
ASYMPTOTIC ANALYSIS OF STOCHASTIC EQUATIONS*

G. C. Papanicolaou

1. INTRODUCTION

The purpose of this survey is to collect together and present with a minimum of mathematical detail some problems and results on asymptotics for stochastic equations. Since the subject is very large and rapidly expanding, the examples presented here deal only with one general problem: the effect of rapid, noisy fluctuations on systems (ordinary differential equations).

More specifically, we consider questions (in the context of examples) like the following. How does one measure the effect of noise on systems (scaling)? How do we identify the situations (asymptotic limits) where noise phenomena are fully developed (white noise)?

* Research supported by the National Science Foundation under Grant No. MCS75-09837.
How do we develop a “fluctuation theory” for solutions of differential equations with random parameters? How do we treat problems where noise phenomena are fully developed but have a weak effect?

In Sections 4 to 16 we give some answers to these and a few other related questions. We also provide references to the extensive literature.

2. BRIEF SURVEY OF SOME FACTS FROM PROBABILITY

We refer to [1]–[5] for systematic exposition of the theory of stochastic processes.

A real valued stochastic process with parameter set \( T = [0, \infty) \) or \( T = \{1, 2, 3, \ldots\} \) is a family \( \{x(t), t \in T\} \) of real valued random variables, i.e., a consistent family of distributions \( \{F_{x_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n)\} \) where for each \( t_1, t_2, \ldots, t_n \in T \) the joint distribution function of \( x(t_1), x(t_2), \ldots, x(t_n) \) is given by the corresponding member of this family of distributions. Consistent means that marginal distributions match when the marginal index set \( \{t_1, t_2, \ldots, t_n\} \) is the same.

Given such a stochastic process one can then construct a probability space \((\Omega, \mathcal{F}, P)\) (where \( \Omega \) is a set, \( \mathcal{F} \) is a \( \sigma \)-algebra of subsets, and \( P \) is a probability measure on \( \mathcal{F} \)) and functions \( x(t, \omega) \) on \( T \times \Omega \) into \( \mathbb{R}^1 \) such that

\[
P(\omega \in \Omega : x(t_1) \in A_1, x(t_2) \in A_2, \ldots, x(t_n) \in A_n) = \int_{A_1} \cdots \int_{A_n} dF_{x(t_1), t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n).
\]

Here, \( t_1, t_2, \ldots, t_n \in T \) and \( A_1, A_2, \ldots, A_n \) are Borel subsets of \( \mathbb{R}^1 \). The space \( \Omega \) can be taken to be the space of all functions on \( T \) in which case for \( \omega \in \Omega \), \( x(t, \omega) \) is the value at \( t \) of the function \( \omega \).

If the process \( \{x(t), t \in T\} \) is stochastically continuous, i.e.,

\[
\lim_{h \downarrow 0} P(|x(t + h) - x(t)| > \delta) = 0, \quad \text{for all } \delta > 0,
\]

then the process can be chosen so that \( x(t, \omega) \) is a jointly measurable function of \( t \) and \( \omega \) and the process is separable. This means that the probability of events that involve uncountably many \( t \)'s is the same when only a countable dense set of \( t \)'s is used.

Let \( \mathcal{F} \) be the \( \sigma \)-algebra generated by events that involve the path \( x(s, \omega) \) only for times \( 0 \leq s \leq t \). Clearly the \( \sigma \)-algebras \( \mathcal{F}_t \) increase with \( t \) and they are contained in \( \mathcal{F} \).

Given a random variable \( Y \) on \( \Omega \), i.e., a real valued measurable function on \( \Omega \) such that

\[
E(|Y|) = \int_{\Omega} |Y(\omega)|P(d\omega) < \infty,
\]

we define the conditional expectation \( E(Y|\mathcal{F}_t) \) of \( Y \) given \( \mathcal{F}_t \) as that \( \mathcal{F}_t \) measurable random variable for which

\[
E(\chi_B Y) = E(\chi_B E(Y|\mathcal{F}_t))
\]

for any event \( B \in \mathcal{F} \). Here \( \chi_B(\omega) \) is the characteristic function of the set \( B \) (equal to one if \( \omega \in B \) and zero if \( \omega \notin B \)). Conditional expectations are always well defined up to \( P \) almost everywhere equivalence. Note that

\[
E(Y) = E(E(Y|\mathcal{F}_0)).
\]

The basic classes of processes are the Markov processes, the martingales, the Gaussian processes and the stationary processes, among others. A process \( \{x(t), t \geq 0\} \) is called Markov if the conditional probability of an event that involves the paths for times \( s > t + h \) given the past up to time \( t \), is a function of the present state \( x(t) \). More formally, if \( B \in \mathcal{F} \) and \( B \notin \mathcal{F}_{t+h} \) then \( E(x_B|\mathcal{F}_t) \), which is in general a functional of \( x(s) \) for \( 0 \leq s \leq t \), is a point function of \( x(t) \) for Markov processes. This means that the measure \( P \) on \((\Omega, \mathcal{F})\) (which is what is meant by the process) has relatively simple structure and is in fact determined by the initial distribution

\[
P(x(0) \leq x) = F_0(x)
\]

and the transition probability distributions

\[
P(x(t) \leq y | x(s) = x) = F(s, t, x, y).
\]

Martingales are processes (measures) for which

\[
\sup_{t \geq 0} E(|x(t)|) < \infty
\]

and

\[
E(x(t)|\mathcal{F}_s) = x(s), \quad 0 \leq s \leq t
\]
almost everywhere with respect to $\mathbb{P}$. The importance of this class of processes will become apparent shortly.

Gaussian processes are characterized by the fact that the finite dimensional distributions $F_{t_1,\ldots,t_n}(x_1,\ldots,x_n)$ are all Gaussian. Hence the measure $P$ is completely characterized by the mean function

$$m(t) = E[x(t)]$$

and the covariance function

$$\rho(t, s) = E[(x(t) - m(t))(x(s) - m(s))], \quad t, s \geq 0.$$

Stationary processes are characterized by the property that for any $t_1, t_2,\ldots,t_n \in T$ and $h > 0$, $x(t_1),\ldots,x(t_n)$, and $x(t_1 + h),\ldots,x(t_n + h)$ have the same distribution. A Gaussian process is stationary if and only if the mean function is independent of $t$ and the covariance a function of $|t - s|$ only.

A positive random variable $\tau$ on $\Omega$, the sample space of a process, is called a stopping time if the events $\{\tau \leq t\}$ are $\mathcal{F}_t$ measurable. If $\tau \wedge t = \min(\tau, t)$, the process $x(\tau \wedge t)$ is called the stopped process corresponding to $x(t)$. If $\tau$ is a stopping time, the $\sigma$-algebra of events $A \cap \{\tau \leq t\}$ with $A \in \mathcal{F}_t$ and $t \geq 0$ is denoted by $\mathcal{F}_\tau$. It represents events associated with the trajectories up to the random time $\tau$. If $\sigma \leq \tau$ are two bounded stopping times and $x(t)$ is a martingale then

$$E[x(\tau) | \mathcal{F}_\sigma] = x(\sigma) \quad \text{a.s.}\ P,$$

i.e., the martingale property holds also for random times. This is the optional stopping theorem.

Let us now consider Markov processes in more detail. Clearly it is not necessary to restrict oneself to real valued processes, so we shall assume that the state space of $\{x(t), t \geq 0\}$ is some set $S$ which is usually a complete separable metric space. A time homogeneous Markov process on $S$ is characterized by the following: (i) the initial distribution $\mu(A)$ which is a probability measure on the Borel subsets $\Sigma$ of $S$ (i.e., $A \in \Sigma$) and (ii) the transition probability function $P(t,x,A)$, $t \geq 0$, $x \in S$, $A \in \Sigma$, which is a measurable function of $t$ and $x$ for each $A$ and a probability measure for each $t$ and $x$. Moreover, the Chapman-Kolmogorov equation holds

$$P(t + s, x, A) = \int_S P(t, x, dy)P(s, y, A).$$

The process $\{x(t), t \geq 0\}$ is now constructed so that

$$P(x(t_1) \in A_1, x(t_2) \in A_2,\ldots,x(t_n) \in A_n)$$

$$= \int \mu(dx) \int_{A_1} P(t_1, x, dy_1) \int_{A_2} P(t_2 - t_1, y_1, dy_2) \cdots$$

$$\times \int_{A_n} P(t_n - t_{n-1}, y_{n-1}, dy_n).$$

Let $B(S, \Sigma)$ be the Banach space of bounded measurable functions on $S$ with the sup norm and let $M(S, \Sigma)$ be the Banach space of finite signed measures with the total variation norm. The transition function induces a semigroup of operators $T(t)$ on $B$ and $M$ as follows:

$$T(t)f(x) = \int_S P(t, x, dy)f(y), \quad f \in B,$$

$$\mu T(t)(A) = \int_S \mu(dx)P(t, x, A), \quad \mu \in M.$$
it is called a Feller semigroup and the corresponding process a
Feller process. If \( T(t) : B(S) \to C(S) \), then it is called a strong
Feller semigroup. If
\[
\sup_{x} P\left( t, x, \{d(x, y) > \delta\} \right) = 0(t), \quad t \downarrow 0, \quad \delta > 0,
\]
where \( d(\cdot, \cdot) \) is the distance function, then the trajectories of the
process do not have, with probability one, discontinuities of the
second kind. They can thus be chosen to be right continuous (with
left-hand limits). A right continuous Feller process is also a strong
Markov process, i.e., the Markov property is valid not only when
conditioning with respect to a fixed time \( t \) but also when conditioning
with respect to a stopping time \( \tau \).

The domain \( D(A) \) of the generator of \( T(t) \) is dense in the strongly
continuous center. Moreover if
\[
R_{\lambda} = \int_{0}^{\infty} e^{-\lambda t} T(t) dt, \quad \lambda > 0
\]
is the resolvent operator then, \( R_{\lambda} : D(A) \to \) the strongly continuous
center and \( \|R_{\lambda}\| \leq 1 \). Conversely (Hille-Yoshida theorem) if \( A \)
A densely defined in a Banach space, \( (\lambda - A) : D(A) \to \) the Banach
space for \( \lambda > 0 \) and \( \lambda\|g\| \leq \|g\| \) for \( g \in D(A) \), then \( A \)
is the infinitesimal generator of a strongly continuous semigroup on
the Banach space (all semigroups are contractions \( \|T(t)\| \leq 1 \)).

Since \( A \), the infinitesimal generator, is what is usually given (the
coefficients in the Fokker-Plank equation, for example, or the birth
and death rates), a theorem such as the above describes when the
“coefficients” determine the process uniquely.

Suppose \( T(t) \) is strongly continuous on \( C(S) \), \( S \) is compact,
and let \( x(t) \), the corresponding process, be right continuous. Suppose \( f(x) \)
is in \( D(A) \) and \( Af(x) = 0 \) (\( f \) is then called harmonic). Then the real
valued process \( f(x(t)) \) is a martingale because
\[
E\{f(x(t + s))|\mathcal{F}_s\} = T(t)f(x(s)) \quad \text{a.s.}
\]
\[
= f(x(s)) + \int_{0}^{t} T(s)Af(x(s)) ds
\]
\[
= f(x(s)).
\]

More generally, if \( f(.) \in D(A) \) then
\[
f(x(t)) = \int_{0}^{t} A f(x(s)) ds
\]
is a martingale. Similarly, if for each \( t \), \( f(t, \cdot) \in D(A) \), \( f(t, x) \)
is bounded and has bounded \( t \) derivative then,
\[
f(t, x(t)) = \int_{0}^{t} \left( \frac{\partial}{\partial s} + A \right) f(s, x(s)) ds
\]
is a martingale. Also, if \( u(x) \) is strictly positive and \( u(.) \in D(A) \),
\[
u(x(t)) \exp\left\{ -\int_{0}^{t} \frac{Au(u(x(s)))}{u(x(s))} ds \right\}
\]
is a martingale.

Let us discuss this last example in more detail. Let \( V(x) \) be a
continuous function on \( S \) and let \( V \) denote the operator corresponding
to multiplication by \( V(x) \). If \( A \) is the infinitesimal generator of a
semigroup \( T(t) \) as above, then \( A + V \) is also the generator of a
semigroup which we denote by \( T_{V}(t) \). Then we have that
\[
T_{V}(t)f(x) = E_{\lambda}\left\{ \exp\left[ \int_{0}^{t} V(x(s)) ds \right] f(x(t)) \right\}
\]
where \( E_{\lambda} \) denotes expectation relative to the measure \( P_{\lambda} \) in the sample
space when the initial measure (at time \( t = 0 \) is concentrated at the
point \( x \in S \). To verify the above formula, we assume \( f \in D(A) \) and,
by direct computation we find that
\[
\lim_{t \uparrow t_{0}} \frac{1}{t} \left( T_{V}(t)f - f \right) = Af + Vf,
\]
where we use the right-continuity of the paths. Since \( A + V \) generates
(uniquely) a semigroup and \( T_{V}(t) \) as defined above is a semigroup,
the identification is complete.

Suppose now that \( u(x) \) is strictly positive and \( u(.) \in D(A) \). Then
\( V(x) = -Au(x)/u(x) \) is continuous and so
\[
T_{V}(t)u(x) = E_{\lambda}\left\{ \exp\left[ -\int_{0}^{t} \frac{Au(u(x(s)))}{u(x(s))} ds \right] u(x(t)) \right\}
\]
\[
= u(x)
\]
Because $dT_{\gamma}(t)\mu/dt = 0$ as is easily shown. From this the martingale property follows immediately.

The above examples show that martingales are produced in profusion in the analysis of Markov processes (and in many other situations).

Among the basic properties of martingales, in addition to the optional stopping theorem, we mention the martingale convergence theorems [2] and various important inequalities for them, the simplest being Kolmogorov’s inequality $(x^* = x \vee 0 = \max(x, 0))$

$$P\left\{ \sup_{0 \leq t \leq T} x(t) > \delta \right\} \leq \frac{E[x^*(T)]}{\delta}, \quad \delta > 0.$$  

The systematic use of martingales to study Markov processes (principally diffusion processes) is carried out in [6], [7].

Diffusion processes on $\mathbb{R}^n$, say, are Markov processes with continuous trajectories such that if $P(t, x, A)$ is their transition probability function, it satisfies

(i) $\lim \sup_{t \to 0} \frac{1}{t} \int_{|x-y| > \delta} P(t, x, dy) = 0$,

(ii) $\lim \sup_{t \to 0} \frac{1}{t} \int_{|x-y| \leq \delta} (x-y)P(t, x, dy) = b(x)$

(iii) $\lim \sup_{t \to 0} \frac{1}{t} \int_{|x-y| \leq \delta} (x-y) \otimes (x-y)P(t, x, dy) = a(x) = 0$.

In (i)–(iii) $\delta > 0$ is arbitrary and $b(x)$ and $a(x)$ are $n$-vector and $n \times n$ matrix functions respectively. Condition (i) is called Lindeberg’s condition and it implies continuity of trajectories along with the Feller property for $P(t, x, dy)$. From (i)–(iii) one finds easily that if $f(x)$ is smooth and bounded, then

$$T(t)f(x) = E_x[f(x(t))] = \int P(t, x, dy)f(y)$$

where

$$A = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^{n} b_j(x) \frac{\partial}{\partial x_j}.$$  

Given infinitesimal diffusion coefficients $a(x) = (a_{ij}(x))$ and infinitesimal drift coefficients $b(x) = (b_j(x))$, when does the operator $A$ above determine uniquely the semigroup $T(t)$ and hence the process $(=\text{measure on the space of continuous trajectories in } \mathbb{R}^n)$? This can be answered by applying the Hille-Yoshida theorem, the difficult thing being the description of the range of the resolvent $(\lambda - A)^{-1}, \lambda > 0$. Such questions can be analyzed by a variety of methods, more or less probabilistic. If the coefficients $(a_{ij}(x))$ are bounded continuous and uniformly positive definite

$$\sum_{i,j=1}^{n} a_{ij}(x) \xi_i \xi_j \geq \alpha \sum_{i=1}^{n} \xi_i^2, \quad \alpha > 0,$$

and if the $b_j(x)$ are bounded and measurable one can construct a unique probability measure $P_x$ on $C([0, \infty), \mathbb{R}^n)$ for each $x \in \mathbb{R}^n$ such that

(i) $P_x(x(0)) = x = 1$,

(ii) $f(x(t)) - \int_{0}^{t} A f(x(s)) \, ds$ is a $P_x$ martingale,

for each $f(x)$ smooth and of compact support [6]. This is the martingale approach to diffusions and it provides, of course, a semigroup $T(t)$ which is Feller. If the coefficients are smooth and bounded, ellipticity, not uniform ellipticity, is enough.

