Diffusion Approximation for Transport Processes with Boundary Conditions

by

Richard S. Ellis*

and

Walter A. Rosenkrantz**

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #478

November 1976

* On leave from Northwestern University, Evanston, Illinois. Partially supported by NSF Grant MPS 71-02838 A04.

**On leave from University of Massachusetts, Amherst, Massachusetts 01002.
I. Introduction.

An interesting open problem in the theory of transport processes is to prove rigorously the diffusion approximation in the presence of general homogeneous boundary conditions. In this paper it will be shown that for the class of processes considered by Pinsky [Pi], Kurtz [Ku], Watanabe [W1,W2], and others, the diffusion approximation is valid on a compact interval with a general class of local homogeneous boundary conditions provided one can establish the a priori bounds (2.23). The bounds refer to the initial-boundary value problem (2.1)-(2.3) and (2.16)-(2.18) for a first order hyperbolic system of partial differential equations, a problem of independent mathematical interest [Kr]. On the whole real line, where no boundary conditions are present, such estimates are easily obtained (see Section V). We are thereby led to a very easy proof of Pinsky's theorem [Pi] and its generalizations to certain position-dependent velocities [Ku] and higher dimensions [WW]. However, on a finite interval, where boundary conditions are imposed, we are able to prove these estimates only for certain special cases (see Section V). We consider as one of our contributions the derivation of direct proofs for the diffusion approximation given that the a priori bounds are known. In general, the a priori estimates hold only if the initial data satisfy certain "compatibility conditions" associated in a natural way with the boundary conditions. See [RM], where this has been noted in a related context.

The transport processes considered in this paper differ in several ways from the more general class of transport processes studied by such authors as Habetler-Matkowsky [HM], Larsen-Keller[LK], Bensoussan-Lions-Papanicolaou [BLP], Williams [Wi], and Wing [Win]. In the works of these authors it is assumed that the possible velocities of the transport particle form a continuum (here the possible velocities are discrete). This leads to an integro-differential equation for which explicit
solutions are difficult to obtain. One can then resort to numerical methods; e.g., one can discretize the velocity space (discrete ordinate method in neutron transport theory [RiMo]). The result of this discretization is the first order hyperbolic system (2.1)-(2.3) with this important difference: the transport equation in [LK], [HM], and [Win] is the adjoint of our system (2.1)-(2.3). The latter is the Kolmogorov backward differential equation of the transport process, for which the initial data are functions; the authors referred to above consider the Kolmogorov forward (or Fokker-Planck) equation, for which the initial data are measures. (This observation may help to explain why the traditional approaches to the neutron transport equation via $L^p$ spaces for $p > 1$ so often lead to technical difficulties [L; Introduction].) One of the advantages in studying the system (2.1)-(2.3) instead of its adjoint is the presence of a maximum principle (Lemma 3.1), which yields a priori bounds on the solution and its derivatives. Such estimates are also of great importance in establishing the stability and consistency of various finite difference schemes for the numerical solution of the neutron transport equation [D].

Our derivation of the diffusion approximation differs markedly from that in [HM], [LK], [Pa], and [Wi]. All of the latter authors use the method of matched asymptotic expansions, the rigorous justification of which entails non-trivial, ergodic-type questions.

Our main reason for treating only the discrete, one-dimensional case is to keep hypotheses to a minimum. Indeed, our methods carry over to a wide class of continuous transport processes, to homogeneous boundary value problems [W2; p. 220] more complicated than (2.3), and to certain higher dimensional problems.

The usual tool for studying the initial-boundary value problems we consider has been the $L^2$-theory [LP,R]. Nevertheless, all the analysis in this paper is done in a setting more familiar to probabilists; namely, in a space of continuous functions
endowed with the supremum norm. This has its advantages. For example, we show the existence of a nice contraction semigroup on this space solving system (2.1)-(2.3) given any choice of boundary parameters satisfying the natural conditions (2.5). This is not true in the $L^2$-case; extra conditions must be met in order to obtain an $L^2$-contraction. On the other hand, it is conceivable that one might be able to prove $L^2$ a priori bounds corresponding to (2.23) in a wider class of cases than we have been able to do. This is of some interest because provided one can obtain such $L^2$ bounds, then all of our work goes over to the $L^2$ setting without change (see Remark at end of part II).

In Section II, we state our main results: Theorem 2.1, on the existence of a Markovian semigroup (i.e., a strongly continuous, positivity-preserving contraction semigroup) corresponding to the discrete, one-dimensional transport process in a compact interval or on the whole real line; Theorem 2.2, on the diffusion approximation for this process. Sections III and IV prove Theorems 2.1 and 2.2, respectively. Section V proves the a priori bounds in special cases.

Notation. All terms $O(1)$, $O(\epsilon)$, $o(1)$, $o(\epsilon)$, etc. are meant as $\epsilon \to 0$ and, if no norm is present, are uniform over $x$ in the interval. The end of a proof is signalled by a $\blacksquare$. 
II. Main results

Let $Q = (q_{ij})$ denote the infinitesimal generator of an $N$ state Markov chain $V(t)$ with state space $\mathcal{V} = \{v_1, \ldots, v_N\}$ and invariant probability measure $\pi = (\pi_1, \ldots, \pi_N)$. It is assumed that $v_1, \ldots, v_k < 0 < v_{k+1}, \ldots, v_N$, $k \geq 1$ and in addition the $v_i$'s are all distinct. $X(t)$ denotes the position of a particle moving in the interval $J = [r_0, r_1]$ with random velocity $V(t)$. If $r_0 = -\infty$ and $r_1 = \infty$ then it is easily shown that $(X(t), V(t))$ is a Markov process whose Kolmogorov backward differential equation is given by the system (2.1), (2.2) (cf. [Pi]).

If, however, one or both of the boundary points $r_0$, $r_1$ are finite then boundary conditions must be imposed which we do in the following way. If the particle hits, say, the left hand boundary $r_0$ with velocity $v_i < 0$ ($i=1,2,\ldots,k$) then with probability $b_{ij}$ it is reflected to the right with velocity $v_j > 0$ ($j=k+1,\ldots,N$) and with probability $1 - \sum_{j=k+1}^{N} b_{ij}$ it disappears. A similar boundary condition is imposed at $r_1$. In this case the corresponding Kolmogorov backward differential equation is an initial-boundary value problem for the vector function $F(t,x) = (F_1(t,x), \ldots, F_N(t,x))$:

\begin{equation}
\frac{\partial F}{\partial t} = v_i \frac{\partial F}{\partial x} + \sum_{i=1}^{N} q_{ij} F_j, \quad i=1,\ldots,N
\end{equation}

(2.1)

\begin{equation}
\lim_{t \to 0} F_i(t,x) = f_i(x), \quad x \in J, \quad i=1,\ldots,N
\end{equation}

(2.2)

(a) $F_i(t,r_0) = \sum_{j=k+1}^{N} b_{ij} F_j(t,r_0), \quad i=1,\ldots,k$

(2.3)

(b) $F_i(t,r_1) = \sum_{j=1}^{k} b_{ij} F_j(t,r_1), \quad i=k+1,\ldots,N.$

Precise conditions under which solutions to the first order hyperbolic system (2.1)-(2.3) exist will be given in Theorem 2.1. In the meantime for future
reference we mention various properties of \( v, Q, \pi \) and \( b_{ij} \) that we shall need.

\[
(2.4) \quad v_1, \ldots, v_k < 0 < v_{k+1}, \ldots, v_N, \quad v_i \text{ all distinct, } k \geq 1.
\]

\[
\begin{align*}
\sum_{j=k+1}^{N} b_{ij} & \leq 1, \text{ for } i=1, \ldots, k, \\
\sum_{j=1}^{k} b_{ij} & \leq 1 \text{ for } i=k+1, \ldots, N
\end{align*}
\]

\[
(2.5) \quad b_{ij} \geq 0 \text{ for } i=1, \ldots, k, j=k+1, \ldots, N \text{ and } i=k+1, \ldots, N, j=1, \ldots, k
\]

\[
\quad b_{ij} = 0 \text{ for all other } i, j.
\]

\[
(2.6) \quad q_{ij} \geq 0, i \neq j; \quad \sum_{j=1}^{N} q_{ij} = 0, i=1, \ldots, N; \text{ dimension of nullspace of } Q \text{ is } 1,
\]

\[
(2.7) \quad Q^*\pi = 0, \text{ each } \pi_i > 0, \quad \sum_{i=1}^{N} \pi_i = 1. \text{ Since } Q \text{ is a generator of a Markov chain we also have } \|\exp(x, Q)\| \leq 1 \text{ all } x \geq 0. \text{ We formulate each of our results first for motion in a compact interval } J, \text{ then note modifications necessary to treat motion on the whole real line. We omit the case of motion in semi-bounded intervals.}
\]

