equiv 0$. The solution ϕ of (2.26) represents the average density of particles at $\bar{x}$ with velocity $\bar{y}$ at time τ, A is the rate of creation or destruction of particles and B represents sources or sinks. The quantity $-\bar{y} \cdot n(\bar{x})\phi(\sigma, \tau, \bar{x}, \bar{y})$, $\bar{x} \in \partial\mathscr{D}$, $\bar{y} \cdot n(\bar{x}) < 0$ is the flux density of particles entering the region $\mathscr{D}$ through the boundary $\partial\mathscr{D}$ and it is a prescribed function $g(\tau, \bar{x}, \bar{y})$ as is the initial particle density $f(\bar{x}, \bar{y})$ at time σ. Particles which exit from $\mathscr{D}$ never return and may be thought of as "killed" at the boundary.

The collision process is characterized by the integral and constant terms of the operator $\mathscr{L}_\tau^*$ given by (2.21). The quantity $q(\tau, \bar{x}, \bar{y})\bar{\pi}(\tau, \bar{x}, \bar{y}, \bar{z})\, d\bar{y}$ is the differential scattering cross-section i.e., the fraction of particles scattered into velocity $\bar{z}$ from velocity $\bar{y}$ within $d\bar{y}$ at time τ and location $\bar{x}$. The quantity $q(\tau, \bar{x}, \bar{y})$ is the total scattering cross-section i.e., the fraction of particles scattered into velocity $\bar{y}$ from all other directions at time τ and location $\bar{x}$. In the context of the transport process $(X(\tau), Y(\tau))$ constructed earlier, q represents the collision frequency (because of (2.2)) and π the jump measure.

When H is not identically equal to zero it represents an external force field. When F is a general function, not identically equal to $\bar{y}$, it represents the local group velocity of wave packets and is therefore appropriate to transport theory for wave processes.

To give a probabilistic representation to the solution of (2.26), let $h(\bar{x}, \bar{y})$ be a bounded measurable function and set

$$(2.27) \qquad \phi_h(\sigma, \tau) = \int_\mathscr{D}\int \phi(\sigma, \tau, \bar{x}, \bar{y})h(\bar{x}, \bar{y})\, d\bar{x}\, d\bar{y}.$$

Knowledge of ϕ_h for arbitrary h is equivalent to knowing ϕ. However, ϕ_h can be constructed under very general conditions in which case (2.26) is satisfied only in a generalized sense. This is a typical advantage of probabilistic representations. Let p_h be defined by

$$
\begin{aligned}
&p_h(\sigma, \tau, x, y) \\
(2.28) \quad &= E\left\{\exp\left(\int_\sigma^\tau A(s, X(s), Y(s))\, ds\right)h(X(\tau), Y(\tau)), \quad \tau \leq \tau_\mathscr{D}\right\}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
\phi_h(\sigma, \tau) &= \int_\sigma^\tau \int_\mathscr{D}\int p_h(s, \tau, x, y)B(s, x, y)\, dy\, dx\, ds \\
(2.29) \qquad &+ \int_\mathscr{D}\int p_h(\sigma, \tau, x, y)f(x, y)\, dy\, dx \\
&+ \int_\sigma^\tau \int_{\partial\mathscr{D}}\int_{F \cdot n < 0} g(s, x, y)p_h(s, \tau, x, y)\, dy\, dS(x)\, ds.
\end{aligned}
$$

One can formally verify that (2.29) gives a representation of the solution of (2.26) via (2.27). For the verification, the identity (2.22) is employed as well as a computation analogous to the one that led to the representation (2.19) for (2.18).

In general, (2.29) is taken as the solution of (2.26) via (2.27). This is the generalized sense in which we interpret (2.26). The question of existence and uniqueness for the basic problem of transport theory is thus settled at least in the sense explained. The representation (2.29) provides, in addition, a tool for studying by probabilistic methods asymptotic problems that arise in linear transport theory.

Let us also view the process (X, Y) in the context of stochastic differential equations. $X(\tau, \sigma, x, y)$ satisfies the equation

$$(2.30) \qquad \frac{dX(\tau)}{d\tau} = F(\tau, X(\tau), Y(\tau)), \quad \tau > \sigma, \qquad X(\sigma, \sigma, x, y) = x.$$

However, the process $Y(\tau)$ is not given independently of X. If we regard $Y(\tau)$ as the random coefficients in the stochastic equation (2.30), then, in the present context, the coefficients depend, in general, on the solution since H, q and π may depend on x. We could write down a generalized equation for Y [12] but we prefer to work directly with the backward Kolmogorov equation (2.10). The dependence of the coefficients $Y(\tau)$ on the solution $X(\tau)$ is reflected in the x-dependence of the operator Q in (2.10) and can be interpreted as feedback i.e., the solution affects the statistical properties of the random coefficients.

3. **Asymptotic problems.**   There are very few problems in linear transport theory that can be solved analytically [19]. These problems are restricted to spatially homogeneous collision operators and, essentially, half-space boundary value problems. The necessity for approximation methods is therefore obvious. In the context of the more general problems (2.10), (2.18), (2.26) or the stochastic equation formulation (2.30), the necessity for approximations is felt even more acutely. We shall focus here on the diffusion approximation which is frequently quite effective. Other approximations, such as the discrete ordinate method, are considered in [4], [20]. Let us now describe the set-up for the diffusion approximation.

Let $\varepsilon > 0$ denote a small parameter and suppose that the collision frequency is q/ε^2 rather than q; i.e., let ε^2 denote the order of magnitude of the time between jumps (or collisions) of Y. Under reasonable conditions, speeding up of the jumps will lead to a uniform equilibrium state. In order to allow for some nontrivial space and time dependence we also speed up the velocity field F and the force field H by replacing them by F/ε and H/ε respectively. It turns out that the speeding up of the velocity

field cannot be done in an arbitrary manner. F must be appropriately "centered" or else the process degenerates in the limit $\varepsilon \to 0$. By centering we mean that the average value of F with respect to the local equilibrium distribution of Y (cf. §6 for precise definitions) should be approximately zero.

With these general remarks in mind we state the transport problem whose asymptotic limit as $\varepsilon \downarrow 0$ will be studied. The transport process $(X^{(\varepsilon)}(\tau, \sigma, x, y), Y^{(\varepsilon)}(\tau, \sigma, x, y))$ is defined as the Markov process with backward equation

$$
\frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial \sigma} + \left[\frac{1}{\varepsilon} F\left(\sigma, \frac{\sigma}{\varepsilon^2}, x, y\right) + G\left(\sigma, \frac{\sigma}{\varepsilon^2}, x, y\right) \right] \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial x}
$$

$$
(3.1) \qquad + \frac{1}{\varepsilon} H\left(\sigma, \frac{\sigma}{\varepsilon^2}, x, y\right) \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial y} + \frac{q(\sigma, x, y)}{\varepsilon^2}
$$

$$
\times \int u^{(\varepsilon)}(\sigma, \tau, x, z) \pi(\sigma, x, y, dz) - \frac{q(\sigma, x, y)}{\varepsilon^2} u^{(\varepsilon)}(\sigma, \tau, x, y) = 0,
$$

$$
\sigma < \tau,
$$

$$
u^{(\varepsilon)}(\tau, \tau, x, y) = f(x, y),
$$

$$
u^{(\varepsilon)}(\sigma, \tau, x, y) = E\{f(X^{(\varepsilon)}(\tau, \sigma, x, y), Y^{(\varepsilon)}(\tau, \sigma, x, y))\}.
$$

Note that we have introduced the additional term G in the velocity field and we have allowed F, G and H to depend on the "fast" time scale σ/ε^2 (or τ/ε^2). The reason for allowing explicit dependence on the fast time comes from the stochastic differential equation interpretation.

The process $X^{(\varepsilon)}(\tau) = X^{(\varepsilon)}(\tau, \sigma, x, y)$ satisfies the stochastic differential equation

$$
\frac{dX^{(\varepsilon)}(\tau)}{d\tau} = \frac{1}{\varepsilon} F(\tau, \tau/\varepsilon^2, X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))
$$

$$
(3.2) \qquad + G(\tau, \tau/\varepsilon^2, X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau)), \quad \tau > \sigma,
$$

$$
X^{(\varepsilon)}(\sigma, \sigma, x, y) = x.
$$

Since $Y^{(\varepsilon)}$ depends on $X^{(\varepsilon)}$, (3.2) is not a closed system of equations as we remarked in §2. The field F is centered and therefore, roughly, the rate of change of $X^{(\varepsilon)}$ is $O(1)$. On the other hand the rate of change of $Y^{(\varepsilon)}$ is, roughly, of the order of magnitude of the jump frequency i.e. $O(1/\varepsilon^2)$. Thus, the scaling we have introduced distinguishes $X^{(\varepsilon)}$ as the slowly varying and $Y^{(\varepsilon)}$ as the rapidly varying part of the process $(X^{(\varepsilon)}, Y^{(\varepsilon)})$ Furthermore, since $Y^{(\varepsilon)}$ is rapidly varying, we expect that it ought to be possible to handle rapid variations of F and G and this leads to the explicit fast-time dependence in (3.2). If $Y^{(\varepsilon)}$ does not depend on $X^{(\varepsilon)}$ then

the analysis of (3.2) in the limit $\varepsilon \downarrow 0$ has been carried out in [21] wherein additional references are given to both theoretical work and to specific applications. In the context of Itô equations a problem similar to (3.2) (cf. §7, Remark 2) has been investigated by Khasminskii [22]. Further information and examples of asymptotic results for processes with fast and slow components are given in [23].

The results stated in §6 tell us that, under appropriate hypotheses, $X^{(\varepsilon)}$ converges weakly as $\varepsilon \downarrow 0$ to a diffusion Markov process whose generator can be computed explicitly. This is the diffusion approximation. Note that for $\varepsilon > 0$ neither $X^{(\varepsilon)}$ nor $Y^{(\varepsilon)}$ are Markov separately but $X^{(\varepsilon)}$ becomes Markov in the limit $\varepsilon \downarrow 0$.

Let us consider the scaled version of the boundary value problem (2.26). It is customary in transport theory to redefine g so that the boundary condition

$$-F(\tau, \bar{x}, \bar{y}) \cdot n(\bar{x})\phi(\sigma, \tau, \bar{x}, \bar{y}) = g(\tau, \bar{x}, \bar{y})$$

is replaced by

$$\phi(\sigma, \tau, \bar{x}, \bar{y}) = g(\tau, \bar{x}, \bar{y}), \qquad \bar{x} \in \partial\mathscr{D}, \quad \tau \geqq \sigma,$$
$$F(\tau, \bar{x}, \bar{y}) \cdot n(\bar{x}) < 0.$$

The problem is then to analyze $\phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{y})$ as $\varepsilon \downarrow 0$ where

$$
\begin{aligned}
&\frac{\partial \phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{y})}{\partial \tau} \\
&\quad + \frac{\partial}{\partial \bar{x}}\left\{\left[\frac{1}{\varepsilon}F\left(\tau, \frac{\tau}{\varepsilon^2}, \bar{x}, \bar{y}\right) + G\left(\tau, \frac{\tau}{\varepsilon^2}, \bar{x}, \bar{y}\right)\right]\phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{y})\right\} \\
&\quad + \frac{\partial}{\partial \bar{y}}\left(\frac{1}{\varepsilon}H(\tau, \tau/\varepsilon^2, \bar{x}, \bar{y})\phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{y})\right) \\
&\quad - \frac{1}{\varepsilon^2}\int q(\tau, \bar{x}, \bar{z})\tilde{\pi}(\tau, \bar{x}, \bar{z}, \bar{y})\phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{z})\,d\bar{z} \\
&\quad + \frac{1}{\varepsilon^2}q(\tau, \bar{x}, \bar{y})\phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{y}) \\
&\quad - A(\tau, \bar{x}, \bar{y})\phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{y}) = B(\tau, \bar{x}, \bar{y}), \\
&\hspace{9cm} \tau > \sigma, \quad \bar{x} \in \mathscr{D}, \\
&\qquad \phi^{(\varepsilon)}(\sigma, \sigma, \bar{x}, \bar{y}) = f(\bar{x}, \bar{y}), \qquad \bar{x} \in \mathscr{D}, \\
&\qquad \phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{y}) = g(\tau, \bar{x}, \bar{y}), \qquad \bar{x} \in \partial\mathscr{D}, \\
&\left(\frac{1}{\varepsilon}F(\tau, \tau/\varepsilon^2, \bar{x}, \bar{y}) + G(\tau, \tau/\varepsilon^2, \bar{x}, \bar{y})\right) \cdot n(\bar{x}) < 0.
\end{aligned}
$$

(3.3)

Let $h(\bar{x}, \bar{y})$ be a bounded measurable function and define $\phi_h^{(\varepsilon)}$ and $p_h^{(\varepsilon)}$ as follows:

$$\phi_h^{(\varepsilon)}(\sigma, \tau) = \iint \phi^{(\varepsilon)}(\sigma, \tau, \bar{x}, \bar{y}) h(\bar{x}, \bar{y}) \, d\bar{y} \, d\bar{x},$$

(3.4) $\qquad p_h^{(\varepsilon)}(\sigma, \tau, x, y) = E\left\{\exp\left(\int_\sigma^\tau A(s, X^{(\varepsilon)}(s), Y^{(\varepsilon)}(s)) \, ds\right)\right.$

$$\left. \times h(X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau)), \tau \leqq \tau_{\mathscr{D}}\right\}.$$

The representation of $\phi^{(\varepsilon)}$ via $\phi_h^{(\varepsilon)}$, analogous to (2.29), now takes the form

$$\phi_h^{(\varepsilon)}(\sigma, \tau) = \int_\sigma^\tau \int_{\mathscr{D}} \int p_h^{(\varepsilon)}(s, \tau, x, y) B(s, x, y) \, dy \, dx \, ds$$

$$+ \int_{\mathscr{D}} \int p_h^{(\varepsilon)}(\sigma, \tau, x, y) f(x, y) \, dy \, dx$$

(3.5)

$$- \int_\sigma^\tau \int_{\partial\mathscr{D}} \int_{(F/\varepsilon+G)\cdot n < 0} \left[\frac{1}{\varepsilon} F\left(s, \frac{s}{\varepsilon^2}, x, y\right) + G\left(s, \frac{s}{\varepsilon^2}, x, y\right)\right]$$

$$\times n(x) g(s, x, y) \cdot p_h^{(\varepsilon)}(s, \tau, x, y) \, dy \, dS(x) \, ds.$$

From this representation it might appear that the weak convergence results of §6 yield the relevant results for $\phi^{(\varepsilon)}$. This however is only partially true because the third term on the right of (3.5) leads to boundary layers near $\partial\mathscr{D}$. The analysis of the boundary layers requires separate considerations which are not given here (see also Remark 3 of §7).

4. Learning theory and other examples. An interesting class of transport processes is obtained from the following mathematical model in learning theory [7], [8]. Let $(\mathscr{X}, \mathscr{B})$ and $(\mathscr{Y}, \mathscr{A})$ be measurable spaces. We consider a subject undergoing observation by an experimenter and suppose temporarily that the times of observation are multiples of a fixed time unit. At the time of the kth observation the state of the subject is a random variable X_k with values in $\mathscr{X}$ whereas the response of the subject, as observed by the experimenter, is a random variable Y_k with values in $\mathscr{Y}$. We assume that functions $P_k(A; X)$, $k \geqq 0$, on $\mathscr{A} \times \mathscr{X}$ are given such that for fixed $X \in \mathscr{X}$ they are probability measures on $\mathscr{A}$ and for fixed $A \in \mathscr{A}$ they are $\mathscr{B}$ measurable functions of X. We also assume that measurable functions $v_k(X, Y)$, $k \geqq 0$, on $\mathscr{X} \times \mathscr{Y}$ are given. The learning model for the subject under observation is now as follows.

Given that the state of the subject at the kth observation $X_k = X$ then,

(4.1) $\qquad P\{Y_k \in A \mid X_k = X\} = P_k(A; X).$

That is, $P_k(A; X)$ is the conditional probability of observing the subject in A given the state of the subject is X. Moreover, given that $Y_k = Y$ and $X_k = X$ we set

$$(4.2) \qquad\qquad X_{k+1} = v_k(X, Y)$$

i.e., the state of the subject at the time of the $k+1$st observation is a given function of the state X and the observation Y at time k. We assume that the distribution of X_0 is also given. The process $(X_k, Y_k), k \geq 0$, so constructed is evidently a discrete time Markov process on $\mathscr{X} \times \mathscr{Y}$. A basic question is: for given P_k and v_k what is the behavior of X_k, the state of the subject, for k large? One may interpret v_k as the penalty or reward that the experimenter offers the subject in order to influence its future state. Similarly, the dependence of P_k on X indicates that the response of the subject is related to its state although in a stochastic manner.

Naturally, in order to obtain a rich asymptotic theory one must make some reasonable assumptions about the spaces $\mathscr{X}$, $\mathscr{Y}$ and P_k and v_k. From the point of view of a literal state-response interpretation of the process (X_k, Y_k), it may well be that $\mathscr{X}$ and $\mathscr{Y}$ should most appropriately be chosen as finite sets. However, we shall interpret the learning model in a broader sense whereby one has a feedback mechanism so that some components of the process can influence the evolution of the remaining components. From this statement the connection with transport processes should be apparent. Let us make this connection somewhat more explicit.

Suppose that $\mathscr{X}$ is the Euclidean space R^n but let $\mathscr{Y}$ be arbitrary. In the context of the broader interpretation of the learning model there is little reason to maintain the time discrete. In continuous time the learning process $(X(\tau), Y(\tau))$ is constructed in exactly the way the transport process was constructed in §2 except that now $H \equiv 0$ in which case the nature of the space $\mathscr{Y}$ becomes irrelevant (cf. §6). Clearly the velocity field $F(\tau, X, Y)$ is the continuous analog of $v_k(X, Y) - X$ i.e., the difference $X_{k+1} - X_k$ goes over to a derivative. The probability measure $P_k(A, X)$ is replaced by the collision operator (2.13) characterized by $q(\tau, X, Y)$ and $\pi(\tau, X, Y, A)$ (note, however, we also allow Y-dependence). Continuous time learning processes are therefore special cases of transport processes. Note that it is important for learning models that (i) $F(\tau, X, Y)$ depend on Y as indicated, or else no control is effected and (ii) $q(\tau, X, Y)$ and $\pi(\tau, X, Y, A)$ depend on X or else the subject responds with the same probabilistic mechanism no matter what its state is.

Let us consider the question about the long-time behavior of the state $X(\tau)$ for the continuous time learning process. In order to make this question meaningful one must interpret the words "long time" appropriately.

This is conveniently done by introducing a small parameter $\varepsilon > 0$ in the following manner.

Consider the process $(X^{(\varepsilon)}(\tau, \sigma, x, y), Y^{(\varepsilon)}(\tau, \sigma, x, y))$ with backward equation

$$
\begin{aligned}
\frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial \sigma} &+ G(\sigma, x, y) \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial x} + \frac{q(\sigma, x, y)}{\varepsilon^2} \\
&\times \int u^{(\varepsilon)}(\sigma, \tau, x, z) \pi(\sigma, x, y, dz) \\
&- \frac{q(\sigma, x, y)}{\varepsilon^2} u^{(\varepsilon)}(\sigma, \tau, x, y) = 0, \quad \sigma < \tau,
\end{aligned}
$$

(4.2)

$$
u^{(\varepsilon)}(\tau, \tau, x, y) = f(x, y),
$$

and corresponding stochastic equations for $X^{(\varepsilon)}$

(4.3) $\dfrac{dX^{(\varepsilon)}(\tau)}{d\tau} = G(\tau, X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau)), \quad \tau > \sigma, \quad X^{(\varepsilon)}(\sigma, \sigma, x, y) = x.$

The asymptotic analysis of $X^{(\varepsilon)}(\tau)$ as $\varepsilon \downarrow 0$ corresponds to slow learning with large drift [7], ε being a measure of the speed of learning. In §6 we show that, for general G, $X^{(\varepsilon)}(\tau)$ tends to the deterministic trajectory $X^{(0)}(\tau)$ which is the solution of an appropriate system of deterministic ordinary differential equations. One then investigates the asymptotic behavior of the deviation process $[X^{(\varepsilon)}(\tau) - X^{(0)}(\tau)]/\varepsilon$. Corollary 1 of §6 gives the asymptotic behavior of this process.

When G in (4.3) is centered i.e., behaves like the F of (3.1), then $X^{(0)}(\tau)$ will be identically equal to the starting point x. This is called slow learning with small drift [7]. To obtain a more interesting asymptotic behavior we rescale the process $X^{(\varepsilon)}(\tau)$ so that it satisfies (3.2) with F centered and taking the place of G above (the G in (3.2) is now omitted to avoid confusion). The asymptotic behavior in this case is given by the theorem of §6. Thus, the asymptotic analysis of slow learning with both large and small drift follows from the asymptotic theory of transport processes.

We have not yet explained the role of the explicit dependence of F and G and H on τ/ε^2 in (3.1) as it did not arise naturally in the learning models above. We give below an example that shows how problems with explicit fast-time dependence arise.

Consider a harmonic oscillator with randomly perturbed frequency

(4.4) $\dfrac{d^2 x(t)}{dt^2} + \omega^2[1 + \varepsilon g(y(t))]x(t) = 0, \quad x(0) = x, \quad \dfrac{dx(0)}{dt} = \dot{x}.$

Here $x(t)$ denotes the position of the oscillator, ω is its angular frequency when $\varepsilon = 0$, ε is a small parameter characterizing the size of the fluctuations

and $y(t)$ is a process that depends on $x(t)$ and $dx(t)/dt$ in a manner that is specified later. Because of the dependence of y on x and dx/dt, (4.4) is a nonlinear stochastic ordinary differential equation. Let $E(t)$ denote the energy of the oscillator

$$(4.5) \qquad E(t) = \tfrac{1}{2}[(dx(t)/dt)^2 + \omega^2 x^2(t)],$$

and define $\theta(t)$ by

$$(4.6) \qquad (dx(t)/dt)[\omega x(t)]^{-1} = -\tan(\omega t + \theta(t)/2).$$

By direct computation we find that $E(t)$ and $\theta(t)$ satisfy the equations

$$\frac{1}{E}\frac{dE(t)}{dt} = \varepsilon\omega g(y(t))\sin(2\omega t + \theta(t)),$$

$$(4.7) \qquad \frac{d\theta(t)}{dt} = \varepsilon\omega g(y(t))[1 + \cos(2\omega t + \theta(t))], \quad t > 0,$$

$$E(0) = E, \qquad \theta(0) = \theta,$$

where E and θ are related to x and $\dot{x}$ via (4.5) and (4.6). Let $\log E(t) = R(t)$ and assume that $y(t)$ is a jump process that depends on $x(t)$ and $dx(t)/dt$ through $R(t)$.