Let $\mathcal{D} \subset \mathbb{R}^n$ be a bounded open set and let $\tau$ be the first time the diffusion process $x(t)$ with generator $A$ reaches $\partial \mathcal{D}$, the boundary of $\mathcal{D}$. Suppose that the boundary value problem

$$\frac{\partial u(t, x)}{\partial t} = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u(t, x)}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_{i}(x) \frac{\partial u(t, x)}{\partial x_j}, \quad t > 0, \quad x \in \mathcal{D},$$  

$$u(0, x) = f(x), \quad x \in \partial \mathcal{D},$$  

$$u(t, x) = g(x), \quad x \in \partial \mathcal{D}, \quad t > 0$$

has a smooth bounded solution when the data $f$ and $g$ are smooth.
On applying the optional stopping theorem to the martingale \( u(t - \tau \wedge s, x(\tau \wedge s)), 0 \leq s \leq t, (t \text{ is fixed}) \) we find that
\[
E_x[u(t - \tau \wedge t, x(\tau \wedge t))] = u(t, x).
\]
Thus,
\[
u(t, x) = E_x[f(x(t))]_{\tau(t \in \cdot)} + E_x[g(x(\tau))]_{\tau(t \in \cdot)}
\]
which shows how the solution of the boundary value problem is expressed probabilistically. Many similar problems can be given a probabilistic representation, including the case of Neumann boundary conditions. For a martingale approach to such problems we refer to [8].

Diffusion processes are of sufficient importance that it is desirable to have many different ways of constructing them from their local parameters \( a(x), b(x) \) (and boundary conditions if any). The continuity of the process and the definition of \( a(x) \) and \( b(x) \) above prompt one to attempt to construct the process as a functional of another, canonical process. The canonical process is Brownian motion where \( a(x) = (\delta_0) \) and \( b(x) = 0 \). Thus, for Brownian motion
\[
P(t, x, dy) = \frac{e^{-|x - y|^2 / 2t}}{(2\pi t)^{-\frac{n}{2}}} dy
\]
where \( |x - y|^2 = \sum_{i=1}^{n} (x_i - y_i)^2 \). Now the defining relations for \( a(x) \) and \( b(x) \) above can be written roughly as
\[
x(t + \Delta t) - x(t) \sim b(x(t))\Delta t + o(x(t))(w(t + \Delta t) - w(t))
\]
where \( w(t) \) is the Brownian motion process on \( \mathbb{R}^n \) and \( o(x) \) is the symmetric square root of the infinitesimal diffusion matrix \( a(x) \). Clearly the above relation should tend to become an exact equation in the limit \( \Delta t \to 0 \). We write
\[
dx(t) = b(x(t))\, dt + o(x(t))\, dw(t)
\]
or in integral form
\[
x(t) = x + \int_0^t b(x(s))\, ds + \int_0^t o(x(s))\, dw(s).
\]

If one can solve the above equation and express \( x(\cdot) \) as a functional of \( w(\cdot) \), then the induced transformation of the measure of the Brownian motion process yields the measure of the diffusion process that we want. But we also have more, namely, a pathwise representation of the process as a functional of Brownian motion.

Before attempting to solve the above integral equation, the Brownian stochastic integral must be defined appropriately. The support of the measure of the Brownian motion is contained in the collection of continuous trajectories, but almost all paths are not of bounded variation. Since local properties of \( w \) will be reflected in local properties of \( x \) the integral
\[
\int_0^t o(x(s))\, dw(s)
\]
requires a special definition and theory, due to Itô [3, 4].

Briefly, one defines the integral for bounded nonanticipating functionals \( f(t, \omega) (\omega \in \Omega \text{ is a continuous path, i.e., } f(t, \omega) \text{ is bounded and measurable jointly in } (t, \omega) \text{ and } f(t, \cdot) \text{ is } \mathcal{F}_t \text{ measurable for each } t \) by first defining it for simple nonanticipating functionals
\[
\int_0^t f(s, \omega)\, dw(s) = \sum f(t_k, \omega)(w(t_{k+1}) - w(t_k))
\]
where the sum is over the finite number of values of the simple functional. Note that the functional is evaluated at the leftmost point of \( (t_k, t_{k+1}) \). This makes the stochastic integral a zero-mean martingale and
\[
E\left\{ \left( \int_0^t f(s, \omega)\, dw(s) \right)^2 \right\} = E\left\{ \int_0^t f^2(s, \omega)\, ds \right\}.
\]

Passage to the limit gives the stochastic integral in general with the martingale property preserved as well as the above identity (if is assumed bounded).

With the stochastic integral and its properties established we return to the stochastic integral equation and attempt to solve it by iteration. As in the case of ordinary differential equations, we must have a Lipschitz condition on \( b(x) \) and \( o(x) \). If we do, the iteration process converges and the representation of the diffusion process as a nonlinear functional of Brownian motion is accomplished.
Not every diffusion can be so represented however, the difficulty being the absence of the Lipschitz continuity for $b$ or $\sigma$. If one wants information contained in the measure only, one does not need the stochastic differential equations. On the other hand, in many problems in modelling and control theory the pathwise representation is important and leads to a number of significant conclusions.

We conclude this brief discussion with some remarks regarding approximations of stochastic processes which is our main topic in Sections 4–16.

One class of approximations concerns behavior of appropriately normalized functionals of a process over $[0, t]$ and then passage to the limit as $t \uparrow \infty$, the convergence being anything from convergence almost everywhere to convergence in the mean to convergence in distribution (weak convergence). For example, if $\{x(t), t \geq 0\}$ is a Markov process on a state space $S$ and if $f(x)$ is a bounded measurable function on $S$, what is the behavior of $\int_0^t f(x(s)) \, ds$ for $t \uparrow \infty$? If $x(t)$ is ergodic and $\bar{P}(dy)$ is its stationary distribution (see next section for more detailed discussion), then

$$\frac{1}{t} \int_0^t f(x(s)) \, ds \to \int f(x) \bar{P}(dy) \equiv \bar{f}, \quad t \uparrow \infty$$

for almost all starting points relative to $\bar{P}$. We may then ask about the behavior of

$$\frac{1}{\sqrt{t}} \int_0^t (f(x(s)) - \bar{f}) \, ds.$$

Since the limit here (in the simplest case) is a Gaussian random variable, the notion of convergence is weak convergence. However, it is usually of interest, and not much more difficult, to ask for a bit more, as follows.

Let $t$ go to infinity as $\epsilon \to 0$ by letting $t = \tau/\epsilon^2$ with $\tau$ an extra parameter. The quantity of interest is then, up to a factor $1/\sqrt{\tau}$,

$$\epsilon \int_0^{t\epsilon^2} (f(x(s)) - \bar{f}) \, ds = \frac{1}{\epsilon} \int_0^t (f(x^{\epsilon}(s)) - \bar{f}) \, ds,$$

where $x^{\epsilon}(t) \equiv x(t/\epsilon^2)$. Now we treat this quantity as a family of processes with parameter $\tau > 0$ and $\epsilon > 0$ labelling the family. That is, we have a family of measures $P^\epsilon$ on $C([0, \infty), R^2)$. The question above generalizes to showing that $P^\epsilon$ converges weakly to Brownian motion. By weak convergence of measures we mean the following. If $X$ is a separable metric space and $P^\epsilon$ is a family of Borel probability measures on $X$, $P^\epsilon$ converges weakly to $P$ if

$$\int_X f(x) P^\epsilon(dx) \to \int_X f(x) P(dx)$$

for each bounded continuous function $f$ on $X$. We shall be concerned with weak convergence in most of what follows.

In the context of Markov processes, if the measures $P^\epsilon$ can be shown to be weakly compact (there are simple sufficient conditions for this [9]) then it suffices to show that the semigroups $T^\epsilon(t)$ converge to a limit semigroup in the strong topology as $\epsilon \to 0$, uniformly on compact time intervals. Let $A^\epsilon$ be the corresponding generator and $R^\epsilon_t$ the resolvents.

By a well-known theorem [10], if for $\lambda > 0$, $R^\epsilon_\lambda$ converges to $R_\lambda f$ for $f \in C(S)$, $S$ compact, say, as $\epsilon \to 0$, then the corresponding semigroups converge. Here it is not actually necessary to assume that $R_\lambda$ is the resolvent of a generator; this follows. On the other hand, one usually has access concretely to $A^\epsilon$ and not $R^\epsilon_t$ (or $T^\epsilon(t)$). If there is a dense set $D \subset \bigcap_{\epsilon > 0} D(A^\epsilon)$ in the Banach space and a generator $A$ of a semigroup (all semigroups are contractions here) $T(t)$ such that the closure of $(\lambda - A)D$ is the whole space and $Af \to Af$ for $f \in D$, then the resolvents converge. For given $f$ in the Banach space there is an $f$ of the form $(\lambda - A)g$, $g \in D$, arbitrarily close to it and, for $\lambda > 0$,

$$\|R^\epsilon_\lambda (\lambda - A)g - R_\lambda (\lambda - A)g\| = \|R^\epsilon_\lambda (\lambda - A)g - g\|$$

$$\leq \|R^\epsilon_\lambda (A^\epsilon - A)g\| \leq \frac{1}{\lambda} \|A^\epsilon g - Ag\| \to 0$$

as $\epsilon \downarrow 0$.

It is actually enough that there exist a set $D$ as above such that for $f \in D$ there is a sequence $f^\epsilon \in D$ and $f^\epsilon \to f$, $A^\epsilon f^\epsilon \to Af$ [11]. To see
this we note that with $g$ as above and $g^\epsilon \to g$,

$$
\| R_\epsilon(\lambda - A)g - R_\lambda(\lambda - A)g \| = \| R_\epsilon(A^\epsilon g^\epsilon - Ag) - R_\lambda(A^\epsilon g^\epsilon - g) \| \\
\leq \frac{1}{\lambda} \| A^\epsilon g^\epsilon - Ag \| + 2 \| g^\epsilon - g \| \to 0
$$
as $\epsilon \downarrow 0$.

In many situations (most of what follows in Sections 4–16) the measures $P^\epsilon$ (for the processes $x^\epsilon(t)$) are not Markovian. However, it may be possible to identify $x^\epsilon(t)$ as the component of a process on a larger state space, this bigger one being Markov. In addition, the limit measure of the $P^\epsilon$ turns out to be itself Markovian. In this case one first shows that the $P^\epsilon$ are a relatively weakly compact family of measures and then one shows, as above, that the semigroup on the bigger space converges as $\epsilon \to 0$ to a semigroup on the subspace when acting on elements of the subspace. In this contraction of space context, the device of choosing $f^\epsilon \to f$ so that $A^\epsilon f^\epsilon \to Af$ is very useful and will be illustrated by many examples in the following.

3. BRIEF REVIEW OF SOME FACTS FROM ERGODIC THEORY

The perturbation analysis of linear and nonlinear equations leads frequently to problems centering on properties of null spaces of linear operators. In the probabilistic context the linear operators are frequently transition operators or infinitesimal generators. The corresponding study of null spaces is, in effect, their ergodic theory. We shall review now some of the basic notions in ergodic theory which are used later.

Let $S$ be a set and $\Sigma$ a $\sigma$-algebra of subsets and let $P(x, A)$ be a transition probability function, i.e., a function on $S \times \Sigma \to [0, 1]$ such that $P(\cdot, A)$ is measurable for each $A \in \Sigma$ and $P(x, \cdot)$ is a probability measure for each $x \in S$. Let $B(S)$ and $M(S)$ be the Banach spaces of bounded measurable functions and finite signed measures on $S$ with the sup and total variation norm respectively.

---

* We consider continuous time problems a little further ahead.
unique invariant probability measure for the transition function $P(x, A)$, i.e.,

$$\bar{P}(A) = \int_A \bar{P}(dx)P(x, A). \quad (3.6)$$

If $\bar{P}$ denotes the operator on $B(S)$ defined by

$$\bar{P}f(x) = \int f(y)\bar{P}(dy) \quad \text{(constant function)}, \quad (3.7)$$

then clearly $(1 - \theta)R_\theta \to \bar{P}$ as $\theta \uparrow 1$ and $\bar{P}$ is a projection operator. Moreover, Poisson's equation (3.4) is solvable if and only if

$$\bar{P}g = \int \bar{P}(dy)g(y) = 0, \quad g \in B(S), \quad (3.8)$$

and the solution is unique up to a constant. The recurrent potential kernel

$$\psi(x, A) = \sum_{n=0}^{\infty} (P^n(x, A) - \bar{P}(A)) \quad (3.9)$$

is well defined and the solution* of (3.4) is

$$u(x) = \int \psi(x, dy)g(y), \quad g \in B(S). \quad (3.10)$$

In the continuous time case we have a transition function $P(t, x, A)$ satisfying the Chapman-Kolmogorov equation as explained in Section 2. Condition (3.5) is replaced by:

There exists a reference probability measure $\phi$ on $S$, a constant $c > 0$ and a $t_0 > 0$ such that

$$P(t_0, x, A) \geq c\phi(A). \quad (3.11)$$

Let $R_\lambda$ be the resolvent operator

$$R_\lambda = \int_0^\infty e^{-\lambda t}T(t) \, dt, \quad \lambda > 0, \quad (3.12)$$

where $T(t)$ is the semigroup on $B(s)$ induced by $P(t, x, A)$. For any $\lambda > 0$ the function $\lambda R_\lambda(x, A)$, defined by

$$\lambda R_\lambda(x, A) = \int_0^\infty \lambda e^{-\lambda t}P(t, x, A) \, dt,$$

the kernel of the resolvent operator, is a transition function and (3.11) implies that

$$\lambda R_\lambda(x, A) \geq c\phi(A)e^{-\lambda t_0}. \quad (3.13)$$

Thus, the discrete time result applies to $\lambda R_\lambda$. The corresponding invariant measure $\bar{P}$ is independent of $\lambda > 0$ as can be seen from the resolvent identity $R_{\lambda_1} = R_{\lambda_2} + (\lambda_2 - \lambda_1)R_{\lambda_1}R_{\lambda_2}$. This in turn implies that $\bar{P}$ is invariant for the transition function $P(t, x, A)$.

Since

$$\lambda R_\lambda(1 - \theta) \sum_{n=0}^{\infty} \theta^n(\lambda R_\lambda)^n = (1 - \theta)\lambda R_{(1 - \theta)\lambda}, \quad (3.14)$$

it follows that $\lim \lambda R_\lambda$ exists as $\lambda \downarrow 0$ and equals the projection operator $\bar{P}$. Moreover the Poisson equation

$$u - \lambda R_\lambda u = R_\lambda g \quad (3.15)$$

has a unique (up to a constant) solution if and only if (3.8) holds. By applying the resolvent identity it follows that $u$ is independent of $\lambda$ in (3.15) and, in fact, we have (3.10) where

$$\psi(x, A) = R_\lambda \sum_{n=0}^{\infty} (\lambda R_\lambda)^n - \bar{P}(x, A)$$

$$= \int_0^\infty [P(t, x, A) - \bar{P}(A)] \, dt. \quad (3.16)$$

If Doeblin's condition does not hold, the situation is considerably more complicated but quite a bit is known [12, 13]. Even if one has access to a more general theory, however, if the recurrent potential kernel maps bounded functions to unbounded ones, the nature of the results in the asymptotics changes (cf. Section 13) rather drastically.

Let us suppose that $S$ is a compact metric space with $\Sigma$ the Borel sets and that the transition function is uniformly Feller, i.e.,

$$\limsup_{x \to y} \sup_{A \in \Sigma} |P(x, A) - P(y, A)| = 0. \quad (3.17)$$

* Up to a constant which will be usually set to zero.
This implies that the transition operator $T$ of (3.1) maps the unit ball in $B(S)$ into a compact set, i.e., it is compact. If in addition $P(x, A)$ has no nontrivial invariant sets (relative to some reference measure), then again we are effectively in Doeblin’s case. A nontrivial invariant set $A$ is such that $P(x, A) = 1$ for all $x \in A$ and $1 > \phi(A) > 0$ ($\phi$ the reference measure).

A Feller transition function on a compact state space $S$ has invariant measures, as can be easily deduced by the relative weak compactness of families of probability measures on $S$. Uniqueness of the invariant measure implies the existence of the projection operator $\bar{P}$. However, in general, uniqueness is not easily shown.