Our first result concerns the existence of a Markovian semi-group solving (2.1)-(2.3) for compact \( J \), (2.1)-(2.2) for \( J = (-\infty, \infty) \). Given \( J \) compact let \( \mathcal{F} \) denote the subset of functions in \( C(J\times \mathcal{Y}) \) (all \( N \)-tuples \( f = (f_1, \ldots, f_N) \) of bounded continuous functions \( f_i(x), x \in J \)) which satisfy the boundary condition (2.3). Thus we define

\[
(2.8) \quad \mathcal{Y} = C(J\times \mathcal{Y}) \cap \mathcal{B}_0 \cap \mathcal{B}_1 \quad \text{where}
\]
\[
\begin{align*}
\beta_0 &= \{ f : \sum_{k=1}^{N} f_k(r_0)n_{1k} = \langle f, n_i \rangle (r_0) = 0, \ i=1,...,k \} \\
\beta_1 &= \{ f : \sum_{k=1}^{N} f_k(r_1)n_{1k} = \langle f, n_i \rangle (r_1) = 0, \ i=k+1,...,N \} \\
\end{align*}
\]

\[
(2.9)
\]

\[
\begin{align*}
n_i &= \left\{ \begin{array}{l}
(0,0,...0,1,0...,0,-b_{i,k+1}, -b_{i,k+2} \ldots -b_{i,N}), \ i=1,...,k \\
(-b_{i,1} - b_{i,2} \ldots - b_{i,k}, 0,0,...,1,0...0), \ i=k+1,...,N,
\end{array} \right.
\end{align*}
\]

with 1 in the ith slot. $\mathcal{J}$ is a Banach space with the norm

\[
(2.11) \quad ||f|| = \sup_{1 \leq i \leq n, \ x \in J} |f_i(x)|.
\]

We define

\[
(2.12) \quad \mathcal{J}^n = \mathcal{J} \cap C^n(J \times \mathbb{R}), \ n=0,1,... \quad \text{where}
\]

\[
C^n(J \times \mathbb{R}) = \{ f \in C(J \times \mathbb{R}) : f_i \in C^n(J) \}. \quad \text{Of course } \mathcal{J}^0 = \mathcal{J}.
\]

We say $f \geq 0$ if each $f_i \geq 0$ on $J$.

Modifications necessary for $J = (-\infty, \infty)$: In this case

\[
\mathcal{J} = \{ f(x) : f_i(x) \in C_0(-\infty, \infty) \} \text{and}
\]

\[
\mathcal{J}^n = \{ f(x) : f_i(x) \in C_0^n(-\infty, \infty), \ i=1,2,...,n \}. \quad \mathcal{J} \text{ is a Banach space}
\]

with the norm 2.11. Let $A$ be defined via the formula

\[
(2.13) \quad Af_i(x) = v_{i1} f_i'(x) + \sum_{j=1}^{N} q_{ij} f_j(x), \ i=1,...,N \text{ or in vector form}
\]

\[
Af(x) = VDf(x) + Qf(x) \text{ where } Df(x) = (f_1'(x),...,f_N'(x))
\]

(in general $D = \partial / \partial x$ and $D^\ell = \partial^\ell / \partial x^\ell$, $\ell = 1,2,...$). $V$ denotes

the $N \times N$ matrix $(v_{i1}\delta_{ij})$. In order to solve (2.1)-(2.2) via semi-
group methods we must define the domain $\mathcal{D}(A)$ of $A$. We set

\begin{equation}
\mathcal{D}(A) = \{ f : f \in \mathcal{S}, Af \in \mathcal{S} \}. \tag{2.14}
\end{equation}

In part III we show that $\mathcal{D}(A)$ is dense in $\mathcal{S}$.

Theorem 2.1. The linear operator $A$ with domain $\mathcal{D}(A)$ as in (2.14) is the infinitesimal generator of a Markovian semi group $T(t) = \exp(tA) : \mathcal{S} \rightarrow \mathcal{S}$. The function $F(t, x) = T(t)f(x), f \in \mathcal{D}(A)$ is the unique bounded solution to the first order hyperbolic system (2.1)-(2.3) for compact $J$, (2.1)-(2.2) for $J = (-\infty, \infty)$.

We now turn our attention to the diffusion approximation or equivalently the central limit theorem for $X + \varepsilon X(t/\varepsilon^2)$ as $\varepsilon \rightarrow 0$. Physically this corresponds to shrinking the order of magnitude of the time between jumps of $V(t)$ to $\varepsilon^2$ and speeding up the velocities by a factor of $\varepsilon^{-1}$. If the associated position process $X^\varepsilon(t)$ is to have a nice limit as $\varepsilon \rightarrow 0$, then the velocities must be appropriately centered. This is expressed by the condition

\begin{equation}
\sum_{i=1}^{N} \pi_i v_i = \pi , v > = 0 \tag{2.15}
\end{equation}

which we shall always assume throughout the remainder of this paper.

The Kolmogorov backward differential equation for the process $(V(t/\varepsilon^2), X^\varepsilon(t))$ is now given by the system

\begin{equation}
\frac{\partial F_i(\varepsilon)}{\partial t} = \varepsilon^{-1} \pi_i DF_i(\varepsilon) + \varepsilon^{-2} \sum_{j=1}^{N} q_{ij} F_j(\varepsilon), i = 1, \ldots, N \tag{2.16}
\end{equation}

\begin{equation}
\lim_{t \rightarrow 0} F_i(\varepsilon)(t, x) = f(x) \tag{2.17}
\end{equation}

\begin{equation}
F(\varepsilon)(t, x) \in \mathcal{S}, t > 0 \tag{2.18}
\end{equation}

Define the operator $A(\varepsilon)$ via the formula
\begin{equation}
A(\varepsilon)f = \varepsilon^{-1}VDF + \varepsilon^{-2}Qf \text{ with the understanding that } A(1) = A \\
\text{and set } \mathcal{B}(A(\varepsilon)) = \mathcal{B}(A). \text{ Then } (2.16)-(2.18) \text{ can be rewritten as an evolution equation in the Banach space } \mathcal{F}:
\end{equation}

\begin{equation}
\begin{cases}
\frac{dF^\varepsilon(t)}{dt} = A(\varepsilon)F^\varepsilon(t) \\
F^\varepsilon(0) = f \\
F^\varepsilon(t) \in \mathcal{F} \text{ all } t > 0
\end{cases}
\end{equation}

Theorem 2.1 can now be applied to (2.16)-(2.18) since all we've done is replace \( v_i \) by \( \varepsilon^{-1}v_i \) and \( q_{ij} \) by \( \varepsilon^{-2}q_{ij} \). In probabilistic language the diffusion approximation corresponds to the weak convergence of the stochastic process \( X^\varepsilon(t) \) to \( b(t) \) where \( b(t) \) is Brownian motion on \( J \) subject to certain boundary conditions if \( J \) is compact. Analytically this is expressed by showing

\begin{equation}
\lim_{\varepsilon \to 0} F^\varepsilon_i(t,x) = G(t,x) \text{ where the limit } G(t,x) \text{ is independent of } i, F(0,x) \text{ is such that } f_1(x) = f_2(x) = \ldots f_N(x) \text{ and } G(t,x) \text{ satisfies the heat equation}
\end{equation}

\begin{equation}
\frac{\partial G}{\partial t} = a^2 \frac{\partial^2 G}{\partial x^2} , \ x \in J \text{ subject to appropriate boundary conditions}
\end{equation}
given in Theorem 2.2 below. For the definition of a see (2.31).