Let $\tau = \varepsilon^2 t$ and set

$$R^{(\varepsilon)}(\tau) = R(\tau/\varepsilon^2), \quad \theta^{(\varepsilon)}(\tau) = \theta(\tau/\varepsilon^2), \quad y^{(\varepsilon)}(\tau) = y(\tau/\varepsilon^2);$$

then from (4.7) we obtain

$$\frac{dR^{(\varepsilon)}(\tau)}{d\tau} = \frac{\omega}{\varepsilon}g(y^{(\varepsilon)}(\tau))\sin\left(\frac{2\omega\tau}{\varepsilon^2} + \theta^{(\varepsilon)}(\tau)\right),$$

$$(4.8) \qquad \frac{d\theta^{(\varepsilon)}(\tau)}{d\tau} = \frac{\omega}{\varepsilon}g(y^{(\varepsilon)}(\tau))\left[1 + \cos\left(\frac{2\omega\tau}{\varepsilon^2} + \theta^{(\varepsilon)}(\tau)\right)\right], \quad \tau > 0,$$

$$R^{(\varepsilon)}(0) = \log E, \qquad \theta^{(\varepsilon)}(0) = \theta.$$

We let $X^{(\varepsilon)}(\tau) = (R^{(\varepsilon)}(\tau), \theta^{(\varepsilon)}(\tau))$ and $Y^{(\varepsilon)}(\tau) = y^{(\varepsilon)}(\tau)$. In this notation (4.8) assumes the form (3.2) with $G = 0$ and $F(\tau/\varepsilon^2, R^{(\varepsilon)}, \theta^{(\varepsilon)}, y^{(\varepsilon)})$ identified in the obvious manner. To specify the problem completely we must prescribe the q and the π that enter the backward equation (3.1), and, of course, the state space of $y^{(\varepsilon)}$. Suppose that $y^{(\varepsilon)}$ takes values in the unit circle S^1 and g is a bounded function on S^1 so that, when ε is small, the frequency in (4.4) does not become negative. If we also assume that q depends on R and σ only, and π depends on R, y and σ but not on θ

then, the backward equation for the process $(R^{(\varepsilon)}, \theta^{(\varepsilon)}, y^{(\varepsilon)})$ is

$$
\begin{aligned}
(4.9) \quad & \frac{\partial u^{(\varepsilon)}(\sigma, \tau, R, \theta, y)}{\partial \sigma} + \frac{\omega}{\varepsilon} g(y) \sin\left(\frac{2\omega\sigma}{\varepsilon^2} + \theta\right) \frac{\partial u^{(\varepsilon)}(\sigma, \tau, R, \theta, y)}{\partial R} \\
& + \frac{\omega}{\varepsilon} g(y) \left[1 + \cos\left(\frac{2\omega\sigma}{\varepsilon^2} + \theta\right)\right] \frac{\partial u^{(\varepsilon)}(\sigma, \tau, R, \theta, y)}{\partial \theta} \\
& + \frac{q(\sigma, R)}{\varepsilon^2} \int_{S^1} u^{(\varepsilon)}(\sigma, \tau, R, \theta, z) \pi(\sigma, R, y, dz) \\
& - \frac{q(\sigma, R)}{\varepsilon^2} u^{(\varepsilon)}(\sigma, \tau, R, \theta, y) = 0, \qquad \sigma < \tau,
\end{aligned}
$$

$$
u^{(\varepsilon)}(\tau, \tau, R, \theta, y) = f(R, \theta, y), \quad R \in R^1, \quad \theta \in [0, 2\pi), \quad y \in S^1.
$$

The example we have just presented shows how the fast time τ/ε^2 enters explicitly into the equations (4.8). Actually, the change of variables that led from (4.4) to (4.8) plays also an essential role because it transforms a problem with noncentered F to one with centered F. We shall return to this example in §8 where we apply the diffusion approximation.

5. **Random evolution.** The purpose of this section is to indicate that the asymptotic analysis of (3.1) is, in fact, a special case of a more general class of problems called random evolutions and also, to refer to some relevant literature. We purposely omit detailed descriptions because these more general problems will not concern us here.

Let $(\mathcal{Y}, \mathcal{A})$ be a measurable space and $\mathcal{X}_1$, $\mathcal{X}_2$ be Banach spaces such that $\mathcal{X}_2$ is continuously imbedded into $\mathcal{X}_1$. Let $V(\tau, t, y)$, $\tau \geq 0$, $t \geq 0$, $y \in \mathcal{Y}$, be bounded linear operators from $\mathcal{X}_2$ to $\mathcal{X}_1$ strongly measurable as functions of their arguments. Let $\mathcal{B}(\mathcal{Y}, \mathcal{X}_1)$ be the space of bounded measurable functions on $\mathcal{Y}$ with values in $\mathcal{X}_1$ and let $Q(\tau)$, $\tau \geq 0$, be a bounded linear operator on $\mathcal{B}(\mathcal{Y}, \mathcal{X}_1)$. Let $\varepsilon > 0$ be a parameter and consider the following Cauchy problem for $u^{(\varepsilon)}(\sigma, \tau, y)$:

$$
\begin{aligned}
(5.1) \quad & \frac{du^{(\varepsilon)}(\sigma, \tau, y)}{d\tau} + \frac{1}{\varepsilon} V\left(\sigma, \frac{\sigma}{\varepsilon^2}, y\right) u^{(\varepsilon)}(\sigma, \tau, y) \\
& \qquad\qquad + \frac{1}{\varepsilon^2} Q(\sigma) u^{(\varepsilon)}(\sigma, \tau, y) = 0, \quad \sigma < \tau,
\end{aligned}
$$

$$
u^{(\varepsilon)}(\tau, \tau, y) = f(y) \in \mathcal{B}(\mathcal{Y}, \mathcal{X}_2).
$$

It is, of course, an important problem to find sufficient conditions so that (5.1) is a well-posed problem. In the notation of §3, the Cauchy problem

$$
\begin{aligned}
(5.2) \quad & \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial \sigma} + \frac{1}{\varepsilon} F\left(\sigma, \frac{\sigma}{\varepsilon^2}, x, y\right) \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial x} + \frac{q(\sigma, x, y)}{\varepsilon^2} \\
& \times \int u^{(\varepsilon)}(\sigma, \tau, x, z) \pi(\sigma, x, y, dz) - \frac{q(\sigma, x, y)}{\varepsilon^2} u^{(\varepsilon)}(\sigma, \tau, x, y) = 0, \quad \sigma < \tau,
\end{aligned}
$$

$$
u^{(\varepsilon)}(\tau, \tau, x, y) = f(x, y),
$$

is a special case of (3.1) and provides a concrete example of the abstract problem (5.1) for which well posedness questions are settled easily.[2] The complete identification of (5.2) as an example of (5.1) is as follows. $\mathcal{X}_1$ is the space of bounded continuous functions on R^n with the sup norm, $\mathcal{X}_2$ is the space of functions with bounded continuous derivatives and normed by the sum of the sup norm of the functions and its derivatives, $V(\tau, t, y)$ are the vector fields $F(\tau, t, x, y)(\partial/\partial x)$, $\mathcal{Y}$ can be an arbitrary measurable space and $Q(\tau)$ is the collision operator.

When V is a general operator and not the vector field $F(\partial/\partial x)$ then one cannot interpret (5.1) as the backward equation of a process $(X^{(\varepsilon)}, Y^{(\varepsilon)})$. To find the appropriate stochastic problem associated with (5.1) we consider again the concrete problem (5.2). Let $f(x, y)$ be a continuously differentiable function of $x \in R^n$ and a bounded measurable function of $y \in \mathcal{Y}$. Define a linear operator $U^{(\varepsilon)}(\tau, \sigma)$ acting on such functions as follows:

$$(5.3) \qquad U^{(\varepsilon)}(\sigma, \tau)f(x, y) = f(X^{(\varepsilon)}(\tau, \sigma, x, y), Y^{(\varepsilon)}(\tau, \sigma, x, y)).$$

Let $Y_0^{(\varepsilon)}(\tau, \sigma, y; x)$ be the Markov process on $\mathcal{Y}$ with backward equation

$$\frac{\partial u_0^{(\varepsilon)}(\sigma, \tau, y; x)}{\partial \sigma} + \frac{1}{\varepsilon^2} q(\sigma, x, y) \int u_0^{(\varepsilon)}(\sigma, \tau, z; x) \pi(\sigma, x, y, dz)$$

$$(5.4) \qquad - (1/\varepsilon^2) q(\sigma, x, y) u_0^{(\varepsilon)}(\sigma, \tau, y; x) = 0, \quad \sigma < \tau,$$

$$u_0^{(\varepsilon)}(\tau, \tau, y; x) = f(x, y),$$

$$u_0^{(\varepsilon)}(\sigma, \tau, y; x) = E\{f(Y_0^{(\varepsilon)}(\tau, \sigma, y; x), x)\}.$$

Let $V(\tau, t)$ denote the operator defined by

$$(5.5) \qquad V(\tau, t)f(x, y) = F(\tau, t, x, y)(\partial f(x, y)/\partial x).$$

It is not difficult to verify that $U^{(\varepsilon)}(\sigma, \tau)$ satisfies, in an appropriate sense,[3] the relations

$$U^{(\varepsilon)}(\sigma, \tau) = U_0^{(\varepsilon)}(\sigma, \tau) + \frac{1}{\varepsilon} \int_\sigma^\tau U^{(\varepsilon)}(\sigma, s) V\left(s, \frac{s}{\varepsilon^2}\right) U_0^{(\varepsilon)}(s, \tau) \, ds;$$

$$(5.6)$$

$$U^{(\varepsilon)}(\sigma, \tau) = U_0^{(\varepsilon)}(\sigma, \tau) + \frac{1}{\varepsilon} \int_\sigma^\tau U_0^{(\varepsilon)}(\sigma, s) V\left(s, \frac{s}{\varepsilon^2}\right) U^{(\varepsilon)}(s, \tau) \, ds$$

where

$$(5.7) \qquad U_0^{(\varepsilon)}(\sigma, \tau)f(x, y) = f(Y_0^{(\varepsilon)}(\tau, \sigma, x, y), x).$$

---

[2] In general one assumes that (5.1) is well posed and has the necessary properties for the asymptotic analysis.

[3] See equations (10.23) and (10.24).

In the abstract set-up we begin with the random operators $U^{(\varepsilon)}$ and $U_0^{(\varepsilon)}$ satisfying the functional equations (5.6) (under appropriate hypotheses). As in the concrete example it is not hard to show that

$$(5.8) \qquad u^{(\varepsilon)}(\sigma, \tau, y) = E\{U^{(\varepsilon)}(\sigma, \tau)f(y)\}$$

and this provides the stochastic background for (5.1). One reason why one wants to investigate (5.6) or (5.1) is that problems of this sort arise in stochastic models in Quantum physics [24], [25 and references therein]. The derivation of master equations for stochastic Quantum problems corresponds to the diffusion limit for the "Classical" problems (5.2) that we consider here.

The study of random evolutions, their asymptotic analysis in particular, has received considerable attention recently and is indeed a subject that has potentially many applications. Starting with the work of Griego and Hersh [26] several works followed which considered various aspects of the Cauchy problem (5.1). We refer to the works of Hersh [9], Pinsky [10] and Kurtz [11] for additional information and references.

6. **The diffusion approximation.** In the construction of transport processes in §2 the Y component of the process was R^m-valued. In the theorem we state here and its proof we require Y to take values in a complete separable and compact metric space S. The compactness is a technical restriction that excludes some interesting examples and is probably unnecessary. For this reason it was not mentioned in §§2 and 3. The other properties of S are required so that the process has a nice path structure [3].

Without a differentiable structure on S we cannot have force fields H. Thus, in the remainder we will assume that $H \equiv 0$. In Remark 1 of §7 we give the form of the results when S is a differentiable manifold and H is not zero.

On $R^n \times S$ we consider the Markov process $(X^{(\varepsilon)}(\tau, \sigma, x, y),$ $Y^{(\varepsilon)}(\tau, \sigma, x, y))$ with backward equation

$$\frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial \sigma} + \mathscr{L}_\sigma^{(\varepsilon)} u^{(\varepsilon)}(\sigma, \tau, x, y) = 0, \quad \sigma < \tau,$$

$$u^{(\varepsilon)}(\tau, \tau, x, y) = f(x, y),$$

$$u^{(\varepsilon)}(\sigma, \tau, x, y) = E\{f(X^{(\varepsilon)}(\tau, \sigma, x, y), Y^{(\varepsilon)}(\tau, \sigma, x, y))\}$$

$$(6.1) \qquad \mathscr{L}_\sigma^{(\varepsilon)} g(x, y) = \left[\frac{1}{\varepsilon} F\left(\sigma, \frac{\sigma}{\varepsilon^2}, x, y\right) + G\left(\sigma, \frac{\sigma}{\varepsilon^2}, x, y\right)\right] \frac{\partial g(x, y)}{\partial x}$$

$$+ \frac{1}{\varepsilon^2} Q_{\sigma, x} g(x, y),$$

$$Q_{\sigma, x} g(x, y) = q(\sigma, x, y) \int_S g(x, z)\pi(\sigma, x, y, dz) - q(\sigma, x, y)g(x, y).$$

Here $f(x, y)$ is a differentiable function of $x \in R^n$ and bounded measurable in $y \in S$. We assume that the vector function $F(\sigma, t, x, y)$ satisfies the following conditions:

    (i) $F(\sigma, t, x, y)$ is a measurable function of its arguments,

    (ii) $|F_i(\sigma, t, x, y)| \leqq C(1 + |x|)$,

    (iii) $|(\partial F_i/\partial x_j)(\sigma, t, x, y)| \leqq C$,

(6.2) (iv) $|(\partial^2 F_i/\partial x_j \, \partial x_k)(\sigma, t, x, y)| \leqq C(1 + |x|^\alpha)$,

$$\left| \frac{\partial^3 F_i(\sigma, t, x, y)}{\partial x_j \, \partial x_k \, \partial x_l} \right| \leqq C(1 + |x|^\alpha), \qquad \left| \frac{\partial^4 F_i(\sigma, t, x, y)}{\partial x_j \, \partial x_k \, \partial x_l \, \partial x_m} \right| \leqq C(1 + |x|^\alpha),$$

$$i = 1, 2, \cdots, n, \quad j, k, l, m = 0, 1, 2, \cdots, n.$$

Here and in the sequel we denote σ (or τ) by x_0 when convenient; C stands for a constant and we adopt throughout the convention that constants, even different constants, are all denoted by the same symbol C. We denote by α a nonnegative integer and by $|\ |$ the Euclidean norm of vectors in R^n or absolute value of scalars. The vector function $G(\tau, t, x, y)$ satisfies the same hypotheses (6.2) as F.

The collision frequency $q(\sigma, x, y)$ and the scattering measure $\pi(\sigma, x, y, A)$ are assumed to satisfy the following hypotheses:

    (i) $q(\sigma, x, y)$ and $\pi(\sigma, x, y, A)$, A a Borel set in S, are measurable functions of $\sigma \geqq 0$, $x \in R^n$, and $y \in S$.

    (ii) $0 \leqq q(\sigma, x, y) \leqq M < \infty, \quad \left| \dfrac{\partial q}{\partial x_i}(\sigma, x, y) \right| \leqq \dfrac{M}{1 + |x|},$

$$\left| \frac{\partial^2 q}{\partial x_i \, \partial x_j}(\sigma, x, y) \right| \leqq M(1 + |x|^\alpha), \quad i, j = 0, 1, 2, \cdots, n.$$

(6.3) (iii) Let $(\partial \pi(\sigma, x, y, A))/\partial x_i$ denote the signed measure obtained by differentiating π with respect to x_i. Let $|\partial \pi/\partial x_i|(\sigma, x, y, A)$ denote the total variation of this signed measure on A. We assume

$$|\partial \pi/\partial x_i|(\sigma, x, y, A) \leqq M/(1 + |x|),$$

$$\left| \frac{\partial^2 \pi}{\partial x_i \, \partial x_j} \right|(\sigma, x, y, A) \leqq M(1 + |x|^\alpha), \quad i, j = 0, 1, 2, \cdots, n.$$

Here M is a constant independent of σ, x, y, A, i and j.

Let $P_0(t, y, A; \sigma, x)$ denote the solution of the t-homogeneous backward equation

(6.4)
$$\partial P_0(t, y, A; \sigma, x)/\partial t = Q_{\sigma,x} P_0(t, y, A; \sigma, x), \quad t > 0,$$
$$P_0(0, y, A; \sigma, x) = \chi_A(y),$$

A a Borel set in S. Here $\chi_A(y)$ denotes the characteristic function of the set A and $\sigma \geqq 0$, $x \in R^n$ are merely parameters. Let $Y_0(t, y; \sigma, x)$ denote the jump Markov process on S with backward equation (6.4) that is,

$$(6.5) \qquad P\{Y_0(t, y; \sigma, x) \in A\} = P_0(t, y, A; \sigma, x), \quad t \geqq 0.$$

This process is called the frozen collision process because it coincides with $(X^{(\varepsilon)}, Y^{(\varepsilon)})$ of (6.1) when the drift is removed; that is, the spatial coordinate does not change with time, the slow time σ is frozen and $t = \tau/\varepsilon^2$.

We require the following hypotheses concerning P_0.

(i) There exists a probability measure $\bar{P}(A; \sigma, x)$ on the Borel sets depending on $x \in R^n$ and $\sigma \geqq 0$ such that for any bounded measurable function $f(y)$

$$(6.6) \qquad \int \bar{P}(d\zeta; \sigma, x)Q_{\sigma, x}f(\zeta) = 0.$$

(ii) Let μ_0 denote the signed measure

$$(6.7) \qquad \mu_0(t, y, A; \sigma, x) = P_0(t, y, A; \sigma, x) - \bar{P}(A; x), \quad t \geqq 0.$$

We assume that there is a monotonically decreasing function $\rho(t)$, $t \geqq 0$, such that for any function $f(x, y)$ with $|f(x, y)| \leqq C(1 + |x|^\alpha)$, α some nonnegative integer,

$$(6.8) \qquad \left| \int \mu_0(t, y, d\zeta; \sigma, x)f(x, \zeta) \right| \leqq \rho(t)C(1 + |x|^\alpha),$$

and

$$(6.9) \qquad \int_0^\infty \rho^{1/5}(s)\, ds < \infty.$$

From (6.6) it follows that $\bar{P}(A; \sigma, x)$ is an invariant measure for Y_0 that is,

$$(6.10) \qquad \int \bar{P}(d\zeta; \sigma, x)P_0(t, \zeta, A; \sigma, x) = \bar{P}(A; \sigma, x), \quad t \geqq 0,$$

and from (6.8) it follows that it is unique and P_0 tends to it sufficiently rapidly as $t \uparrow \infty$. For each $\sigma \geqq 0$ and $x \in R^n$ the condition (6.8) can be translated into a condition on the spectrum of the collision operator $Q_{\sigma, x}$. From (6.6) it follows that zero belongs to the spectrum.

In §3 we described in rough terms the centering condition that the vector field F must satisfy. We now state this condition precisely. We assume that

$$(6.11) \qquad \left| \int_{t_0}^{t_0+t} \int F(\tau, s, x, y)\bar{P}(dy; \tau, x)\, ds \right| \leqq \frac{C}{1 + t^4}(1 + |x|), \quad t, t_0 \geqq 0.$$

This is the centering condition. When $F(\tau, s, x, y)$ is independent of s,

(6.11) is replaced by

$$\int F(\tau, x, y)\bar{P}(dy; x) = 0.$$

This form justifies the terminology we use. Along with (6.11), we need also to assume that

(6.12)
$$\left| \frac{\partial}{\partial x_i} \int_{t_0}^{t_0+t} \int F(\tau, s, x, y)\bar{P}(dy; \tau, x)\, ds \right| \leqq C,$$

$$t, t_0 \geqq 0, \quad i = 0, 1, \cdots, n.$$

This hypothesis is compatible with the properties of $\bar{P}$ that follow from (6.3) and (6.6) as the results of §10 show.

The theorem to be stated shortly tells us that $X^{(\varepsilon)}(\tau, \sigma, x, y)$ converges to a diffusion Markov process as $\varepsilon \to 0$, $0 \leqq \sigma \leqq \tau \leqq T$. We define now the drift vector $(b^j(\tau, x))$ and the diffusion matrix $(a^{ij}(\tau, x))$, $i, j = 1, \cdots, n$, of the limiting diffusion process.

(6.13)
$$a^{ij}(\tau, x) = \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} \int \int_{t_0}^{t_0+t} \int \bar{P}(dy; \tau, x)\mu_0(|s - \sigma|, y, dz; \tau, x)$$
$$\times F^i(\tau, \sigma, x, y)F^j(\tau, s, x, z)\, d\sigma\, ds.$$

(6.14)
$$b^j(\tau, x) = \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} \int \int \int_{t_0}^{s} \int \bar{P}(dy; \tau, x)\sum_{i=1}^{n} F^i(\tau, \sigma, x, y)$$
$$\times (\partial/\partial x_i)[\mu_0(s - \sigma, y, dz; \tau, x)F^j(\tau, s, x, z)]\, d\sigma\, ds$$
$$+ \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} \int \bar{P}(dy; \tau, x)G^j(\tau, s, x, y)\, ds,$$

$$0 \leqq \tau \leqq T, \quad t_0 \geqq 0, \quad x \in R^n, \quad i, j = 1, 2, \cdots, n.$$

We require that the limits (6.13) and (6.14) be approached sufficiently rapidly and uniformly in x and t_0. The precise manner of approach is specified by the following inequalities.