In a locally compact space the question of existence of invariant measures is again not too difficult to settle [14]. However uniqueness is difficult.

We shall conclude this section with a brief comment on the connection of notions of mixing [15] to the above.

When dealing with first order perturbation theory (for example laws of large numbers, etc.) the existence and uniqueness of an invariant measure is enough. Actually the existence of the projection operator $\bar{P}$ is enough.

For second order perturbation theory (central limit theorem for example) it is necessary to have a solvability theory for Poisson’s equation which usually is called the Fredholm alternative. Now Doeblin’s condition leads to a solvability theory that is very strong: $\psi(x, A)$ exists and maps bounded functions to bounded functions. In the asymptotics less will do. For example: $\psi$ maps bounded functions to $\bar{P}$ integrable ones. Mixing conditions, in one way or another, affect the existence and general behavior of the recurrent potential operator.

The mixing condition

$$\sup_{\|f\| \leq 1} \int \bar{P}(dx) \left| \int [P(x, dy) - \bar{P}(dy)]f(y) \right| = \rho < 1,$$

due to Rosenblatt [15], is considerably weaker than Doeblin’s condition and still allows the validity of the central limit theorem. Its connections with recurrent potential theory have not been explored yet, however. For more information about notions of mixing we refer to [16].

4. OSCILLATORY PROBLEMS, AVERAGING, AND LAW OF LARGE NUMBERS FOR STOCHASTIC EQUATIONS

Differential equations whose coefficients are rapidly oscillating functions of time are frequently analyzed by the method of averaging [17]. Since noise fluctuations are to a certain extent similar to oscillations in general (the differences will be elucidated below), averaging provides good motivation for the analysis of many fluctuation phenomena.

We begin with the following frequently used example. Let $x(t)$ be the amplitude of an oscillator with equation of motion*

$$\ddot{x} + \epsilon(\omega^2 x^2 + \dot{x}^2 - 1)\dot{x} + \omega^2 x = 0, \quad t > 0, \quad x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0. \quad (4.1)$$

The damping term is nonlinear in such a way that we expect a limit cycle to develop when $\epsilon \ll 1$ and time is large. To analyze (4.1) we introduce polar coordinates

$$x = r \cos(\omega t + \theta),$$
$$\dot{x} = -\omega r \sin(\omega t + \theta), \quad (4.2)$$

where $r(t)$ and $\theta(t)$ are functions of time satisfying the system

$$\dot{r} = \epsilon r (1 - \omega^2 r^2) \sin^2(\omega t + \theta),$$
$$\dot{\theta} = \frac{\epsilon}{2} (1 - \omega^2 r^2) \sin(2\omega t + 2\theta), \quad (4.3)$$
$$r(0) = r_0, \quad \theta(0) = \theta_0.$$

Here $r_0$ and $\theta_0$ are obtained from (4.1) and (4.2).

We note that (4.3) has the following general form:

$$\frac{dx^*(t)}{dt} = \epsilon F(x^*(t), t), \quad x^*(0) = x, \quad (4.4)$$

* $x(t) = dx(t)/dt$. 
where \( x^e(t) \) is an \( n \)-dimensional vector function of time and \( F(x, t) \) is an \( n \)-dimensional vector function on \( \mathbb{R}^n \times [0, \infty) \), periodic in \( t \) (in this case, almost periodic more generally). Averaging its simplest form amounts to this. Let

\[
\tilde{F}(x) = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau F(x, s) \, ds
\]  

and let

\[
\frac{d\tilde{x}(\tau)}{d\tau} = F(\tilde{x}(\tau)), \quad \tilde{x}(0) = x. \tag{4.6}
\]

Then \( x^e(\tau/e) \) is well approximated by \( \tilde{x}(\tau) \) for \( \tau \) in any finite time interval if \( e \) is small.

We first apply this to (4.3). We obtain the system

\[
\frac{d\tilde{r}(\tau)}{d\tau} = \frac{1}{\lambda} \tilde{r}(\tau)(1 - \omega^2 \tilde{r}(\tau)), \quad \tilde{r}(0) = r_0,
\]

\[
\frac{d\tilde{p}(\tau)}{d\tau} = 0, \quad \tilde{p}(0) = \theta_0. \tag{4.7}
\]

Clearly

\[
\tilde{r}(\tau) = \frac{1}{\omega} \left[ 1 + e^{-\omega \tau} (\omega^2 \tilde{r}_0^2 - 1) \right]^{1/2}. \tag{4.8}
\]

Thus, the approximation of \( r(\tau/e) \) by \( \tilde{r}(\tau) \) does capture the qualitative features one may expect from (4.1).

For a more exact statement of the averaging principle we redefine the time scale \( t \) so that

\[
\frac{dx^e(t)}{dt} = F\left(x^e(t), \frac{t}{e}\right), \quad x^e(0) = x, \tag{4.9}
\]

with first approximation

\[
\frac{d\tilde{x}(t)}{dt} = F(\tilde{x}(t)), \quad \tilde{x}(0) = x, \tag{4.10}
\]

\( \tilde{F}(x) \) being given by (4.5). Suppose that \( F(x, t) \) and \( \tilde{F}(x) \) are smooth functions of \( x \) and bounded. Suppose that (4.5) holds in a stronger sense

\[
\left| \int_0^t [F(x, s) - \tilde{F}(x)] \, ds \right| \leq c < \infty,
\]

\[
\left| \int_0^t [F_\alpha(x, s) - \tilde{F}_\alpha(x)] \, ds \right| \leq c < \infty, \tag{4.11}
\]

for all \( \alpha \in \mathbb{R}^n \) and \( t \geq 0 \). Then given \( T < \infty \) there is an \( \epsilon_0 \) and a constant \( c' \) such that

\[
\sup_{0 \leq t \leq T} |x^e(t) - \tilde{x}(t)| < c' \epsilon, \quad \text{if } \epsilon \leq \epsilon_0. \tag{4.12}
\]

The proof of this theorem is simple and instructive so we shall give it.

Since (4.11) holds, it suffices to show that

\[
w^e(t) = x^e(t) - \tilde{x}(t) - \epsilon \int_0^{t/e} [F(x^e(s), s) - \tilde{F}(x^e(t))] \, ds \tag{4.13}
\]

has small norm. But

\[
\frac{dw^e(t)}{dt} = F(x^e(t)) - \tilde{F}(\tilde{x}(t))
\]

\[
- \epsilon \int_0^{t/e} [F_\alpha(x^e(s), s) - \tilde{F}_\alpha(x^e(t))]F\left(x^e(t), \frac{t}{e}\right) \, ds,
\]

\[
w^e(0) = 0 \tag{4.14}
\]

hence*

\[
|w^e(t)| \leq c_1 \int_0^t |w^e(s)| \, ds + c_2 \epsilon t
\]

from which the result follows.

What should be noted in the above simple argument is that the passage from \( F(x, t) \) to \( \tilde{F}(x) \), with \( x \) fixed, is the projection (4.5). If \( F(x, t) \) is almost periodic uniformly in \( x \) then the limit (4.5) exists uniformly in \( x \) and uniformly in \( \tau \) if \( s \) is replaced by \( s + \tau \) in the integrand. This is all that is necessary for mere convergence of \( x^e(t) \)

\* \( c_1, c_2, \ldots, \) etc., are constants.
to $\tilde{x}(t)$. Hypothesis (4.11) concerns the "recurrent potential" and is strong enough to give the error estimate (4.12).

In the event $\tilde{F}(x) \equiv 0$ and it is necessary to go to higher order perturbations, then the recurrent potential will actually appear in the final result, so we must assume that it exists, clearly. Higher order corrections can be constructed even if $\tilde{F}(x) \neq 0$ but then the approximations will themselves depend on $\epsilon$. The algorithm for constructing them is well known [17] and we will give it in Section 7.

The literature on problems associated with averaging is huge. Typical questions range from implementation to specific examples, to extensions so that the approximations are valid uniformly in time ($T = \infty$ above), to questions of existence and stability of periodic solutions, etc. We refer to the survey [18] and Hale's book [19] for more information.

We turn now to the stochastic problem which is analogous to (4.9). To make the connections with Section 3 more transparent we shall deal with the following model.

Let $\{y(t), t \geq 0\}$ be a Markov process* on a state space $(S, \Sigma)$ and let $P(t, y, A)$ be its transition function. Assume that it is ergodic with invariant measure $P(A)$ and such that

$$\lim_{\lambda \to 0} \mathcal{R}_\lambda f(y) = \lim_{\lambda \to 0} \int_0^\infty e^{-\lambda t} T(t)f(y) dt$$

$$= \tilde{F} f = \int \tilde{P}(dy) f(y),$$

uniformly in $y$ for $f \in \mathcal{B}(S)$. Here $T(t)$ is the semigroup associated with $P(t, y, A)$.

Let $F(x, y)$ be a bounded measurable $n$-vector function on $\mathbb{R}^n \times S$, with bounded $x$ derivatives. Consider the system

$$\frac{dx^\epsilon(t)}{dt} = F(x^\epsilon(t), y^\epsilon(t)), \quad x^\epsilon(0) = x, \quad (4.16a)$$

where we have defined

$$y^\epsilon(t) = y(t/\epsilon). \quad (4.16b)$$

When $S$ is the real line, $F(x, y)$ is almost periodic in $y$ and the transition function corresponds to deterministic motion at unit speed, we recover (4.9).

It is natural that $\tilde{F}(x)$ in (4.5) be replaced by

$$\tilde{F}(x) = \int \tilde{P}(dy) F(x, y)$$

and that

$$\frac{dx(t)}{dt} = \tilde{F}(x(t)), \quad x(0) = x \quad (4.18)$$

should in some sense approximate (4.16).

A simple result along these lines is that if the recurrent potential exists (analog of (4.11)), i.e., (3.16) is well defined, then

$$\lim_{\epsilon \to 0} \sup_{0 \leq t \leq T} \left| x^\epsilon(t) - x^\epsilon(0) \right| > \delta = 0, \quad T < \infty \quad (4.19)$$

for any $\delta > 0$ and uniformly in $(x, y) \in \mathbb{R}^n \times S$. We may call this the law of large numbers.

The terminology is justified because for $n = 1$ and $F(x, y) = x_\epsilon(y)$ independent of $x$ with $A \in \Sigma$, we have

$$x^\epsilon(t) = x + \int_0^t x_\epsilon(y(s)) ds$$

$$= x + \int_0^t x_\epsilon(y(s/\epsilon)) ds$$

$$= x + \epsilon \int_0^{t/\epsilon} x_\epsilon(y(s)) ds,$$

which converges with probability one in fact (by the ergodic theorem) to $x + tP(A)$ as $\epsilon \to 0$.

The proof is similar to the one for averaging. First we define

$$\psi_F(x, y) = \int_0^m \int \left[ P(t, y, dz) - \tilde{P}(dz) \right] F(x, z) dt$$

so that

$$-A\psi_F = F \quad (4.21)$$

* Thus, $x(t) = x(t, \omega), \omega \in \Omega, (\Omega, \mathcal{F}, P)$ a probability space. Similarly in (4.16) $x^\epsilon(t) = x^\epsilon(t, \omega)$. We shall not indicate the $\omega$ explicitly in the sequel.
$A$ being the infinitesimal generator of $[y(t), t \geq 0]$ (cf. (3.15), (3.16)). Then we note that $(x'(t), y'(t))$ is Markovian on $R^a \times S$ with generator

$$
\frac{1}{\epsilon} A + F(x, y) \frac{\partial}{\partial x}
$$

with the dot denoting the inner product and $\partial/\partial x$ the gradient operator. From the discussion of Section 2 we know that

$$
x'(t) + \epsilon \psi(x'(t), y'(t)) - \int_0^t \left( \frac{1}{\epsilon} A + F \frac{\partial}{\partial x} \right) (x + \epsilon \psi)(x'(s), y'(s)) \, ds
$$

is a martingale. In fact* the expression

$$
x'(t) + \epsilon \psi(x'(t), y'(t)) - x - \epsilon \psi(x, y) - \int_0^t \left( \frac{1}{\epsilon} A + F \frac{\partial}{\partial x} \right) (x + \epsilon \psi)(x'(s), y'(s)) \, ds \equiv M'(t) \tag{4.22}
$$

is a zero mean martingale. Using (4.21) and the fact that $A1 = 0$ ($A$ is a Markovian generator), we find that

$$
x'(t) - x - \int_0^t \frac{\partial}{\partial x} \psi(x(s), y(s)) \, ds
$$

$$
= M'(t) + \epsilon \left( \psi(x(t), y) - \psi(x'(t), y'(t)) \right) + \epsilon \int_0^t F \frac{\partial}{\partial x} \psi(x'(s), y'(s)) \, ds.
$$

Subtracting the integrated version of (4.18) from this yields

$$
x'(t) - \bar{x}(t) = \int_0^t [F(x'(s)) - F(\bar{x}(s))] \, ds + M'(t)
$$

$$
+ \epsilon \left( \psi(x(t), y) - \psi(x'(t), y'(t)) \right) + \epsilon \int_0^t F \frac{\partial}{\partial x} \psi(x'(s), y'(s)) \, ds. \tag{4.23}
$$

This is the analog of (4.13) and (4.14) (combined).

* The notation in the integral means that the operator acts on the function shown and the result is evaluated at the argument shown.

To get the result (4.19) and (4.23) is now an easy matter. It is roughly the same process as in the deterministic case except now the martingale term $M'(t)$ must also be estimated. Since the second moments of $M'(t)$ are proportional to $\epsilon$ an application of Kolmogorov's inequality is all that is necessary.

As a final note we should point out that most of the extensions and side issues associated with averaging make sense for the stochastic problem also. However, what is more interesting is that in the stochastic case one can ask many additional questions that have no real analogs in the deterministic case; for example, the fluctuation theory of $x'(t)$ about $\bar{x}(t)$ (properly scaled) as discussed in Section 7. The second order perturbation theory associated with (4.16), when $F(x) \equiv 0$ in (4.17), is discussed in Section 7. Further information regarding averaging can be found in the papers of Kurtz [20], for example, [21], and the next section. Some extensions and applications of averaging can be found in [22] and [23], as well as in [24].

5. SPATIAL AVERAGING AND HOMOGENIZATION

We now consider an example treated by Freidlin [25], Khasminskii [26], and in [27]. We pose the problem in a deterministic way at first. Let $(a_i(y))$ and $(b_i(y))$ be a symmetric positive definite $n \times n$ matrix and $n$-vector function of $y \in S$, respectively, where $S$ is the unit torus on $R^n$, assumed smooth. Consider the diffusion equation

$$
\frac{\partial u^\epsilon(t, x)}{\partial t} = \frac{1}{2} \sum_{i, j = 1}^n a_i(x) \frac{\partial^2 u^\epsilon(t, x)}{\partial x_i \partial x_j} + \sum_{i = 1}^n b_i(x) \frac{\partial u^\epsilon(t, x)}{\partial x_i}, \tag{5.1}
$$

$$
t > 0, \quad x \in \mathcal{D},
$$

$$
u^\epsilon(0, x) = f(x), \quad \nu^\epsilon(t, x) = 0, \quad x \in \partial \mathcal{D}, \quad t > 0.
$$

Here $\mathcal{D}$ is a bounded open set in $R^n$ and $\partial \mathcal{D}$ is its assumed smooth boundary.

The question is: how does $\nu^\epsilon$ behave as $\epsilon \to 0$ and the coefficients, being periodic, oscillate more and more rapidly?

Let us assume that the operator

$$
A = \frac{1}{2} \sum_{i, j = 1}^n a_i(y) \frac{\partial^2}{\partial y_i \partial y_j}, \tag{5.2}
$$
is uniformly elliptic on $S$ and let $[y(t), t \geq 0]$ be the diffusion process* it generates on $S$. The uniform ellipticity implies clearly that $y(t)$ is ergodic and, by smoothness of $a_i(y)$, the invariant measure has a density $\hat{p}(y)$ relative to Lebesgue measure on $S$. Furthermore, the recurrent potential kernel exists since Doeblin's theory applies via the strong maximum principle.