Our main result is that the diffusion approximation is valid provided the vector function \( U^\varepsilon(\lambda, x) = (U_1^\varepsilon(\lambda, x), \ldots, U_N^\varepsilon(\lambda, x)) \) satisfies the a priori estimate

\begin{equation}
\varepsilon \| D^\varepsilon U^\varepsilon(\lambda, x) \| = 0(1), \varepsilon = 1, 2, 3, \text{ all } f \in \mathcal{F}^2, \lambda > 0
\end{equation}

where

\[ U_i^\varepsilon(\lambda, x) = \int_0^\infty \exp(-\lambda t) F_i^\varepsilon(t,x) dt, i = 1, \ldots, N. \]

Remark: In the course of proving theorem 2.1 we shall show that \( f \in \mathcal{F}^n \) implies
\( \tilde{u}(\varepsilon)(\lambda, x) = (\lambda - A(\varepsilon))^{-1} f \in \mathcal{J}^{n+1} \). If instead of (2.23) one could prove the stronger estimate \( \|D^\ell \tilde{u}(\varepsilon)(\lambda, x)\| = O(1) \), \( \ell = 1, 2, 3 \), some \( \lambda > 0 \) then all the \( O(1) \) error terms in section IV would be replaced by \( O(\varepsilon) \). For the special cases in section V we actually verify this stronger estimate.

Theorem 2.2 yields the diffusion approximation under a wide variety of boundary conditions, the specifications of which depend on certain constants \( \alpha_i (i = 1, \ldots, N) \), \( \eta(r_\lambda) \) defined as follows:

\[
\alpha_i = \begin{cases} 
1 - \sum_{j=k+1}^{N} b_{ij}, & i = 1, \ldots, k, \\
1 - \sum_{j=1}^{k} b_{ij}, & i = k+1, \ldots, N;
\end{cases}
\]  

(2.24)

\[
\eta(r_\lambda) = \langle (Q^*)^{-1} \gamma(r_\lambda), v \rangle,
\]  

(2.25)

\[
\theta(r_\lambda) = \langle (Q^*)^{-1} (V(Q^*)^{-1} (\gamma(r_\lambda)), v \rangle \quad (\text{if } \eta(r_\lambda) = 0),
\]  

(2.26)

where \( \gamma(r_\lambda) = (\gamma_m(r_\lambda); m = 1, \ldots, N) \) are given by

\[
\gamma_m(r_0) = \begin{cases} 
1, \text{ for } m = 1, \ldots, k \\
- \sum_{j=1}^{k} b_{jm}, \text{ for } m = k+1, \ldots, N,
\end{cases}
\]

(2.27)

\[
\gamma_m(r_1) = \begin{cases} 
- \sum_{j=k+1}^{N} b_{jm}, \text{ for } m = 1, \ldots, k,
1, \text{ for } r = k+1, \ldots, N.
\end{cases}
\]

(2.28)

For the definition of \( (Q^*)^{-1} \) in (2.25), (2.26), see (4.10). In terms of these numbers, we define the following three cases at \( r_0 \) (with a similar definition at \( r_1 \)):
Case 1: at least one \( \alpha_i > 0, \ i=1, \ldots, k; \)

Case 2: \( \alpha_i = 0, \ i=1, \ldots, k, \) and \( \eta(r_0) \neq 0; \)

Case 3: \( \alpha_i = 0, \ i=1, \ldots, k, \) \( \eta(r_0) = 0, \theta(r_0) \neq 0. \)

We show that in Case 1 the limiting diffusion \( b(t) \) is absorbed at \( r_0, \) in Case 2 \( b(t) \) is reflected at \( r_0 \) back into \( J, \) and in Case 3 \( b(t) \) adheres at \( r_0 \) (similarly at \( r_1 \)). To see, for example, the absorption note that if Case 1 holds at \( r_0, \) then the transport particle has a positive probability of being absorbed at \( r_0 \) each time it hits \( r_0. \) Since in any fixed time interval the number of such hits should tend to infinity as \( \varepsilon \to 0, \) it is plausible that \( b(t) \) will be absorbed at \( r_0 \) with probability one.

Given \( g \in \mathcal{L}^2(J), \) we define \((k=0,1)\)

\[
H_k g = \begin{cases} 
  g(r_k) & \text{if Case 1 holds at } r_k, \\
  Dg(r_k) & \text{if Case 2 holds at } r_k, \\
  D^2g(r_k) & \text{if Case 3 holds at } r_k.
\end{cases}
\]

We set

\[
\mathcal{A} = \{ g \in \mathcal{L}^2(J) \mid H_k g = 0, \ k = 0,1 \};
\]

\( \mathcal{G} = \mathcal{A} \) (which is \( \mathcal{L}(J) \) if either Case 2 or Case 3 holds at both \( r_0 \) and \( r_1 \)); \( \mathcal{G}^n = \mathcal{L}^n(J) \cap \mathcal{A}, \ n = 1,2,\ldots \).

**Modifications necessary for \( J = (-\infty, \infty): \)** We define \( \mathcal{G} = \mathcal{L}_0(-\infty, \infty), \) \( \mathcal{G}^n = \mathcal{L}^n_0(-\infty, \infty), \ n = 1,2,\ldots \)

For \( J \) compact or \( J = (-\infty, \infty), \) we assume that the velocities are centered as in (2.15) and define the operator

\[
(2.30) \quad \Omega = aD^2,
\]

\[
(2.31) \quad a = -\langle (Q^*)^{-1} \pi v, v \rangle.
\]
where \(<-,-\rangle\) denote the \(\ell_2\) inner product and \((Q^*)^{-1}\) is defined in (4.10). Provided
(2.15) holds, the positivity of \(a\) is known [Pi]. \(\hat{\Omega}\) is the infinitesimal generator
of a Markovian semi group \(\exp(t\Omega)\) on \(\mathcal{G}\) corresponding to a Brownian motion with
variance \(2a\) (which satisfies the appropriate boundary conditions at \(r_0\) and at \(r_1\)
for \(J\) compact), see [Ma].

**Theorem 2.2.** Assume the a priori bounds (2.23) and the centering (2.15).
Define the map \(P: \mathcal{G} \to \mathcal{G}\) by \(Pg = (g, \ldots, g)\). Then for each \(g \in \mathcal{G}\), \(0 < T < \infty\),

\[
(2.32) \sup_{0 \leq t \leq T} \left\| \exp(tA_e)Pg - P\exp(t\Omega)g \right\| = o(1).
\]

The a priori bounds (2.23) are proved in the following cases:

(i) (2.1)-(2.2), \(J = (-\infty, \infty)\) (see [Pi]);

(ii) (2.1)-(2.3), \(N = 2, k = 1, b_{ij}\) satisfy (2.5),

\[
(2.33) \text{the centering (2.15) holds;}
\]

(iii) (2.1)-(2.3), arbitrary \(N\) even, \(k = N/2\),

(a) \(v_i = -v_{N+1-i}, i = N/2 + 1, \ldots, N\),

(b) \(b_{i,N+1-i} = 1, i=1, \ldots, N, \) all other \(b_{ij} = 0\),

(c) \(QR = RQ\),

where \(R\) is the \(N\times N\) matrix \((\delta_{i,N+1-i}) =

\[
\begin{bmatrix}
0 & 0 & \ldots & 0 & 1 \\
0 & 0 & \ldots & 1 & 0 \\
\vdots & \vdots & & \vdots & \vdots \\
\vdots & \vdots & & \vdots & \vdots \\
1 & 0 & \ldots & 0 & 0
\end{bmatrix}
\]

Conditions (2.33) (iii) (a)-(b) imply that the scattering rule at \(r_0\) and \(r_1\) is
\(v_i \to -v_i\). If \(Q = Q^*\) in addition to (2.33) (iii) (c), then \(Q\) is symmetric about
both its diagonals (alone) (c) says that \(Q\) is symmetric about its midpoint); thus
for any \(1 \leq i, j \leq N\) the transitions \(v_i \to v_j, v_j \to v_i, -v_i \to -v_j, -v_j \to -v_i\) in
the interior of \(J\) would all have equal probability. (See [E1], [E2; Note] for
such matrices \( Q \). In the course of treating (iii), we reduce the problem of obtaining the \textit{a priori} estimates in the general case of (2.1)-(2.3) to a matrix problem.

We end this section with several remarks about Theorems 2.1-2.2.

\textbf{Remarks 2.3.} (i) In Section IV we show that \( \eta(r_{i}) \neq 0, \ i = 0, 1 \), in the case of isotropic scattering; i.e., for

\begin{equation}
Q = \Gamma - I, \quad \Gamma = \frac{1}{N} \begin{bmatrix}
1 & \ldots & 1 \\
\vdots & & \vdots \\
1 & \ldots & 1
\end{bmatrix}.
\end{equation}

Hence only Cases 1 or 2 in (2.27) may arise at \( r_{0} \) and at \( r_{1} \). We also give a numerical example to show that Case 3 in (2.27) can arise (although we have not been able to verify (2.32) for such a case). We do not know if for all choices of \( b_{ij}, v_{i}, \) and \( Q \) one of the three cases in (2.27) must hold. However, we have not found a counterexample.