(6.15)
$$\left| \frac{1}{t} \int_{t_0}^{t_0+t} \int \int_{t_0}^{t_0+t} \int \bar{P}(dy; \tau, x)\mu_0(|s - \sigma|, y, dz; \tau, x) \right.$$
$$\left. \times F^i(\tau, \sigma, x, y)F^j(\tau, s, x, z)\, d\sigma\, ds - a^{ij}(\tau, x) \right| \leqq \frac{C(1 + |x|^2)}{1 + t},$$

(6.16)
$$\left| \frac{1}{t} \int_{t_0}^{t_0+t} \int \int \int_{t_0}^{s} \int \bar{P}(dy; \tau, x)\sum_{i=1}^{n} F^i(\tau, \sigma, x, y) \right.$$
$$\times \frac{\partial}{\partial x_i}[\mu_0(s - \sigma, y, dz; \tau, x)F^j(\tau, s, x, z)]\, d\sigma\, ds$$
$$\left. + \frac{1}{t} \int_{t_0}^{t_0+t} \int \bar{P}(dy; \tau, x)G^j(\tau, s, x, y)\, ds - b^j(\tau, x) \right| \leqq \frac{C(1 + |x|)}{1 + t}.$$

When $F(\tau, s, x, y)$ and $G(\tau, s, x, y)$ are independent of s then (6.13) and (6.14) simplify to

$$(6.17) \qquad a^{ij}(\tau, x) = 2 \iint \bar{P}(dy; \tau, x) \int_0^\infty \mu_0(s, y, dz; \tau, x) \, ds$$
$$\times F^i(\tau, x, y) F^j(\tau, x, z),$$

$$b^j(\tau, x) = \iint \bar{P}(dy; \tau, x) \sum_{i=1}^n F^i(\tau, x, y) \frac{\partial}{\partial x_i}$$
$$(6.18) \qquad \times \left[\int_0^\infty \mu_0(s, y, dz; \tau, x) \, ds F^j(\tau, x, z) \right]$$
$$+ \int \bar{P}(dy; \tau, x) G^j(\tau, x, y).$$

Note that the s integral in (6.17) and (6.18) is convergent in view of (6.8) and (6.9).

Let $C^{k,\alpha}(R^n)$, with k and α nonnegative integers, denote the collection of functions $f(x)$ on R^n with continuous partial derivatives with respect to x up to order k inclusive, such that there is a nonnegative integer α and

$$(6.19) \qquad \left| \frac{\partial^k f(x, y)}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_n^{k_n}} \right| \leqq C(1 + |x|^\alpha),$$

k_i nonnegative integers, $k_1 + \cdots + k_n = k$. On $C^{2,\alpha}(R^n)$ we define the second order, possibly degenerate, elliptic partial differential operator $\mathscr{L}_\sigma$ by

$$(6.20) \qquad \mathscr{L}_\sigma g(x) = \frac{1}{2} \sum_{i,j=1}^n a^{ij}(\sigma, x) \frac{\partial^2 g(x)}{\partial x_i \partial x_j} + \sum_{j=1}^n b^j(\sigma, x) \frac{\partial g(x)}{\partial x_j},$$
$$\sigma \geqq 0, \quad x \in R^n.$$

From (6.13) it follows that $(a^{ij}(\sigma, x))$ is nonnegative definite. We do not assume that it is positive definite. $\mathscr{L}_\sigma$ is referred to as the infinitesimal generator of the limiting diffusion Markov process.

As in [21] we require that the final value problem

$$(6.21) \qquad \frac{\partial u(\sigma, \tau, x)}{\partial \sigma} + \mathscr{L}_\sigma u(\sigma, \tau, x) = 0, \quad \sigma < \tau, \quad u(\tau, \tau, x) = f(x) \in C^{2,\alpha},$$

has a unique solution $u(\sigma, \tau, x)$ which is sufficiently regular when f is sufficiently regular. In this respect the method of analysis of (6.1) that we follow here is similar to the one Khinchine employed in [5] in establishing diffusion approximations. The results we need about (6.21) are summarized in Lemma 13.1. With hypotheses (6.22) stated below and Itô's calculus [12], [29], these results can be easily verified. However, the assumption that $a(\tau, x)$ has a smooth square root is sometimes difficult to verify from

the assumptions on F, G, π and q. In that case one must employ the theory of Oleinik [27], [28] where (6.22), (i) is not required and the remaining hypotheses below are suitably modified.

$$\text{(i)} \quad a^{ij}(\tau, x) = \sum_{k=1}^{n} c^{ik}(\tau, x)c^{jk}(\tau, x),$$

$$\text{(ii)} \quad |c^{ij}(\tau, x)| \leq C(1 + |x|), \quad |b^{j}(\tau, x)| \leq C(1 + |x|),$$

(6.22) (iii) $\quad \left| \dfrac{\partial c^{ij}}{\partial x_k}(\tau, x) \right| \leq C, \quad \left| \dfrac{\partial b^{j}(\tau, x)}{\partial x_k} \right| \leq C,$

$$i, j = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots, n,$$

(iv) $\quad c^{ij}(\tau, x) \in C^{4,\alpha}(R^{n+1}), \quad b^{j}(\tau, x) \in C^{4,\alpha}(R^{n+1}),$

$$i, j = 1, 2, \cdots, n, \quad \alpha \geq 0 \text{ integer.}$$

Note the similarity of (6.22) with (6.2).

We state now the theorem characterizing the limiting behavior of $X^{(\varepsilon)}$ as $\varepsilon \downarrow 0$.

THEOREM. *Let $(X^{(\varepsilon)}(\tau, \sigma, x, y), Y^{(\varepsilon)}(\tau, \varepsilon, x, y))$ be the Markov process on $R^n \times S$ with backward equation (6.1). Then $X^{(\varepsilon)}$ converges weakly as $\varepsilon \to 0$, $0 \leq \sigma \leq \tau \leq T$, to a diffusion Markov process on R^n with generator $\mathscr{L}_\sigma$ given by (6.20).*

Furthermore, let $f(x, y)$ be bounded measurable and in $C^{4,\alpha}(R^n)$, $\alpha \geq 0$, as a function of x and let $u(\sigma, \tau, x)$ denote the solution of

(6.23)
$$\partial u(\sigma, \tau, x)/\partial \sigma + \mathscr{L}_\sigma u(\sigma, \tau, x) = 0, \quad 0 \leq \sigma < \tau \leq T,$$

$$u(\tau, \tau, x) = \bar{f}(x; \tau) = \int f(x, y)\bar{P}(dy; \tau, x).$$

Then, there exists an $\bar{\alpha} > \alpha$ such that for $0 \leq \sigma < \tau \leq T$

(6.24)
$$\lim_{\varepsilon \downarrow 0} \sup_{x, y} \frac{|u^{(\varepsilon)}(\sigma, \tau, x, y) - u(\sigma, \tau, x)|}{1 + |x|^{\bar{\alpha}}} = 0.$$

Let $\bar{u}^{(\varepsilon)}(\sigma, \tau, x, y)$ denote the solution of (6.1) with $f = \bar{f}(x; \tau)$. Then, for $0 \leq \sigma \leq \tau \leq T$ we have

(6.25)
$$|\bar{u}^{(\varepsilon)}(\sigma, \tau, x, y) - u(\sigma, \tau, x)| \leq \varepsilon^{1/4}C(1 + |x|^{\bar{\alpha}}).$$

Here C is a constant that depends on T, f and other quantities but does not depend on ε. Note that we have $\sigma < \tau$ in order for (6.24) to hold. This nonuniformity constitutes an initial layer for the singular perturbation problem (6.1).

Suppose that F in (6.1) is identically zero. Then according to the above theorem $X^{(\varepsilon)}(\tau, \sigma, x, y)$ converges to the solution $X^{(0)}(\tau, \sigma, x)$ of the

deterministic system of differential equations

$$(6.26) \qquad \frac{dX^{(0)}(\tau)}{d\tau} = \bar{G}(\tau, X^{(0)}(\tau)), \quad \tau > \sigma, \qquad X^{(0)}(\sigma, \sigma, x) = x,$$

where, from (6.18), we have

$$(6.27) \qquad \bar{G}(\tau, x) = \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} \int P(dy; \tau, x) G(\tau, s, x, y) \, ds.$$

Let $Z^{(\varepsilon)}(\tau, \sigma, x^0, z, y)$ be the process defined by

$$(6.28) \qquad Z^{(\varepsilon)}(\tau, \sigma, x^0, z, y) = \frac{1}{\varepsilon} [X^{(\varepsilon)}(\tau, \sigma, x, y) - X^{(0)}(\tau, \sigma, x^0)], \quad \tau > \sigma,$$

$$Z^{(\varepsilon)}(\sigma, \sigma, x^0, z, y) = z,$$

with (x^0, x, z) satisfying the relation

$$(6.29) \qquad\qquad\qquad x = x^0 + \varepsilon z.$$

The process $Z^{(\varepsilon)}$ is called the fluctuation process of $X^{(\varepsilon)}$ about $X^{(0)}$. We shall apply the above theorem to determine the asymptotic behavior of this fluctuation process.

Consider the joint process

$$(X^{(0)}(\tau, \sigma, x^0), Z^{(\varepsilon)}(\tau, \sigma, x^0, z, y), Y^{(\varepsilon)}(\tau, \sigma, x^0 + \varepsilon z, y))$$

which is a transport process. Evidently $X^{(0)}$ is deterministic and independent of ε but it is carried along in our formulation so as to allow us to obtain a convenient form for the generator of the limit of $Z^{(\varepsilon)}$.[4] The backward equation for the joint process follows easily from (6.1), (6.28) and (6.29), and is

$$\frac{\partial \tilde{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, y)}{\partial \sigma} + \bar{G}(\sigma, x^0) \frac{\partial \tilde{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, y)}{\partial x^0}$$

$$+ \frac{1}{\varepsilon} \left[G\left(\sigma, \frac{\sigma}{\varepsilon^2}, x^0 + \varepsilon z, y\right) - \bar{G}(\sigma, x^0) \right] \frac{\partial \tilde{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, y)}{\partial z}$$

$$(6.30) \qquad + \frac{q(\sigma, x^0 + \varepsilon z, y)}{\varepsilon^2} \int \tilde{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, \zeta) \pi(\sigma, x^0 + \varepsilon z, y, d\zeta)$$

$$- \frac{q(\sigma, x^0 + \varepsilon z, y)}{\varepsilon^2} \tilde{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, y) = 0, \qquad \sigma < \tau,$$

$$\tilde{u}^{(\varepsilon)}(\tau, \tau, x^0, z, y) = f(x^0, z, y),$$

$$\tilde{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, y) = E\{f(X^{(0)}(\tau, \sigma, x^0), \ Z^{(\varepsilon)}(\tau, \sigma, x^0, y, z),$$

$$Y^{(\varepsilon)}(\tau, \sigma, x^0 + \varepsilon z, y))\}.$$

---

[4] This device has also been employed by McKean in [30].

Equation (6.30) is almost of the same form as (6.1). To apply the theorem we need the analogs of (6.11), (6.13) and (6.14) for the present problem. We state now these hypotheses.

We assume that

$$(6.31) \quad \left| \int_{t_0}^{t_0+t} \int [G(\tau, s, x^0, y) - \bar{G}(\tau, x^0)] \bar{P}(dy; \tau, x^0) \, ds \right| \leqq \frac{C(1 + |x^0|)}{1 + t^4},$$

$$
\begin{aligned}
(6.32) \quad a^{ij}(\tau, x^0) = \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} &\int \int_{t_0}^{t_0+t} \int \bar{P}(dy; \tau, x^0) \mu_0(|s - \sigma|, y, d\zeta; \tau, x^0) \\
&\times [G^i(\tau, \sigma, x^0, y) - \bar{G}^i(\tau, x^0)] \\
&\times [G^j(\tau, s, x^0, \zeta) - \bar{G}^j(\tau, x^0)] \, d\sigma \, ds,
\end{aligned}
$$

$$(6.33) \quad b^{ij}(\tau, x^0) = \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} \int \bar{P}(dy; \tau, x^0) \frac{\partial}{\partial x_j^0} G^i(\tau, s, x^0, y) \, ds,$$

$$i, j = 1, \cdots, n.$$

As with (6.13) and (6.14) the limits (6.32) and (6.33) are assumed to satisfy the analogs of (6.15), (6.16) as well as the analog of (6.12). Define $\tilde{\mathscr{L}}_\sigma$ on $C^{2,\alpha}(R^n \times R^n)$ as follows:

$$
\begin{aligned}
(6.34) \quad \tilde{\mathscr{L}}_\sigma g(x^0, z) = &\frac{1}{2} \sum_{i,j=1}^n a^{ij}(\sigma, x^0) \frac{\partial^2 g(x^0, z)}{\partial z_i \, \partial z_j} \\
&+ \sum_{i,j=1}^n b^{ij}(\sigma, x^0) z_j \frac{\partial g(x^0, z)}{\partial z_i} + \sum_{i=1}^n \bar{G}^i(\sigma, x^0) \frac{\partial g(x^0, z)}{\partial x_i^0},
\end{aligned}
$$

$$\sigma \geqq 0, \quad x^0 \in R^n, \quad z \in R^n.$$

We have the following result.

COROLLARY 1. *Let* $(X^{(0)}(\tau, \sigma, x^0), Z^{(\varepsilon)}(\tau, \sigma, x^0, z, y))$ *be as above. Then, as* $\varepsilon \to 0$, $0 \leqq \sigma \leqq \tau \leqq T$, $(X^{(0)}, Z^{(\varepsilon)})$ *converges weakly to a Gauss-Markov process with generator* $\tilde{\mathscr{L}}_\sigma$ *given by* (6.34).

Furthermore, let $f(x^0, z, y)$ *be bounded measurable and in* $C^{4,\alpha}(R^n \times R^n)$ *as a function of* x^0 *and* z. *Let* $\tilde{u}(\sigma, \tau, x^0, z)$ *denote the solution of*

$$
\begin{aligned}
(6.35) \quad &\frac{\partial \tilde{u}(\sigma, \tau, x^0, z)}{\partial \sigma} + \tilde{\mathscr{L}}_\sigma \tilde{u}(\sigma, \tau, x^0, z) = 0, \quad 0 \leqq \sigma < \tau \leqq T, \\
&\tilde{u}(\tau, \tau, x^0, z) = f(x^0, z, \tau) = \int f(x^0, z, y) \bar{P}(dy; \tau, x^0).
\end{aligned}
$$

Then there exists an integer $\bar{\alpha} > \alpha$ *such that for* $0 \leqq \sigma < \tau \leqq T$,

$$(6.36) \quad \lim_{\varepsilon \downarrow 0} \sup_{x^0, y, z} \frac{|\tilde{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, y) - \tilde{u}(\sigma, \tau, x^0, z)|}{1 + |x^0|^{\bar{\alpha}} + |z|^{\bar{\alpha}}} = 0.$$

Let $\bar{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, y)$ *denote the solution of* (6.30) *with* $f = \bar{f}(x^0, z, \tau)$
Then there is an integer $\bar{\alpha} > \alpha$ *such that for* $0 \leq \sigma \leq \tau \leq T$,

$$(6.37) \quad |\bar{u}^{(\varepsilon)}(\sigma, \tau, x^0, z, y) - \tilde{u}(\sigma, \tau, x^0, z)| \leq \varepsilon^{1/4} C(1 + |x^0|^{\bar{\alpha}} + |z|^{\bar{\alpha}}).$$

Let $A(\sigma, x)$ and $g(\sigma, x)$ be bounded uniformly continuous functions and
consider the transport equation

$$
\begin{aligned}
(6.38) \quad & \frac{\partial w^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial \sigma} + \mathcal{L}_\sigma^{(\varepsilon)} w^{(\varepsilon)}(\sigma, \tau, x, y) + A(\sigma, x) w^{(\varepsilon)}(\sigma, \tau, x, y) \\
& \qquad\qquad = g(\sigma, x), \qquad \sigma < \tau, \\
& \quad w^{(\varepsilon)}(\tau, \tau, x, y) = f(x, y).
\end{aligned}
$$

Let $(X^{(\varepsilon)}(\tau, \sigma, x, y), Y^{(\varepsilon)}(\tau, \sigma, x, y))$ denote as usual the transport proc-
ess with backward equation (6.1). Then the solution of (6.38) can be
written in the form

$$
\begin{aligned}
(6.39) \quad w^{(\varepsilon)}(\sigma, \tau, x, y) = {} & E\left\{ \exp\left(\int_\sigma^\tau A(s, X^{(\varepsilon)}(s)) \, ds \right) f(X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau)) \right\} \\
& - E\left\{ \int_\sigma^\tau g(s, X^{(\varepsilon)}(s)) \exp\left(\int_\sigma^s A(\gamma, X^{(\varepsilon)}(\gamma)) \, d\gamma \right) ds \right\}.
\end{aligned}
$$

Thus, $w^{(\varepsilon)}$ is a bounded and continuous functional of $X^{(\varepsilon)}$ if we substitute
in (6.39) the average $\bar{f}(x, \tau)$ in place of $f(x, y)$. In fact, for the representa-
tion (6.39) to hold, A and g can also depend on y. Independence of y is
required for the following result.

COROLLARY 2. *For* $0 \leq \sigma \leq \tau \leq T$ *the solution* $w^{(\varepsilon)}(\sigma, \tau, x, y)$ *of* (6.38)
with $f = \bar{f}$ *converges as* $\varepsilon \to 0$ *to* $w^{(0)}(\sigma, \tau, x)$, *uniformly on compact sets, where*

$$
(6.40) \quad \frac{\partial w^{(0)}(\sigma, \tau, x)}{\partial \sigma} + \mathcal{L}_\sigma w^{(0)}(\sigma, \tau, x) + A(\sigma, x) w^{(0)}(\sigma, \tau, x) = g(\sigma, x).
$$

$$\sigma < \tau, \qquad w^{(0)}(\tau, \tau, x) = \bar{f}(x, \tau).$$

This corollary is an immediate consequence of the weak convergence
of $X^{(\varepsilon)}$ obtained in the theorem above.

Let $\mathscr{D} \subset R^n$ be a bounded open set with smooth boundary and assume
that $f(x)$ satisfies the condition that

$$(6.41) \qquad\qquad \{x \in R^n : f(x) \neq 0\}$$

is contained in a compact subset of $\mathscr{D}$. Consider the following boundary-value problem:[5]

$$
\begin{aligned}
&\text{(i)} \quad \text{(6.1) holds for } x \text{ in the interior of } \mathscr{D}, \; y \in S, \; \sigma < \tau. \\
&\text{(ii)} \quad u^{(\varepsilon)}(\tau, \tau, x, y) = f(x). \\
(6.42) \quad &\phantom{\text{(ii)}} \quad u^{(\varepsilon)}(\sigma, \tau, x, y) = 0, \quad \sigma \leqq \tau, \quad x \in \partial\mathscr{D}, \quad y \in S, \\
&\text{(iii)} \quad \left(\frac{1}{\varepsilon} F\left(\sigma, \frac{\sigma}{\varepsilon^2}, x, y \right) + G\left(\sigma, \frac{\sigma}{\varepsilon^2}, x, y \right) \right) \cdot n(x) > 0.
\end{aligned}
$$

Here, as in §2, $\partial\mathscr{D}$ denotes the boundary of $\mathscr{D}$ and $n(x)$ the unit outer normal to $\partial\mathscr{D}$ at x. Let $\tau_{\mathscr{D}}^{(\varepsilon)}$ denote the first time $X^{(\varepsilon)}$ reaches $\partial\mathscr{D}$ starting from (x, y) at time σ. From (2.19) it follows that

$$(6.43) \quad u^{(\varepsilon)}(\sigma, \tau, x, y) = E\{f(X^{(\varepsilon)}(\tau, \sigma, x, y)), \tau \leqq \tau_{\mathscr{D}}^{(\varepsilon)}(\sigma, x, y)\}.$$

As we pointed out in §2, we take (6.43) as defining the (generalized) sense in which (6.42) is understood here. In particular, boundary values are assumed only where this is compatible with (6.43). We assume also that the limiting Markov process is such that the exit time $\tau_{\mathscr{D}}^0$ is a.s. a continuous functional of the path.

Note that this boundary-value problem is simple enough so that boundary layers do not arise.

From (6.43) and the theorem of this section we obtain the following.

COROLLARY 3. *Let $u^{(\varepsilon)}(\sigma, \tau, x, y)$ be defined by (6.43). Then as $\varepsilon \downarrow 0$, $0 \leqq \sigma \leqq \tau \leqq T$, $u^{(\varepsilon)}$ converges, uniformly on compact subsets, to $u^{(0)}(\sigma, \tau, x)$ which satisfies (6.23) with $\hat{f}(x, \tau) \equiv f(x)$ and is zero for $x \in \partial\mathscr{D}$.*

Here, $u^{(0)}$ assumes the boundary values in the generalized sense again i.e., at those points for which the representation

$$u^{(0)}(\sigma, \tau, x) = E\{f(X^{(0)}(\tau, \sigma, x)), \tau \leqq \tau_{\mathscr{D}}^{(0)}(\sigma, x)\}$$

permits it. We note that the exit time $\tau_{\mathscr{D}}^{(\varepsilon)}$ is not a continuous functional of the trajectories (X, Y) and therefore the theorem does not apply directly to (6.43). However, the set of trajectories for which $\tau_{\mathscr{D}}^{(0)}$ is discontinuous has zero measure and so the theorem is applied after a small detour.

Corollaries 2 and 3 deal, essentially, with the first two terms on the right side of (3.5). As we mentioned in §3 the third term requires additional consideration.

7. Remarks on the theorem and corollaries. We collect in this section some remarks concerning the results stated in the preceding section. Remark 1 points out the features of special interest in the results as well

---

[5] Corrected from an earlier version after a remark by R. Hersh.

as some deficiencies. In remark 2 we give references to previous work and in remark 3 we examine some related problems and give references. Remark 4 concerns Corollary 1.

1. Briefly, the interesting features of the theorem of §6 are: (i) we show weak convergence of $X^{(\varepsilon)}$ to a diffusion Markov process, not just convergence at one time point, (ii) from (6.25) it follows that moments of all orders of $X^{(\varepsilon)}$ tend to the corresponding moments of the limiting diffusion, (iii) (6.1) allows for fast-time, σ/ε^2, dependence and we therefore obtain, simultaneously with the diffusion approximation, an averaging approximation [31], (iv) from (6.1) it follows that the centering need only hold approximately, which allows us to obtain Corollary 1 from the theorem with no additional effort. As we have remarked already, another important feature is that F, G, π and q are allowed to depend on x, y, and σ.

The principal defect of the results is the requirement that S, the state space of $Y^{(\varepsilon)}$ be compact. This requirement simplifies the mathematical treatment but it is not essential and can be removed. For example, it is not required in Kurtz' set-up [11]. We feel that Grad's approach [32] can be used effectively to treat (6.1) when $S=R^m$ and q and π satisfy appropriate hypotheses. In [32], $S=R^3$, q and π are independent of σ and x, $H \equiv 0$, $F \equiv 0$, $G \equiv y$ and the null-space of $Q_{\sigma,x}$ is five-dimensional, not one-dimensional as in (6.1). However, in this case (6.1) is not positivity preserving and hence not the backward equation of a transport process.

When S is a compact differentiable manifold without boundary then, with minor modifications, the results of §6 can be extended to include force fields H. It is actually of interest to recast (6.1) and the theorem of §6 into the appropriate form when $(X^{(\varepsilon)}, Y^{(\varepsilon)})$ takes values in $S_X \times S_Y$ where S_X is an n-dimensional differentiable, that is, a C^∞ manifold and S_Y is a compact differentiable manifold without boundary of dimension m.