Define

$$\bar{a}_i = \int_S a_i(y)\hat{p}(y)\,dy,$$

$$\bar{b}_{ij} = \int_S b_{ij}(y)\hat{p}(y)\,dy, \quad i, j = 1, 2, \ldots, n. \quad (5.3)$$

Then $u'(t, x)$ converges uniformly† in $x \in \partial E$, $0 \leq t < T$ to the solution $\bar{u}(t, x)$ of

$$\frac{\partial \bar{u}(t, x)}{\partial t} = \frac{1}{2} \sum_{i,j=1}^{n} \bar{a}_{ij} \frac{\partial^2 \bar{u}(t, x)}{\partial x_i \partial x_j} + \sum_{i=1}^{n} \bar{b}_i \frac{\partial \bar{u}(t, x)}{\partial x_i}, \quad (5.4)$$

$$u(0, x) = f(x), \quad \bar{u}(t, x) = 0, \quad t > 0, \quad x \in \partial E.$$

Thus the limit problem has constant coefficients which are appropriate averages of the original ones. This explains the terminology.

If $a_i = a_i(x, y)$, $b_{ij} = b_{ij}(x, y)$, $x \in \partial E$, $y \in S$, then $x$ is carried along as a parameter in (5.2) and (5.3) and $\bar{a}_i, \bar{b}_{ij}$ depend now on $x$.

In terms of stochastic differential equations, the process generated by the operator on the right side of (5.1) satisfies‡

$$dx(t) = b\left(\frac{x(t)}{\epsilon}\right) dt + \sigma\left(\frac{x(t)}{\epsilon}\right) dw(t), \quad x^c(0) = x, \quad (5.5)$$

where $\sigma(y)$ is the symmetric (smooth, since $\sigma$ is smooth and uniformly elliptic) square root of $a(y)$ and $w(t)$ is the standard $n$-dimensional Brownian motion. The convergence corresponds to weak convergence on $C([0, T]; R^n)$ of the corresponding measures. The limit measure is, essentially, Brownian motion.

---

* Thus $y(t) = y(t, \omega)$, $\omega \in \Omega, (\Omega, \mathcal{F}, P)$ a probability space.
† We also assume that $f(x) = 0$ outside a compact subset of $E$ and is smooth.
‡ The argument $\omega \in \Omega, (\Omega, \mathcal{F}, P)$ a probability space, is omitted.
From (5.8) and the ergodicity of $A$ in (5.2) we conclude that $v_0 = v_0(t, x)$. We may thus take $v_1 = 0$ in (5.9). Thus (5.10) takes the form

$$A v_2 + \frac{1}{2} \sum_{i,j=1}^{n} a_i(y) \frac{\partial v_0}{\partial x_i x_j} \sum_{i=1}^{n} b_i(y) \frac{\partial v_0}{\partial x_i} - \frac{\partial v_0}{\partial t} = 0. \tag{5.11}$$

In order that (5.11) have a bounded solution the inhomogeneous term must integrate to zero relative to the invariant measure $\bar{\rho}(y) \, dy$; (5.11) is just Poisson's equation (3.15)–(3.16). Applying this solvability condition we find that $v_0(t, x)$ must satisfy (5.4) (and now we insert initial and boundary condition).

After this formal step the proof goes as follows. Let $\bar{u}(t, x)$ solve (5.4) and be smooth. Let $v_2$ be any bounded solution of (5.11). Then if

$$w^e(t, x) = u^e(t, x) - \bar{u}(t, x) - \varepsilon^2 v_2\left(t, x, \frac{x}{\varepsilon}\right),$$

we find that by construction

$$\left[ \frac{\partial}{\partial t} - \frac{1}{2} \sum_{i,j=1}^{n} a_i\left(\frac{x}{\varepsilon}\right) \frac{\partial^2}{\partial x_i \partial x_j} - \sum_{i=1}^{n} b_i\left(\frac{x}{\varepsilon}\right) \frac{\partial}{\partial x_i} \right] w^e(t, x) = O(\varepsilon),$$

$$w^e(0, x) = O(\varepsilon^2), \quad x \in \mathcal{D},$$

$$w^e(t, x) = O(\varepsilon^2), \quad x \in \partial \mathcal{D}, \quad t > 0.$$ The maximum principle now yields the desired result.

Homogenization is a subject with many ramifications and a surprisingly diverse field of applications. A systematic treatment will be given in a forthcoming book [28]. We also refer to the papers of Babuška [29] for further results and applications.

6. THE GAUSS-MARKOV APPROXIMATION AND APPLICATIONS

We continue the analysis of (4.16) under the assumptions stated there and with (4.19) established. To analyze fluctuations* set

$$x^e(t) = \bar{x}(t) + \sqrt{\varepsilon} \omega^e(t). \tag{6.1}$$

* Again the $\omega \in \Omega, (\Omega, \mathcal{F}, P)$ a probability space, is omitted.
where \( \psi(y, A) \) is the recurrent potential kernel (3.16), and \( (\sigma(y, x)) \) is the symmetric square root of \( (\sigma_j) \), then the formal limit of \( z^\varepsilon(t) \) satisfies

\[
z(t) = \int_0^t \sigma(\bar{x}(s)) \, dw(s) + \int_0^t \frac{\partial \bar{F}(\bar{x}(s))}{\partial x} \, z(s) \, ds, \quad (6.6)
\]

where \( w(t) \) is the standard \( n \)-dimensional Brownian motion.

Thus \( z(t) \) is a time-inhomogeneous Gaussian process with independent increments satisfying the linear stochastic differential equation (6.6) where \( \bar{x}(t) \) satisfies (4.17).

Let \( U(t, s) \) be the fundamental solution of the linear variational equation to (4.18)

\[
\frac{dU(t, s)}{dt} = \frac{\partial \bar{F}(\bar{x}(t))}{\partial x} \, U(t, s), \quad t > s, \quad (6.7)
\]

\[U(s, s) = I \text{ (identity)}.\]

Let

\[
v(t) = \int_0^t \sigma(\bar{x}(s)) \, dw(s)
\]

which is a Gaussian process with independent increments. Then,

\[
z(t) = \int_0^t U(t, s) \, dv(s)
\]

and so

\[
E\{z(t)z^T(t)\} = \int_0^t ds \, U(t, s) \sigma(\bar{x}(s)) U^T(t, s)
\]

where \( T \) denotes transpose.

The above description points to the fact that the fluctuation process has a nice and relatively simple asymptotic behavior. We should stress that (i) the approximation (which can be easily proved) is valid only on finite but arbitrary time intervals and this limits its potential usefulness and (ii) the approximation is a fairly crude one and makes sense in great generality (in infinite dimensional, for example, partial differential equations, problems). For a proof we refer to [3] and [30] as well as [31–33].

A simple proof can be given along the following lines. We consider jointly \((z^\varepsilon(t), \bar{x}(t), y^\varepsilon(t))\) which is a Markov process on \( R^n \times R^n \times S \) with infinitesimal generator*

\[
\mathcal{L}^\varepsilon = \frac{1}{\varepsilon} A + \frac{1}{\sqrt{\varepsilon}} (F(\bar{x} + \sqrt{\varepsilon} z, y) - \bar{F}(\bar{x}) \frac{\partial}{\partial z} \bar{F}(\bar{x}) \frac{\partial}{\partial \bar{x}}) \quad (6.10)
\]

Evidently \( \bar{x}(t) \) is deterministic and decouples from \((z^\varepsilon(t), y^\varepsilon(t))\) but it is carried along to make things time homogeneous. Now note that \( \mathcal{L}^\varepsilon \) has the form

\[
\mathcal{L}^\varepsilon = \frac{1}{\varepsilon} \mathcal{L}_1 + \frac{1}{\sqrt{\varepsilon}} \mathcal{L}_2 + \mathcal{L}_3 + O(\sqrt{\varepsilon}) \quad (6.11)
\]

where \( \mathcal{L}_1 = A, \mathcal{L}_2 = (F(\bar{x}, y) - \bar{F}(\bar{x})) \cdot \bar{\sigma} / \bar{\sigma} \bar{x} \) and

\[
\mathcal{L}_3 = \bar{F}(\bar{x}) \frac{\partial}{\partial \bar{x}} + \frac{\partial F(\bar{x}, y)}{\partial \bar{x}} - z \frac{\partial}{\partial \bar{x}}.
\]

If \( T^*(t) \) denotes the semigroup generated by \( \mathcal{L}^\varepsilon \) on \( C(R^n \times R^n \times S) \) and if

\[
\mathcal{T} = \sum_{i,j=1}^n \alpha_i(\bar{x}) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i,j=1}^n \frac{\partial F(\bar{x})}{\partial x_i} z_j \frac{\partial}{\partial z_j} + \sum_{i=1}^n F(\bar{x}) \frac{\partial}{\partial \bar{x}_i}
\]

is the generator of the limit process \((z(t), \bar{x}(t))\) with semigroup \( T(t) \) on \( C(R^n \times R^n) \), then we must show that \( T^*(t)f \xrightarrow{\varepsilon \to 0} T(t)f \) for each \( f \in C(R^n \times R^n) \) and smooth. Compactness of the corresponding measures can also be established without difficulty. We are therefore in the situation described at the end of Section 2.

Because of the singular form of \( \mathcal{L}^\varepsilon \) in (6.11) the generators do not converge when acting on fixed functions. Following Kurtz [11], for given \( f(\bar{x}, \bar{x}) \) smooth we define \( f_1(z, \bar{x}, y) \) and \( f_2(z, \bar{x}, y) \) as solutions of the following Poisson’s equations

\[
A f_1 + (F(\bar{x}, y) - \bar{F}(\bar{x})) \frac{\partial f_1(z, \bar{x})}{\partial z} = 0,
\]

\[
A f_2 + (F(\bar{x}, y) - \bar{F}(\bar{x})) \frac{\partial f_1(z, \bar{x}, y)}{\partial z} + \frac{\partial F(\bar{x}, y)}{\partial \bar{x}} z \frac{\partial f_2(z, \bar{x})}{\partial z} + \bar{F}(\bar{x}) \frac{\partial f_2(z, \bar{x})}{\partial \bar{x}} - \mathcal{L} f(z, \bar{x}) = 0.
\]

* \( A \) is the infinitesimal generator of \( y(t) \).
Note that by definition of $F$ in (4.17), and of $\mathcal{P}$ in (6.12) the solvability condition for the equations (6.13), under the strong ergodicity assumptions we have here, $f_1$ and $f_2$ are well defined. Now set

$$f^\varepsilon(z, x, y) = f(z, x) + \sqrt{\varepsilon}f_1(z, x, y) + \varepsilon f_2(z, x, y).$$

(6.14)

By construction

$$\mathcal{P}f^\varepsilon = \mathcal{P}f + O(\sqrt{\varepsilon})$$

and since $f^\varepsilon \to f$ clearly the result follows as was explained at the end of Section 2.

It should be noted that the above construction of $f_1$ and $f_2$ can be motivated simply by inserting power series and equating coefficients as in (5.7)–(5.11). We refer to the above as a second order result because one has to carry perturbation theory one order further than in Section 4 which is the first order result.

7. DIFFUSION APPROXIMATIONS

We now consider the problem

$$\frac{dx^\varepsilon(t)}{dt} = \frac{1}{\varepsilon} F(x^\varepsilon(t), y^\varepsilon(t)), \quad x^\varepsilon(0) = x,$$

$$y^\varepsilon(t) = y(t/\varepsilon),$$

(7.1)

where $F(x, y)$ is a bounded measurable $n$-vector function on $\mathbb{R}^n \times S$ with bounded $x$-derivatives and $y(t)$ is an ergodic Markov process on a state space $(S, \Sigma)$ with transition function $P(t, y, B)$, infinitesimal generator $A$, invariant measure $\mathcal{P}(B)$ and recurrent potential $\psi(x, B)$ (cf. (3.16)). We assume in this section that

$$\int_S F(x, y) P(dy) = 0$$

(7.2)

and we have changed the scaling in (7.1) as compared to (4.16), (4.17) in order to deal with integral powers of $\varepsilon$ and with the limit being $O(1)$.

From the result (4.19) it follows that on time intervals of the form $0 \leq t \leq \varepsilon C$, $C$ is a constant, the process $x^\varepsilon(t)$ does not move much from its initial value. On the larger time intervals $0 \leq t \leq T < \infty$ we shall show that $x^\varepsilon(t)$ behaves like a diffusion Markov process.

First we consider $(x^\varepsilon(t), y^\varepsilon(t))$ jointly as a Markov process on $\mathbb{R}^n \times S$. The infinitesimal generator of this process is easily seen to have the following form on smooth functions

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon} A + \frac{1}{\varepsilon} F(x, y) \cdot \frac{\partial}{\partial x}$$

(7.3)

The semigroup $T^\varepsilon(t)$ is given by

$$T^\varepsilon(t)f(x, y) = E_{x, y}(f(x^\varepsilon(t), y^\varepsilon(t))),$$

(7.4)

and the process $x^\varepsilon(t)$ ($y^\varepsilon(t)$ is given) is constructed pathwise.

To see how $T^\varepsilon(t)f$ behaves when $\varepsilon \downarrow 0$ and $f = f(x)$ (we look at the $x^\varepsilon(t)$ process only) and we set

$$u^\varepsilon(t, x, y) = T^\varepsilon(t)f(x, y),$$

(7.5)

with $f(x)$ smooth, and thus

$$\frac{\partial u^\varepsilon(t, x, y)}{\partial t} = \frac{1}{\varepsilon} Au^\varepsilon(t, x, y) + \frac{1}{\varepsilon} F(x, y) \cdot \frac{\partial u^\varepsilon(t, x, y)}{\partial x}$$

(7.6)

$$u^\varepsilon(0, x, y) = f(x).$$

We analyze (7.6) by the usual methods of asymptotic expansions. Put

$$u^\varepsilon = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \cdots,$$

(7.7)

insert this in (7.6), and equate coefficients of equal powers of $\varepsilon$ to obtain the following sequence of problems

$$Au_0 = 0$$

(7.8)

$$Au_1 + F \cdot \frac{\partial u_0}{\partial x} = 0$$

(7.9)

$$Au_2 + F \cdot \frac{\partial u_1}{\partial x} = \frac{\partial u_0}{\partial t} = 0,$$
Since $A$ acts on $u_0$ as a function of $y$ alone and $A$ is ergodic ($A1 = 0$ and 1 is the only such solution), $u_0 = u_0(t, x)$ but it is not yet determined further.

Now we look at (7.9). It is a Poisson's equation in $S$ space with $x$ and $t$ being parameters. Since $A$ is ergodic (7.9) will have a solution provided the solvability condition
\begin{equation}
\int \overline{P}(dy) F_x u_0 = 0 \tag{7.11}
\end{equation}
holds. But this is (7.2) and therefore $u_1$ exists and is given by
\begin{equation}
u_1(t, x, y) = \int_S \phi(y, dz) F_x u_0(t, x). \tag{7.12}
\end{equation}

Next we consider (7.10). The solvability condition for it is
\begin{equation}
\int \overline{P}(dy) \left[ F_x u_0(t, x, y) - \frac{\partial u_0(t, x)}{\partial t} \right] = 0. \tag{7.13}
\end{equation}

Using (7.12) in (7.13) yields the following diffusion equation for $u_0(t, x)$.
\begin{equation}
\frac{\partial u_0(t, x)}{\partial t} = \int \overline{P}(dy) \int \phi(y, dz) \sum_{i,j=1}^n F_i(x, y) \frac{\partial}{\partial x_j} \left( F_j(x, z) \frac{\partial u_0(t, x)}{\partial x_j} \right). \tag{7.14}
\end{equation}

The diffusion and drift coefficients are given by
\begin{equation}
a_i(x) = \int \overline{P}(dy) \int \phi(y, dz) F_i(x, y) F_j(x, z), \tag{7.15}
\end{equation}
\begin{equation}
b_i(x) = \int \overline{P}(dy) \int \phi(y, dz) \sum_{i,j=1}^n F_j(x, y) \frac{\partial F_i(x, z)}{\partial x_j}, \tag{7.16}
\end{equation}
and hence (7.14) becomes
\begin{equation}
\frac{\partial u_0(t, x)}{\partial t} = \overline{D} u_0 = \sum_{i,j=1}^n a_i(x) \frac{\partial^2 u_0(t, x)}{\partial x_i \partial x_j} + \sum_{j=1}^n b_j(x) \frac{\partial u_0(t, x)}{\partial x_j}, \tag{7.17}
\end{equation}
\begin{equation}
u_0(t, x) = f(x). \tag{7.18}
\end{equation}
On the other hand, the drift coefficient \(7.16\), which has also the form
\[
b(x) = \int_0^\infty dt E \left\{ \sum_{i=1}^n F(x, y(t)) \frac{\partial F_i(x, y(t))}{\partial x_i} \right\},
\]
(7.19')
can be nonzero even for “deterministic” stationary processes; for example \(F(x, y(t))\) almost periodic in \(t\) in which case \(7.19'\) is actually given by a Cesaro limit. In the almost periodic case we recover the second order averaging formulas [17].