(ii) Theorems 2.1-2.2 can be generalized to the case where the \( v_{i} \) and the \( q_{ij} \) depend on \( x \) and the \( b_{ij} \) on \( \varepsilon \). When \( v_{i} = v_{i}(x) \) and \( q_{ij} = q_{ij}(x) \), we obtain limiting (elliptic) operators more complicated than \( \Omega \). When \( b_{ij} = b_{ij}(\varepsilon) \) we obtain limiting boundary conditions which are linear combinations of those in (2.29) (even for constant \( v_{i} \) and \( q_{ij} \)). For example, let \( N = 2, k = 1, b_{ij} = \beta_{ij}(1+\varepsilon z_{ij})^{-1}, \) where

\[ \beta_{ij} = \begin{cases}
1, & \text{for } i \neq j, \\
0, & \text{otherwise}.
\end{cases} \]

and \( z_{ij} > 0 \); take \( v_{i} \) and \( q_{ij} \) as at the beginning of this section and satisfying (2.15). As discovered by Watanabe [W1], this leads to elastic Brownian motion. The \textit{a priori} bounds in Section V cover this choice of \( b_{ij} \), and so we obtain a simple proof of the result in [W1].
(iii) Concerning the $L^2$ theory, one can prove the existence of a smooth $L^2$-contraction semigroup solving (2.1)-(2.3) for a subset of the $b_{ij}$ satisfying (2.5). We claim that if in these cases $L^2$ a priori bounds corresponding to (2.23) can be obtained, then all of our results can be proved in an $L^2$ setting. Indeed, the key interior estimate (4.8) would follow since this depends only on the a priori bounds while the (pointwise) boundary estimates (4.9) would be a consequence of Sobolev inequalities. The final step would be the $L^2$ theory of the limiting diffusion semigroup, which has been studied in [Mc].

Acknowledgements. We thank Alan Glei it for the numerical example at the end of Section IV.
Part III. Proof of Theorem 2.1

We verify the hypotheses of the Hille-Yosida theorem [Ma; p. 2]:

\[
\begin{align*}
(i) & \quad \mathcal{D}(A(\varepsilon)) \text{ is dense in } \mathcal{D} \\
(ii) & \quad \text{For every } f \in \mathcal{D}, \lambda > 0, \text{ the equation} \\
& \quad \lambda F(\varepsilon) - A(\varepsilon)F(\varepsilon) = f \text{ has a unique solution} \\
& \quad F(\varepsilon) \in \mathcal{D}(A(\varepsilon)) \text{ such that} \\
& \quad \lambda \|F(\varepsilon)\| \leq \|f\| \\
(iii) & \quad \text{if } f \geq 0 \text{ then } F(\varepsilon) \geq 0. \\
(iv) & \quad \text{If } f \in \mathcal{D}^n \text{ then } F(\varepsilon) \in \mathcal{D}^{n+1}, n=0,1,\ldots
\end{align*}
\]

(3.1)

An examination of the proof of (ii)-(v) yields the fact (needed later) that

(3.2) equation (3.1(ii)) has a unique solution \( F(\varepsilon) \in \mathcal{D}^1 \) such that (iii), (iv), (v) hold provided only \( f \in C(J,J') \). In proving (i)-(v) we may take \( \varepsilon = 1 \), writing \( F \) for \( F^{(1)} \) (not to be confused with \( DF \)), \( A \) for \( A(1) \). We first take \( J \) compact then \( J = (-\infty, \infty) \).

Proof of (i)-(v) for \( J \) compact.

(i) It suffices to prove \( \mathcal{D}(A) \) dense in \( \mathcal{D}^1 \) since the latter is dense in \( \mathcal{D} \). This will follow from the observation, proved below, that given any \( \varepsilon > 0 \) and any vectors \( \beta = (\beta_1, \ldots, \beta_N), \delta = (\delta_1, \ldots, \delta_N) \) we can construct a function \( g \in \mathcal{D}^1 \) with the properties (i) \( g(r_i) = f(r_i) \), \( i=0,1 \) (ii) \( ||f-g|| = O(\varepsilon) \) and (iii) \( Dg(r_0) = \beta, Dg(r_1) = \delta \). Assuming this to be the case let us apply it to the problem at hand. In particular we choose \( \beta \) and \( \delta \) so that the boundary conditions

\[
\begin{align*}
(i) & \quad \langle A \beta, n_i \rangle (r_0) = 0 \quad i=1,\ldots,k \\
(ii) & \quad \langle A \beta, n_i \rangle (r_1) = 0 \quad i=k+1,\ldots,N
\end{align*}
\]

(3.3)
This will be so if

\begin{align}
\langle \beta, V^{-1}n_i \rangle &= -\langle f, Q^*n_i \rangle \quad (r_0), \ i=1, \ldots, k \\
\langle \delta, V^{-1}n_i \rangle &= -\langle f, Q^*n_i \rangle \quad (r_1), \ l=k+1, \ldots, N.
\end{align}

We are using the hypothesis that \( f(r_0) = g(r_1) \ i=0,1 \). In the first instance we have \( k \) equations in \( N \) unknowns and in the second we have \( N-k \) equations in \( N \) unknowns. These equations are easily solved by choosing \( \beta_{k+1}, \ldots, \beta_N \) arbitrarily and then solving explicitly for \( \beta_1, \ldots, \beta_k \) and a similar statement is valid for the vector \( \delta \), mutatis mutandis. Since \( g \in \mathcal{F} \), \( Ag \in \mathcal{F} \) (by construction) and

\[ ||g-f|| = O(\varepsilon) \]

it follows that \( g \in \mathcal{B}(A) \) and hence \( \mathcal{B}(A) \) is dense in \( \mathcal{F} \). We now sketch the construction of the function \( g \). Given \( \varepsilon > 0 \) set \( g_i(x) = f_i(x) \), for \( r_0 + \varepsilon \leq x \leq r_1 - \varepsilon \), \( g_i(r_\chi) = f_i(r_\chi) \), \( \chi = 0,1 \). Then interpolate on the intervals \([r_0, r_0 + \varepsilon]\), \([r_1 - \varepsilon, r_1]\) by a continuously differen- tieable function \( g \) such that

\[ Dg(r_0) = \beta, \quad Dg(r_0 + \varepsilon) = Df(r_0 + \varepsilon), \quad Dg(r_1) = \delta, \quad Dg(r_1 - \varepsilon) = Df(r_1 - \varepsilon). \]

A straight forward but tedious computation shows that one can choose \( g \) so that

\[ ||f-g|| = o(\varepsilon) \]

on the subintervals \([r_0, r_0 + \varepsilon]\), \([r_1 - \varepsilon, r_1]\). It's perhaps easiest to just draw a picture. This completes the proof.

The proofs of (ii)-(v) all depend on the next three lemmas.

We say that \( f \in C(J \times \mathcal{U}) \) has a local maximum at \((i,x)\), \( 1 \leq i \leq N \), \( x \in J \) if for some neighborhood \( \mathcal{U} \) of \( x \), \( \mathcal{U} \subseteq J \), \( f_i(x) > f_j(y) \), \( j=1, \ldots, N \), all \( y \in \mathcal{U} \) (similarly for a local minimum).

Lemma 3.1. Assume \( f \in C^1(J \times \mathcal{U}) \) has an interior local maximum (resp., minimum) at \((i,x)\) \( (r_0 < x < r_1) \). Then

\begin{equation}
A f_i(x) \leq 0 \quad \text{(resp.,} > 0) \quad . \text{The same conclusion holds even at a boundary point} \ r_0 \text{or} \ r_1 \text{provided} \ f \in \mathcal{F} \ i.e., \text{provided} \ f \text{satisfies the boundary conditions} (2.9).
\end{equation}

Remark: Lemma (3.1) is also valid for the operator \( A(\varepsilon) \).
Proof: At an interior maximum \( f_i'(x) = 0 \), so that
\[
A f_i(x) = \sum_{j=1}^{N} q_{ij} f_j(x) \leq f_i(x) \sum_{j=1}^{N} q_{ij} = 0,
\]
where we've used the property (2.6) of the \( q_{ij} \)'s. Now assume a maximum occurs at
\( x = r_0 \) (a similar argument works at \( r_1 \)). Since
\[
f_m(r_0) = \sum_{j=k+1}^{N} b_{mj} f_j(r_0) \leq \max_{k+1 \leq j \leq N} f_j(r_0), \text{ for } m=1,\ldots, k
\]
we may assume that the maximum occurs at \( f_i(r_0) \) where \( k+1 \leq i \leq N \), hence \( v_i > 0 \).
Thus
\[
A f_i(r_0) = v_i f_i'(r_0) + \sum_{j=1}^{N} q_{ij} f_j(r_0) \leq 0,
\]
because \( f_i'(r_0) \leq 0 \). The case of a minimum is handled similarly.