Let $F(\tau, t, y)$ and $G(\tau, t, y)$, $\tau \geq 0$, $t \geq 0$, $y \in S_Y$ be two vector fields on S_X which are smooth functions of τ, t and y. Let $H(\tau, t, x)$, $\tau \geq 0$, $t \geq 0$, $x \in S_X$ be a vector field on S_Y which is a smooth function of τ, t, and x. We shall not denote explicitly the dependence of F and G on x and H on y. Let $(X^{(\varepsilon)}, Y^{(\varepsilon)})$ be the Markov process on $S_X \times S_Y$ with backward equation

$$
\frac{\partial u^{(\varepsilon)}(\sigma, \tau)}{\partial \sigma} + \frac{1}{\varepsilon} F\left(\sigma, \frac{\sigma}{\varepsilon^2}\right) u^{(\varepsilon)}(\sigma, \tau) + G\left(\sigma, \frac{\sigma}{\varepsilon^2}\right) u^{(\varepsilon)}(\sigma, \tau)
$$

(7.1)
$$
+ \frac{1}{\varepsilon} H\left(\sigma, \frac{\sigma}{\varepsilon^2}\right) u^{(\varepsilon)}(\sigma, \tau) + \frac{1}{\varepsilon^2} Q_\sigma u^{(\varepsilon)}(\sigma, \tau) = 0, \quad \sigma < \tau,
$$

$$
u^{(\varepsilon)}(\sigma, \tau) = f \in C^\infty(S_X \times S_Y).
$$

The frozen process $Y_0(t, y; \sigma, x)$ and P_0 of (6.4) are defined as before. In place of (6.11) we assume

$$
(7.2) \quad \sup_{x,y} \left| \int_{t_0}^{t_0+t} \int F(\tau, s) f(x, y) \bar{P}(dy; \tau, x) \, ds \right| \leqq \frac{C}{1 + t^4}, \quad t \uparrow \infty,
$$

$$
t_0 \geqq 0, \quad f \in C^\infty.
$$

Here C denotes a constant that depends on f. We also assume that the analog of (6.12) holds.

Define $\mathscr{L}_\tau$ on $C^\infty(S_X)$ by the following limit.

$$
\begin{aligned}
(7.3) \quad \mathscr{L}_\tau f(x) = {} & \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} \int \int_{t_0}^{s} \int \bar{P}(dy; \tau, x) F(\tau, \sigma, y) \\
& \qquad\qquad \times \mu_0(s - \sigma, y, dz; \tau, \cdot) F(\tau, s, z) f \cdot (x, y) \, d\sigma \, ds \\
& + \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} \int \bar{P}(dy; \sigma, x) G(\tau, s, y) f \cdot (x, y) \, ds \\
& + \lim_{t \uparrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} \int \int_{t_0}^{s} \int \bar{P}(dy; \sigma, x) H(\tau, \sigma, x) \mu_0(s - \sigma, \cdot, dz; \sigma, x) \\
& \qquad\qquad \times F(\tau, s, \cdot) f \cdot (x, y) \, d\sigma \, ds.
\end{aligned}
$$

Here the limit is uniform on $x \in S_X$, $\tau \in [0, T]$ and $t_0 \geqq 0$. Note that the first two terms in (7.3) correspond to the operator $\mathscr{L}_\tau$ of §6 and the last term is due to the force field. The theorem of §6 reads now the same way here provided sufficient smoothness is assumed on the y-dependence of the various quantities.

2. The early work by physicists on problems of interest here is contained in the paper of Uhlenbeck and Ornstein [33] (with references to earlier work of Smoluchovski and Fürth), Kramers [6], Chandrasekhar [10] and Wang Chang and Uhlenbeck [34]. The work of Kramers is closest to the path we follow here. In all the above works the collision operator Q is almost immediately replaced by a second order differential operator in y. This corresponds to the assumption that the mass of the moving particle is much larger than the mass of the particles in the medium through which it is moving. The mathematical elucidation of this approximation is given in the first few sections of the paper of Il'in and Khasminskii [17]. The approximations of interest to us here are all carried out after this preliminary approximation in the above works. Emphasis is placed on what is termed the Smoluchovski or large viscosity approximation. This corresponds to our diffusion limit when in (3.1), $X^{(\varepsilon)} \in R^1$, $Y^{(\varepsilon)} \in R^1$, $F = y$, $G = 0$, $H = H(x)$ and Q is independent of σ and x and not an integral operator but a second order differential operator i.e., the backward

equation is:

$$(7.4) \quad \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial \sigma} + \frac{1}{\varepsilon} y \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial x} + \frac{1}{\varepsilon} H(x) \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial y}$$

$$+ \frac{1}{\varepsilon^2} \left(\frac{1}{2} \frac{\partial^2 u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial y^2} - y \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial y} \right) = 0.$$

Let $\beta(\tau)$ denote the standard Brownian motion process on R^1. Then $(X^{(\varepsilon)}, Y^{(\varepsilon)})$ satisfy the Itô equations

$$dX^{(\varepsilon)}(\tau) = \frac{1}{\varepsilon} Y^{(\varepsilon)}(\tau) \, d\tau,$$

$$(7.5) \quad dY^{(\varepsilon)}(\tau) = \frac{1}{\varepsilon} H(X^{(\varepsilon)}(\tau)) \, d\tau - \frac{1}{\varepsilon^2} Y^{(\varepsilon)}(\tau) \, d\tau + \frac{1}{\varepsilon} d\beta(\tau), \qquad \tau > \sigma,$$

$$X^{(\varepsilon)}(\sigma) = x, \qquad Y^{(\varepsilon)}(\sigma) = y.$$

The usual existence and uniqueness theory guarantees solutions for (7.5) and (7.4) under mild restrictions on $H(x)$. Equation (7.5) can be written formally as a second order equation for $X^{(\varepsilon)}(\tau)$,

$$(7.6) \qquad \varepsilon^2 \frac{d^2 X^{(\varepsilon)}(\tau)}{d\tau^2} + \frac{dX^{(\varepsilon)}(\tau)}{d\tau} = H(X^{(\varepsilon)}(\tau)) + \frac{d\beta(\tau)}{d\tau}$$

and the limit $\varepsilon \to 0$ should formally lead to

$$\frac{dX^{(0)}(\tau)}{d\tau} = H(X^{(0)}(\tau)) + \frac{d\beta(\tau)}{d\tau},$$

which means that $X^{(\varepsilon)}$ should tend to a diffusion Markov process with generator

$$(7.7) \qquad \frac{1}{2} \frac{\partial^2}{\partial x^2} + H(x) \frac{\partial}{\partial x}.$$

This is in fact the Smoluchovski approximation. Note however that this result is not covered by our analysis because Q is not of the form we have assumed and $S = R^1$ is not compact. The mathematical analysis of the limit $\varepsilon \to 0$ in (7.4) is given by Il'in and Khasminskii in [17]. An elegantly simple analysis of the limit $\varepsilon \to 0$ in (7.5) is given by Nelson [18] who shows, in fact, that the Smoluchovski approximation holds with probability one. What Nelson terms pseudotheorem 10.2 [18, p. 77] is actually the theorem of Il'in and Khasminskii [17].

No doubt one does not have to assume that Q is a second order differential operator for the Smoluchovski limit to hold. All that is required is that the collision operator have an invariant measure that is approached sufficiently rapidly and the velocity be centered or approximately centered. With the obvious interpretation of the various quantities, one can verify

that, indeed, our formula (7.3) yields the correct result (7.7), although, at present, it is known to hold only when S_Y is compact.

Recent developments in the asymptotic analysis of problems of the same general form as (6.1) where Q, or its equivalent, is x-dependent, i.e. with feedback, are contained in the works of Kurtz [11] and Khasminskii [22]. Kurtz has a general operator framework but does not allow for fast-time dependence and does not show weak convergence. The weak convergence is probably not difficult to obtain by his methods but it is of interest to also obtain the averaging of the fast time.

In [22] Khasminskii considers a problem that translates to our notation as follows. Let $\beta(\tau)$ denote the r-dimensional standard Brownian motion and let $\phi^{(k)}(x, y)$ and $\psi^{(k)}(x, y)$ be n and m-vector functions of $x \in R^n$, $y \in R^m$ respectively, $k = 1, 2, \cdots, r$. Let $(X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))$ be the R^{n+m}-dimensional diffusion process defined by the Itô equations

$$dX^{(\varepsilon)}(\tau) = G(X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))\, d\tau$$

$$+ \sum_{k=1}^{r} \phi^{(k)}(X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))\, d\beta_k(\tau),$$

(7.8) $$dY^{(\varepsilon)}(\tau) = (1/\varepsilon^2)K(X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))\, d\tau$$

$$+ \frac{1}{\varepsilon} \sum_{k=1}^{r} \psi^{(k)}(X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))\, d\beta_k(\tau), \qquad \tau > 0,$$

$$X^{(\varepsilon)}(0) = x, \quad Y^{(\varepsilon)}(0) = y.$$

The vector functions G, K, $\phi^{(k)}$ and $\psi^{(k)}$ satisfy mild regularity conditions. This process $(X^{(\varepsilon)}, Y^{(\varepsilon)})$ is different from the processes considered here in several respects. To make the comparison somewhat more explicit we introduce a system of Itô equations that corresponds to (3.1), (3.2):

$$dX^{(\varepsilon)}(\tau) = \frac{1}{\varepsilon} F\left(\tau, \frac{\tau}{\varepsilon^2}, X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau)\right) d\tau$$

$$+ G\left(\tau, \frac{\tau}{\varepsilon^2}, X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau)\right) d\tau,$$

(7.9) $$dY^{(\varepsilon)}(\tau) = (1/\varepsilon^2)K(\tau, X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))\, d\tau$$

$$+ \frac{1}{\varepsilon} \sum_{k=1}^{r} \psi^{(r)}(\tau, X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))\, d\beta_k(\tau), \qquad \tau > \sigma,$$

$$X^{(\varepsilon)}(\sigma, \sigma, x, y) = x, \quad Y^{(\varepsilon)}(\sigma, \sigma, x, y) = y.$$

Clearly, instead of a jump process $Y^{(\varepsilon)}$ we have a diffusion here and the feedback is perhaps more clearly visible. Note that the $X^{(\varepsilon)}$ equation in (7.8) has Brownian motion terms but no F terms whereas the opposite is true in (7.9). This is the essential difference between the set-up here and Khasminskii's problem (7.8); the fact that $Y^{(\varepsilon)}$ is a jump process or a

diffusion is not so important. In particular, because (7.8) has no F term, no centering problems arise and no assumptions about the existence of an invariant measure for the "frozen" Y process need be introduced. Indeed, no such hypotheses are introduced in [22] where the result is that $X^{(\varepsilon)}$ converges weakly as $\varepsilon \downarrow 0$ to a diffusion Markov process whose generator is obtained in terms of averages of G and $\phi^{(k)}$ with respect to the distribution of the frozen process (which is defined by equation (1.5) in [22]). It should be added that Khasminskii motivates (7.8) as a generalization to a problem studied previously [25] and [12, p. 335 where references to earlier work of Gihman are given] whereby the $Y^{(\varepsilon)}$ equation in (7.8) is simply

$$(7.10) \qquad dY^{(\varepsilon)}(\tau) = (1/\varepsilon^2)\, d\tau.$$

The simpler problem corresponds to the method of averaging [31] for Itô equations whereas the more general problem (7.8) corresponds to Volosov's generalization of the method of averaging [36] for the Itô equations.

An interesting problem is the analysis of the combined version of (7.8) and (7.9) in the limit $\varepsilon \downarrow 0$ for jump processes or for Itô equations.

3. We return now to problem (6.1). When Q does not depend on σ and x we have the random evolution problem; we refer to [9] and [10] for additional information. When $F=0$, $G=y$ and Q is independent of σ and x the problem can be dealt with directly by Fourier transforms. Ellis and Pinsky [37 and references therein] have studied this problem when Q has a multidimensional null-space. This is a linearized version of the hydrodynamical limit for the Boltzmann equation [38 and references therein] that leads to Euler's equations. The problem $F=y$, $G=0$ has also been considered and the connection of this and related problems with the Hilbert and Chapman-Enskog expansions has been studied in [37].

From the point of view of linear transport theory the basic problem is the adjoint of (6.1) with $F=y$, $G=0$, Q independent of σ and x, $y \in S^2$, the unit sphere in R^3, and the boundary conditions (3.3). Asymptotic expansions for this problem have been constructed by Larsen and Keller in [39], studied further in [40] and presented in a physical context in [4]. The results of §6 are not sufficient to deal with the boundary conditions (3.3) because of the presence of boundary layers. The analysis of the transport-theoretic boundary value problem in the diffusion limit by probabilistic methods will be given elsewhere. The analysis given here of the free space problem or the boundary value problem (6.42) which does not have boundary layers, agrees with the results of [39] and [4] including in particular the analysis of the initial layer.

4. Corollary 1, as presented here, is a simple consequence of a more general result but it is of substantial independent interest. Let

$(X^{(0)}(\tau, \sigma, x^0, z), Z^{(0)}(\tau, \sigma, x^0, z))$ be the diffusion Markov process with generator (6.34). Define $\alpha(\sigma, x^0)$ as the symmetric square root of $a(\sigma, x^0)$

$$(7.11) \qquad a^{ij}(\sigma, x^0) = \sum_{k=1}^{n} \alpha^{ik}(\sigma, x^0)\alpha^{jk}(\sigma, x^0).$$

Then $(X^{(0)}, Z^{(0)})$ satisfy the Itô equations

$$(7.12) \qquad dX^{(0)}(\tau) = \check{G}(\tau, X^{(0)}(\tau))\, d\tau,$$

$$(7.13) \qquad \begin{aligned} dZ^{(0)}(\tau) &= b(\tau, X^{(0)}(\tau))Z^{(0)}(\tau)\, d\tau + \alpha(\tau, x^0(\tau))\, d\beta(\tau), \qquad \sigma > \tau, \\ X^{(0)}(\sigma, \sigma, x^0) &= x^0, \quad Z^{(0)}(\sigma, \sigma, x^0, z) = z. \end{aligned}$$

Here $\beta(\tau)$ is the standard n-dimensional Brownian motion and, as we mentioned in §6 already, $X^{(0)}$ is deterministic but is carried along for convenience. Note that (7.13) is a linear Itô equation so $Z^{(0)}$ is Gaussian. We can solve (7.13) as follows. Let $U(\tau, \sigma; x^0)$ be the solution matrix of the equation

$$\frac{dU(\tau, \sigma; x^0)}{d\tau} = b(\tau, X^{(0)}(\tau, \sigma; x^0))U(\tau, \sigma; x^0), \quad \tau > \sigma,$$

$$U(\sigma, \sigma; x^0) = I, \quad \text{the identity.}$$

Then,

$$(7.14) \quad Z^{(0)}(\tau, \sigma, x^0, z) = U(\tau, \sigma; x^0)z + \int_{\sigma}^{\tau} U(\tau, s; x^0)\alpha(s, X^{(0)}(s, \sigma, x^0))\, d\beta(s)$$

and hence

$$(7.15) \qquad E\{Z^{(0)}(\tau, \sigma, x^0, z)\} = U(\tau, \sigma; x^0)z.$$

Since z is the fluctuation at the beginning of the motion, it is natural to assume $z=0$. In this case we can compute the covariance matrix of $Z^{(0)}$ and obtain

$$(7.16) \qquad \begin{aligned} E&\{Z^{(0)}(\tau, \sigma, x^0, 0)Z^{(0)T}(\tau, \sigma, x^0, 0)\} \\ &= \int_{\sigma}^{\tau} U(\tau, s; x^0)a(s, X^{(0)}(s, \sigma, x^0))U^T(\tau, s; x^0)\, ds. \end{aligned}$$

This is a simple but useful formula which illustrates the kind of information that is readily accessible.

Corollary 1 has been obtained previously by Khasminskii [13] and Gihman-Skorohod [12, p. 342 where references to earlier work are given] when there is no feedback present, i.e. when the collision operator $Q_{\sigma,x}$ in (6.30) is independent of x. Actually, in the no-feedback case, the results of [13] and [12] are more general with regard to the nature of $Y^{(t)}$.

Recently, B. White [41] has analyzed stochastic differential-difference equations in a manner analogous to [13] and has applied the results to several interesting examples.

A theorem similar to Corollary 1 is given by Norman in [7, p. 118]. The limiting process in this result is of the form (7.12), (7.13) but the hypotheses and conclusions differ from the ones of §6 in several respects. Norman gives a learning-theoretic interpretation of his results and should be consulted for additional information (cf. also the remarks concerning slow learning in §4).

8. **Some applications of the diffusion approximation.** In this section we consider two applications. The first is an example studied by Stroock [14] and Baggett and Stroock [15] and, in a special case, corresponds to the diffusion approximation for one-speed neutron transport [4], [39]. This example can also be interpreted as a slow learning model. The second example is the harmonic oscillator of §4. We assume that smoothness and boundedness conditions on coefficients, as stated in §6, hold here and we do not state them explicitly.

We consider (6.1) when $X^{(\varepsilon)} \in R^3$ and $Y^{(\varepsilon)} \in S^2$ the unit sphere in R^3. Let $dS(y)$, $y \in S^2$, denote normalized uniform measure on S^2. We assume that $\pi(\sigma, x, y, A)$ has a density with respect to this uniform measure which is rotation invariant that is,

$$(8.1) \qquad \pi(\sigma, x, y, dz)/dS(z) = \tilde{\pi}(\sigma, x, y \cdot z).$$

Here $y \cdot z$ denotes the dot product of y and $z \in S^2$ and we denote the density with $\tilde{\pi}$. We shall also assume that $q = q(\sigma, x)$ is independent of y. The case where q depends on y can be treated in principle but the formulas are unwieldy; we remark further on this below. We rewrite the version of (6.1) which concerns us here.

$$
\begin{aligned}
& \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial \sigma} + \left[\frac{1}{\varepsilon} F(\sigma, x, y) + G(\sigma, x, y) \right] \frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial x} \\
(8.2) \qquad & \qquad + \frac{q(\sigma, x)}{\varepsilon^2} \int_{S^2} u^{(\varepsilon)}(\sigma, \tau, x, z) \tilde{\pi}(\sigma, x, y \cdot z) \, dS(z) \\
& \qquad - (q(\sigma, x)/\varepsilon^2) u^{(\varepsilon)}(\sigma, \tau, x, y) = 0, \quad \sigma < \tau, \\
& u^{(\varepsilon)}(\tau, \tau, x, y) = f(x, y), \quad x \in R^3, \quad y \in S^2.
\end{aligned}
$$

Note that we have assumed that F and G do not depend on the fast time σ/ε^2 explicitly. This is because we have a learning theoretic interpretation in mind (or, when $F = y$, $G = 0$, the one-speed neutron transport). This interpretation is as follows [14].

A bacterium, whose position at time τ is $X^{(\varepsilon)}(\tau)$ and its velocity $Y^{(\varepsilon)}(\tau)$, moves along the trajectories of the vector field $F/\varepsilon + G$ for a random length of time at the end of which the velocity changes direction. The probability distribution of the time between jumps and of the new velocity

immediately after the jump is given by (2.4) (adapted to the present notation). The position and velocity of the bacterium $(X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau))$ is thus a transport process. It is constructed in exactly the way such processes were constructed in the beginning of §2. Interest in (8.2) centers on the dependence of q, the collison frequency, on x, because then one can think of the motion as adapting to environmental conditions reflected through the x dependence of q and π.

We assume here that $q(\sigma, x) \geq \delta > 0$ and that F, G and $\tilde{\pi}$ have convergent expansions in spherical harmonics:

$$F^i(\sigma, x, y) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} F_{lm}^i(\sigma, x) Y_l^m(y),$$

(8.3) $\qquad G^i(\sigma, x, y) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} G_{lm}^i(\sigma, x) Y_l^m(y), \quad i = 1, 2, 3,$

$$\tilde{\pi}(\sigma, x, y \cdot z) = \frac{1}{4\pi} \sum_{l=0}^{\infty} \tilde{\pi}_l(\sigma, x)(2l + 1) P_l(y \cdot z).$$

In the expansion for $\tilde{\pi}$, $\pi = 3.14 \cdots$ and should not be confused with previous notation. Since $\tilde{\pi}$ is a density, $\tilde{\pi}_0 \equiv 1$ and we assume $\sup_{l \neq 0} \tilde{\pi}_l < 1$. The meaning of this condition is discussed at the end of this section. Using the addition theorem for spherical harmonics we rewrite the expansion $\tilde{\pi}$ in the form ($* = $ complex conjugate)

(8.4) $\qquad \tilde{\pi}(\sigma, x, y \cdot z) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \tilde{\pi}_l(\sigma, x) Y_l^{m*}(z) Y_l^m(y).$

The spherical harmonics are orthonormal

(8.5) $\qquad \int_{S^2} Y_l^{m*}(z) Y_{l'}^{m'}(z) \, dS(z) = \delta_{mm'} \delta_{ll'}$

and complete in $L^2(S^2)$.