The above can be extended a bit to deal with processes \(y(t)\) that are ergodic in a strong sense (0 is an isolated point of the spectrum of the generator \(\mathcal{A}\)) but they can still have isolated point eigenvalues on the imaginary axis (corresponding to cyclically moving sets in the ergodic decomposition of the state space \(\mathcal{S}\)). The simplest instance of this is when \(\mathcal{O}\) is an \(n \times n\) skew-symmetric matrix with isolated eigenvalues and \(7.1\) is replaced by
\[
\frac{dx^e(t)}{dt} = \frac{1}{\epsilon^2} Q x^e(t) + \frac{1}{\epsilon} F(x^e(t), y^e(t)), \quad x^e(0) = x.
\]
(7.20)

Then, if
\[
x^e(t) = e^{-\epsilon t} x^e(t), \quad \tau^e(t) = \tau + t/\epsilon^2
\]
(7.21)
the process \((x^e(t), y^e(t), \tau^e(t))\) on \((\mathbb{R}^n \times \mathcal{S} \times \mathbb{R})\) is a Markov process with generator
\[
\mathcal{L}^e = A + \frac{\partial}{\partial \tau} + \frac{1}{\epsilon} e^{-\epsilon t} F(\epsilon^2 x, y) \frac{\partial}{\partial x}.
\]
(7.22)

Now the “\(y\) process” is defined on \(\mathcal{S} \times \mathbb{R}^1\) with generator \(A + \partial/\partial \tau\).

Since \(A\) and \(\partial/\partial \tau\) commute and \(\partial/\partial \tau\) (acting on almost periodic functions if \(F(x, y)\) is analytic in \(x\), for example) has discrete pure imaginary spectrum we are in the situation described above.

Note that the transformation \(7.21\) picks out the envelope (slowly varying part) of the process \(x^e(t)\). The operator \(\mathcal{L}^e\) has now the form
\[
\mathcal{E} f(x) = \lim_{T \to \infty} \frac{1}{T} \int_0^T dt \int_0^\infty ds \int P(dy) [P(s, y, dz) - \bar{P}(dz)]
\]
\[
e^{-\epsilon t} F(\epsilon^2 x, y) \frac{\partial}{\partial x} \left( e^{-\epsilon t} + n F(\epsilon^2 + n x, z) \frac{\partial f(x)}{\partial x} \right).
\]
(7.23)

and the result is that \(x^e(t)\) converges (weakly) as \(\epsilon \to 0\) to the diffusion with generator \(\mathcal{E}\) given by \(7.23\). An application is given in Section 15.

References for the above are [34–37] with applications in [38]. Operator treatment is given in [37], [40], with diverse applications, and in [39], [41].

We note finally that if instead of \(7.20\) we have
\[
\frac{dx^e(t)}{dt} = \frac{1}{\epsilon^2} Q x^e(t) + \frac{1}{\epsilon} F(x^e(t), y^e(t)), \quad x^e(0) = x,
\]
(7.24)
and if
\[
x^e(t) = e^{-\epsilon t} x^e(t),
\]
(7.25)
then \(x^e(t)\) converges (weakly) as \(\epsilon \to 0\) to the diffusion process generated by
\[
\mathcal{E} f(x) = \lim_{T \to \infty} \frac{1}{T} \int_0^T dt \int \mathcal{P}(dy) \int \psi(y, dz)
\]
\[
\cdot e^{-\epsilon t} F(\epsilon^2 x, y) \frac{\partial}{\partial x} \left( e^{-\epsilon t} F(\epsilon^2, z) \frac{\partial f(x)}{\partial x} \right).
\]
(7.26)

We shall use this in Section 16.

8. Diffusion Approximation of Markov Chains

The term diffusion approximation is used in at least two different ways. One is in the context of problems described in Section 7, i.e., diffusion that emerges as a result of more and more rapidly varying fluctuations. The other is the passage from a motion by discontinuous movements, a random walk, to a continuous motion, a diffusion. In the latter problems ergodic theory does not enter. There are actually two ways in which diffusion approximations of the second kind arise. One is the random walk (or Markov chain) to diffusion limit when the random walk is given at first. In the other the diffusion is given, a consistent difference approximation is defined, and then one studies the convergence of the difference approximation. The literature on these problems is huge. We mention here the work of Khinchine [42],
functions are preserved by $T(t)$ the proof is elementary and amounts to a two-term Taylor expansion in the expression ($f$ smooth)

$$\frac{1}{\epsilon}\left(T_{\epsilon} - I\right)f(x) - \mathcal{D}f(x)$$

as $\epsilon \downarrow 0$. Here $\mathcal{D}$ is the diffusion operator with coefficients $a(x)$ and $b(x)$.

The "nonsmooth" case requires compactness arguments with best results given in [6].

9. STABILITY, INVARIANT MEASURES AND LARGE DEVIATIONS FOR STOCHASTIC EQUATIONS

Consider again problem (4.9) and its first approximation (4.10) under hypothesis (4.11) so that for any $T < \infty$

$$\limsup_{\epsilon \downarrow 0, \theta \to T} |x^\epsilon(x) - \bar{x}(t)| = 0,$$

for all starting points $x \in \mathbb{R}^n$. Suppose that

$$\bar{F}(0) = 0, \quad \bar{F}(0, t) = 0, \quad t \geq 0 \quad (9.1)$$

and that the matrix $\partial F(0)/\partial x$ has eigenvalues with negative real parts. $F(x, t)$ and $\bar{F}(x)$ are smooth vector functions of $x$ here. Hypothesis (9.1) implies that $x = 0$ is an equilibrium point of the limit problem (4.10); in fact an asymptotically stable equilibrium point, since the variational matrix has eigenvalues with negative real parts. The question is now whether the solution $x^\epsilon(x)$, with $\epsilon$ fixed but sufficiently small, converges as $t \uparrow \infty$ to the origin if the initial point is near enough to the origin.

The answer to this question is yes; it is a result of Bogoliubov which can be found in [17] (cf. also [18]).

Before returning to this problem we mention some other related, more difficult questions. If the limit problem (4.10) has an orbitally stable periodic solution, does (4.9) also have a periodic solution for $\epsilon$ sufficiently small? If for $\epsilon > 0$ there are no periodic solutions, are there solutions that remain on some specific (invariant) manifold? The latter is the problem of bifurcation of a periodic solution of
(4.10) to an invariant manifold. We refer to [19] for consideration of these questions and further references to the sizable literature.

Let us consider now Bogoliubov's problem. Since
\[
\frac{dx(t)}{dt} = F(x(t)), \quad x(0) = x, \tag{9.2}
\]
has \(x = 0\) as a stable solution, there is a smooth function \(V(x)\) in \(\mathbb{R}^n\) that is positive definite
\[
V(x) \geq 0, \quad V(x) = 0 \iff x = 0 \tag{9.3}
\]
and
\[
F(x) \cdot \frac{\partial V(x)}{\partial x} = \sum_{i=1}^{n} \frac{\partial F(x)}{\partial x_i} \frac{\partial V(x)}{\partial x_i} \leq -\gamma V(x), \quad \gamma > 0, \tag{9.4}
\]
for \(x\) is a fixed neighborhood \(\mathcal{D}\) of zero having compact closure.

Let \(V_1(x, \tau)\) be defined by
\[
V_1(x, \tau) = -\int_{0}^{\tau} \left[ F(x, s) - F(x) \right] \frac{\partial V(x)}{\partial x} \, ds, \tag{9.5}
\]
and put
\[
V^*(x, t) = V(x) + \epsilon V_1(x, t/\epsilon). \tag{9.6}
\]
Since \(V_1\) is bounded in \(\tau\) (uniformly), if \(\epsilon\) is sufficiently small and \(x\) is in a fixed neighborhood of zero with compact closure there are constants positive \(c_1\) and \(c_2\) such that for \(\epsilon \leq \epsilon_0,\)
\[
c_1 V(x) \leq V^*(x, t) \leq c_2 V(x), \quad x \in \mathcal{D}, \tag{9.7}
\]
Let \(\tilde{\gamma}\) be a constant to be chosen below and consider
\[
\left( \frac{\partial}{\partial t} + F\left( x^*(t), \frac{t}{\epsilon} \right) \cdot \frac{\partial}{\partial x} + \tilde{\gamma} \right) V^*(x, t)
\]
\[
= \tilde{\gamma} V^*(x, t) + F(x) \cdot \frac{\partial V(x)}{\partial x} + \epsilon F\left( x, \frac{t}{\epsilon} \right) \cdot \frac{\partial V_1(x, t/\epsilon)}{\partial x}
\]
\[
\leq \left( \tilde{\gamma} c_2 - \gamma + c_0 \epsilon \right) V(x)
\]
where \(c_3\) is a constant. If \(\epsilon\) is sufficiently small there is a \(\tilde{\gamma} > 0\) such that
\[
\left( \frac{\partial}{\partial t} + F\left( x^*(t), \frac{t}{\epsilon} \right) \cdot \frac{\partial}{\partial x} + \tilde{\gamma} \right) V^*(x, t) \leq 0, \quad x \in \mathcal{D}. \tag{9.8}
\]
Thus,
\[
e^{\tilde{\gamma} t} V(x^*(t)) \leq e^{\tilde{\gamma} t} \frac{1}{c_1} V^*(x^*(t), t) \leq \frac{c_0}{c_1} V(x), \quad x \in \mathcal{D}, \quad t \geq 0, \tag{9.9}
\]
from which Bogoliubov's result follows since \(V(x)\) is positive definite.

The above result can be generalized immediately to the stochastic problem (4.16) or (7.1).

Consider (4.16) under the hypothesis that the recurrent potential kernel (4.20) is well defined and with \(\mathcal{F}(x)\) of (4.17) satisfying
\[
\mathcal{F}(0) = 0, \quad \mathcal{F}(0, y) = 0. \tag{9.10}
\]
Suppose there is a positive definite Lyapounov function \(V(x)\) satisfying (9.3) and (9.4). Then there is an \(\epsilon_0 > 0\) such that for given \(\eta_1 > 0\) and \(\eta_2 > 0\)
\[
P_{x, \epsilon}\{|x^*(t)| \leq \eta_2 e^{-\tilde{\gamma} t}, \quad t \geq 0\} \geq 1 - \eta_1, \quad \epsilon < \epsilon_0,
\]
provided \(|x(0)| = |x| < \delta\) and where \(\tilde{\gamma} > 0\) is a constant and \(y \in \mathcal{S}.
\]
The proof of this is almost identical to the one just given taking into account the modifications necessary to estimate the extra martingale term (as in (4.22)) using Kolmogorov's inequality.

Ilyaponov methods, like the above, for Itô stochastic equations are discussed in [3], [46] and [47], among other places. Extension of the method to stochastic Bogoliubov-like problems are given in [48]. In [48] only (7.1) is treated, it being the more difficult of the two ((4.16), (7.1)) situations.

In the stochastic context, results concerning stability and bifurcation of periodic solutions have not been analyzed as far as we know.
Let us now consider (6.2) and its limit equation (6.6). Suppose* that \( \bar{F}(0) = 0 \) again and take the solution \( \bar{x}(t) \equiv 0 \) in (6.6) so that the fluctuation process is time homogeneous. Since the forcing term
\[
\int_0^t \sigma(0) \, dw(s)
\]
is present, the process \( z(t) \) cannot tend to a fixed limit as \( t \to \infty \). At best, we may expect that as \( t \to \infty \) it reaches a stationary distribution. We can easily assess the situation since \( z(t) \) is Gaussian and time homogeneous. In fact if \( Q = \partial \bar{F}(0)/\partial x \) has eigenvalues with negative real parts, \( z(t) \) has a unique stationary distribution which is Gaussian with mean zero and covariance (cf. (6.9))
\[
\int_0^\infty e^{\epsilon t} u(0) e^{\epsilon t} \, dt.
\]
The problem is now to find out if for \( \epsilon \) sufficiently small but fixed, \( z'(t) \), the solution of (6.2) (with \( \bar{x}(t) \equiv 0 \) and \( F(0, y) \equiv 0 \)) also has a stationary or invariant distribution as \( t \to \infty \) (which will not be Gaussian, however).

From results in [48] one can conclude that this is sometimes true, but the stationary distribution when \( \epsilon \) is positive may not be unique. Any sequence of such invariant distributions tends to the Gaussian limit as \( \epsilon \to 0 \), however.

In many important problems \( z'(t) \) of (6.2) fails to behave as \( t \to \infty \), \( \epsilon > 0 \), like the limit as \( t \to \infty \) of (6.6), no matter how small \( \epsilon > 0 \) is chosen. Stated another way, the limits \( t \to \infty \) and \( \epsilon \to 0 \) cannot be interchanged without changing the result. The reasons for this are clear from the nature of the approximation (6.6) which is local—in the neighborhood of a fixed, stable limit orbit. When \( t \to \infty \) with \( \epsilon > 0 \) the global motion governed by the vector field \( F \) enters the picture and large deviations (which are not negligible when \( t \to \infty \)) may distort the local picture entirely.

In Section 12 we return to the question of large deviations in a different context.

* The \( \omega \in \Omega, (\Omega, \mathcal{F}, P) \) a probability space, is omitted.

10. TRANSPORT PROBLEMS

Transport problems are of the form (7.1) with \( y(t) \) a jump Markov process so that its infinitesimal generator \( A \) on \( B(S) \) has the form
\[
Af(y) = q(y) \int_S \pi(y, dz) f(z) = q(y) f(y).
\]
(10.1)

Here we assume that
\[
0 < q_t < q(y) \leq q_u < \infty
\]
(10.2)
and that the transition probability function \( \pi(y, B) \) satisfies Doeblin's condition relative to some reference measure on \( S \). Thus the jump process \( y(t) \) on \( S \) governed by \( A \) is strongly ergodic. The function \( q(y) \) is the local rate with which jumps take place and \( \pi(y, B) \) is the jump transition probability at the instant that a jump occurs.

The process \((x^\epsilon(t), y^\epsilon(t))\) governed by (7.1) is a Markov process on \( \mathcal{Q}^\epsilon \times S \) with generator
\[
\mathcal{Q}^\epsilon = \frac{1}{\epsilon^2} A + \frac{1}{\epsilon} F(x, y) \frac{\partial}{\partial x}.
\]
(10.3)

From the point of view of differential equations we are dealing with the integrodifferential Kolmogorov equation
\[
\frac{\partial u^\epsilon(t, x, y)}{\partial t} = \frac{1}{\epsilon} F(x, y) \frac{\partial u^\epsilon(t, x, y)}{\partial x} + \frac{q(y)}{\epsilon^2} \int_S \pi(y, dz) u^\epsilon(t, x, z) - \frac{q(y)}{\epsilon^2} u^\epsilon(t, x, z),
\]
(10.4)
\[
t > 0, \quad x \in \mathcal{Q}^\epsilon, \quad y \in S,
\]
\[
u^\epsilon(0, x, y) = f(x, y).
\]

This is a transport equation, a basic equation in many areas of physics (cf. [52] and the references cited there).

In the physical contexts one deals with particle densities, hence forward equations, but in any case the underlying probability measure is what one wants to analyze. In physical contexts, \( S \) is usually \( \mathcal{Q}^\epsilon \) or a subset thereof and \( y \) is thought of as velocity of a moving
particle. One also takes, usually, $F(x, y) = y$ if the medium without the scattering is not refracting. The parameter $\epsilon$ in (10.4) corresponds to the mean free time (or mean free path) between collisions or discontinuous changes in the velocity. Equation (10.4) subject to (7.2) is scaled for passage to the diffusion limit.

Equation (10.4) is also called a conservative transport equation because $A_1 = 0$. In many interesting problems the operator $A$ generates a semigroup $T(t)$ which is positivity preserving (like a Markov semigroup) but $T(t)1 \neq 1$ so it is not a Markov semigroup. Such problems are called nonconservative and the generator $A$ can be written a sum of a Markov generator $\bar{A}$ plus a potential term $V$ (multiplication operation by $V$). The sign of $V$ corresponds to creation or destruction (locally) of particles.