Lemma 3.2 (Uniqueness). \( F(x) = (F_1(x), \ldots, F_N(x)) \) be a solution to (3.1) such that \( F \in \mathcal{D} \). Then \( \lambda ||F|| \leq ||f|| \).

Proof. There must exist \((i, x) \) \( 1 \leq i \leq N, x \in J \) such that either \( F_i(x) = ||F|| \) or \( -F_i(x) = ||F|| \). In the first case a maximum occurs at \((i, x) \) so by (3.5)
\[
\lambda ||F|| = \lambda F_i(x) \leq f_i(x) - A F_i(x) = f_i(x) \leq ||f||.
\]

In the second case we consider \(-F \) and \(-f \) to get the same conclusion.

Lemma 3.3 (existence): To every \( f \in C(J \times \mathcal{Y}), \lambda > 0 \), there exists a unique \( F \in \mathcal{D}(A) \) satisfying the equation \((\lambda - A)F = f \) and the boundary conditions
\[
<F, n_i> (r_0) = 0, i=1,\ldots,k; \quad <F, n_i> (r_1) = 0, i=k+1,\ldots,N.
\]

Proof: Uniqueness is a consequence of lemma 3.2 since the a priori estimate implies \( F \equiv 0 \) if \( f \equiv 0 \). We rewrite the boundary conditions as

\[
\Lambda_0 F(r_0) - \Lambda_1 F(r_1) = 0, \quad \text{where the rows}
\]
\[
r_j(\Lambda_0), r_j(\Lambda_1), j=1,\ldots,N \text{ of the matrices } \Lambda_0 \text{ and } \Lambda_1
\]
are given by

\[\text{(3.6)}\]
\[ r_j(A_0) = \begin{cases} n_j & j = 1, \ldots, k \\ (0, \ldots, 0) & j = k+1, \ldots, N \end{cases} \]

\[ r_j(A_1) = \begin{cases} n_j & j = 1, \ldots, k \\ (0, \ldots, 0) & j = k+1, \ldots, N \end{cases} \text{ see [H; pp407-408]} \]

Let \( Y(x) = \exp\{xV^{-1}(\lambda I - Q)\} \) denote the fundamental matrix for the first order system \((\lambda - A)F = f\). Then

\[ F(x) = Y(x)c - \int_{r_0}^{x} Y^{-1}(y)V^{-1}f(y)dy \]

is a solution. We now choose the vector \( c \) so as to satisfy the boundary condition (3.6). This leads to the condition

\[ \{A_0Y(r_0) - A_1Y(r_1)\} c = A_1Y(r_1) \int_{r_0}^{r_1} Y^{-1}(y)V^{-1}f(y)dy. \]

We now claim that \( A_0Y(r_0) - A_1Y(r_1) \) is an invertible matrix. If not there would exist a non zero vector \( c_0 \) such that

\[ \{A_0Y(r_0) - A_1Y(r_1)\}c_0 = 0. \]

If we set \( f = 0 \) this states that \( F(x) = Y(x)c_0 \) is a nontrivial solution to the homogeneous equation \((\lambda - A)F = 0\) satisfying the boundary condition (3.6). But \( F = 0 \) by lemma 3.2. Hence (3.9) can be solved for \( c \) to yield a solution to (3.1).

Since \( AF = \lambda F - f \) with \( F \) and \( f \) satisfying the boundary conditions it follows that \( AF \) satisfies them as well. Hence \( F \in \mathcal{D}(A) \). We have thus completed the proof of (i), (ii), (iii) of (3.1). We now show that \( f \geq 0 \) implies \( F \geq 0 \).

Assume to the contrary that \( F \) takes on a strictly negative value at \((i,x)\) which we may assume to be a negative local minimum. Now \( F_i(x) < 0 \) and \( AF_i(x) \geq 0 \)
imply $f_i(x) = \lambda F_i(x) - Af_i(x) < 0$, a contradiction to the hypothesis that $f_i(x) \geq 0$. (v) of (3.1) is an immediate consequence of representation (3.8)

Proof of (3.1)(i)-(v) for $J = (-\infty, \infty)$:

Since $C_0^\infty(J)$ is dense in $C_0(J)$ (in the supnorm) assertion (i) is trivial. To prove (ii)-(v) we shall use lemmas 3.1 and 3.2 which apply as stated to $J = (-\infty, \infty)$ see 2.12 and the following paragraph for the definitions of $\mathcal{J}$ and $\mathcal{J}^1$ in this case.

We now prove the existence of a solution to $(\lambda-A)f(x) = f(x), \ -\infty < x < \infty$ with the property that $F_i(x) \in C_0^1(-\infty, \infty)$ provided $F_i(x) \in C_0(-\infty, \infty), i = 1, \ldots, N$.

First we make a change of variables with $U_i(x) = F_i(v_i x), g_i(x) = f_i(v_i x)$.

Let $B U_i(x) = U_i'(x) + \sum_{j=1}^N q_{ij} U_j(x)$. Now $U$ is a solution to

$$(3.10) \quad (\lambda - B)U = g$$

if and only if $(\lambda-A)f = f$. Moreover $U$ and $g$ vanish at $\pm \infty$ if and only if $F$ and $f$ do. Let $g$ have compact support. Then it is easy to see that $U(x) = \exp(x(\lambda I - Q)) \int_x^\infty \exp(-y(\lambda I - Q))g(y)dy$ satisfies (3.10). Since $g$ has compact support $\exp(x(\lambda I - Q)) = \exp x \lambda I \exp(-xQ)$ and for $x < 0 \ ||\exp(-xQ)|| < 1$ since $Q$ is an infinitesimal generator of a Markov chain. Thus we've produced a function $U(x) = (\lambda - B)^{-1}g(x)$ satisfying the equation (3.10) and such that $U_i(x) \in C_0(-\infty, \infty), i = 1, \ldots, N$ provided $g$ has compact support. From lemma 3.2 we deduce that $\lambda ||U|| \leq ||g||$.

Suppose now only that $g_i(x) \in C_0(-\infty, \infty)$. Let $g^n(x)$ denote a sequence of continuous functions with compact support such that $\lim_{n \to \infty} ||g^n - g|| = 0$ and $U^n = (\lambda - B)^{-1}g^n$.

Then $U^n - U^m = (\lambda - B)^{-1}(g^n - g_m)$ and lemma 3.2 together imply

$$(3.11) \quad ||U^n - U^m|| \leq \lambda^{-1}||g^n - g^m||.$$
It also follows from the equation (3.10) that

\[(3.12) \quad ||D^n - D^m|| \leq (2 + ||Q||)||g^n - g^m||.\]

Since \(\lim_{n \to \infty} ||g^n - g^m|| = 0\) we conclude that \(u^n\) and \(D^u^n\) form Cauchy sequences and \(\lim_{m \to \infty} \)

hence there exists a function \(U \in \mathcal{I}\) such that \(\lim_{n \to \infty} ||U - u^n|| = 0\),

\(\lim_{n \to \infty} ||D^U - D^u^n|| = 0\), and therefore \((\lambda - B)U = g\). This completes the proof of existence. The proof of (iv) and (v) are straightforward and omitted.
IV. Proof of Theorem 2.2.