To implement the theorem of §6 (we assume that all smoothness and growth conditions for F, G, q and π hold here) we must compute $\tilde{P}$ and P_0 from (6.4) and (6.6). Let $P_0(t, y, \bar{y}; \sigma, x)$ denote the density of P_0 with respect to the uniform measure on S^2. Using the above expansion, (6.4) takes the form

$$\frac{\partial P_0(t, y, \bar{y}; \sigma, x)}{\partial t} = q(\sigma, x) \int_{S^2} P_0(t, z, \bar{y}; \sigma, x)$$

(8.6) $\qquad \qquad \times \sum_{l,m} \tilde{\pi}_l(\sigma, x) Y_l^{m*}(z) Y_l^m(y) \, dS(z)$

$$- q(\sigma, x) P_0(t, y, \bar{y}; \sigma, x) = 0 \qquad t > 0$$

$$P_0(0, y, \bar{y}; \sigma, x) = \delta(y - \bar{y}).$$

From the orthonormality relation (8.5) we find that

$$(8.7) \qquad P_{0,lm}(t, \bar{y}; \sigma, x) = \int P_0(t, y, \bar{y}; \sigma, x) Y_l^{m^*}(y)\, dS(y)$$

satisfies the equation

$$(8.8) \qquad \frac{\partial P_{0,lm}(t, \bar{y}; \sigma, x)}{\partial t} = q(\sigma, x)(\tilde{\pi}_l(\sigma, x) - 1)P_{0,lm}(t, \bar{y}; \sigma, x), \quad t > 0,$$

$$P_{0,lm}(0, \bar{y}; \sigma, x) = Y_l^{m^*}(\bar{y}).$$

Therefore we have the expansion

$$(8.9) \quad P_0(t, y, \bar{y}; \sigma, x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \exp(-t[1 - \tilde{\pi}_l(\sigma, x)]q(\sigma, x))Y_l^m(y)Y_l^{m^*}(\bar{y}).$$

From (8.9) and the assumption $\tilde{\pi}_l < 1$, $l \geqq 1$ it follows that

$$(8.10) \qquad \bar{P}(y; \sigma, x) = 1$$

i.e., the invariant measure is the uniform measure which is of course what one should expect with $q > 0$ and $\pi_l < 1$, $l \geqq 1$. The centering hypothesis for F is:

$$(8.11) \qquad \int_{S^2} F(\sigma, x, y)\, dS(y) = 0.$$

We are now ready to calculate the diffusion and drift coefficients for the limit diffusion Markov process to which $X^{(\varepsilon)}$ tends. We use the expansions (8.3), (8.9), and (8.10) in (6.17), (6.18) to obtain

$$(8.12) \quad a^{ij}(\tau, x) = 2 \sum_{l=1}^{\infty} \sum_{m=-l}^{l} \frac{F_{lm}^{i^*}(\tau, x) F_{lm}^j(\tau, x)}{q(\tau, x)[1 - \tilde{\pi}_l(\tau, x)]}, \quad i, j = 1, 2, 3,$$

$$(8.13) \quad \begin{aligned} b^j(\tau, x) = &\sum_{i=1}^{3} \sum_{l=1}^{\infty} \sum_{m=-l}^{l} F_{lm}^{i^*}(\tau, x) \frac{\partial}{\partial x_i}\left[\frac{F_{lm}^j(\tau, x)}{q(\tau, x)[1 - \tilde{\pi}_l(\tau, x)]}\right] \\ &+ G_{00}^j(\tau, x). \end{aligned}$$

The infinitesimal generator $\mathscr{L}_\tau$ of the diffusion Markov process to which $X^{(\varepsilon)}$ converges can thus be written in the form

$$(8.14) \quad \begin{aligned} \mathscr{L}_\tau f \cdot (x) = &\sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{l=1}^{\infty} \sum_{m=-l}^{l} F_{lm}^{i^*}(\tau, x) \\ &\times \frac{\partial}{\partial x_i}\left[\frac{F_{lm}^j(\tau, x)}{q(\tau, x)[1 - \tilde{\pi}_l(\tau, x)]} \frac{\partial f(x)}{\partial x_j}\right] \\ &+ \sum_{j=1}^{3} G_{00}^j(\tau, x) \frac{\partial f(x)}{\partial x_j}. \end{aligned}$$

When $F = y$, $G = 0$, then (8.14) reduces to

$$
\begin{aligned}
\mathscr{L}_\tau f \cdot (x) &= \frac{4\pi}{3} \sum_{i=1}^{3} \frac{\partial}{\partial x_i} \left(\frac{1}{q(\tau, x)[1 - \tilde{\pi}_1(\tau, x)]} \frac{\partial f(x)}{\partial x_i} \right) \\
&= \frac{4\pi}{3} \nabla \cdot \left(\frac{1}{q(\tau, x)[1 - \tilde{\pi}_1(\tau, x)]} \nabla f(x) \right).
\end{aligned}
$$
(8.15)

Let us also consider the case $F \equiv 0$ and apply Corollary 1. From (8.14) it is clear that when $F \equiv 0$, $X^{(\varepsilon)}$ converges to a deterministic process $X^{(0)}$ satisfying the system of equations

$$
\frac{dX^{(0)}(\tau, \sigma, x^0)}{d\tau} = G_{00}(\tau, X^{(0)}(\tau, \sigma, x^0)), \quad \tau > \sigma,
$$
(8.16)

$$
X^{(0)}(\sigma, \sigma, x^0) = x^0.
$$

We proceed now to find the generator of the limit of the fluctuation process $[X^{(\varepsilon)} - X^{(0)}]/\varepsilon = Z^{(\varepsilon)}$ by using (6.32) and (6.33) along with (8.3), (8.9) and (8.10). A short calculation yields for $\tilde{\mathscr{L}}_\tau$ of (6.34) the expression

$$
\begin{aligned}
\tilde{\mathscr{L}}_\tau g(x^0, z) &= \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{l=1}^{\infty} \sum_{m=-l}^{l} \frac{G_{lm}^{i*}(\tau, x^0) G_{lm}^{j}(\tau, x^0)}{q(\tau, x^0)[1 - \tilde{\pi}_l(\tau, x^0)]} \frac{\partial^2 g(x^0, z)}{\partial z_i \, \partial z_j} \\
&\quad + \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial G_{00}^{i}(\tau, x^0)}{\partial x_j^0} z_j \frac{\partial g(x^0, z)}{\partial z_i} \\
&\quad + \sum_{i=1}^{3} G_{00}^{i}(\tau, x^0) \frac{\partial g(x^0, z)}{\partial x_i^0} .
\end{aligned}
$$
(8.17)

Note that here the diffusion coefficients are independent of z and the drift is linear in z. According to remark 4 of §7, $Z^{(0)}(\tau, \sigma, x^0, z)$ is a Gauss-Markov process with covariance given by (7.16) when $z = 0$. Note further that the effect of feedback is manifested through q and $\tilde{\pi}_l$ in (8.17) which depend on x^0.

When $q = q(\sigma, x, y)$ depends on y, the above analysis is complicated in going from (8.6) to (8.8) and hence in obtaining (8.9). The calculation can be carried out using the Clebsch-Gordan coefficients [42] but the formulas are unwieldy.

The second example we consider is the harmonic oscillator (4.4). We shall apply the theorem of §6 to (4.9). We must first, however, introduce some hypotheses. Suppose S^1 is parametrized by $\phi \in [0, 2\pi)$ and π

has a density $\tilde{\pi}(\sigma, R, \phi - \phi')$. We rewrite (4.9) as follows.

$$\frac{\partial u^{(\varepsilon)}(\tau, \sigma, R, \theta, \phi)}{\partial \sigma} + \frac{\omega}{\varepsilon} g(\phi) \sin\left(\theta + \frac{2\omega\sigma}{\varepsilon^2}\right) \frac{\partial u^{(\varepsilon)}(\sigma, \tau, R, \theta, \phi)}{\partial R}$$

$$+ \frac{\omega}{\varepsilon} g(\phi)\left[1 + \cos\left(\theta + \frac{2\omega\sigma}{\varepsilon^2}\right)\right]\frac{\partial u^{(\varepsilon)}(\sigma, \tau, R, \theta, \phi)}{\partial \theta}$$

$$\text{(8.18)} \qquad + \frac{q(\sigma, R)}{\varepsilon^2} \int_0^{2\pi} u^{(\varepsilon)}(\sigma, \tau, R, \theta, \phi')\tilde{\pi}(\sigma, R, \phi - \phi')\frac{d\phi'}{2\pi}$$

$$- \frac{q(\sigma, R)}{\varepsilon^2} u^{(\varepsilon)}(\sigma, \tau, R, \theta, \phi) = 0, \qquad \sigma < \tau,$$

$$u^{(\varepsilon)}(\tau, \tau, R, \theta, \phi) = f(R, \theta, \phi).$$

In a manner analogous to the preceding example we assume that

$$\text{(8.19)} \qquad \tilde{\pi}(\sigma, R, \phi) = \sum_{l=-\infty}^{\infty} \tilde{\pi}_l(\sigma, R)e^{il\phi},$$

$$\text{(8.20)} \qquad g(\phi) = \sum_{l=-\infty}^{\infty} g_l e^{il\phi}.$$

Again we have $\tilde{\pi}_0 = 1$ and we assume $q > 0$ and $\sup_{l \neq 0} \tilde{\pi}_l < 1$. The meaning of this condition is discussed at the end of this section. For centering, we assume $g_0 = 0$.

The invariant measure here is $d\phi/2\pi$ and P_0 is easily found to be

$$P_0(t, \phi, \bar{\phi}; \sigma, R) \, d\bar{\phi}$$

$$\text{(8.21)} \qquad = \sum_{l=-\infty}^{\infty} \exp(-t[1 - \tilde{\pi}_l(\sigma, R)]q(\sigma, R))\exp(il(\phi - \bar{\phi}))\frac{d\bar{\phi}}{2\pi}.$$

We proceed now to compute the diffusion and drift coefficients of the diffusion Markov process to which $(R^{(\varepsilon)}, \theta^{(\varepsilon)})$ tends as $\varepsilon \downarrow 0$. On using (6.13) and (6.14) we find that all coefficients are independent of θ and depend upon τ and R only. Thus $R^{(\varepsilon)}$, the logarithm of the energy, converges to a diffusion process on the real line with generator

$$\text{(8.22)} \quad \mathcal{L}_\tau f(R) = \frac{\partial}{\partial R}\left[\sum_{l=1}^{\infty} |g_l|^2 \frac{[1 - \tilde{\pi}_l(\tau, R)]q(\tau, R)\omega^2}{[1 - \tilde{\pi}_l(\tau, R)]^2 q^2(\tau, R) + 4\omega^2} \frac{\partial f(R)}{\partial R}\right].$$

We omit here the calculations which are analogous to the ones for the previous example. Note that if q and π are independent of R, that is, (4.4) is a linear oscillator, then the diffusion and drift coefficients in (8.22) are only functions of τ and therefore $R^{(\varepsilon)}$ converges in this case to a time-inhomogeneous Brownian motion on the real line.

We shall now examine the assumption $\sup_{l \neq 0} \tilde{\pi}_l < 1$ introduced above.

Since the density $\tilde{\pi}(\sigma, x, y \cdot z)$ has a convergent expansion (8.3), $\tilde{\pi}_l \to 0$, $l \to \infty$ and so the assumption above is equivalent to $\tilde{\pi}_l < 1$ for all $l \geq 1$.

Let us show that $\sup_{l \neq 0} \tilde{\pi}_l < 1$ implies that (6.8) and (6.9) are satisfied and $\rho(t)$ is, in fact, proportional to a decreasing exponential. We shall treat (8.9) since (8.21) can be treated in the identical manner. From (8.9) we have the following.

$$P_0(t, y, \bar{y}; \sigma, x)$$

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \exp(-t[1 - \tilde{\pi}_l(\sigma, x)]q(\sigma, x)) Y_l^m(y) Y_l^{m^*}(\bar{y})$$

(8.22)

$$= \sum_{\nu=0}^{\infty} \exp(-tq(\sigma, x)) \frac{[tq(\sigma, x)]^\nu}{\nu!} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \tilde{\pi}_l^\nu(\sigma, x) Y_l^m(y) Y_l^{m^*}(\bar{y})$$

$$= \sum_{\nu=0}^{\infty} e^{-tq(\sigma, x)} \frac{[tq(\sigma, x)]^\nu}{\nu!} \tilde{\pi}^{(\nu)}(\sigma, x, y \cdot \bar{y}).$$

Here $\tilde{\pi}^{(\nu)}(\sigma, x, y \cdot \bar{y})$ is the νth convolution of the density $\tilde{\pi}(\sigma, x, y \cdot \bar{y})$ with itself and for $\nu = 0$ it is the delta function concentrated at $\bar{y}$.

The hypothesis $\sup_{l \neq 0} \tilde{\pi}_l(\sigma, x) < 1$ implies that

(8.23) $$|\tilde{\pi}^{(\nu)}(\sigma, x, y \cdot \bar{y}) - 1| \leq C\rho^\nu, \quad \nu \geq 1,$$

where C is a constant and $0 < \rho < 1$. This is seen from the expression

$$\tilde{\pi}^{(\nu)}(\sigma, x, y \cdot \bar{y}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \tilde{\pi}_l^\nu(\sigma, x) Y_l^m(y) Y_l^{m^*}(\bar{y}),$$

recalling that $\tilde{\pi}_0 \equiv 1$ here.

Let $f(y)$ be a bounded measurable function. Then,

$$\int [P_0(t, y, \bar{y}; \sigma, x) - 1] f(\bar{y}) \, dS(\bar{y})$$

$$= \sum_{\nu=0}^{\infty} e^{-tq(\sigma, x)} \frac{[tq(\sigma, x)]^\nu}{\nu!} \int [\tilde{\pi}^{(\nu)}(\sigma, x, y \cdot \bar{y}) - 1] f(\bar{y}) \, dS(\bar{y}).$$

Using (8.23) we obtain

$$\left| \int [P_0(t, y, \bar{y}; \sigma, x) - 1] f(\bar{y}) \, dS(\bar{y}) \right|$$

$$\leq C' \sum_{\nu=0}^{\infty} e^{-tq(\sigma, x)} \frac{[tq(\sigma, x)]^\nu}{\nu!} \rho^\nu \sup_{\bar{y}} |f(\bar{y})|,$$

where C' is a constant. Thus

$$\left| \int [P_0(t, y, \bar{y}; \sigma, x) - 1] f(\bar{y}) \, dS(\bar{y}) \right| \leq C' \sup_y |f(y)| e^{-tq(\sigma, x)[1-\rho]},$$

and hypotheses (6.8), (6.9) have been verified.

The physical meaning of the condition $\sup_{t \neq 0} \tilde{\pi}_t(\sigma, x) < 1$ is that it makes the "frozen" velocity process i.e., the velocity of the bacterium in the absence of drift, a mixing Markov process in the sense of (6.8), (6.9).

9. **Outline of the proof of the theorem.** The proof of the theorem of §6 is patterned after the proof of Theorem 1 in [43] which was also followed in [21]. The essential steps are perhaps most transparent in [43] which is recommended for this reason. We review briefly the argument here.

The interval $[0, \tau]$ is broken up into intervals of length $\Delta \sim \varepsilon^{7/4}$. The power 7/4 is of no special significance; any exponent in [1, 2) would do if the peripheral estimates could be obtained for it. The intervals Δ are small in the τ time scale so that the $X^{(\varepsilon)}$ process does not change much over one Δ interval. On the other hand the $Y^{(\varepsilon)}$ process has intrinsic time scale τ/ε^2 so the interval Δ is $\Delta/\varepsilon^2 \sim \varepsilon^{-1/4}$ for $Y^{(\varepsilon)}$ i.e., it is large. It is actually large enough so that $Y^{(\varepsilon)}$ in this interval can be treated with $X^{(\varepsilon)}$ frozen to its value at the beginning of the interval and so that $Y^{(\varepsilon)}$, with $X^{(\varepsilon)}$ frozen, is equilibrating to its invariant measure $\bar{P}$. This is the standard way of thinking that one finds in the physics literature, for example in [6], and our approach simply follows this argument. An intuitive understanding in terms of multiple time scales may be obtained from [44].

In §12 we derive an a priori estimate for the moments of $X^{(\varepsilon)}$. This estimate is necessary for the proof of weak convergence. §13 sets up the decomposition of the problem into local estimates over intervals of length Δ as described above. §14 provides the estimate that corresponds to the statement: Δ/ε^2 is large enough so that the process $X^{(\varepsilon)}$ is statistically independent over nonoverlapping Δ intervals. This explains the title for this section. §15 provides the estimate that corresponds to: Δ is small enough so that $X^{(\varepsilon)}$ does not change much over a Δ interval. This section is entitled Local Taylor Expansion for obvious reasons.

We remark again that the hypothesis that S, the state space of $Y^{(\varepsilon)}$, is compact, is used heavily in what follows but we feel the argument can be refined, perhaps along the lines of [32], to remove this hypothesis.

Throughout we employ the convention that C denotes a constant, not necessarily the same constant. Similarly we use $\bar{\alpha}$ to denote a sufficiently large nonnegative integer.

We also assume that $G \equiv 0$ in (6.1) since the treatment of G is elementary by comparison to F, and $G \neq 0$ would only clutter the formulas. It can be verified that the proof that follows carries over without essential changes to the analysis of (6.30) and Corollary 1. Furthermore, because we have assumed that F, G, q and π depend on σ the same way as on x (with the

same degree of smoothness) and since $0 \leq \sigma \leq \tau \leq T < \infty$, we may treat σ as another x-coordinate and hence eliminate it from explicit considerations. In fact, with the notation $\sigma = x_0$ in (6.2) and (6.3) we have anticipated this simplification. With no loss of generality then, F ($G \equiv 0$, we assume), q and π will be assumed independent of explicit slow-time dependence in the sequel.

10. **Preliminary results and identities.** We begin with an estimate on P_0 the solution of (6.4). Since x is a parameter and everything depends smoothly on it $P_0(t, y, A; x)$ is differentiable in x. We need the following result.

LEMMA 10.1.   *Let* $f(x, y)$, $x \in R^n$, $y \in S$, *be measurable and such that*

(10.1) $|f(x, y)| \leq C(1 + |x|^\alpha)$,

(10.2) $\left| \dfrac{\partial f(x, y)}{\partial x_i} \right| \leq C(1 + |x|^{\alpha-1})$, $i = 1, 2, \cdots, n$.

Then,

(10.3) $\left| \dfrac{\partial}{\partial x_i} \int \bar{P}(dy; x) f(x, y) \right| \leq C(1 + |x|^{\alpha-1})$,

and

(10.4)

$$\left| \frac{\partial}{\partial x_i} \int [P_0(t, y, d\zeta; x) - \bar{P}(d\zeta; x)] f(x, \zeta) \right|$$

$$\leq \rho\left(\frac{t}{2}\right) C(1 + |x|^{\alpha-1}), \qquad t > 0, \quad i = 1, 2, \cdots, n.$$

PROOF.   Let $u_0(t, y; x)$ denote the solution of

(10.5)

$$\frac{\partial u_0(t, y; x)}{\partial t}$$

$$= q(x, y) \int u_0(t, z; x) \pi(x, y; dz) - q(x, y) u_0(t, y; x), \quad t > 0,$$

$$u_0(0, y; x) = f(x, y),$$

and let $\bar{f}(x)$ be defined by

(10.6) $\bar{f}(x) = \int \bar{P}(dy; x) f(x, y)$.

We denote derivatives with respect to x_i by a comma and subscript i.

Differentiating (10.5) and using this notation we obtain

$$\frac{\partial}{\partial t} u_{0,i}(t, y; x)$$

$$= q(x, y) \int u_{0,i}(t, z; x)\pi(x, y, dz) - q(x, y)u_{0,i}(t, y; x)$$

(10.7)
$$+ q_{,i}(x, y) \int u_0(t, z; x)\pi(x, y, dz) - q_{,i}(x, y)u_0(t, y; x)$$

$$+ q(x, y) \int u_0(t, z; x)\pi_{,i}(x, y, dz), \quad t > 0,$$

$$u_{0,i}(0, y; x) = f_{,i}(x, y).$$

In the last three terms on the right side of (10.7) we may replace u_0 by $u_0 - f$ because the f cancels. Integrating (10.7) yields

$$u_{0,i}(t, y; x) = \int P_0(t, y, d\zeta; x)f_{,i}(x, \zeta) + \int_0^t \int P_0(t - s, y, d\zeta; x)$$

$$\times \left[q_{,i}(x, \zeta) \int (u_0(t, z; x) - f(x))\bar{\pi}(x, \zeta, dz) \right.$$

(10.8)
$$- q_{,i}(x, \zeta)(u_0(t, \zeta; x) - f(x))$$

$$\left. + q(x, \zeta) \int (u_0(t, z; x) - f(x))\bar{\pi}_{,i}(x, \zeta, dz) \right] ds.$$

We estimate $u_{0,i}$ from (10.8) by using the hypotheses of the lemma and hypotheses (6.3) and (6.8) as follows.

(10.9)
$$|u_{0,i}(t, y; x)| \leq C(1 + |x|^{\alpha-1}) + C(1 + |x|^{\alpha-1}) \int_0^t \rho(s)\, ds$$

$$\leq C(1 + |x|^{\alpha-1}).$$

In the last inequality we used (6.9).[6]

From (10.9) we obtain (10.3) by an elementary argument. We proceed therefore with the proof of (10.4). To simplify the notation we introduce along with

(10.10) $Q_x f(y) = q(x, y) \int f(z)\pi(x, y, dz) - q(x, y)f(y)$

of (6.1), the operator

$$\tilde{Q}_x^i f(y) = q_{,i}(x, y) \int [f(z) - f(y)]\pi(x, y, dz)$$

(10.11)
$$+ q(x, y) \int f(z)\pi_{,i}(x, y, dz).$$

_____

[6] Note that when $f(x, y) = \chi_A(y)$ then, $|P_{0,i}(t, y, A; x)| \leq C(1 + |x|)^{-1}$ and the same estimate holds for $P_{,i}(A; x)$. This is used in (10.19).

Let

(10.12) $$v(t, y; x) = u_0(t, y; x) - \bar{f}(x).$$

Then

(10.13)
$$\frac{\partial v(t, y; x)}{\partial t} = Q_x v(t, y; x), \quad t > 0,$$

$$v(0, y; x) = f(x, y) - \bar{f}(x).$$

Differentiating (10.13) with respect to x_i and integrating the resulting equation yields

(10.14)
$$v_{,i}(t, y; x) = \int P_0(t, y, d\zeta; x)[f_{,i}(x, \zeta) - \bar{f}_{,i}(x)]$$
$$+ \int_0^t \int P_0(t - s, y, d\zeta; x)\tilde{Q}_x^i v(s, \zeta; x)\, ds.$$

From (6.6) we have the identity

(10.15) $$\int \bar{P}(d\zeta; x)Q_x v(t, \zeta; x) = 0,$$

and hence by differentiation we obtain

(10.16) $$\int \bar{P}_{,i}(d\zeta; x)Q_x v(t, \zeta; x) + \int \bar{P}(d\zeta; x)\tilde{Q}_x^i v(t, \zeta; x) = 0,$$

where we have used (10.15) to eliminate the $v_{,i}$ term in (10.16). Now we return to (10.14) and rewrite it as follows.