Depending on whether or not the term $V$ is large (of $O(1)$) or small, the corresponding asymptotic analysis of (10.4) changes. If $V = O(\epsilon^2)$ there are no essential changes in (10.4) as it stands. If $V = O(1)$ then in general there will be no limit as $\epsilon \to 0$ in the usual sense. Let $T_\epsilon(t)$ be the semigroup generated by $\bar{A} + V$ on $S$, suppose that

$$\lim_{t \to t^+} \frac{1}{\epsilon} \log \|T_\epsilon(t)\| = 0^*$$

and that it satisfies a Doeblin condition (cf. (3.11)). Then [51] there is a positive function $\phi(y)$ such that $(\bar{A} + V)\phi(y) = 0$. If we now consider not $u'(t, x, y)$ but $v'(t, x, y) = u'(t, x, y)\phi(y)$ then it satisfies a conservative transport equation and we are back to (10.4).

Note that in (10.4) the initial data depends on both $x$ and $y$ while in (7.6) we did not allow this. Since in the diffusion limit dependence on $y$ disappears, it is clear that an initial layer develops near $t = 0$.

Let $\mathcal{D}$ denote the operator defined by (7.17) with $(a_d)$ and $(b_d)$ as in (7.15) and (7.16) and with $A$ now as in (10.1). Let

$$f(x) = \int P(dy)f(x, y),$$

$$u'(\tau, y; x) = \int P(\tau, y, dz)[f(z, x) - f(x)].$$

The ergodic properties of $y(t)$ generated by $A$ imply that the third term in (10.7), the initial layer correction, is negligible away from an $O(\epsilon^2)$ neighborhood of $t = 0$.

Nonuniform behavior near spatial boundaries, or boundary layers, is discussed in the next section.

11. APPROXIMATIONS IN BOUNDED DOMAINS, BOUNDARY CONDITIONS AND BOUNDARY LAYERS

Let $\mathcal{D} \subset \mathbb{R}^n$ be a bounded open set with smooth boundary $\partial \mathcal{D}$. For each $x \in \partial \mathcal{D}$ let $S^+$ and $S^-$ be a decomposition of the state space $S$ as follows

$$S^\pm = \{y \in S : \hat{n}(x) \cdot F(x, y) \equiv 0, x \in \partial \mathcal{D}\}. \quad (11.1)$$

Here $\hat{n}(x)$ denotes the unit outward normal to $\partial \mathcal{D}$ at $x$.

We consider now the boundary value problem

$$\frac{\partial u^e(x)}{\partial x} - \frac{1}{\epsilon} F(x, y) \cdot \frac{\partial u^e(x, y)}{\partial x} + \frac{g(y)}{\epsilon^2} \int_S \pi(y, dz)u^e(x, z)$$

$$= \frac{g(y)u^e(x, y)}{\epsilon^2} = 0,$$

$$x \in \mathcal{D}, \ y \in S^+ \quad \text{and} \quad x \in \partial \mathcal{D}, \ y \in S^- \quad \text{if} \quad u^e(x, y) = f(x, y), \ x \in \partial \mathcal{D}, \ y \in S^+. \quad \text{if} \quad (11.2)$$

Then, it is easily seen that

$$\left| u'(t, x, y) - u(t, x) - u'(t, \frac{x}{\epsilon^2}, y; x) \right| \to 0 \quad (10.7)$$

uniformly in $(x, y) \in \mathcal{D} \times S$ and $0 \leq t \leq T < \infty$ as $\epsilon \to 0$. Here $u(t, x)$ satisfies the diffusion equation

$$\frac{\partial u(t, x)}{\partial t} = \mathcal{L}u(t, x), \quad t > 0$$

$$u(0, x) = f(x). \quad (10.8)$$

Note that $u'(\tau, y; x)$ satisfies the initial layer equation

$$\frac{\partial u'(\tau, y; x)}{\partial \tau} = Au'(\tau, y; x), \quad \tau > 0$$

$$u'(0, y; x) = f(x, y) - f(x). \quad (10.9)$$

This limit exists by subadditivity of $\log\|T_\epsilon(t)\|$.

* This limit exists by subadditivity of $\log\|T_\epsilon(t)\|$.
This is a time independent problem and we are interested in the asymptotic behavior of the solution \(u_\epsilon(x, y)\) as \(\epsilon \to 0\).

First we discuss the probabilistic representation of the solution of (11.2) and its well-posedness.

Let \((x^\prime(t), y^\prime(t))\) be the Markov process on \(R^n \times S\) generated by (10.3) under the usual ergodicity hypotheses on \(A\) of (10.1). Let \(D\) be defined by (7.17) (via (7.15)–(7.16) with \(A\) given by (10.1)) and assume that its coefficients are smooth with at least one diagonal entry of \((a_{ij})\) strictly positive in \(D\). We assume (7.2) holds.

Let \(\tau\) be the first exit time of \(x^\prime(t)\) from \(D\). From our assumptions it follows that for all \(\epsilon\) sufficiently small \(E_{x, y}(\tau) \leq c < \infty\). Now on the basis of the remarks in Section 2 regarding the use of the optional stopping theorem for probabilistic representations, we conclude that if

\[
u_\epsilon(x, y) = E_{x, y} [f(x^\prime(\tau), y^\prime(\tau))],
\]

then this is well defined and is the unique solution of (11.2) (its integral equation version) with \(f\) bounded.

If \(f\) was a function of \(x\) only then, it is clear that \(u_\epsilon(x, y) \to u(x)\) uniformly in \((x, y) \in D \times S\) as \(\epsilon \to 0\) where

\[
Du(x) = 0, \quad x \in D
\]

\[
u(x) = f(x), \quad x \in \partial D
\]

is an elliptic boundary value problem. The proof of this is a straightforward extension of the one in Section 7 given that \(E \{\tau\} \leq c < \infty\).

The latter is proved by an argument like the one in Section 9.

How does \(u_\epsilon(x, y)\) behave in the general case? We must consider more closely what happens near the boundary \(\partial D\).

We define stretched boundary layer coordinates near \(\partial D\) so that \(x \to (y, \eta)\) where (for \(x \in D\))

\[
x = \xi(y) + c\eta \hat{h}(y), \quad \eta \leq 0.
\]

Here \(\xi(y)\) is the local parametric representation of the surface \(\partial D\), \(y = (y_1, y_2, \ldots, y_{n-1}) \in R^{n-1}\), and \(\eta \leq 0\) is distance from the boundary. With each \(x\) near enough to \(\partial D\), \(\xi(y) = \xi(y(x))\) since \(y = y(x)\) is the unique point on \(\partial D\) with \(x - \xi(y)\) parallel to the normal at the same point. In local boundary layer coordinates,

\[
\frac{\partial}{\partial x_i} \to \frac{1}{\epsilon} \hat{h}_i(y) \frac{\partial}{\partial y_i} + \sum_{k=1}^{n-1} \Gamma_k(y) \frac{\partial}{\partial y_k}, \quad i = 1, 2, \ldots, n,
\]

where \(\hat{h}_i(y)\) or \(\hat{h}_i(x)\), \(x \in \partial D\) are the components of the normal vector and \(\Gamma_k(y)\) are functions that depend on the geometry of \(\partial D\) (its principal curvatures) and are regular in \(\eta \geq 0\).

Next we express the generator \(L^\epsilon\) near \(\partial D\) in terms of \((y, \eta)\). It has the form \((1/\epsilon^2)L_{BL} + \) less singular in \(\epsilon\) terms where \((x \in \partial D)\)

\[
L_{BL} = \hat{h}(y) \cdot F(x, y) \frac{\partial}{\partial y} + q(y) \int \pi(y, dz) - q(y).
\]

(11.6)

It is easy to see that \(L_{BL}\) generates a Markov process \((H(t), y(t))\) on \((-\infty, \infty) \times S\). In fact, since \(x \in \partial D\) is merely a parameter in \(L_{BL}\), if we let

\[
v(y; x) = \hat{h}(y) \cdot F(x, y),
\]

then \(y(t)\) is the usual process on \(S\) and

\[
H(t) = \eta + \int_0^t v(y(s); \xi) ds,
\]

(11.7)

(11.8)

with \(\eta\) the starting point.

Let \(\tau\) be the first exit time of \(H(t)\) from \((-\infty, 0]\) given that \(H(0) = \eta < 0\) or \(\eta = 0\) and \(y \in S^-\) (\(a(y) \leq 0\) on \(S^-\)). Hypothesis (7.2) implies that \(\tau\) is a proper random variable and hence we may consider the exit random variable \(y(\tau)\) which takes values in \(S^+\). Let

\[
P_{BL}(\eta, y, B), y \in S^+, B \subset S^+, \eta \geq 0 \text{ be defined by}
\]

\[
P_{BL}(\eta, y, B) = P[y(\tau) \in B \mid H(0) = -\eta, y(0) = y].
\]

(11.9)

What we have just pointed out and the fact that the operator \(L_{BL}\) has \(\eta\)-independent coefficients, implies that \(P_{BL}\) is a Markov transition probability function on \(S^+\) with \(\eta\) (=distance from the origin) playing the role of the time variable. The boundary layer analysis is intimately related to the ergodic properties of \(P_{BL}\) as \(\eta \to \infty\).
Suppose in fact that $P_{BL}$ is ergodic and let $\bar{P}_{BL}$ be its invariant measure (it is not necessary to assume existence of the recurrent potential unless we want exponential decay of the boundary layer corrections away from $\partial \Omega$). $\bar{P}_{BL}$ has the following meaning: it is the equilibrium (or stationary) distribution of the "velocity" variable $y(t)$ at the instant $x(t)$ touches $\partial \Omega$ while the process begins further and further in the interior.

Define

$$f(x) = \int_{\eta^+} \bar{P}_{BL}(dy)f(x, y),$$

and, for $\eta \leq 0$,

$$u^{BL}(\eta, y; x) = \int_{\eta^+} P_{BL}(-\eta, y, dz)[f(x, z) - f(x)], \quad x \in \partial \Omega. \quad (11.11)$$

We may let $y \in S$ in (11.11) by extending the above definition. If the recurrent potential of $P_{BL}$ exists then (11.11) tends to zero exponentially fast as $\eta \to -\infty$, uniformly in $y$ (and in the parameter $x \in \partial \Omega$).

With this information at hand we return to (11.2) and (11.3). We have that

$$\lim_{\epsilon \to 0} \sup_{x \in \partial \Omega} \left| u^\epsilon(x, y) - u(x) - \xi(x)u^{BL}\left(\frac{x - \xi(x)}{\epsilon}, y; \xi(x)\right)\right| = 0.$$  

Here $\xi(x) = 1$ in a sufficiently small neighborhood of $\partial \Omega$ and zero outside a larger neighborhood and it is smooth. The function $u(x)$ is the solution of (11.5).

The proof of the above result and many other examples (reflecting boundary conditions, etc.) can be found in [52].

12. APPROXIMATION OF PROCESSES WITH SMALL DIFFUSION

We shall consider the effect of a small amount of noise on a system, when the noise is fully developed. The problems in this section are not related to the rapid fluctuation limit. They have been analyzed by a number of authors, for example in [49], [50], [53]-[58], by a variety of methods, probabilistic and analytical.

Let $(a_i(x))$ and $(b_i(x))$ be a symmetric $n \times n$ positive definite matrix function of $x \in R^2$ and an $n$-vector function of $x \in R^n$, both assumed smooth. Let $\Omega \subset R^n$ be a bounded region with smooth boundary. We consider first the boundary value problem ($\epsilon > 0$)

$$\mathcal{L}u^\epsilon(x) = \epsilon \sum_{i=1}^n a_{ii}(x) \frac{\partial u^\epsilon(x)}{\partial x_i} + \sum_{j=1}^n b_j(x) \frac{\partial u^\epsilon(x)}{\partial x_j} - c(x) u^\epsilon(x) + f(x) = 0, \quad x \in \partial \Omega.$$  

$$u^\epsilon(x) = g(x), \quad x \in \partial \Omega. \quad (12.1)$$

Here $f(x)$, $x \in \partial \Omega$ and $g(x)$, $x \in \partial \Omega$ are given smooth functions and we assume that

$$c(x) \geq c_0 > 0, \quad (12.2)$$

although this is not necessary for what follows.

It is clear that (12.1) has a unique smooth solution $u^\epsilon$ and the problem is to describe its asymptotic behavior as $\epsilon \to 0$. Let $(u_0(x))$ be the smooth symmetric square root of $(a_i(x))$ and let $x^\epsilon(t)$ be the diffusion process on $R^n$ solving the stochastic differential equation

$$dx^\epsilon(t) = b(x^\epsilon(t)) dt + \sqrt{2\epsilon} \sigma(x^\epsilon(t)) dw(t), \quad x^\epsilon(0) = x. \quad (12.3)$$

Let $\tau$ be the first exit time from $\Omega$ of $x^\epsilon(t)$ given that $x^\epsilon(0) = x \in \partial \Omega$. Then,

$$u^\epsilon(x) = E_x \left\{ \int_0^{\tau} \left( \exp \left( -\int_0^s c(x^\epsilon(\gamma)) d\gamma \right) \right) f(x^\epsilon(s)) ds \right\} + E_x \left\{ \left( \exp \left( -\int_0^{\tau} c(x^\epsilon(\gamma)) d\gamma \right) \right) g(x^\epsilon(\tau)) \right\}. \quad (12.4)$$

Depending on the nature of the trajectories of

$$dx(t) = b(x(t)) dt, \quad x(0) = x, \quad (12.5)$$

the formal limit of (12.3) when $\epsilon$, the noise intensity, is zero, different types of behavior can arise. We begin with the simplest.

Assume that $b(x) \neq 0$ for $x \in \partial \Omega$ and let

$$\Gamma^+ \subset \partial \Omega = \{ x \in \partial \Omega : \exists \tau > 0 \text{ and } x \in \partial \Omega \text{ such that } x(0) = x \text{ and } x = x(\tau) \in \partial \Omega \}. \quad (12.6)$$
Let $\Gamma^- = \partial D - \Gamma^+$. The set $\Gamma^-$ consists of points that can be reached along orbits of (12.5) when the time is running backward (negative) and of points that are on orbits themselves. Note also that on $\Gamma^+$, $b(x) \cdot \hat{n}(x) > 0$, where $\hat{n}(x)$ is the unit outward normal, and on $\Gamma^-$, $b(x) \cdot \hat{n}(x) \leq 0$. To avoid complications we assume that $b(x) \cdot \hat{n}(x) = 0$ only at two isolated points of the boundary.

It is clear that at $x \in D$, $u^\epsilon(x) \to u_0(x)$ where

$$u_0(x) = \int_0^T \left( \exp \left( - \int_0^s c(x(y)) \, dy \right) \right) f(x(s)) \, ds + \left( \exp \left( - \int_0^T c(x(y)) \, dy \right) \right) g(x(T)).$$

(12.7)

Here $T$ is the exit time of the deterministic orbits (12.5). Evidently $u_0(x) \neq g(x)$, $x \in \Gamma^-$ (except the two points) and $u_0(x)$ does not satisfy the boundary condition in the formal limit $\epsilon = 0$ in (12.1).

We must introduce a boundary layer correction.

We employ the notation of Section 11. Near $\partial D$ we represent $x$ by $(\gamma, \eta)$ coordinates where

$$x = \gamma \cdot \delta \gamma + \exp \hat{n}(\gamma), \quad \eta < 0$$

(12.8)

(cf. (11.4), (11.5)). In these coordinates the operator $\mathcal{L}^\epsilon$ on the left side of (12.1) takes the form

$$\frac{1}{\epsilon} \left[ \sum_{i,j=1}^n a_{ij}(\gamma) \delta \eta \frac{\partial}{\partial \eta} + \sum_{j=1}^n b_j(\gamma) \delta \eta \frac{\partial}{\partial \eta} \right] + O(1),$$

(12.9)

where $O(1)$ stands for a differential operator in $\partial / \partial \eta$, $\partial / \partial \gamma_j$ with coefficients that are regular in $\epsilon > 0$. Let

$$\delta(x) = \sum_{i,j=1}^n a_{ij}(x) \delta \eta \delta \eta_j, \quad x \in \partial D$$

$$\delta(x) = \sum_{j=1}^n b_j(\gamma) \delta \eta_j, \quad x \in \partial D$$

(12.10)

and

$$\mathcal{L}^BL = \delta(x) \frac{\partial^2}{\partial \eta^2} + \delta(x) \frac{\partial}{\partial \eta}.$$  

(12.11)

Let $u^{BL}(\eta; x)$ be the solution of

$$\mathcal{L}^BL \eta \mathcal{L}^BL(\eta; x) = 0, \quad \eta < 0$$

$$u^{BL}(0; x) = g(x),$$

$$u^{BL}(x) = g(x), \quad x \in \Gamma^-.$$  

(12.12)

Clearly

$$u^{BL}(\eta; x) = g(x) e^{-\frac{\xi(x)}{\epsilon} \eta}, \quad x \in \Gamma^-.$$  

(12.13)

where, we recall, $\xi(x) < 0$ on $\Gamma^-$.