We prove Theorem 2.2 for $J$ compact. For $J = (-\infty, \infty)$, the proof follows from Steps 1, 2, 4 below. We shall use the fact that given $g \in C(J)$, the equation

$$\lambda W - \Omega W = g, \lambda > 0,$$

has a unique solution $W \equiv (\lambda I - \Omega)^{-1} g \in \mathcal{G}$ and

$$\| (\lambda I - \Omega)^{-1} g \| \leq \frac{1}{\lambda} \| g \|, \quad [\text{Ma}]$$

To prove (2.32), we show that for all $g \in \mathcal{G}$, $\lambda > 0$,

$$\| P (\lambda I - \Omega)^{-1} g - (\lambda I - A(\varepsilon))^{-1} P g \| = o(1).$$

A simple extension of the Trotter-Kato theorem [Y; p. 269] gives (2.32) for all $g \in \mathcal{D}(\Omega^2) = \{ g : g \in \mathcal{A}, D^2 g \in \mathcal{A} \}$ (see [S; Remark, p. 255]). Now (2.32) follows for all $g \in \mathcal{G}$ since $\mathcal{D}(\Omega^2)$ is dense in $\mathcal{G}$ and all operators in (2.32) are contractions.

To prove (4.2), we need $(\lambda I - A(\varepsilon))^{-1} P g \in \mathcal{F}^3$, so we first prove (4.2) for $g \in \mathcal{G}^2$ (see (3.1(v)) (4.2) easily extends to all $g \in \mathcal{G}$. The proof comes in four steps. Step 1 shows that to prove (4.2) one may replace $(\lambda I - A(\varepsilon))^{-1} P g$ by $<\pi, (\lambda I - A(\varepsilon))^{-1} P g>$; Steps 2 and 3 give the interior and boundary estimates for the resolvent, respectively. Step 4 completes the proof. Fixing $g \in \mathcal{G}^2$, we define

$$W(\varepsilon) = <\pi, U(\varepsilon)> \quad \text{where } \pi \text{ is given in (2.6) and } U(\varepsilon) = (\lambda I - A(\varepsilon))^{-1} P g; \text{i.e.}$$

$$\begin{align*}
(\lambda I - \frac{1}{\varepsilon} V D - \frac{1}{\varepsilon^2} Q)U(\varepsilon) &= P g, \\
(\lambda I - \varepsilon)^{-1} P g &= W(\varepsilon).
\end{align*}$$

(a) $U_0^{(\varepsilon)}(r_0) = \sum_{j=1}^{N} b_{i,j} U_j^{(\varepsilon)}(r_0)$,

(b) $U_1^{(\varepsilon)}(r_1) = \sum_{j=1}^{N} b_{i,j} U_j^{(\varepsilon)}(r_1)$,

We set

$$W = (\lambda I - \Omega)^{-1} g.$$
Step 1. For each \( i = 1, \ldots, N, \)

\[
U_i(\varepsilon) - W(\varepsilon) = o(1).
\]

(4.7)

Step 2. For \( \lambda > 0, \)

\[
(\lambda I - \Omega)W(\varepsilon) = g + o(1).
\]

(4.8)

Step 3. (See (2.27) for the definition of Cases 1-3.)

(i) \( W(\varepsilon)(r_0) = o(1) \) if Case 1 holds at \( r_0, \)

(ii) \( DW(\varepsilon)(r_0) = o(1) \) if Case 2 holds at \( r_0, \)

(iii) \( D^2W(\varepsilon)(r_0) = o(1) \) if Case 3 holds at \( r_0, \)

with a similar statement for \( r_1. \)

Step 4. \( W(\varepsilon) - W = o(1), \)

which combined with Step 1 gives (4.2).
We define e as the constant vector \((1/N, \ldots, 1/N)\) and note that [Y; p. 205]

\[
R(Q^*) = N(Q)^\perp, \text{ where }
R(Q^*) \equiv \{ s \in \mathbb{R}^N | \exists r \in \mathbb{R}^N \text{ with } Q^*r = s \},
\]

\[
N(Q)^\perp \equiv \{ c \in \mathbb{R}^N | \langle c, e \rangle = 0 \}.
\]

Given \(\omega \in N(Q)^\perp\), we define the pseudo-inverse \((Q^*)^{-1}\) of \(Q^*\) by

\[
(Q^*)^{-1} \omega = -\int_0^\infty (e^{tQ^*} - \langle \omega, e^t \rangle e^t) \ dt.
\]

One can show that \(r \equiv (Q^*)^{-1} \omega\) is well-defined and is the unique solution of \(Q^*r = \omega\) which belongs to \(N(Q)^\perp\). Given \(\omega \in R(Q^*)\), we shall say \(\omega, D^k U(\varepsilon) > \in R(Q^*), k = 0, 1, 2, \ldots\). For the rest of this section, we write \(U\) for \(U(\varepsilon)\).

**Lemma 4.1.** Assume (2.23) For any \(c' \in N(Q)^\perp\), we have

\[
\langle c', D^m U \rangle = o(1), \ m = 0, 1, 2.
\]

Given \(c \in N(Q)^\perp\), we have

\[
\langle c, D^m U \rangle = -\varepsilon \langle (Q^*)^{-1} c, V D^{m+1} U \rangle + o(\varepsilon^2), m = 0, 1, 2,
\]

and

\[
\langle c, D^m U \rangle = -\varepsilon \langle (Q^*)^{-1} c, v \rangle < u, D^{m+1} U > + o(\varepsilon), m = 0, 1,
\]

where \(u\) is any vector \((u_1, \ldots, u_N)\) satisfying \(\Sigma u_i = 1\).

**Proof.** From (4.4), \(D^m U\) satisfies the equation

\[
(\lambda I - \frac{1}{\varepsilon} V D - \frac{1}{\varepsilon^2} Q)D^m U = D^m pg.
\]
Hence, defining $\psi = (Q^*)^{-1}c$, we have

\begin{equation}
\langle c, D^m U \rangle = \langle \psi, QD^m U \rangle \\
\quad = -\varepsilon \langle \psi, VD^{m+1} U \rangle + \lambda \varepsilon^2 \langle \psi, D^m U \rangle - \varepsilon^2 \langle \psi, D^m Pg \rangle.
\end{equation}

Now (4.11) is a consequence of (4.14) and (2.23). To prove (4.12), we note that the term $\langle \psi, D^m U \rangle$ is $o(1)$ by (4.11) (since $\psi \in R(Q^*)$) and $\langle \psi, D^m Pg \rangle = ND^m g \langle \psi, e \rangle = 0$. To prove (4.13), we have by (4.12)

\[ \langle c, D^m U \rangle = -\varepsilon \langle (Q^*)^{-1}c, \nu \rangle \langle \mu, D^{m+1} U \rangle + \varepsilon \phi + o(\varepsilon^2), \]

where

\[ \phi \equiv \langle (Q^*)^{-1}c, \nu \rangle \langle \mu, D^{m+1} U \rangle - \langle (Q^*)^{-1}c, VD^{m+1} U \rangle. \]

If we can show that $\phi \in R(Q^*)$, then (4.13) will follow by (4.12) applied to $D^{m+1} U$. But

\[ \phi = \langle \langle (Q^*)^{-1}c, \nu \rangle \mu - V(Q^*)^{-1}c, D^{m+1} U \rangle, \]

and

\[ \sum_{i=1}^N \langle (Q^*)^{-1}c, \nu \rangle \mu_i - \sum_{i=1}^N \langle V(Q^*)^{-1}c \rangle_i = \langle (Q^*)^{-1}c, \nu \rangle - \langle (Q^*)^{-1}c, \nu \rangle = 0. \]

**Proof of Step 1.** We have

\[ V_1(\varepsilon) - W(\varepsilon) = \langle c', V(\varepsilon) \rangle, \]

where $(c')_j = \delta_{ji} - \pi_j, j=1,\ldots,N$. Since $\sum_{j=1}^N (c')_j = 0$, (4.7) follows from (4.11).

**Proof of Step 2.** Take the inner product of both sides of (4.4) with $\pi$ and use $Q^*\pi = 0$ to derive

\begin{equation}
\lambda \langle \pi, U \rangle - \frac{1}{\varepsilon} \langle \pi, VDU \rangle = g.
\end{equation}

By (2.15), $\langle \pi, VDU \rangle$ is in $R(Q^*)$, so by (4.11) we have

\begin{equation}
\lambda \langle \pi, U \rangle + \langle Q^* \rangle^{-1} \langle \pi \nu \rangle \langle \nu, D^2 U \rangle = g + o(1).
\end{equation}

This is (4.8).
Proof of Step 3. We prove (4.9) (i)-(iii) using

\( \tilde{W}(\varepsilon) \equiv \langle e, U \rangle \) rather than \( W(\varepsilon) \equiv \langle \pi, U \rangle \). Since \( e - \pi \in N(Q)^1 \),
\( W(\varepsilon) - \tilde{W}(\varepsilon) = o(1) \) by (4.11). We write \( \gamma, \eta, \) and \( \theta \) for \( \gamma(r_0) \),
\( \eta(r_0) \), and \( \theta(r_0) \), respectively.