(10.17)
$$v_{,i}(t, y; x) = \int [P_0(t, y, d\zeta; x) - \bar{P}(d\zeta; x)]f_{,i}(x, \zeta)$$
$$- \int \bar{P}_{,i}(d\zeta; x)f(x, \zeta)$$
$$+ \int_0^t \int [P_0(t - s, y, d\zeta; x) - \bar{P}(d\zeta; x)]\tilde{Q}_x^i v(s, \zeta; x)\, ds$$
$$+ \int_0^t \int \bar{P}(d\zeta; x)\tilde{Q}_x^i v(s, \zeta; x)\, ds.$$

Let us use in (10.17) the identity (10.16) and also

(10.18) $$v(t, y; x) = f(x, y) - \bar{f}(x) + \int_0^t Q_x v(s, y; x)\, ds.$$

We obtain

(10.19)
$$v_{,i}(t, y, x) = \int [P_0(t, y, d\zeta; x) - \bar{P}(d\zeta; x)]f_{,i}(x, \zeta)$$
$$+ \int_0^t \int [P_0(t - s, y, d\zeta; x) - \bar{P}(d\zeta; x)]\tilde{Q}_x^i v(s, \zeta; x)$$
$$- \int \bar{P}_{,i}(d\zeta; x)v(t, \zeta; x).$$

From the hypotheses of the lemma (10.9), (6.3), (6.8) and (6.9) follows the estimate

$$(10.20) \quad |v_{,i}(t, y, x)| \leqq C(1 + |x|^{\alpha-1})\rho(t) + C(1 + |x|^{\alpha-1}) \int_0^t \rho(t-s)\rho(s)\,ds$$
$$+ C(1 + |x|^{\alpha-1})\rho(t).$$

Since $\rho(t)$ is monotonically decreasing and (6.9) holds,

$$\int_0^t \rho(t-s)\rho(s)\,ds \leqq \rho\left(\frac{t}{2}\right)\int_0^{t/2}\rho(s)\,ds + \rho\left(\frac{t}{2}\right)\int_{t/2}^t \rho(t-s)\,ds$$

and thus (10.4) follows. The proof of the lemma is complete.

Let us rewrite (6.1) without explicit σ-dependence, with $G \equiv 0$ and the notation (10.10):

$$\frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial \sigma} + \frac{1}{\varepsilon} F\left(\frac{\sigma}{\varepsilon^2}, x, y\right)\frac{\partial u^{(\varepsilon)}(\sigma, \tau, x, y)}{\partial x}$$

$$(10.21) \qquad\qquad\qquad + \frac{1}{\varepsilon^2} Q_x u^{(\varepsilon)}(\sigma, \tau, x, y) = 0, \quad \sigma < \tau,$$

$$u^{(\varepsilon)}(\tau, \tau, x, y) = f(x, y).$$

Recall that P_0 satisfies (6.4) which we rewrite here

$$(10.22) \qquad \frac{\partial P_0(t, y, A; x)}{\partial t} = Q_x P_0(t, y, A; x), \quad t > 0,$$

$$P_0(0, y, A; x) = \chi_A(y).$$

In the sequel we employ repeatedly two identities which we call the parametrix[7] identities. They are as follows:

$$u^{(\varepsilon)}(\sigma, \tau, x, y) = u_0((\tau - \sigma)/\varepsilon^2, y; x)$$

$$(10.23) \qquad\qquad + \frac{1}{\varepsilon} \int_\sigma^\tau \int P_0\left(\frac{s-\sigma}{\varepsilon^2}, y, d\zeta; x\right)$$

$$\times F_i\left(\frac{s}{\varepsilon^2}, x, \zeta\right) u_{,i}^{(\varepsilon)}(s, \tau, x, \zeta)\,ds,$$

$$u^{(\varepsilon)}(\sigma, \tau, x, y) = u_0((\tau - \sigma)/\varepsilon^2, y; x)$$

$$(10.24) \qquad\qquad + \frac{1}{\varepsilon} \int_\sigma^\tau \int P^{(\varepsilon)}(\sigma, s, x, y; d\xi, d\eta)$$

$$\times F_i\left(\frac{s}{\varepsilon^2}, \xi, \eta\right) u_{0,i}\left(\frac{\tau - s}{\varepsilon^2}, \eta; \xi\right) ds.$$

---

[7] Terminology suggested by H. P. McKean.

The first one is called the forward and the second the backward parametrix identity and we employ subscript-comma notation for derivatives with respect to x and the summation convention. We also denote by $P^{(\varepsilon)}(\sigma, s, x, y; A, B)$ the solution of (10.21) when $f(x, y) = \chi_{A \times B}(x, y)$, the characteristic function of the Borel sets $A \subset R^n$, $B \subset S$, and we denote by $u_0(t, y; x)$ the solution of (10.22) with $f(x, y)$ as initial function i.e.,

$$u^{(\varepsilon)}(\sigma, \tau, x, y) = \int P^{(\varepsilon)}(\sigma, \tau, x, y, d\xi, d\eta) f(\xi, \eta),$$

(10.25)

$$u_0(t, y; x) = \int P_0(t, y, d\eta; x) f(x, \eta).$$

Both identities can be verified easily by direct computation. Local (in time) differentiability is, of course, immediate from the smoothness hypotheses on the coefficients and data. Better estimates for derivatives of $u^{(\varepsilon)}$ with respect to x, the only ones that require consideration, are obtained in the next section. It should be observed that (10.23) and (10.24) are the analogs of (5.6) after expectations are taken and the Markov property of $(X^{(\varepsilon)}, Y^{(\varepsilon)})$ is used.

11. **Estimates for derivatives of the transport equation.** Under the hypotheses of §6 the existence, uniqueness and local smoothness of solutions of (10.21) (or (6.1)) follow from the usual iteration arguments. We need a bound for x derivatives of $u^{(\varepsilon)}$ independent of ε and valid in an interval of length $\Delta = \varepsilon^{7/4}$. This bound is obtained in this section.

From the differential equation

(11.1) $\dfrac{dX^{(\varepsilon)}(\tau)}{d\tau} = \dfrac{1}{\varepsilon} F\left(\dfrac{\tau}{\varepsilon^2}, X^{(\varepsilon)}(\tau), Y^{(\varepsilon)}(\tau)\right), \quad \tau > \sigma, \quad X^{(\varepsilon)}(\sigma, \sigma, x, y) = x,$

and from (6.2), it follows readily that

(11.2) $|X^{(\varepsilon)}(\tau, \sigma, x, y)| \leqq C(1 + |x|), \quad 0 \leqq \tau - \sigma \leqq \varepsilon.$

Therefore, if $g(x, y)$ is measurable and

$$|g(x, y)| \leqq C(1 + |x|^{\alpha}), \quad \alpha \geqq 0 \text{ integer},$$

we have

(11.3) $\left| \int P^{(\varepsilon)}(\sigma, \tau, x, y, d\xi, d\eta) g(\xi, \eta) \right| \leqq C(1 + |x|^{\alpha}), \quad 0 \leqq \tau - \sigma \leqq \varepsilon.$

We begin with the following result.

LEMMA 11.1.   *Let* $u^{(\varepsilon)}(\sigma, \tau, x, y)$ *denote the solution of* (10.21) *with data* $f(x)$ *independent of* y *and in* $C^{1,\alpha}(R^n)$, $\alpha \geqq 0$. *Let* $\Delta = \Delta(\varepsilon) = \varepsilon^{7/4}$. *Then there*

exists an integer $\bar{\alpha} > \alpha$ such that

(11.4) $|u_{,i}^{(\varepsilon)}(\sigma, \tau, x, y)| \leqq C(1 + |x|^{\bar{\alpha}}), \quad i = 1, \cdots, n, \quad 0 \leqq \tau - \sigma \leqq \Delta,$

where C is independent of τ, σ and ε.

PROOF. Define $H^{(\varepsilon)}$ by

(11.5) $H^{(\varepsilon)}(\sigma, \tau, x, y) = 1 - \exp\left[-\frac{1}{\varepsilon^2}\int_\sigma^\tau q(\xi^{(\varepsilon)}(s, \sigma, x, y), y)\, ds\right],$

where $\xi^{(\varepsilon)}(s, \sigma, x, y) = \xi^{(\varepsilon)}(s)$ denotes the solution of

(11.6) $\dfrac{d\xi^{(\varepsilon)}(s)}{ds} = \dfrac{1}{\varepsilon} F\left(\dfrac{s}{\varepsilon^2}, \xi^{(\varepsilon)}(s), y\right), \quad s > \sigma, \quad \xi^{(\varepsilon)}(\sigma, \sigma, x, y) = x.$

We rewrite (10.21) as an integral equation using the above notation as follows.

$$u^{(\varepsilon)}(\sigma, \tau, x, y) = f(\xi^{(\varepsilon)}(\tau))(1 - H^{(\varepsilon)})$$

(11.7)
$$+ \int_\sigma^\tau \int u^{(\varepsilon)}(s, \tau, \xi^{(\varepsilon)}(s), z)$$
$$\times \pi(\xi^{(\varepsilon)}(s), y, dz)\, d_s H^{(\varepsilon)}(\sigma, s, x, y).$$

Employing comma-subscript notation and the summation convention we differentiate (11.7) with respect to x_i and obtain

$$u_{,i}^{(\varepsilon)}(\sigma, \tau, x, y)$$
$$= f_{,j}(\xi^{(\varepsilon)}(\tau))\xi_{j,i}^{(\varepsilon)}(\tau)(1 - H^{(\varepsilon)}) - f(\xi^{(\varepsilon)}(\tau))H_{,i}^{(\varepsilon)}$$
$$+ \int_\sigma^\tau \int u_{,j}^{(\varepsilon)}(s, \tau, \xi^{(\varepsilon)}(s), z)\xi_{j,i}^{(\varepsilon)}(s)\pi(\xi^{(\varepsilon)}(s), y, dz)\, d_s H^{(\varepsilon)}(\sigma, s, x, y)$$

(11.8)
$$+ \int_\sigma^\tau \int u^{(\varepsilon)}(s, \tau, \xi^{(\varepsilon)}(s), z)\pi_{,j}(\xi^{(\varepsilon)}(s), y, dz)\xi_{j,i}^{(\varepsilon)}(s)\, d_s H^{(\varepsilon)}(\sigma, s, x, y)$$
$$+ \int_\sigma^\tau \int u^{(\varepsilon)}(s, \tau, \xi^{(\varepsilon)}(s), z)\pi(\xi^{(\varepsilon)}(s), y, dz)\, d_s H_{,i}^{(\varepsilon)}(\sigma, s, x, y),$$
$$i = 1, 2, \cdots, n.$$

From (11.6) we obtain by differentiation

(11.9) $\xi_{j,i}^{(\varepsilon)}(\tau) = \delta_{ij} + \dfrac{1}{\varepsilon}\int_\sigma^\tau F_{j,k}\left(\dfrac{s}{\varepsilon^2}, \xi^{(\varepsilon)}(s), y\right)\xi_{k,i}^{(\varepsilon)}(s)\, ds.$

We also have the easily verified identity, for smooth f,

(11.10) $f(\xi^{(\varepsilon)}(\tau)) = f(x) + \dfrac{1}{\varepsilon}\int_\sigma^\tau F_k\left(\dfrac{s}{\varepsilon^2}, \xi^{(\varepsilon)}(s), y\right)f_{,k}(\xi^{(\varepsilon)}(s))\, ds.$

Now we employ (11.9), (11.10) and (10.23) to rewrite (11.8) as follows.

$$u_{,i}^{(\varepsilon)} = f_{,i}(\xi^{(\varepsilon)}(\tau))(1 - H^{(\varepsilon)}) + \int_\sigma^\tau \int u_{,i}^{(\varepsilon)}(s, \tau, \xi^{(\varepsilon)}(s), z)$$
$$\times \pi(\xi^{(\varepsilon)}(s), y, dz) \, d_s H^{(\varepsilon)}(\sigma, s, x, y)$$
$$+ \frac{1}{\varepsilon} f_{,j}(\xi^{(\varepsilon)}(\tau)) \int_\sigma^\tau F_{j,k}\left(\frac{s}{\varepsilon^2}, \xi^{(\varepsilon)}(s), y\right) \xi_{k,i}^{(\varepsilon)}(s) \, ds(1 - H^{(\varepsilon)})$$
$$+ \frac{1}{\varepsilon} \int_\sigma^\tau \int u_{,j}^{(\varepsilon)}(s, \tau, \xi^{(\varepsilon)}(s), z) \pi(\xi^{(\varepsilon)}(s), y, dz)$$

(11.11)
$$\times \int_\sigma^s F_{j,k}\left(\frac{s_1}{\varepsilon^2}, \xi^{(\varepsilon)}(s_1), y\right) \xi_{k,i}^{(\varepsilon)}(s_1) \, ds_1 \, d_s H^{(\varepsilon)}(\sigma, s, x, y)$$
$$+ \frac{1}{\varepsilon} \int_\sigma^\tau \int \int_s^\tau \int P_0\left(\frac{s_1 - s}{\varepsilon^2}, z, d\zeta; \xi^{(\varepsilon)}(s)\right) F_k\left(\frac{s_1}{\varepsilon^2}, \xi^{(\varepsilon)}(s), \zeta\right)$$
$$\times u_{,k}^{(\varepsilon)}(s_1, \tau, \xi^{(\varepsilon)}(s), \zeta) \pi_{,j}(\xi^{(\varepsilon)}(s), y, dz)$$
$$\times ds_1 \xi_{j,i}^{(\varepsilon)}(s) \, d_s H^{(\varepsilon)}(\sigma, s, x, y)$$
$$+ \frac{1}{\varepsilon} \int_\sigma^\tau \int \int_s^\tau \int P_0\left(\frac{s_1 - s}{\varepsilon^2}, z, d\zeta; \xi^{(\varepsilon)}(s)\right) F_k\left(\frac{s_1}{\varepsilon^2}, \xi^{(\varepsilon)}(s), \zeta\right)$$
$$\times u_{,k}^{(\varepsilon)}(s_1, \tau, \xi^{(\varepsilon)}(s), \zeta) \pi(\xi^{(\varepsilon)}(s), y, dz) \, ds_1 \, d_s H_{,i}^{(\varepsilon)}$$
$$- \frac{1}{\varepsilon} \int_\sigma^\tau \int_s^\tau F_k\left(\frac{s_1}{\varepsilon^2}, \xi^{(\varepsilon)}(s), y\right) f_{,k}(\xi^{(\varepsilon)}(s_1)) \, ds_1 \, dH^{(\varepsilon)},$$
$$i = 1, 2, \cdots, n.$$

To simplify notation we write (11.11) in the form
$$u_{,i}^{(\varepsilon)} = f_{,i}(\xi^{(\varepsilon)}(\tau))(1 - H^{(\varepsilon)})$$

(11.12)
$$+ \int_\sigma^\tau \int u_{,i}^{(\varepsilon)}(s, \tau, \xi^{(\varepsilon)}(s), z) \pi(\xi^{(\varepsilon)}(s), y, dz) \, d_s H^{(\varepsilon)}$$
$$+ (1/\varepsilon) W_i^{(\varepsilon)}(\sigma, \tau, x, y),$$

where $W_i^{(\varepsilon)}$ is identified by comparison with (11.11) and depends on $u^{(\varepsilon)}$ and derivatives $u_{,j}^{(\varepsilon)}, j = 1, \cdots, n$.

We can use the fundamental solution $P^{(\varepsilon)}$ of (10.21) to write (11.12) as follows:

$$u_{,i}^{(\varepsilon)}(\sigma, \tau, x, y) = (1/\varepsilon) W_i^{(\varepsilon)}(\sigma, \tau, x, y)$$
$$+ \int P^{(\varepsilon)}(\sigma, \tau, x, y, d\bar{x}, d\bar{y}) f_{,i}(\bar{x})$$

(11.13)
$$+ \frac{1}{\varepsilon} \int_\sigma^\tau \int P^{(\varepsilon)}(\sigma, s, x, y, d\bar{x}, d\bar{y}) q(\bar{x}, \bar{y})$$
$$\times \int W_i^{(\varepsilon)}(s, \tau, \bar{x}, z) \pi(\bar{x}, \bar{y}, dz).$$

Now we use the hypotheses $f \in C^{1,\alpha}(R^n)$, $0 \leqq \tau - \sigma \leqq \Delta = \varepsilon^{7/4}$, (6.2), (6.3) and (11.2), (11.3) to estimate the right side of (11.13). To the resulting inequality for $\max_i |u_{,i}^{(\varepsilon)}|$ we apply Gronwall's lemma (Gronwall's inequality) and (11.4) follows. The constant C depends on the various constants in the hypotheses of §6 and the function f. We omit further details.

LEMMA 11.2. *In Lemma* 11.1 *assume* $f \in C^{2,\alpha}(R^n)$. *Then,*

(11.14)
$$|u_{,ij}^{(\varepsilon)}(\sigma, \tau, x, y)| \leqq C(1 + |x|^{\tilde{\alpha}}),$$
$$i, j = 1, \cdots, n, \quad 0 \leqq \tau - \sigma \leqq \Delta,$$

where $\Delta = \varepsilon^{7/4}$.

We omit the proof of this lemma which is similar to the one above but involves more computation.

12. **Estimate for moments.** In this section we derive the following estimate which does not depend on the lemmas of §11.

LEMMA 12.1. *Let* $f(x) = |x|^p$, $p \geqq 0$ *integer. Then the solution of* (10.21) (*or* 6.1), *with this f as data, satisfies*

(12.1)
$$u^{(\varepsilon)}(\sigma, \tau, x, y) = E\{|X^{(\varepsilon)}(\tau, \sigma, x, y)|^p\}$$
$$\leqq C(1 + |x|^p), \quad 0 \leqq \sigma \leqq \tau \leqq T,$$

where C is independent of ε *but depends on T and p* (*and other quantities*).

PROOF. We may assume that p is even and denote it by $2p$. The case of odd p follows by Schwarz' inequality from the even one.

Let $f(x) = |x|^{2p}$. From the differential equation (11.1) it follows that

(12.2)
$$E\{f(X^{(\varepsilon)}(\tau, \sigma, x, y))\} = E\{f(X^{(\varepsilon)}(\tau))\}$$
$$= f(x) + \frac{1}{\varepsilon} \int_\sigma^\tau E\left\{F_i\left(\frac{s}{\varepsilon^2}, X^{(\varepsilon)}(s, \sigma, x, y), Y^{(\varepsilon)}(s, \sigma, x, y)\right)\right.$$
$$\left. \times f_{,i}(X^{(\varepsilon)}(s, \sigma, x, y))\right\} ds.$$

We recall that we are employing the summation convention and comma subscript notation for x-derivatives.

We introduce the following notation:

(12.3) $g(t, x, y) = F_i(t, x, y)f_{,i}(x), \quad t \geqq 0,$

(12.4) $v_0^{(\varepsilon)}(\sigma, s, y; x) = \int P_0\left(\frac{s - \sigma}{\varepsilon^2}, y, d\eta; x\right) g\left(\frac{s}{\varepsilon^2}, x, \eta\right),$

(12.5) $v^{(\varepsilon)}(\sigma, s, x, y) = \int P^{(\varepsilon)}(\sigma, s, x, y, d\xi, d\eta) g\left(\frac{s}{\varepsilon^2}, \xi, \eta\right), \quad s \geqq \sigma.$

With this notation (12.2) becomes

(12.6) $$E\{f(X^{(\varepsilon)}(\tau))\} = f(x) + \frac{1}{\varepsilon}\int_\sigma^\tau v^{(\varepsilon)}(\sigma, s, x, y)\, ds.$$

We now use the identity (10.24) to express $v^{(\varepsilon)}$ in terms of $v_0^{(\varepsilon)}$.

(12.7)
$$
\begin{aligned}
v^{(\varepsilon)}(\sigma, s, x, y) &= v_0^{(\varepsilon)}(\sigma, s, x, y) \\
&+ \frac{1}{\varepsilon}\int_\sigma^s\int P^{(\varepsilon)}(\sigma, \gamma, x, y, d\xi, d\eta)F_j\left(\frac{\gamma}{\varepsilon^2}, \xi, \eta\right)v_{0,j}^{(\varepsilon)}(\gamma, s, \eta; \xi)\, d\gamma.
\end{aligned}
$$

Substituting $v^{(\varepsilon)}$ from (12.7) into (12.6) and recalling that $f(x)=|x|^{2p}$ we obtain the following identity.

(12.8)
$$
\begin{aligned}
E\{|X^{(\varepsilon)}(\tau)|^{2p}\} &= |x|^{2p} + \frac{1}{\varepsilon}\int_\sigma^\tau v_0^{(\varepsilon)}(\sigma, s, y; x)\, ds \\
&+ \frac{1}{\varepsilon^2}\int_\sigma^\tau\int_\sigma^s E\left\{F_j\left(\frac{\gamma}{\varepsilon^2}, X^{(\varepsilon)}(\gamma, \sigma, x, y), Y^{(\varepsilon)}(\gamma, \sigma, x, y)\right)\right. \\
&\left.\times\, v_{0,j}^{(\varepsilon)}(\gamma, s, Y^{(\varepsilon)}(\gamma, \sigma, x, y); X^{(\varepsilon)}(\gamma, \sigma, x, y))\right\} d\gamma\, ds.
\end{aligned}
$$

Let us rewrite $v_0^{(\varepsilon)}$ of (12.4) in the form[8]

(12.9)
$$
\begin{aligned}
v_0^{(\varepsilon)}(\sigma, s, y; x) &= \int \mu_0\left(\frac{s-\sigma}{\varepsilon^2}, y, d\eta; x\right)g\left(\frac{s}{\varepsilon^2}, x, \eta\right) \\
&+ \int \bar{P}(d\eta; x)g\left(\frac{s}{\varepsilon^2}, x, \eta\right) \\
&\equiv u_0^{(\varepsilon)}(\sigma, s, y; x) + w_0(s/\varepsilon^2; x).
\end{aligned}
$$

From (12.3) and hypothesis (6.2) it follows that

(12.10)
$$
\begin{aligned}
|g(t, x, y)| &\leq C(1 + |x|^{2p}), \\
|g_{,j}(t, x, y)| &\leq C(1 + |x|^{2p-1}), \quad t \geq 0, j = 1, \cdots, n.
\end{aligned}
$$

These inequalities, (6.8) and Lemma 10.1 yield

(12.11)
$$
\begin{aligned}
|u_0^{(\varepsilon)}(\sigma, s, y; x)| &\leq \rho((s-\sigma)/\varepsilon^2)C(1 + |x|^{2p}), \\
|u_{0,j}^{(\varepsilon)}(\sigma, s, y; x)| &\leq \rho((s-\sigma)/2\varepsilon^2)C(1 + |x|^{2p-1}).
\end{aligned}
$$

Hypotheses (6.11), (6.12) and (12.10) yield

(12.12)
$$
\begin{aligned}
\left|\int_{t_0}^{t_0+T} w_0(s; x)\, ds\right| &\leq \frac{C(1 + |x|^{2p})}{1 + T^4}, \\
\left|\int_{t_0}^{t_0+T} w_{0,j}(s; x)\, ds\right| &\leq C(1 + |x|^{2p-1}), \quad T, t_0 \geq 0.
\end{aligned}
$$

---

[8] μ_0 is defined by (6.7).