Let $\xi(x)$ be equal to one in a neighborhood $\theta_1$ of $\Gamma^-$, equal to zero outside of $\theta \supset \theta_1$ and smooth. Let $\theta_2$ be small enough so that the coordinates (12.8) are valid. Then

$$\limsup_{\epsilon \to 0} \sup_{x \in D} \left| u^\epsilon(x) - u_0(x) - \xi(x) u^{BL} \left( \frac{x - \xi(x) \cdot \hat{n}(x)}{\epsilon} ; x \right) \right| = 0.$$  

(12.14)

To prove this result we let

$$u^\epsilon = u^\epsilon - u_0 - \xi(x) u^{BL}.$$  

(12.15)

Then $u^\epsilon(x) = 0$, $x \in \partial D$ by construction. Moreover, with $\mathcal{L}^\epsilon$ the operator on the left side of (12.1), we have

$$\mathcal{L}^\epsilon u^\epsilon = -f - \mathcal{L}^BL u_0 - \mathcal{L}^BL(\xi u^{BL})$$

$$= -\epsilon \sum_{i,j=1}^n a_{ij} \frac{\partial^2 u_0}{\partial x_i \partial x_j} - \xi \left( \mathcal{L}^\epsilon - \frac{1}{\epsilon} \mathcal{L}^BL \right) u^{BL}$$

$$- \left[ \mathcal{L}^\epsilon(\xi u^{BL}) - \xi \mathcal{L}^BL u^{BL} \right].$$

(12.16)

By the decay of $u^{BL}$ to zero as $\eta \to -\infty$ it follows that the right-hand side of (12.16) is uniformly bounded and at each $x \in \partial D$ it tends to zero as $\epsilon \to 0$. In view of (12.2) and the maximum principle the result (12.14) follows (actually one must look in a bit more detail at what happens near the two special points of $\Gamma^-$ where $\hat{n} \cdot b = 0$; continuity, however, prevents difficulties).

As another example we consider a Neumann problem as follows

$$\mathcal{L}^\epsilon u^\epsilon(x) + f(x) = 0, \quad x \in D,$$

$$\gamma(x) \cdot \nabla u^\epsilon(x) + a(x) u^\epsilon(x) - g(x) = 0, \quad x \in \partial D.$$  

(12.17)
Here $\gamma(x)$ is a vector field defined on $\partial \mathcal{D}$, smooth, so that
\[ \gamma(x) \cdot \hat{n}(x) \leq \gamma_0 < 0, \quad x \in \partial \mathcal{D}, \] (12.18)
i.e., $\gamma(x)$ points inward and is uniformly nontangential. In addition we assume that
\[ \alpha(x) \geq \alpha_0 > 0. \] (12.19)

The stochastic process associated with (12.17) is a reflected diffusion process. In terms of stochastic differential equations it can be defined as the solution of
\[ dx^t(t) = b(x^t(t)) \, dt + \sqrt{2} \sigma(x^t(t)) \, dw(t) + \gamma(x^t(t)) \, dN^x(t), \tag{12.20} \]
\[ x^t(0) = x \in \mathcal{D}. \]

Here $N^x(t)$ is a continuous, nondecreasing, nonanticipating functional of Brownian motion such that $N^x(t)$ increases only when $x^t(t) \in \partial \mathcal{D}$. In terms of martingales [8], the approach we follow, the measure $P^x_\epsilon$ on $C([0, \infty); \mathcal{D})$ associated with $(L^\gamma, \gamma)$ of (12.17) is the solution of the submartingale problem
\[ (i) \ P^x_\epsilon[x(0) = x], \]
\[ (ii) \ h(x(t)) - \int_0^t L^\gamma h(x(s)) \, ds \text{ is a submartingale relative to } P^x_\epsilon \text{ for each } h \in C^2(\mathcal{D}) \text{ with } \gamma(x) \cdot \nabla h(x) \geq 0. \] (12.21)

For each $h \in C^2(\mathcal{D})$ there is an increasing process $N^h_\epsilon(t)$ so that
\[ h(x(t)) - \int_0^t L^\gamma h(x(s)) \, ds - N^h_\epsilon(t) \]
is a martingale relative to $P^x_\epsilon$ (when the measure is explicitly displayed we do not use superscript $\epsilon$ on $\alpha(\cdot)$). To normalize this increasing process we let $\phi(x)$ be a support function for the smooth domain $\mathcal{D}$, i.e.,
\[ \mathcal{D} = \{ x \mid \phi(x) > 0 \}, \]
\[ \partial \mathcal{D} = \{ x \mid \phi(x) = 0 \}. \] (12.22)

and $|\nabla \phi(x)| \geq 1$, $x \in \partial \mathcal{D}$. Let $N^\epsilon(t)$ be the increasing process corresponding to $\hat{n} = \hat{\phi}$ above. For any other smooth $\hat{n}$ we have that
\[ N^\epsilon(t) = \int_0^t \gamma(x(s)) \cdot \nabla \phi(x(s)) \, dN^\epsilon(s), \] (12.23)

which is well defined by (12.18). Let us, in particular, normalize the choice of $\phi$ so that $\gamma \cdot \nabla \phi = 1$.

Now the solution of (12.17) can be expressed as follows.
\[ u^\epsilon(x) = E_x \left\{ \int_0^\infty \left( \exp \left[ - \int_0^t c(x^s(s)) \, ds - \int_0^t \alpha(x^s(s)) \, dN^\epsilon(s) \right] \right) f(x^t(s)) \, ds \right\} \]
\[ + E_x \left\{ \int_0^\infty \left( \exp \left[ - \int_0^t c(x^s(s)) \, ds - \int_0^t \alpha(x^s(s)) \, dN^\epsilon(s) \right] \right) \right\} \times g(x^t(s)) \, dN^\epsilon(s). \] (12.24)

That this is a well-defined expression follows easily from the boundedness of $f$ and $g$ and from (12.19), (12.2).

Let us return to the perturbation expansion. We decompose $\partial \mathcal{D}$ into $\Gamma^+$ and $\Gamma^-$ as before. The boundary layer near $\Gamma^-$ is essentially the same as before. What is interesting now is the interior limit solution $u_\gamma(x)$, valid on $\mathcal{D} \cup \Gamma^+$ and satisfying
\[ b(x) \cdot \nabla u_\gamma(x) - c(x)u_\gamma(x) + f(x) = 0, \quad x \in \mathcal{D}, \]
\[ \gamma(x) \cdot \nabla u_\gamma(x) - a(x)u_\gamma(x) + g(x) = 0, \quad x \in \partial \mathcal{D}. \] (12.25)

Any vector on $\partial \mathcal{D}$ can be decomposed into the sum of a vector along the normal $\hat{n}$ and a vector tangent to $\partial \mathcal{D}$. For $\gamma(x)$ and $b(x)$ we write
\[ \gamma(x) = \gamma(x) \cdot \hat{n}(x) \hat{n}(x) + \gamma_t(x), \quad x \in \partial \mathcal{D}, \] (12.26)
\[ b(x) = b(x) \cdot \hat{n}(x) \hat{n}(x) + b_t(x), \quad x \in \partial \mathcal{D}, \] (12.27)

where $\gamma_t(x)$ and $b_t(x)$ are defined by (12.26) and (12.27). Using these formulas in (12.25) we obtain
\[ \gamma_t(x) \cdot \hat{n}(x) \frac{\partial u_\gamma(x)}{\partial \hat{n}(x)} + \gamma_t(x) \cdot \nabla u_\gamma(x) - a(x)u_\gamma(x) + g(x) = 0, \quad x \in \Gamma^+, \]
\[ b_t(x) \cdot \hat{n}(x) \frac{\partial u_\gamma(x)}{\partial \hat{n}(x)} + b_t(x) \cdot \nabla u_\gamma(x) - c(x)u_\gamma(x) + f(x) = 0, \quad x \in \Gamma^+. \] (12.28)
Since \( b \cdot \hat{n} > 0 \) and \( y \cdot \hat{n} < 0 \), for \( x \in \Gamma^+ \) we may eliminate \( \partial u_0(x)/\partial \hat{n}(x) \) and obtain an equation for a motion entirely confined to \( \Gamma^+ \):

\[
\left( \frac{b(x)}{b(x) \cdot \hat{n}(x)} - \frac{\gamma(x)}{\gamma(x) \cdot \hat{n}(x)} \right) \cdot \nabla u_0(x) - \left( \frac{c(x)}{b(x) \cdot \hat{n}(x)} - \frac{\alpha(x)}{\gamma(x) \cdot \hat{n}(x)} \right) u_0 + \left( \frac{f(x)}{b(x) \cdot \hat{n}(x)} - \frac{g(x)}{\gamma(x) \cdot \hat{n}(x)} \right) = 0, \quad x \in \Gamma^+. \tag{12.29}
\]

The validity of the first approximation \( u_0(x) \) satisfying the first equation in (12.25) in the interior and (12.29) on \( \Gamma^+ \) (+ boundary layers on \( \Gamma^- \)) depends crucially on the nature of the tangential vector field

\[
F_i(x) = \frac{b_i(x)}{b(x) \cdot \hat{n}(x)} - \frac{\gamma_i(x)}{\gamma(x) \cdot \hat{n}(x)}, \quad x \in \Gamma^+. \tag{12.30}
\]

If \( F_i(x) \) has one stable equilibrium point in \( \Gamma^+ \), then the expansion is still valid as in the case of the Dirichlet problem.

If, however, \( F_i(x) \) has unstable equilibrium points or other singular behavior, interior boundary layers will develop. We shall not pursue this further here but we shall pass to our third and last example.

Consider the problem

\[
\mathcal{L}^e u^e(x) = \epsilon \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 u^e(x)}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial u^e(x)}{\partial x_i} + f(x) = 0, \quad x \in \mathcal{D},
\]

\[
u^e(x) = g(x), \quad x \in \partial \mathcal{D}, \tag{12.31}
\]

where \( \mathcal{D} \subset \mathbb{R}^n \) is bounded and \( \partial \mathcal{D} \) is smooth and \( f \) and \( g \) are given. Assume that

\[
\sum_{i,j=1}^n a_{ij}(x) \xi_i \xi_j \geq a_0 |\xi|^2, \quad a_0 > 0, \tag{12.32}
\]

and that

\[
b(0) = 0, \quad \frac{\partial b(x)}{\partial x} \Bigg|_{x=0} = B \tag{12.33}
\]

has eigenvalues with negative real points. Assume finally that

\[
b(x) \cdot \hat{n}(x) < 0, \quad x \in \partial \mathcal{D} \tag{12.34}
\]

where \( \hat{n}(x) \) is the unit outward normal. Clearly, if \( x^e(t) \) is the solution of (12.3) and \( \tau^e \) the first exit time (a proper random variable) then,

\[
u^e(x) = E_x \left( \int_0^{\tau^e} f(x^e(s)) \, ds \right) + E_x \{ g(x^e(\tau^e)) \}. \tag{12.35}
\]

The problem here is different, and much more difficult, than the previous ones because all of the boundary is the set \( \Gamma^- \); there is no exit possible in the limit. In the case \( f \equiv 0 \), \( u^e(x) \) tends as \( \epsilon \to 0 \) to a constant independent of \( x \) (for each \( x \) in the interior) which is a functional of the data \( g(x) \) on the boundary. We shall consider here a formal analysis that leads to the correct result in this case [58]. Proofs can be found in [55], in a special case, which can be adapted in general; they are much more difficult than the above.

The solution \( u^e \) tends to a constant \( c \) in the interior. Near the boundary, in local boundary layer coordinates we have the boundary layer correction (cf. (12.10)–(12.13)).

\[
(g(x) - c)e^{-\beta(x)/\eta(x)}, \quad x \in \partial \mathcal{D}, \quad (\beta(x) \leq 0, \eta \leq 0). \tag{12.36}
\]

Thus,

\[
u^e(x) \sim c + (g - c)e^{-\beta(x)/\eta(x)} \tag{12.37}
\]

is the first term of a composite expansion with the constant \( c \) still undetermined. Carrying out the expansion to higher order will not allow for the determination of \( c \) so we determine it as follows [58].

Let \( \mathcal{L}^e \) be the operator adjoint to \( \mathcal{L}^e \) in (12.31). By the WKB (or geometrical optics) method [58] we construct an asymptotic solution \( w^e \) of

\[
\mathcal{L}^e w^e = 0
\]

with \( w^e \) normalized to equal one at the origin. It has the form

\[
w^e = e^{\delta \varepsilon}(w_0 + \ldots), \tag{12.38}
\]

where both \( \phi \) and \( w_0 \) can be computed readily by solving the eikonal
(for \( \varphi \)) equation and the transport equation for \( w_\epsilon \). We also have
\[
0 = \int_{\mathbb{R}} \left[ w_\epsilon \Delta^\epsilon u_\epsilon - u_\epsilon \Delta^\epsilon w_\epsilon \right]
= \int_{\mathbb{R}} \left[ w_\epsilon \sum_{n>1} a_n \frac{\partial u_\epsilon}{\partial x_j} - u_\epsilon \sum_{n>1} a_n \frac{\partial w_\epsilon}{\partial x_j} + \sum_{n>1} b_n(x) \hat{\eta}_n(x) u_\epsilon w_\epsilon \right] dx.
\]
(12.39)

Now we insert in the identity (12.39) the expansions (12.38) and (12.37), solve for \( c_\epsilon \), and take the limit \( \epsilon \downarrow 0 \). Thus \( c_\epsilon \) is obtained as the limit of the ratio of two Laplace-type integrals and this solves (formally) the problem of determining the distribution of exit points in cases (12.33), (12.34).

13. NONERGODIC DRIVING PROCESSES AND LOCAL TIME

We return now to the analysis of differential equations with rapidly fluctuating components. We shall examine briefly what happens when in (4.16) the process \( \gamma(t) \) (with \( \gamma_\epsilon(t) \) scaled in a manner other than (4.17)) on \( S \) is not ergodic but null recurrent.\[59\]. We restrict attention to the case \( S = \mathbb{R}^1, \gamma(t) = \) Brownian motion and \( F(x, y) \) has compact support in \( y \) for all \( x \in \mathbb{R} \).

Let
\[
F(x) = \int_{-\infty}^{\infty} F(x, y) \, dy \quad (13.1)
\]
and let \( \bar{x}(t) \) satisfy the equation
\[
\frac{d\bar{x}(t)}{dt} = F(\bar{x}(t)), \quad \bar{x}(0) = x. \quad (13.2)
\]

Let \( l_\epsilon(t) \) be the local time of Brownian motion \( \gamma(t) \) with \( \gamma(0) = y \). That is, formally,
\[
l_\epsilon(t) = \int_0^t \delta(\gamma(s)) \, ds,
\]
or by Tanaka's formula [4]
\[
l_\epsilon(t) = 2\gamma^+(t) - 2 \int_0^t H(\gamma(s)) \, dy(s), \quad (13.3)
\]
where \( y^+ = y \vee 0, H(y) = 1 \) if \( y \geq 0, H(y) = 0, y < 0 \). The formula (13.3) may serve as definition of (13.3) while the formal expression involving the delta function explains the terminology.

Let \( \gamma(t) \) be the process which is the solution of (7.1) with the above assumptions and with \( y(t) = \gamma(t)/\epsilon^2 \).

Then \( \gamma(t), \gamma(t) \) converges (weakly) to \( (\bar{x}(l_\epsilon(t)), \gamma(t)) \) as \( \epsilon \to 0 \) and \( 0 \leq t \leq T < \infty \).

To see what this result implies we look at the case \( F(x, y) = \chi_d(y) \) with \( 0 \in \mathcal{A} \). Then,
\[
\gamma(t) = x + \int_0^t \frac{1}{\epsilon} \chi_d(\gamma(s)) \, ds
\]
\[
= x + \int_0^t \frac{1}{\epsilon} \chi_d(\gamma(s)/\epsilon^2) \, ds
\]
\[
\sim x + \int_0^t \frac{1}{\epsilon} \chi_d\left(\frac{1}{\epsilon} \gamma(s)\right) \, ds \quad \text{(Brownian scaling)}
\]
\[
= x + \int_0^t \frac{1}{\epsilon} \chi_d(\gamma(s)) \, ds
\]
\[
\sim x + \int_0^t \delta(\gamma(s)) \, ds \quad \text{(formally)}.
\]

The appearance of the local time is clear in this example. The general result has locally the same character.