(i) Adding (4.5) (a) over \( i = 1, \ldots, k \) and inserting
\( \Sigma_{j=k+1}^N U_j(r_0) \) on both sides of the resulting equation, we find

\[
\langle e, U \rangle(r_0) = N^{-1} \left[ \Sigma_{i=1}^k \Sigma_{j=k+1}^N b_{ij} U_j(r_0) + \Sigma_{j=k+1}^N U_j(r_0) \right] = N^{-1} T.
\]

Hence for any real \( w \),

\[
(1-w)\langle e, U \rangle(r_0) = N^{-1} T - w\langle e, U \rangle(r_0).
\]

By hypothesis on the \( a_i \)'s, the number

\[
w \equiv N^{-1} \left[ \Sigma_{i=1}^k \Sigma_{j=k+1}^N b_{ij} + N - k \right] \in [0,1),
\]

and this \( w \) puts the right-hand side of (4.17) in \( R(Q^*) \). Hence by (4.11), we conclude \( \langle e, U \rangle(r_0) = o(1) \).

(ii) Equations (4.5) (a) imply that

\[
0 = \Sigma_{i=1}^k U_i(r_0) - \Sigma_{i=1}^k \Sigma_{j=k+1}^N b_{ij} U_j(r_0) = \langle \gamma, U \rangle(r_0),
\]

where \( \gamma \) is defined after (2.26). We claim that \( \gamma \in N(Q)^1 \):

\[
\Sigma_{i=1}^N y_i = \Sigma_{i=1}^k (1 - \Sigma_{j=k+1}^N b_{jm}) = \Sigma_{i=1}^k 1 = 0
\]

by hypothesis on the \( a_i \)'s. Hence by (4.13), \( \langle e, DU \rangle(r_0) = o(1) \) provided \( \eta \equiv \langle (Q^*)^{-1} \gamma, v \rangle \not= 0 \).
(iii) From (4.18), since \( \gamma \in \mathcal{R}(Q^*) \), we have using (4.12) that

\[ (4.19) \quad \langle V(Q^*)^{-1} \gamma, DU \rangle (r_0) = o(\varepsilon). \]

But by hypothesis \( \sum_{i=1}^{N} [V(Q^*)^{-1} \gamma]_i \equiv n = 0 \), and so by (4.13) and (4.19)

\[ o(\varepsilon) = \langle V(Q^*)^{-1} \gamma, DU \rangle (r_0) = -\varepsilon \langle Q^*^{-1} (V(Q^*)^{-1} \gamma), v \rangle < e, D^2 U \rangle (r_0) + o(\varepsilon). \]

Provided \( \theta \equiv \langle Q^*^{-1} (V(Q^*)^{-1} \gamma), v \rangle \neq 0 \), we see that \( < e, D^2 U \rangle (r_0) = o(1) \).

**Proof of Step 4.** We denote by \( s \) the \( o(1) \) error in (4.8) and by \( h_0, h_1 \) the \( o(1) \) errors in (4.9) (at \( r_0 \) and \( r_1 \), respectively). The function \( X(\varepsilon) \equiv W(\varepsilon) - W \) satisfies

\[ (\lambda I - \Omega)X(\varepsilon) = s, \quad H_i X(\varepsilon) = h_i, \quad i = 0,1. \]

Let \( Z(\varepsilon) \) solve

\[ (4.20) \quad (\lambda I - \Omega)Z(\varepsilon) = 0, \quad H_i Z(\varepsilon) = h_i, \quad i = 0,1. \]

We show below that \( Z(\varepsilon) \) exists and that

\[ (4.21) \quad ||Z(\varepsilon)|| \leq \text{const}(|h_0| + |h_1|). \]

Hence \( Y(\varepsilon) \equiv X(\varepsilon) - Z(\varepsilon) \) satisfies

\[ (\lambda I - \Omega)Y(\varepsilon) = k, \quad H_i Y(\varepsilon) = 0, \quad i = 0,1. \]

Since \( s \in \mathcal{C}(J) \), the bound \( ||Y(\varepsilon)|| \leq \lambda^{-1}||s|| \) is valid (see (4.1)).

This, combined with (4.21), completes Step 4. Concerning \( Z(\varepsilon) \), let \( u, v, w \) be a fundamental set of solutions for \( (\lambda I - \Omega)u = 0 \). Since
\( \lambda > 0 \) is in the resolvent set of \( \Omega \), we have

\[
\det \begin{bmatrix}
H_0u & H_0w \\
H_1u & H_1w
\end{bmatrix} \neq 0
\]

and so we can find constants \( c_1 \), \( c_2 \) so that \( z(\varepsilon) = c_1u + c_2w \).

Hence (4.21) holds. \( \blacksquare \)

We end this section by discussing Remark 2.3 (i).

**Proposition 4.2.** Assume that \( Q \) has the form (2.33)(iii). Then

\[
(-1)^i \eta(r_i) > 0, \quad i = 0, 1.
\]

**Proof.** We prove (4.22) for \( i = 0 \), the proof for \( i = 1 \) being similar. We write \( \gamma \) for \( \gamma(r_0) \). From (2.34), \( (Q^*)^{-1} = -I \) on \( N(Q)^\perp \), and since \( \gamma \in N(Q)^\perp \) we have

\[
\eta(r_0) = -\langle \gamma, v \rangle = \sum_{m=1}^{k} |v_j| + \sum_{m=k+1}^{N} (r_{j=1}^{k} b_{jm}) v_m > 0. \quad \blacksquare
\]

To show that Case 3 in (2.18) can arise, it can be checked that if

\[
Q = \begin{bmatrix}
-1.6 & .1 & 1.4 & .1 \\
.1 & -1.35 & .1 & 1.15 \\
1.4 & .1 & -1.6 & .1 \\
.1 & 1.15 & .1 & -1.35
\end{bmatrix}, \quad \gamma = (1, 1, -5, -1.5),
\]

\( v = (1, 4, -4, -1) \),

then

\[
Q^{-1} \gamma = (0.875, -1.25, -0.375, 1.125), \quad \kappa = 0, \quad \theta \in [5, 16].
\]
V. A Priori Bounds

We prove the bounds (see remark after (2.23))

$$||D^\lambda U(\varepsilon)(\lambda, \cdot)|| = O(1) \quad \text{for } \lambda > 0, \quad \lambda = 1, 2, 3,$$

in the three cases mentioned in Section II (see (2.27)). We assume $f \in \mathcal{E}^\lambda$ and write $U$ for $U(\varepsilon)(\lambda, \cdot)$. Note that it suffices to prove (5.1) for say $\lambda = 2.4$ since we may then interpolate with the $O(1)$ bound on $U$ derived in Section III [CS].

**Case (2.27)(i) (i).** (2.1) - (2.2), $J = (-\infty, \infty)$.

Since the $v_i$ and $Q$ are independent of $x$, we have for any $\lambda \geq 1$

$$||D^\lambda U|| \leq \lambda^{-1}||D^\lambda f||.$$

and so by (3.2)

$$||D^\lambda U|| \leq \lambda^{-1}||D^\lambda f||.$$

**Case (2.27(ii)).** (2.1)-(2.3), $N = 2$, $k = 1$, $b_{ij}$ satisfy (2.5), the centering (2.15 holds.)
We prove (5.1) inductively for \( l \) even given \( f \in F^l \). We write
\[
Q = \begin{bmatrix} -q_1 & q_1 \\ q_2 & -q_2 \end{bmatrix}, \quad q_1, q_2 > 0. \quad \text{The centering (2.15) implies}
\]
(5.4) \[ q_2 v_1 + q_1 v_2 = 0. \]

Say we know (5.1) for some even \( l \) (it holds for \( l = 0 \)). If a maximum of \( D^{l+2} \) should occur in the interior of \( J \), then we would have \( ||D^{l+2}U|| \leq \lambda^{-1} ||D^{l+2}f|| \). Thus it suffices to show
(5.5) \[ D^{l+2}U(x_i) = o(1), \quad i = 0, 1. \]