We now insert (12.9) into the right side of (12.8) and estimate terms with the help of (12.11), (12.12) and hypotheses (6.2) and (6.9) as follows.

$$
E\{|X^{(\varepsilon)}(\tau)|^{2p}\} = |x|^{2p} + \frac{1}{\varepsilon}\int_\sigma^\tau u_0^{(\varepsilon)}(\sigma, s, y; x)\, ds + \frac{1}{\varepsilon}\int_\sigma^\tau w_0\left(\frac{s}{\varepsilon^2}; x\right) ds
$$

$$
+ \frac{1}{\varepsilon^2}\int_\sigma^\tau\int_\sigma^s\int P^{(\varepsilon)}(\sigma, \gamma, x, y, d\xi, d\eta)
$$

$$
\times F_j\left(\frac{\gamma}{\varepsilon^2}, \xi, \eta\right) u_{0,j}^{(\varepsilon)}(\gamma, s, \eta; \xi)\, d\gamma\, ds
$$

(12.13)

$$
+ \frac{1}{\varepsilon^2}\int_\sigma^\tau\int_\sigma^s\int P^{(\varepsilon)}(\sigma, \gamma, x, y, d\xi, d\eta)
$$

$$
\times F_j\left(\frac{\gamma}{\varepsilon^2}, \xi, \eta\right) w_{0,j}\left(\frac{s}{\varepsilon^2}; \xi\right) d\gamma\, ds
$$

$$
\leqq |x|^{2p} + \varepsilon C(1 + |x|^{2p})
$$

$$
+ C\int_\sigma^\tau\int P^{(\varepsilon)}(\sigma, \gamma, x, y, d\xi, d\eta)(1 + |\xi|^{2p})\, d\gamma.
$$

We recall here the convention about constants stated in §9. Thus,

$$
(12.14) \quad E\{|X^{(\varepsilon)}(\tau)|^{2p}\} \leqq C(1 + |x|^{2p}) + C\int_\sigma^\tau E\{1 + |X^{(\varepsilon)}(\gamma)|^{2p}\}\, d\gamma,
$$

and from this and Gronwall's lemma (12.1) follows.

We note that hypothesis (6.12) is used only in Lemma 12.1 and nowhere else in the proof of the theorem of §6. Hypothesis (6.11) is not used fully in Lemma 12.1 but will be in §15. Similarly, (6.9) will be used fully in §14.

13. **Decomposition into local problems.** In order to effect the desired decomposition of $u^{(\varepsilon)}(\sigma, \tau, x, y)$, the solution of (10.21), and reduce the estimation problem (6.25) into local problems over time intervals of length $\Delta = \varepsilon^{7/4}$, we need some a priori information about the behavior of solutions of (6.21). This information is summarized in the following.

LEMMA 13.1.   *Let $\mathcal{L}_\sigma$ be defined by (6.20) and assume that the hypotheses (6.22) hold. Then (6.21) has a unique classical solution $u(\sigma, \tau, x)$ for $f(x) \in C^{4,\alpha}(R^n)$, $\alpha \geqq 0$, and there is an $\bar{\alpha} > \alpha$ such that $u(\sigma, \tau, x) \in C^{4,\bar{\alpha}}(R^n)$, $0 \leqq \sigma \leqq \tau \leqq T$.*

The proof of this lemma can be obtained by the Itô calculus [12]. Without the factorization hypothesis (6.22 (i)) one must use the theory of Oleinik [27], [28].

Before proceeding with the decomposition into local problems we prove the following lemma which says that, except in an initial layer, $f(x, y)$ may

be replaced by

$$\tilde{f}(x) = \int \bar{P}(dy; x) f(x, y)$$

in studying (6.1) (or (10.21)). This lemma along with (6.25) disposes of (6.24).

LEMMA 13.2. *Let $0 \leqq \sigma < \tau \leqq T$ and define $u^{(\varepsilon)}$ and $\bar{u}^{(\varepsilon)}$ as the solutions of (10.21) with data $f(x, y)$ and $\tilde{f}(x)$ respectively. Suppose $f(x, y)$ is measurable and in $C^{1,\alpha}(R^n)$, $\alpha \geqq 0$ as a function of x. Then, for some $\bar{\alpha} > \alpha$,*

$$(13.1) \qquad \lim_{\varepsilon \to 0} \sup_{x, y} \frac{|u^{(\varepsilon)}(\sigma, \tau, x, y) - \bar{u}^{(\varepsilon)}(\sigma, \tau, x, y)|}{1 + |x|^{\bar{\alpha}}} = 0.$$

PROOF. For this lemma we require that $\Delta = \varepsilon^{7/4}$ which implies that $\Delta/\varepsilon \to 0$, $\Delta/\varepsilon^2 \to \infty$ as $\varepsilon \to 0$. Assume $\sigma < \tau - \Delta$ for $\varepsilon < \varepsilon_0$ say.

The following inequalities are easily verified.

$$
\begin{aligned}
& |u^{(\varepsilon)}(\sigma, \tau, x, y) - \bar{u}^{(\varepsilon)}(\sigma, \tau, x, y)| \\
&= \left| \int P^{(\varepsilon)}(\sigma, \tau - \Delta, x, y, d\bar{x}, d\bar{y}) \right. \\
& \qquad \left. \times [u^{(\varepsilon)}(\tau - \Delta, \tau, \bar{x}, \bar{y}) - \bar{u}^{(\varepsilon)}(\tau - \Delta, \tau, \bar{x}, \bar{y})] \right| \\
&\leqq \left| \int P^{(\varepsilon)}(\sigma, \tau - \Delta, x, y, d\bar{x}, d\bar{y})[u^{(\varepsilon)}(\tau - \Delta, \tau, \bar{x}, \bar{y}) - \tilde{f}(\bar{x})] \right| \\
& \quad + \left| \int P^{(\varepsilon)}(\sigma, \tau - \Delta, x, y, d\bar{x}, d\bar{y})[\bar{u}^{(\varepsilon)}(\tau - \Delta, \tau, \bar{x}, \bar{y}) - \tilde{f}(\bar{x})] \right|.
\end{aligned}
$$

(13.2)

From (10.24) we obtain

$$
\begin{aligned}
& \left| \int P^{(\varepsilon)}(\sigma, \tau - \Delta, x, y, d\bar{x}, d\bar{y})[u^{(\varepsilon)}(\tau - \Delta, \tau, \bar{x}, \bar{y}) - \tilde{f}(\bar{x})] \right| \\
&\leqq \left| \int P^{(\varepsilon)}(\sigma, \tau - \Delta, x, y, d\bar{x}, d\bar{y}) \left[u_0 \left(\frac{\Delta}{\varepsilon^2}, \bar{y}; \bar{x} \right) - \tilde{f}(\bar{x}) \right] \right| \\
& \quad + \frac{1}{\varepsilon} \left| \int P^{(\varepsilon)}(\sigma, \tau - \Delta, x, y, d\bar{x}, d\bar{y}) \right. \\
& \qquad \left. \times \int_{\tau - \Delta}^{\tau} \int P^{(\varepsilon)}(\tau - \Delta, \tau, \bar{x}, \bar{y}, d\xi, d\eta) \right. \\
& \qquad \left. \times F_i \left(\frac{s}{\varepsilon^2}, \xi, \eta \right) u_{0,i} \left(\frac{s - \tau + \Delta}{\varepsilon^2}, \xi, \eta \right) ds \right| \\
&\leqq \left| \int P^{(\varepsilon)}(\sigma, \tau - \Delta, x, y, d\bar{x}, d\bar{y})(1 + |\bar{x}|^\alpha) \right| C\rho \left(\frac{\Delta}{\varepsilon^2} \right) \\
& \quad + C \cdot \frac{\Delta}{\varepsilon} \left| \int P^{(\varepsilon)}(\sigma, \tau - \Delta, x, y, d\bar{x}, d\bar{y})(1 < |\bar{x}|^{\bar{\alpha}}) \right| \\
&\leqq C \left[\rho \left(\frac{\Delta}{\varepsilon^2} \right) + \frac{\Delta}{\varepsilon} \right] (1 + |x|^{\bar{\alpha}}).
\end{aligned}
$$

(13.3)

Here we have employed (6.8), (6.2), (10.9), (11.2) and Lemma 12.1. The second term on the right side of (13.2) is estimated in the same way as (13.3). From (6.9) and the assumption $\Delta = \varepsilon^{7/4}$ the assertion of the lemma follows.

We proceed now with the proof of (6.25) that is, we shall show that the solution $\bar{u}^{(\varepsilon)}$ of the final value problem (10.21) with data $\bar{f}(x)$ and the solution $u(\sigma, \tau, x)$ of (6.21) with data $\bar{f}(x)$ also, (recall $\mathscr{L}$ is independent σ here; cf. remarks in §9) satisfy for some $\bar{\alpha} > \alpha$ the estimate

$$
(13.4) \qquad |\bar{u}^{(\varepsilon)}(\sigma, \tau, x, y) - u(\sigma, \tau, x)| \leqq \varepsilon^{1/4} C(1 + |x|^{\bar{\alpha}}),
$$
$$
0 \leqq \sigma \leqq \tau \leqq T,
$$

where $\bar{f}(x) \in C^{4,\alpha}(R^n)$, $\alpha \geqq 0$. For this purpose we introduce the notation with $\Delta = \varepsilon^{7/4}$,

$$
(13.5) \qquad \sigma_k = \sigma + k\Delta, \qquad k = 0, 1, 2, \cdots, m, \qquad \sigma_0 = \sigma, \qquad \sigma_m = \tau,
$$
$$
X^{(\varepsilon)}(\sigma_k, \sigma, x, y) = X_k^{(\varepsilon)},
$$

where we assume, without loss in generality, that $(\tau - \sigma)/\Delta = m$ takes integer values. This notation is employed throughout in the sequel. The following can be verified easily.

$$
|\bar{u}^{(\varepsilon)}(\sigma, \tau, x, y) - u(\sigma, \tau, x)|
$$
$$
= \left| \sum_{k=1}^{m} E\{u(\sigma_k, \tau, X_k^{(\varepsilon)}) - u(\sigma_{k-1}, \tau, X_{k-1}^{(\varepsilon)})\} \right|
$$
$$
= \left| \sum_{k=1}^{m} \int u(\sigma_k, \tau, \bar{x}) P^{(\varepsilon)}(\sigma, \sigma_k, x, y, d\bar{x}, d\bar{y}) \right.
$$
$$
(13.6) \qquad \left. - \int u(\sigma_{k-1}, \tau, \bar{x}) P^{(\varepsilon)}(\sigma, \sigma_{k-1}, x, y, d\bar{x}, d\bar{y}) \right|
$$
$$
= \left| \sum_{k=1}^{m} \int P^{(\varepsilon)}(\sigma, \sigma_{k-1}, x, y, d\bar{x}, d\bar{y}) \right.
$$
$$
\left. \times \left[\int P^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, \bar{x}, \bar{y}, d\xi, d\eta) u(\sigma_k, \tau, \xi) - u(\sigma_{k-1}, \tau, \bar{x}) \right] \right|.
$$

Define $\bar{u}$ by

$$
(13.7) \quad \bar{u}(\sigma_k, \tau, x) = \int \bar{P}(dy; x) \int P^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, y, d\xi, d\eta) u(\sigma_k, \tau, \xi).
$$

Then, from (13.6) and (13.7) we obtain the decomposition

$$|\bar{u}^{(\varepsilon)}(\sigma, \tau, x, y) - u(\sigma, \tau, x)|$$

$$\leq \sum_{k=1}^{m} \left| \int P^{(\varepsilon)}(\sigma, \sigma_{k-1}, x, y, d\bar{x}, d\bar{y}) \right.$$

(13.8)

$$\times \left. \left[\int P^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, \bar{x}, \bar{y}, d\xi, d\eta) u(\sigma_k, \tau, \xi) - \bar{u}(\sigma_k, \tau, \bar{x}) \right] \right|$$

$$+ \sum_{k=1}^{m} \int P^{(\varepsilon)}(\sigma, \sigma_{k-1}, x, y, d\bar{x}, d\bar{y}) |\bar{u}(\sigma_k, \tau, \bar{x}) - u(\sigma_{k-1}, \tau, \bar{x})|.$$

Thus, to obtain the estimate (13.4) we shall estimate separately the two sums on the right side of (13.8). For $k = 1, 2, \cdots, m$ let $I_{1,k}^{(\varepsilon)}$ and $I_{2,k}^{(\varepsilon)}$ be defined by

$$I_{1,k}^{(\varepsilon)}(x, y) = \left| \int P^{(\varepsilon)}(\sigma, \sigma_{k-1}, x, y, d\bar{x}, d\bar{y}) \right.$$

(13.9)

$$\times \left. \left[\int P^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, \bar{x}, \bar{y}, d\xi, d\eta) u(\sigma_k, \tau, \xi) - \bar{u}(\sigma_k, \tau, \bar{x}) \right] \right|$$

(13.10) $I_{2,k}^{(\varepsilon)}(x) = |\bar{u}(\sigma_k, \tau, x) - u(\sigma_{k-1}, \tau, x)|.$

The estimation of (13.9) and (13.10) is carried out in §§14 and 15. The titles of these sections derive from the significance of the two sums on the right side of (13.8) as we mentioned in §9. At the end of §15 we shall return to (13.8) and complete the derivation of the estimate (13.4).

14. **Local independence.** In this section we prove the following result.

LEMMA 14.1. $I_{1,k}^{(\varepsilon)}(x, y)$ of (13.9) satisfies

(14.1)
$$I_{1,k}^{(\varepsilon)}(x, y) \leq \varepsilon^2 C(1 + |x|^{\bar{\alpha}}), \quad k = 2, 3, \cdots, m,$$

$$I_{1,1}^{(\varepsilon)}(x, y) \leq \varepsilon C(1 + |x|^{\bar{\alpha}}), \quad \bar{\alpha} > \alpha, \text{ integer.}$$

PROOF. Let $g^{(k)}(x) = u(\sigma_k, \tau, x)$ with $\bar{f}(x) \in C^{4,\alpha}(R^n)$ so that $g^{(k)}(x) \in C^{4,\bar{\alpha}}(R^n)$ for all $k = 1, 2, \cdots, m$ and $0 \leq \sigma \leq \tau \leq T$. Define $w_k^{(\varepsilon)}$ by

(14.2) $w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, y)$

$$= \int P^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, y, d\bar{x}, d\bar{y}) g^{(k)}(\bar{x}).$$

From (13.7) we have

$$\bar{u}(\sigma_k, \tau, x) = \int \bar{P}(dy; x) \int P^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, y, d\bar{x}, d\bar{y}) g^{(k)}(\bar{x})$$

(14.3)

$$= \int \bar{P}(dy; x) w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, y).$$

Thus,

$$I_{1,k}^{(\varepsilon)}(x, y) = \left| \int P^{(\varepsilon)}(\sigma, \sigma_{k-1}, x, y, d\bar{x}, d\bar{y}) \right.$$

(14.4)

$$\left. \times \left[w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, \bar{x}, \bar{y}) - \int \bar{P}(d\zeta; \bar{x}) w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, \bar{x}, \zeta) \right] \right|.$$

We introduce additional notation as follows.

(14.5)
$$v_k^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y) = \int P^{(\varepsilon)}(\sigma, \sigma_{k-1}, x, y, d\bar{x}, d\bar{y})$$

$$\times w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, \bar{x}, \bar{y}).$$

(14.6) $v_{k0}^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y) = \int P_0\left(\frac{\sigma_{k-1} - \sigma}{\varepsilon^2}, y, d\bar{y}; x\right) w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, \bar{y}),$

$$\bar{v}_k^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y) = \int P^{(\varepsilon)}(\sigma, \sigma_{k-1}, x, y, d\bar{x}, d\bar{y})$$

(14.7)

$$\times \int \bar{P}(d\zeta; \bar{x}) w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, \bar{x}, \zeta).$$

From the parametrix identity (10.24) we obtain the identities

$$v_k^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y) = v_{k0}^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y)$$

(14.8)

$$+ \frac{1}{\varepsilon} \int_\sigma^{\sigma_{k-1}} \int P^{(\varepsilon)}(\sigma, s, x, y, d\bar{x}, d\bar{y}) F_i\left(\frac{s}{\varepsilon^2}, \bar{x}, \bar{y}\right)$$

$$\times v_{k0,i}^{(\varepsilon)}(s, \sigma_{k-1}, \sigma_k, \bar{x}, \bar{y}),$$

$$\bar{v}_k^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y)$$

$$= \bar{u}(\sigma_k, \tau, x)$$

(14.9)

$$+ \frac{1}{\varepsilon} \int_\sigma^{\sigma_{k-1}} \int P^{(\varepsilon)}(\sigma, s, x, y, d\bar{x}, d\bar{y}) F_i\left(\frac{s}{\varepsilon^2}, \bar{x}, \bar{y}\right) \bar{u}_{,i}(\sigma_k, \tau, \bar{x}).$$

Similarly, from (10.23) we obtain the identity

$$w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, y) = g^{(k)}(x) + \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \int P_0\left(\frac{s - \sigma_{k-1}}{\varepsilon^2}, y, d\bar{y}; x\right)$$

(14.10)

$$\times F_j\left(\frac{s}{\varepsilon^2}, x, \bar{y}\right) w_{k,j}^{(\varepsilon)}(s, \sigma_k, x, \bar{y}) ds.$$

From (14.4), (14.5)–(14.7) and the identities (14.8) and (14.9) we have

$$I_{1,k}^{(\varepsilon)}(x, y) = |v_k^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y) - \bar{v}_k^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y)|$$

$$\leqq |v_{k0}^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y) - \bar{u}(\sigma_k, \tau, x)|$$

(14.11)
$$+ \left| \frac{1}{\varepsilon} \int_\sigma^{\sigma_{k-1}} \int P^{(\varepsilon)}(\sigma, s, x, y, d\bar{x}, d\bar{y}) F_i \left(\frac{s}{\varepsilon^2}, \bar{x}, \bar{y} \right) \right.$$

$$\left. \times \; [v_{k0}^{(\varepsilon)}(s, \sigma_{k-1}, \sigma_k, \bar{x}, \bar{y}) - \bar{u}(\sigma_k, \tau, \bar{x})]_{,j} \, ds \right|.$$

Furthermore, from the definition of $v_{k0}^{(\varepsilon)}$ and $\bar{u}$ and from the identity (14.10) it follows that

$$|v_{k0}^{(\varepsilon)}(\sigma, \sigma_{k-1}, \sigma_k, x, y) - \bar{u}(\sigma_k, \tau, x)|$$

$$= \left| \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \int \left[P_0 \left(\frac{\sigma - \sigma_{k-1}}{\varepsilon^2}, y, d\bar{y}; x \right) - \bar{P}(d\bar{y}; x) \right] \right.$$

$$\left. \times \int P_0 \left(\frac{s - \sigma_{k-1}}{\varepsilon^2}, \bar{y}, d\zeta; x \right) F_j \left(\frac{s}{\varepsilon^2}, x, \zeta \right) w_{k,j}^{(\varepsilon)}(s, \sigma_k, x, \zeta) \, ds \right|$$

$$= \left| \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \int \left[P_0 \left(\frac{s - \sigma}{\varepsilon^2}, y, d\zeta; x \right) - \bar{P}(d\zeta; x) \right] \right.$$

$$\left. \times \; F_j \left(\frac{s}{\varepsilon^2}, x, \zeta \right) w_{k,j}^{(\varepsilon)}(s, \sigma_k, x, \zeta) \, ds \right|.$$

Substituting this into (14.11) and using (14.10) again we obtain finally the desired inequality for estimating $I_{1,k}^{(\varepsilon)}$:

$$I_{1,k}^{(\varepsilon)}(x, y) \leqq \left| \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \int \left[P_0 \left(\frac{s - \sigma}{\varepsilon^2}, y, d\zeta; x \right) - \bar{P}(d\zeta; x) \right] \right.$$

$$\left. \times \; F_j \left(\frac{s}{\varepsilon^2}, x, \zeta \right) w_{k,j}^{(\varepsilon)}(s, \sigma_k, x, \zeta) \, ds \right|$$

(14.12)
$$+ \frac{1}{\varepsilon^2} \left| \int_\sigma^{\sigma_{k-1}} \int_{\sigma_{k-1}}^{\sigma_k} P^{(\varepsilon)}(\sigma, s, x, y, d\bar{x}, d\bar{y}) F_i \left(\frac{s}{\varepsilon^2}, \bar{x}, \bar{y} \right) \right.$$

$$\times \left\{ \int \int \left[P_0 \left(\frac{t - s}{\varepsilon^2}, \bar{y}, d\zeta; x \right) - \bar{P}(d\zeta; \bar{x}) \right] \right.$$

$$\left. \left. \times \; F_j \left(\frac{t}{\varepsilon^2}, \bar{x}, \zeta \right) w_{k,j}^{(\varepsilon)}(t, \sigma_k, \bar{x}, \zeta) \right\}_{,i} \, dt \, ds \right|.$$

Note that $k=1$ the second term on the right side of (14.12) is zero.

Let us estimate the term with the factor $1/\varepsilon$ on the right side of (14.7). Since $\Delta = \varepsilon^{7/4}$, Lemma 11.1 tells us that $|w_{k,j}^{(\varepsilon)}(s, \sigma, x, \zeta)|$ is bounded by a

power of $|x|$. On using hypotheses (6.2) and (6.8) we find that

$$\left| \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \int \mu_0\left(\frac{s-\sigma}{\varepsilon^2}, y, d\zeta; x\right) F_j\left(\frac{s}{\varepsilon^2}, x, \zeta\right) w_{k,j}^{(\varepsilon)}(s, \sigma_k, x, \zeta) \, ds \right|$$

$$(14.13) \quad \leqq \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \rho\left(\frac{s-\sigma}{\varepsilon^2}\right) ds C(1+|x|^{\bar{\alpha}}), \quad \bar{\alpha} \text{ some nonnegative integer,}$$

$$\leqq \varepsilon \rho^{4/5}(\Delta/\varepsilon^2) C(1+|x|^{\bar{\alpha}}), \quad \text{for } k = 2, 3, \cdots, m,$$

$$\leqq \varepsilon^2 C(1+|x|^{\bar{\alpha}}).$$

Here we have used hypothesis (6.9) and the monotonicity of ρ. It is at this point that the full strength of (6.9) is used. For $k=1$ we have the obvious estimate $\leqq \varepsilon C(1+|x|^{\bar{\alpha}})$.

The second term on the right side of (14.12) is estimated using Lemma 10.1, Lemma 11.1 and Lemma 12.1. It is less than or equal to ($k \geqq 2$ here)

$$\frac{1}{\varepsilon^2} \int_{\sigma}^{\sigma_{k-1}} \int_{\sigma_{k-1}}^{\sigma_k} \rho\left(\frac{t-s}{2\varepsilon^2}\right) dt \, ds C(1+|x|^{\bar{\alpha}}), \quad \bar{\alpha} \text{ some nonnegative integer,}$$

$$\leqq \varepsilon^2 C(1+|x|^{\bar{\alpha}}).$$

The last inequality follows from (6.9) and the monotonicity of ρ. The proof of the lemma is complete.