In the event \( F(x) = 0 \) one can obtain results analogous to those of Section 7 where the limit process is a diffusion run at the local time of a suitably scaled driving process. More details can be found in [59].

14. STOCHASTIC TWO-POINT BOUNDARY VALUE PROBLEMS

The asymptotic methods that deal with initial value problems can be adapted to deal with two-point boundary value problems. We shall do this for a general linear problem, an example of which follows
in the next two sections. The considerations here are deterministic and of a formal algebraic type.

Let \( x(t) \) be a vector of dimension \( 2n \) with the first \( n \) components denoted by \( x_1(t) \) and the last \( n \) components denoted by \( x_2(t) \). It satisfies the two-point boundary value problem

\[
\frac{d}{dt} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, \quad t_1 < t < t_2, \tag{14.1}
\]

\[
B_{11}x_1(t_1) + B_{12}x_2(t_2) = b_1,
\]

\[
C_{11}x_1(t_2) + C_{12}x_2(t_2) = c_1. \tag{14.2}
\]

Here the \( 2n \times 2n \) matrix \( A(t) \) is given along with the \( n \times n \) matrices \( B_{11}, B_{12}, C_{11}, C_{12} \) and the \( n \) vectors \( b_1 \) and \( c_1 \). We assume that there exist \( n \times n \) matrices \( B_{21}, B_{22}, C_{21}, C_{22} \) such that the \( 2n \times 2n \) matrices

\[
B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \quad C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},
\]

are nonsingular. We may therefore write the boundary conditions (14.2) in the form

\[
Bx(t_1) = b, \quad Cx(t_2) = c \tag{14.3}
\]

where

\[
b = \begin{pmatrix} b_1 \\ b_1 \end{pmatrix}, \quad c = \begin{pmatrix} c_1 \\ c_1 \end{pmatrix} \tag{14.4}
\]

and \( \tilde{b}_1, \tilde{c}_1 \) are unknown \( n \)-vectors.

Let \( Y(t, s) \) be the fundamental solution of (14.1)

\[
\frac{dY(t, s)}{dt} = A(t)Y(t, s), \quad t > s, \quad Y(t, t) = I, \tag{14.5}
\]

where \( I \) is the \( 2n \times 2n \) identity. We have

\[
\begin{pmatrix} x_1(t_2) \\ x_2(t_2) \end{pmatrix} = \begin{pmatrix} Y_{11}(t_2, t_1) & Y_{12}(t_2, t_1) \\ Y_{21}(t_2, t_1) & Y_{22}(t_2, t_1) \end{pmatrix} \begin{pmatrix} x_1(t_1) \\ x_2(t_1) \end{pmatrix}. \tag{14.6}
\]

Using (14.3) and (14.4) in (14.6) we obtain \( (B \) is nonsingular)

\[
\begin{pmatrix} c_1 \\ \tilde{c}_1 \end{pmatrix} = CY(t_{12}, t_1)B^{-1} \begin{pmatrix} b_1 \\ \tilde{b}_1 \end{pmatrix}. \tag{14.7}
\]

If we define

\[
D = CY(t_{12}, t_1)B^{-1} \tag{14.8}
\]

and assume that \( D_{12}^{-1} \) exists then we find that

\[
\tilde{b}_1 = D_{12}^{-1}[c_1 - D_{11}b_1],
\]

\[
\tilde{c}_1 = (D_{21} - D_{22}D_{12}^{-1}D_{11})b_1 + D_{22}D_{12}^{-1}c_1. \tag{14.9}
\]

Once \( \tilde{b}_1 \) and \( \tilde{c}_1 \) have been obtained we can compute the solution vector \( x(t) \) at any point \( t_1 < t < t_2 \) using the fundamental matrices \( Y(t, t_1) \) or \( Y(t_2, t) \). Thus, from the point of view of stochastic problems if one has information on the fundamental solution matrix \( Y(t, s) \), then the solution of two-point (or multipoint) boundary value problems follows by algebraic considerations.

15. WAVE PROPAGATION IN A ONE-DIMENSIONAL RANDOM MEDIUM

Consider the two-point boundary value problem

\[
\frac{d^2u(t)}{dt^2} + ku^2(1 + \gamma(t))u(t) = 0, \quad 0 < t < L,
\]

\[
u(t) = e^{\lambda t} + Re^{-\lambda t}, \quad t < 0,
\]

\[
u(t) = Te^{\lambda t}, \quad t > L,
\]

\[
u(t), \quad \frac{du(t)}{dt} \quad \text{continuous.} \tag{15.1}
\]

Here \( \gamma(t) \) is a given real valued stochastic process, \( k > 0 \) is a constant, and the solution \( u(t) \) is a complex valued process. The problem is in the form (15.1) in order to demonstrate its physical meaning. Namely, \( u(t) \) is the wave function at location \( t \) when an incident wave of unit amplitude and with wave number \( k \) impinges on a slab of random medium whose index of refraction is \( (1 + \gamma(t))^{1/2} \).
\[ R = -\frac{b}{\dot{a}} \quad T = \frac{1}{\dot{a}} \quad |T|^2 + |R|^2 = 1. \] (15.8)

To emphasize dependence on parameters we may write
\[ R = R(L, k; y(t)), \quad T = T(L, k; y(t)). \]

We come now to the stochastic problem. From the way the index of refraction \((1 + y(t))^{1/2}\) is written, it is clear that \(y(t)\) is intended to play the role of fluctuations. So we assume it is stationary and that
\[ E[y(t)] = 0, \quad t \geq 0. \] (15.9)

Let \(\rho(t)\) be the covariance function
\[ \rho(t) = E[y(t + s)y(s)]/\alpha^2, \quad \alpha^2 = E[y(t)^2], \] (15.10)

and let
\[ I = \int_0^\infty \rho(t) \, dt, \] (15.11)

assumed finite. Note that \(I\) has the dimensions of a length (\(t\) has dimension of length) and that \(\alpha\), the standard deviation of the fluctuations, is dimensionless. Thus the transmission coefficient, the object of principal interest, is a complex valued random variable with \(|T| \leq 1\) and
\[ T = T(L, k, l, \alpha), \] (15.12)

where \(L, k, l\) and \(\alpha\) are parameters, with \(k, L, l\) dimensionless.

We wish to find the probability distribution of \(T\) as a function of its parameters. This turns out to be an impossible job even when one allows \(y(t)\) to be a most convenient, for calculations, process like white noise. There are at least 4 interesting asymptotic limits that one can study. We shall introduce them next and comment on their physical significance.

The first, and perhaps the simplest, is to let \(l\) be proportional to a small parameter \(e\), say \(l \to le^{\beta}\) and let \(e \to 0\). Clearly
\[ T(L, k, le^{\beta}, \alpha) \to 1 \quad \text{as} \quad e \downarrow 0 \] (15.13)
and this is not hard to show. Next one considers the fluctuation
\[ \frac{1}{\varepsilon^2} (T(L, k, \varepsilon^2 l, \alpha) - 1), \] (15.14)
and one can show [60] that it tends asymptotically as \( \varepsilon \downarrow 0 \) to a Gaussian random variable. This is the Gauss-Markov limit of Section 6. It is particularly useful when the index of refraction is not simply \((1 + \varepsilon y(t))^{1/2}\), but, say, \((\zeta(t) + \varepsilon y(t))^{1/2}\) where \(\zeta(t)\) is deterministic (the mean profile). In this case the right-hand side of (15.13) is not 1 but the deterministic transmission coefficient.

The second case is the white noise limit (cf. (7.1)) where \(\alpha \rightarrow \alpha/\varepsilon\) and \(l \rightarrow \varepsilon l\). Then \(T(L, k, \varepsilon^2 l, \alpha/\varepsilon)\) tends (weakly) as \(\varepsilon \rightarrow 0\) to a random variable that can be determined in principle by solving a complicated diffusion equation on a 3-dimensional manifold. Physically, this case corresponds to large fluctuations and small correlation length relative to the wavelength of the incident wave \((2\pi/k)\) and the width \(L\). There are no explicit results for this case as far as we know.

The third case is to take \(\alpha \rightarrow \alpha/\varepsilon^2, L \rightarrow L/\varepsilon^2\). It is easy to see that
\[ T\left(\frac{L}{\varepsilon^2}, k, l, \alpha\varepsilon\right) = T\left(\frac{L}{\varepsilon}, \frac{k}{\varepsilon}, \varepsilon^2 l, \varepsilon \alpha\right) = T\left(\frac{L}{\varepsilon}, k, l, \varepsilon \alpha\right). \] (15.15)
Here the limit can be computed explicitly [23] as we shall describe below. It corresponds to (7.20) and (7.21). Physically we have (i) small fluctuations and (ii) either large width with \(k\) and \(l\) fixed, or large \(k\) and small \(l\) with \(L\) fixed, or large \(L\), large \(k\) and small \(l\) as shown in (15.15). In particular, the wavelength and the correlation length are comparable while the width is large, which is a physically interesting problem.

The fourth case is \(L \rightarrow L/\varepsilon^2, k \rightarrow \varepsilon k\), which leads to
\[ T\left(\frac{L}{\varepsilon}, \varepsilon k, l, \alpha\right) = T\left(\frac{L}{\varepsilon}, k, \varepsilon^2 l, \alpha\right) = T\left(\frac{L}{\varepsilon}, k, l, \alpha\right). \] (15.16)
Physically, we treat the case of fluctuations that are of order one, correlations that are of order one, low frequency and large width (first scaling in (15.16)). This limit corresponds to (7.24), (7.25) and is important in the heat conduction problem that we discuss in Section 16.

Now we return to the third scaling (15.15). According to the above discussion it is all a question of finding the asymptotic behavior of the fundamental solution process \(Y^e(t, s)\) which satisfies the stochastic equation
\[ \frac{dY^e(t, s)}{dt} = \frac{iky(t/e^2)}{2} \left( \begin{array}{cc} 1 & e^{-2it/k\varepsilon^2} \\ -e^{2it/k\varepsilon^2} & -1 \end{array} \right) Y^e(t, s), \] (15.17)
\[ Y^e(s, s) = I. \]
This is a problem of the form (7.20) with the substitution (7.21), but matrix-valued. The formula (7.23) works just as well here and the problem is to solve the diffusion equation with \(\mathcal{D}\) as the diffusion operator. This is done in [23] and references therein, and the ideas are simple but there are a lot of details. We shall only give the results for the quantity \(T\) here:
\[ \lim_{\varepsilon \downarrow 0} \mathbb{E} \left[ T\left(\frac{L}{\varepsilon^2}, k, l, \alpha\varepsilon\right) \right] = e^{-L^2/k} \int_{-\infty}^{\infty} e^{-\varepsilon^2 \beta t} \pi t \sinh \pi t \cosh^2 \pi t \ dt, \]
\[ \lim_{\varepsilon \downarrow 0} \mathbb{E} \left[ T\left(\frac{L}{\varepsilon^2}, k, l, \alpha\varepsilon\right)^2 \right] = e^{-L^2/k} \int_{-\infty}^{\infty} e^{-\varepsilon^2 \beta t} \left( t^2 + \frac{1}{4} \right) \pi t \sinh \pi t \cosh^2 \pi t \ dt, \] (15.18)
where
\[ \beta = \frac{(k\varepsilon)^2}{2} \int_0^\infty \rho(s) \cos 2ks \ ds. \] (15.19)
Note that the \(l\) of (15.11) does not enter directly but rather through the expression (15.19). Of course, the discussion concerning the scaling did not take into account finer details which necessarily enter (and affect the result).

16. HEAT CONDUCTION IN A ONE-DIMENSIONAL RANDOM MEDIUM

The formulation of this problem is as follows. We begin with (15.1) but now assume that the incident wave is not a time harmonic wave of unit amplitude and with wave number \(k\). We assume that
the incident wave at location \( t = 0 \) is time dependent*; specifically a stationary random function of time with decomposition

\[
v_1(t) = \int_{-\infty}^{\infty} e^{-i \omega t} \tilde{A}(\omega) \, d\omega
\]

\[
\tilde{A}(-\omega) = \tilde{A}(\omega) \quad \text{(for \( v_1 \) to be the real-valued).}
\]

Moreover

\[
\mathbb{E}(v_1(t + \sigma) v_1(\tau)) = \int_{-\infty}^{\infty} e^{-i \omega} \theta(\omega) \, d\omega
\]

where \( \theta(\omega) \geq 0 \) is the spectral density and \( \mathbb{E}(\cdot) \) denotes average and is distinct from \( E(\cdot) \) which involves the coefficient \( y(t) \) (we assume \( y(\cdot) \) and \( v_1(\cdot) \) are independent).

The transmitted pulse at location \( t = L \) is given by

\[
v^t_L = \int_{-\infty}^{\infty} e^{-i \omega t} \frac{1}{c} T \left( \frac{\omega}{c}, t, \alpha \right) \, d\omega
\]

where \( c \) is the speed in vacuum (say \( c = 1 \)). Thus

\[
E\mathbb{E}(v^t_L \cdot v^t_L) = \int_{-\infty}^{\infty} e^{-i \omega} E\left( \left| T \left( \frac{\omega}{c}, t, \alpha \right) \right|^2 \right) \theta(\omega) \, d\omega.
\]

To model a heat bath on the left side we may take \( \theta(\omega) = 1 \), in which case the quantity of primary interest is the total amount of energy transmitted (on the average) by the medium at all frequencies

\[
J(L, l, \alpha) = E(v^t_L v^t_L) = 2c \int_{-\infty}^{\infty} E\left( \left| T \left( L, k, l, \alpha \right) \right|^2 \right) dk.
\]

The problem is to find [62, 63] \( J(L, l, \alpha) \) as \( L \to \infty \). Let \( L = 1/\epsilon^2 \) with \( \epsilon \to 0 \). Then

\[
J(L, l, \alpha) = 2c \epsilon \int_{-\infty}^{\infty} E\left( \left| T \left( \frac{L}{\epsilon^2}, \epsilon \cdot \frac{k}{\epsilon}, l, \alpha \right) \right|^2 \right) \, dk.
\]

Hence

\[
\lim_{L \to \infty} \sqrt{L} J(L, l, \alpha) = \lim_{\epsilon \to 0} 2c \epsilon \int_{0}^{\infty} E\left( \left| T \left( \frac{1}{\epsilon^2}, \epsilon k, l, \alpha \right) \right|^2 \right) \, dk,
\]

provided the limits exist. We note that the integrand on the right is scaled in exactly the form (15.16). The problem is now the passage of the limit \( \epsilon \to 0 \) under the integral sign. For this one needs an estimate so as to use the dominated convergence theorem. This estimate follows, essentially, from the work of Pastur and Fed'man [61] but the details require further attention and will be given elsewhere (since they are also of independent interest). Assuming this interchange we have now

\[
\lim_{L \to \infty} \sqrt{L} J(L, l, \alpha) = 2c \int_{0}^{\infty} \lim_{\epsilon \to 0} E\left( \left| T \left( \frac{1}{\epsilon^2}, \epsilon k, l, \alpha \right) \right|^2 \right) \, dk.
\]

We compute the \( \epsilon \) limit in (16.4) exactly as in [23] but with the formula (7.26) instead of (7.23) which leads to essentially the same result up to a redefinition of \( \beta \) in (15.19). We find, in fact, that

\[
\lim_{L \to \infty} \sqrt{L} J(L, l, \alpha) = \frac{2c \pi^{3/2}}{(\alpha^2/2)^{1/2}} \int_{0}^{\infty} \left( t^2 + 1 \right)^{-1/2} \frac{t \sinh \pi t}{\cosh^2 \pi t} \, dt.
\]

The physical significance of this result is the following. The integrated (over frequencies) or total transmittance of the medium decreases as \( L \to \infty \) (all other parameters being fixed) like \( 1/\sqrt{L} \). Moreover, most of the transmission occurs in the neighborhood of frequencies proportional to \( 1/\sqrt{L} \), i.e., low frequencies. The medium acts asymptotically like a low pass filter. For further information on the physics of the problem we refer to [62, 63].

REFERENCES

16. ———, Dependence and asymptotic independence for random processes, this study.
29. I. Babuška, Several reports from the Department of Mathematics, University of Maryland.