By the inductive hypothesis on \( D^l U \), we know
(5.6) \[ \epsilon ||D^{l+1}U|| = o(1), \]
since \( D^{l+1}U \) satisfies (from (5.2))
(5.7) \[ D^{l+1}U = -\frac{1}{\epsilon} V^{-1} QD^l U + \epsilon V^{-1} \lambda D^l U - \epsilon V^{-1} D^l f \]
\[ = -\frac{1}{\epsilon} V^{-1} QD^l U + o(\epsilon). \]

We claim that (5.5) will follow once we have shown
(5.8) \[ D^{l+1}U_1(x_i) - D^{l+1}U_2(x_i) = o(\epsilon), \quad i = 0, 1. \]

Indeed, then
\[
QD^{l+1}U(x_i) = (-q_1, q_2)(D^{l+1}U_1(x_i) - D^{l+1}U_2(x_i)) = o(\epsilon),
\]
and so, (5.5) is a consequence of (5.6) and the equation
(5.9) \[ D^{l+2}U = -\frac{1}{\epsilon} V^{-1} QD^{l+1}U + \epsilon V^{-1} \lambda D^{l+1}U - \epsilon V^{-1} D^{l+1}f. \]
We prove (5.8). From (5.7) we have

\[ D^{l+1}U_1(r_i) = \frac{1}{\varepsilon} \frac{q_1}{v_1} (D^lU_1(r_i) - D^lU_2(r_i)) + o(\varepsilon), \]

(5.10)

\[ D^{l+1}U_2(r_i) = -\frac{1}{\varepsilon} \frac{q_2}{v_2} (D^lU_1(r_i) - D^lU_2(r_i)) + o(\varepsilon). \]

We obtain (5.8) from (5.10) and (5.4).

Case (2.33(iii)). N even, k = N/2, v_1, b_{ij}, Q satisfy (2.33(iii)) (a)-(c).

We prove that (5.1) holds for any l even provided that f \in F^l satisfies the compatibility conditions

(5.11) \[ D^l f_i(r_0) = 0 = D^l f_i(r_1), i = 1, \ldots, N. \]

First, assuming (5.11) we show for general l, N, k, v_i, b_{ij}, and Q how to derive boundary conditions for D^lU of the form

(a) \[ D^l U_i(r_0) = \sum_{j=k+1}^{N} b_{ij}^{(l)} D^l U_j(r_0), i = 1, \ldots, k, \]

(5.12)

(b) \[ D^l U_i(r_1) = \sum_{j=1}^{k} b_{ij}^{(l)} D^l U_j(r_1), i = k+1, \ldots, N, \]

for certain b_{ij}^{(l)}. Since D^lU satisfies (5.2) on J, a sufficient condition for (5.1) is that the b_{ij}^{(l)} satisfy (2.5) for l = 2 and l = 3 or 4. Under the hypotheses of case (iii), we show that for any l = 1,2,3,...,
(5.13) \[ b_{i,N+1-i}^{(l)} = (-1)^l, \quad i = 1, \ldots, N; \text{ all other } b_{ij}^{(l)} = 0; \]

for \( l \) even the \( b_{ij}^{(l)} \) have the right form. Note that the requirement (5.11) on \( f \) does not affect our proof of Theorem 2.1 since the set of \( g \in G \) such that \( f \equiv Pg \) satisfies (5.11) for \( l \geq 1 \) is dense in \( G \).

We derive (5.12) (a). By (5.2) for \( l = 1,2,\ldots \)

\( D^lU \) at \( r_0 \) satisfies

\[
(5.14) \quad <[(\lambda - \frac{Q}{\varepsilon^2})^{-1}V]^lD^lU,n_i>(r_0) = <(\lambda - \frac{Q}{\varepsilon^2})^{-l}D^lU,f,n_i>(r_0)
\]

\[ = 0, \quad i = 1,\ldots,k, \]

since \( f \) satisfies (5.11).

Defining \( H \) as the \( k \times N \) matrix with \( i \text{th} \) row \( n_i \) and setting

\[ E = [(\lambda - \frac{Q}{\varepsilon^2})^{-1}V]^l, \]

we rewrite (5.14) as

\[ \sum_{j=1}^{k} (HE)_{ij} D^j U_j = -\sum_{j=k+1}^{N} (HE)_{ij} D^j U_j, \quad i = 1,\ldots,k. \]

This can be expressed as

\[
\begin{bmatrix}
D^l U_1 \\
\vdots \\
D^l U_k
\end{bmatrix}
- HES
\begin{bmatrix}
D^l U_{k+1} \\
\vdots \\
D^l U_N
\end{bmatrix}
= 0,
\]

where \( S \) is the \( N \times k \) matrix and \( \tilde{S} \) the \( N \times (N-k) \) matrix with partitioned forms

\[ S = \begin{bmatrix}
I \\
0
\end{bmatrix}^{k} \begin{bmatrix}
N-k \\
\end{bmatrix}, \quad \tilde{S} = \begin{bmatrix}
0 \\
I
\end{bmatrix}^{k} \begin{bmatrix}
N-k \\
\end{bmatrix}. \]
(I and O are identity and zero matrices, respectively, with the indicated number of rows and columns.) Thus, if \((HES)^{-1}\) exists, we obtain (5.12)(a):

\[
\begin{bmatrix}
D^LU_1 \\
\vdots \\
D^LU_k \\
\end{bmatrix} = -(HES)^{-1} HES
\]

\[
\begin{bmatrix}
D^LU_{k+1} \\
\vdots \\
D^LU_N \\
\end{bmatrix}
\]

(5.15)

This can be further simplified by writing

\[
E = \begin{bmatrix}
E_1 & E_2 \\
E_3 & E_4 \\
\end{bmatrix}^k, \quad H = \begin{bmatrix} I & -T \end{bmatrix}^k \quad (i.e. \ T_{ij} = b_{i,k+j}).
\]

(5.16)

Then (5.15) becomes (provided \((E_1 - TE_3)^{-1}\) exists)

\[
\begin{bmatrix}
D^LU_1 \\
\vdots \\
D^LU_k \\
\end{bmatrix} = -(E_1 - TE_3)^{-1}(E_2 - TE_4)
\]

\[
\begin{bmatrix}
D^LU_{k+1} \\
\vdots \\
D^LU_N \\
\end{bmatrix}
\]

(5.17)

We now assume the hypotheses of case (iii). From (2.33(iii) a, we have that \(RVR = -V\) and thus

5.18 \(RER = (-1)^6E\).

Writing \(R\) in the form

\[
\begin{bmatrix}
0 & \bar{R} \\
\bar{R} & 0 \\
\end{bmatrix}
\]

(\(\bar{R}\) an \(N/2 \times N/2\) reversal matrix),

we substitute into (5.18) the Form of \(E\) from (5.16) and obtain

5.19 \(E_3 = (-1)^6\bar{R}E_2 \bar{R}, \quad E_4 = (-1)^6\bar{R}E_1 \bar{R}\).

Since (2.33(iii))(b) states that \(T = \bar{R}\), the matrix \(E_1 - TE_3\) in (5.17) becomes (as \(\bar{R}^2 = I\))

5.20 \(E_1 - TE_3 = E_1 + E_2 \bar{R},\)
where we take the upper sign for \( l \) even and the lower sign for \( l \) odd.

**Lemma 5.1.** \( E_1 \bar{+} E_2 \bar{R} \) is invertible.

**Proof.** If not, there exists \( x \in \mathbb{R}^{N/2} \), \( x \neq 0 \), so that 
\[
(E_1 \bar{+} E_2 \bar{R})x = 0.
\]
But then by (5.19)
\[
E \begin{bmatrix} x \\ \bar{R}x \end{bmatrix} = \begin{bmatrix} E_1 & E_2 \\ E_3 & E_4 \end{bmatrix} \begin{bmatrix} x \\ \bar{R}x \end{bmatrix} = \begin{bmatrix} E_1 x \bar{+} E_2 \bar{R}x \\ \bar{R}(E_2 \bar{R}x \bar{+} E_1 x) \end{bmatrix} = 0.
\]
This violates the invertibility of \( E \).

By the lemma and (5.19) - (5.20), (5.17) becomes
\[
\begin{bmatrix} D^l U_1 \\ \vdots \\ D^l U_{N/2} \\ D^l U_N \end{bmatrix} = -(E_1 \bar{+} E_2 \bar{R})^{-1}(E_2 \bar{+} E_1 \bar{R}) \begin{bmatrix} D^l U_{N/2} + 1 \\ \vdots \\ D^l U_N \end{bmatrix}.
\]
This proves (5.13).
References