Note that the results of this section are independent of the centering conditions (6.11) and (6.12).

15. **Local Taylor expansion.** In this section we prove the following estimate.

LEMMA 15.1. *For $k=1, 2, \cdots, m$ and for some $\bar{\alpha} > \alpha$ integer,*

$$(15.1) \quad I_{2,k}^{(\varepsilon)}(x) = |\bar{u}(\sigma_k, \tau, x) - u(\sigma_{k-1}, \tau, x)| \leqq \varepsilon^2 C(1+|x|^{\bar{\alpha}}).$$

PROOF. Again we denote $u(\sigma_k, \tau, x)$ by $g^{(k)}(x)$. We also recall the remarks of §9 concerning explicit slow-time dependence namely, that without loss in generality we may suppress it in the proof. Thus, the operator $\mathscr{L}$ in (6.21) is independent of σ. By integrating (6.21) and iterating the integral equation once,[9] and recalling the definition (13.7) of $\bar{u}$, it follows that

$$I_{2,k}^{(\varepsilon)}(x) = \left| \int \bar{P}(dy; x) \int P^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, y, d\bar{x}, d\bar{y}) g^{(k)}(\bar{x}) \right.$$

$$(15.2) \qquad \left. - g^{(k)}(x) - \Delta\mathscr{L} g^{(k)}(x) - \int_{\sigma_{k-1}}^{\sigma_k} (s - \sigma_{k-1})\mathscr{L}^2 u(s, \tau, x) \, ds \right|.$$

Here we employ the notation introduced by (13.5) with $\Delta = \varepsilon^{7/4}$.

---

[9] Iteration is a legitimate procedure in view of (6.22) and Lemma 13.1.

Let $w_k^{(\varepsilon)}$ be defined by (14.2). By iterating (10.24) twice we obtain the identity

$$
w_k^{(\varepsilon)}(\sigma_{k-1}, \sigma_k, x, y)
$$

$$
= g^{(k)}(x) + \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} F_i\left(\frac{s}{\varepsilon^2}, x, y\right) g_{,i}^{(k)}(x)\, ds
$$

$$
+ \frac{1}{\varepsilon^2} \int_{\sigma_{k-1}}^{\sigma_k} \int_{\sigma_{k-1}}^{s_1} F_i\left(\frac{s_2}{\varepsilon^2}, x, y\right) \left[\int P_0\left(\frac{s_1 - s_2}{\varepsilon^2}, y, d\bar{y}; x\right) \right.
$$

(15.3)
$$
\left. \times F_j\left(\frac{s_1}{\varepsilon^2}, x, \bar{y}\right) g_{,j}^{(k)}(x) \right]_{,i} ds_2\, ds_1
$$

$$
+ \frac{1}{\varepsilon^3} \int_{\sigma_{k-1}}^{\sigma_k} \int_{\sigma_{k-1}}^{s_1} \int_{\sigma_{k-1}}^{s_2} \int P^{(\varepsilon)}(\sigma_{k-1}, s_3, x, y, d\bar{x}, dy_3) F_i\left(\frac{s_3}{\varepsilon^2}, \bar{x}, y_3\right)
$$

$$
\times \left\{ \int P_0\left(\frac{s_2 - s_3}{\varepsilon^2}, y_3, dy_2; \bar{x}\right) F_j\left(\frac{s_2}{\varepsilon^2}, \bar{x}, y^2\right) \right.
$$

$$
\left. \times \left[\int P_0\left(\frac{s_1 - s_2}{\varepsilon^2}, y_2, dy_1; \bar{x}\right) F_l\left(\frac{s_1}{\varepsilon^2}, \bar{x}, y_1\right) g_{,l}^{(k)}(\bar{x}) \right]_{,j} \right\}_{,i} ds_3\, ds_2\, ds_1.
$$

Hypothesis (6.2) and Lemma 13.1 as well as a mild extension of the estimates of §10 concerning P_0, yield the result that the last term in (15.3) is $\leq C(\Delta/\varepsilon)^3(1+|x|^{\bar{\alpha}})$ for some $\bar{\alpha} > \alpha$. Using this estimate and (15.3) in (15.2) we obtain

$$
I_{2,k}^{(\varepsilon)}(x) \leq \left| \frac{1}{\varepsilon} \int_{\sigma_{k-1}}^{\sigma_k} \int \bar{P}(dy; x) F_i\left(\frac{s}{\varepsilon^2}, x, y\right)\, ds g_{,i}^{(k)}(x) \right|
$$

$$
+ \Delta \left| \frac{1}{\varepsilon^2 \Delta} \int_{\sigma_{k-1}}^{\sigma_k} \int_{\sigma_{k-1}}^{s_1} \int \bar{P}(dy; x) F_i\left(\frac{s_2}{\varepsilon^2}, x, y\right) \right.
$$

$$
\times \left[\int \mu_0\left(\frac{s_1 - s_2}{\varepsilon^2}, y, d\bar{y}; x\right) F_j\left(\frac{s_1}{\varepsilon^2}, x, \bar{y}\right) g_{,j}^{(k)}(x) \right]_{,i} ds_2\, ds_1
$$

(15.4)
$$
\left. - \mathscr{L} g^{(k)}x)) \right|
$$

$$
+ \left| \frac{1}{\varepsilon^2} \int_{\sigma_{k-1}}^{\sigma_k} \int_{\sigma_{k-1}}^{s_1} \int \bar{P}(dy; x) F_i\left(\frac{s_2}{\varepsilon^2}, x, y\right)\, ds_2 \right.
$$

$$
\left. \times \left[\int \bar{P}(d\bar{y}; x) F_j\left(\frac{s_1}{\varepsilon^2}, x, \bar{y}\right) g_{,j}^{(k)}(x) \right]_{,i} ds_1 \right|
$$

$$
+ (\Delta/\varepsilon)^3 C(1 + |x|^{\bar{\alpha}}) + \Delta^2 C(1 + |x|^{\bar{\alpha}}).
$$

Here we have also used Lemma 13.1 and hypothesis (6.22) to estimate $\mathscr{L}^2 u$.

From (6.11) it follows that the first term on the right side of (15.4) is $\leq \varepsilon(\varepsilon^2/\Delta)^4 C(1+|x|^{\bar{\alpha}})=\varepsilon^2 C(1+|x|^{\bar{\alpha}})$. It is here that the full strength of (6.11) is used. Hypotheses (6.15) and (6.16) yield for the second term on the right side of (15.4) the estimate $\leq \varepsilon^2 C(1+|x|^{\bar{\alpha}})$. The third term is also $\leq \varepsilon^2 C(1+|x|^{\bar{\alpha}})$, a conclusion following from (6.11) and Lemma 10.1. The proof of Lemma 15.1 is complete.

Note that hypothesis (6.9) was not used in this lemma as should be expected.

Let us now complete the proof of (13.4). Returning to (13.8) we see that the lemmas of §§14 and 15 yield the estimate

$$|u^{(\varepsilon)}(\sigma, \tau, x, y) - u(\sigma, \tau, x)|$$

(15.5)
$$\leq [(\tau - \sigma)/\Delta]\varepsilon^2 C(1 + |x|^{\bar{\alpha}}) + \varepsilon C(1 + |x|^{\bar{\alpha}})$$
$$= (\tau - \sigma)\varepsilon^{1/4} C(1 + |x|^{\bar{\alpha}}) + \varepsilon C(1 + |x|^{\bar{\alpha}}).$$

The $O(\varepsilon)$ term follows from the $I_{1,1}^{(\varepsilon)}$ estimate in (14.1). Since $0 \leq \tau - \sigma \leq T$ (15.5) gives us the desired estimate (13.4).

16. **Weak convergence.** From the estimates of the preceding section it follows that the finite dimensional distributions of $X^{(\varepsilon)}$ converge to those of the diffusion Markov process $X^{(0)}$ with generator $\mathscr{L}_\sigma$ given by (6.20). Suppose that $x \in U \subset R^n$ a compact set. Recall that $X^{(\varepsilon)}(\tau)$ is continuous. In order to show that it converges weakly to $X^{(0)}$ it is sufficient to show, according to a well-known theorem [3, p. 450], that for any $0 \leq \sigma \leq \tau \leq T$ there are constants C and $\beta > 0$ such that

(16.1) $E\{|X^{(\varepsilon)}(\tau) - X^{(\varepsilon)}(\sigma)|^4\} \leq C(\tau - \sigma)^{1+\beta}.$

Here $X^{(\varepsilon)}(\tau)$ denotes the process $X^{(\varepsilon)}(\tau, 0, x, y)$ with $x \in U$. We shall use the results of the preceding sections to prove (16.1) with $\beta = \frac{1}{7}$.

From the identity

(16.2) $E\{|X^{(\varepsilon)}(\tau) - X^{(\varepsilon)}(\sigma)|^4\} = \int P^{(\varepsilon)}(0, \sigma, x, y, d\bar{x}, d\bar{y})$
$$\times E\{|X^{(\varepsilon)}(\tau, \sigma, \bar{x}, \bar{y}) - \bar{x}|^4\}$$

and Lemma 12.1, it follows that it is sufficient to show that, for $0 \leq \sigma \leq \tau \leq T$,

(16.3) $E\{|X^{(\varepsilon)}(\tau, \sigma, x, y) - x|^4\} \leq C(\tau - \sigma)^{1+\beta}(1 + |x|^{\bar{\alpha}}),$

where $\bar{\alpha}$ is some nonnegative integer. To prove (16.3) we consider the cases $\tau - \sigma \geq \varepsilon^{7/4}$ and $\tau - \sigma < \varepsilon^{7/4}$ separately.

Consider first the case $\tau - \sigma \geq \varepsilon^{7/4}$. We shall use (15.5) to obtain (16.3) with $\beta = \frac{1}{7}$. For this purpose let $f(x) = |x - z|^4$, regarding z as a parameter, and denote the solution of (6.21), with this f as data, by $u(\sigma, \tau, x; z)$. Denote by $u^{(\varepsilon)}(\sigma, \tau, x, y; z)$ the solution of (10.21) with the same data

$f(x) = |x - z|^4$. From (15.5), when $\tau - \sigma \geq \varepsilon^{7/4}$, we obtain

$$E\{|X^{(\varepsilon)}(\tau, \sigma, x, y) - x|^4\}$$
$$= u^{(\varepsilon)}(\sigma, \tau, x, y; x)$$

(16.4)
$$\leq |u^{(\varepsilon)}(\sigma, \tau, x, y; x) - u(\sigma, \tau, x; x)| + u(\sigma, \tau, x; x)$$
$$\leq (\tau - \sigma)^{1+1/7}C(1 + |x|^{\bar{a}}) + (\tau - \sigma)^2 C(1 + |x|^{\bar{a}}) + I_{1,1}^{(\varepsilon)}(x, y; x).$$

Here we have used the fact that $u(\sigma, \tau, x; x) \leq (\tau-\sigma)^2 C(1+|x|^{\bar{a}})$ which can be obtained by Itô's calculus [12], [29] or otherwise. We have denoted by $I_{1,1}^{(\varepsilon)}(x, y; z)$ the quantity defined by (14.4) when $k=1$ and $f(x)=|x-z|^4$. Note that the estimate (14.1) is too crude for our purposes here so we shall reconsider the estimation of $I_{1,1}^{(\varepsilon)}$. We shall show that

(16.5)
$$I_{1,1}^{(\varepsilon)}(x, y; x) \leq C(\tau - \sigma)^{1+2/7}(1 + |x|^{\bar{a}}).$$

This estimate and (16.4) yield (16.3) in the case $\tau - \sigma \geq \varepsilon^{7/4}$.

Before proceeding to the proof of (16.5), let us show that (16.3) is valid also when $\tau - \sigma < \varepsilon^{7/4}$. For this purpose we employ the identity (15.3) with $g^{(k)}(x) = |x-z|^4$, $\sigma_{k-1} = \sigma$, $\sigma_k = \tau$. On the left side of (15.3) we obtain $E\{|X^{(\varepsilon)}(\tau, \sigma, x, y) - x|^4\}$ by setting $z = x$. On the right side the first three terms drop out on setting $z = x$ and the last term is less than or equal to

(16.6)
$$((\tau - \sigma)/\varepsilon)^3 C(1 + |x|^{\bar{a}}) \leq (\tau - \sigma)^{1+2/7} C(1 + |x|^{\bar{a}}),$$

where we have used the hypothesis $\tau - \sigma < \varepsilon^{7/4}$. Thus, (16.3) is valid with $\beta = \frac{1}{7}$.

Let us now prove (16.5). From (14.2) and (14.3) we obtain

$$I_{1,1}^{(\varepsilon)}(x, y; x) =$$

(16.7)
$$\left| \int \int P^{(\varepsilon)}(\sigma, \sigma + \Delta, x, y, d\bar{x}, d\bar{y}) u(\sigma + \Delta, \tau, \bar{x}; x) \right.$$
$$\left. - \int \bar{P}(d\zeta; x) \int P^{(\varepsilon)}(\sigma, \sigma + \Delta, x, \zeta, d\bar{x}, d\bar{y}) u(\sigma + \Delta, \tau, \bar{x}; x) \right|.$$

The elementary inequality

(16.8)
$$u(\sigma + \Delta, \tau, \bar{x}; x) = E\{|X^{(0)}(\tau, \sigma + \Delta, \bar{x}) - x|^4\}$$
$$\leq 8u(\sigma + \Delta, \tau, \bar{x}; \bar{x}) + 8|\bar{x} - x|^4,$$

along with the estimate $u(\sigma, \tau, x; x) \leq C(\tau-\sigma)^2(1+|x|^{\bar{a}})$ yield

$$I_{1,1}^{(\varepsilon)}(x, y; x)$$
$$\leq C(\tau - \sigma)^2(1 + |x|^{\bar{a}})$$

(16.9)
$$+ \left| 8 \int P^{(\varepsilon)}(\sigma, \sigma + \Delta, x, y, d\bar{x}, d\bar{y}) |\bar{x} - x|^4 \right|$$
$$+ \left| 8 \int \bar{P}(d\zeta; x) \int P^{(\varepsilon)}(\sigma, \sigma + \Delta, x, y, d\bar{x}, d\bar{y}) |\bar{x} - x|^4 \right|.$$

The last two terms on the right side of (16.8) can be estimated with the use of (15.3) in the same manner that (16.6) was obtained. Both of these terms are less than or equal to

$$(\Delta/\varepsilon)^3 C(1 + |x|^{\tilde{\alpha}}) = \varepsilon^{9/4} C(1 + |x|^{\tilde{\alpha}}) \leqq (\tau - \sigma)^{1+2/7} C(1 + |x|^{\tilde{\alpha}}),$$

where we have used the hypothesis $\tau - \sigma \geqq \varepsilon^{7/4}$. Thus, (16.5) holds. The proof of the theorem of §6 is complete.

REFERENCES

1. A. N. Kolmogorov, *Über die analytischen Methoden in der Wahrscheinlichkeits-rechnung*, Math. Ann. **104** (1931), 415–458.

2. W. Feller, *Zur Theorie der stochastischen Prozesse*, Math. Ann. **113** (1936), 113–160.

3. I. I. Gihman and A. V. Skorohod, *Introduction to the theory of random processes*, "Nauka", Moscow, 1965, English transl., Saunders, Philadelphia, Pa., 1969. MR **33**, #6689; **40** #923.

4. K. M. Case and P. F. Zweifel, *Linear transport theory*, Addison-Wesley, Reading, Mass., 1967. MR **37** #1140.

5. A. Khinchine, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933.

6. H. A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Physica **7** (1940), 284–304. MR **2**, 140.

7. M. F. Norman, *Markov processes and learning models*, Academic Press, New York, 1972.

8. M. Iosifescu and R. Theodorescu, *Random processes and learning*, Die Grundlehren der math. Wissenschaften, Band 150, Springer-Verlag, New York, 1969. MR **45** #2781.

9. R. Hersh, *Random evolutions: a survey of results and problems*, Rocky Mountain Math. J. (to appear).

10. M. Pinsky, *Multiplicative operator functionals and their asymptotic properties*, Advances in Probability, vol. 3, Dekker, New York (to appear).

11. T. G. Kurtz, *A limit theorem for perturbed operator semigroups with application to random evolutions*, J. Functional Analysis **12** (1973), 55–67.

12. I. I. Gihman and A. V. Skorohod, *Stochastic differential equations*, "Naukova Dumka", Kiev, 1968; English transl., Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 72, Springer-Verlag, Berlin and New York, 1972. MR **41** #7777.

13. R. Z. Has'minskiĭ (Khasminskii), *On stochastic processes defined by differential equations with a small parameter*, Teor. Verojatnost. i Primenen. **11** (1966), 240–259= Theor. Probability Appl. **11** (1966), 211–228. MR **34** #3636.

14. D. W. Stroock, *Some stochastic processes which arise from a model of the motion of a bacterium* (to appear).

15. L. Baggett and D. W. Stroock, *An ergodic theorem for Poisson processes on a compact group with applications to random evolutions*, J. Functional Analysis (to appear).

16. S. Chandrasekhar, *Stochastic problems in Physics and Astronomy*, Rev. Modern Phys. **15** (1943), 1–89. MR **4**, 248.

17. A. M. Il'in and R. Z. Has'minskiĭ (Khasminskiĭ), *On the equations of Brownian motion*, Teor. Verojatnost. i Primenen. **9** (1964), 421–444. (Russian) MR **29** #5283.

18. E. Nelson, *Dynamical theories of Brownian motion*, Princeton Univ. Press, Princeton, N.J., 1967. MR **35** #5001.

19. K. M. Case, "The soluble boundary value problems of transport theory," in *The Boltzmann equation*, F. A. Grunbaum (editor), Courant Inst. Lecture Notes, 1972.

20. R. D. Richtmyer and K. W. Morton, *Difference methods for initial-value problems*, 2nd ed., Interscience Tracts in Pure and Appl. Math., no. 4, Interscience, New York, 1967. MR **36** #3515.

21. G. C. Papanicolaou and W. Kohler, *Asymptotic theory of mixing stochastic stochastic ordinary differential equations*, Comm. Pure Appl. Math. (to appear).

22. R. Z. Has'minskiĭ (Khasminskii), *On the principle of averaging for Itô's stochastic differential equations*, Kybernetika (Prague) **4** (1968), 260–279. (Russian) MR **41** #4681.

23. W. Kohler and G. C. Papanicolaou, *Limit theorems for stochastic equations with rapidly varying components* (to appear).

24. R. Kubo, *Stochastic Liouville equations*, J. Mathematical Phys. **4** (1963), 174–183. MR **26** #7370.

25. G. S. Agarwal, "Master equation methods in quantum optics," in *Progress in Optics*. Vol. 11, E. Wolf (editor), North-Holland, Amsterdam, 1973.

26. R. J. Griego and R. Hersh, *Theory of random evolutions with applications to partial differential equations*, Trans. Amer. Math. Soc. **156** (1971), 405–418. MR **43** #1261.

27. O. A. Oleĭnik and E. V. Radkevič, *Equations of second order with nonnegative characteristic form*, Itogi Nauki Mat. Anal. 1969, VINITI, Moscow, 1971; English transl., Plenum Press, New York, 1973.

28. D. W. Stroock and S. R. S. Varadhan, *On degenerate elliptic-parabolic operators of second order and their associated diffusions*, Comm. Pure Appl. Math. **25** (1972), 651–714.

29. H. P. McKean, Jr., *Stochastic integrals*, Probability and Math. Statist., no. 5, Academic Press, New York, 1969. MR **40** #947.

30. ———, Jr., *Fluctuations in the kinetic theory of gases* (to appear).

31. N. N. Bogoljubov and Ju. A. Mitropol'skiĭ, *Asymptotic methods in the theory of nonlinear oscillations*, 2nd rev. ed., Fizmatgiz, Moscow, 1958; English transl., Gordon and Breach, New York, 1961. MR **20** #6812; **25** #5242.

32. H. Grad, *Solution of the Boltzmann equation in an unbounded domain*, Comm. Pure Appl. Math. **18** (1965), 345–354. MR **32** #8913.

33. G. E. Uhlenbeck and L. S. Ornstein, *On the theory of Brownian motion*, Phys. Rev. **36** (1930), 823–841.

34. Ming Chen Wang and G. E. Uhlenbeck, *On the theory of Brownian motion. II*, Rev. Modern Phys. **17** (1945), 323–342. MR **7**, 130.

35. R. Z. Has'minskiĭ (Khasminskii), *The averaging principle for parabolic and elliptic differential equations and Markov processes with small diffusion*, Teor. Verojatnost. i Primenen. **8** (1963), 3–25=Theor. Probability Appl. **8** (1963), 1–21. MR **28** #4253.

36. V. M. Volosov, *Averaging in systems of ordinary differential equations*, Uspehi Mat. Nauk **17** (1962), no. 6 (108), 3–126=Russian Math. Surveys **17** (1962), no. 6, 1–126. MR **26** #3976.

37. R. S. Ellis and M. A. Pinsky, *Asymptotic nonuniqueness of the Navier-Stokes equation in kinetic theory*, Bull. Amer. Math. Soc. **80** (1974), 1160–1164.

38. H. Grad, *Singular and nonuniform limits of solutions of the Boltzmann equation*, SIAM-AMS Proc., vol. 1, Amer. Math. Soc., Providence, R.I., 1969, pp. 269–308. MR **40** #8410.

39. E. W. Larsen and J. B. Keller, *Asymptotic solution of neutron transport problems for small mean free paths*, J. Mathematical Phys. **15** (1974), 75–81.

40. E. W. Larsen, *Solutions of the steady, one-speed neutron transport equation for small mean free paths*, J. Mathematical Phys. **15** (1974), 299–305.

41. B. White, Dissertation, New York University, Sept. 1974.

42. N. Ja. Vilenkin, *Special functions and the theory of group representations*, "Nauka", Moscow, 1965; English transl., Transl. Math. Monographs, vol. 22, Amer. Math. Soc., Providence, R.I., 1968. MR **35** #420; **37** #5429.

43. G. C. Papanicolaou and S. R. S. Varadhan, *A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations*, Comm. Pure Appl. Math. **26** (1973), 497–524.

44. G. C. Papanicolaou and J. B. Keller, *Stochastic differential equations with applications to random harmonic oscillators and wave propagation in random media*, SIAM J. Appl. Math. **21** (1971), 287–305. MR **46** #8304.