The Annals of Applied Probability
1994, Vol. 4, No. 2, 390–413

FLUID MODELS IN QUEUEING THEORY AND WIENER–HOPF FACTORIZATION OF MARKOV CHAINS[1]

BY L. C. G. ROGERS

University of London

This paper applies the earlier work of Barlow, Rogers and Williams on the Wiener–Hopf factorization of finite Markov chains to a number of questions in the theory of fluid models of queues. Specifically, the invariant distribution for an infinite-buffer model and for a finite-buffer model are derived. The laws of other functionals of the fluid models can be easily derived and compactly expressed in terms of the fundamental Wiener–Hopf factorization.

1. Introduction. Fluid models of queues have been intensively studied during the 1980's, although the origins date back further (see [21, 14, 7, 6, 9, 1, 16, 17, 18, 5, 22 and 2]).

A simple example illustrates many of the features of this class of models. Consider a large water tank of capacity a, in the bottom of which there are some taps, each of which allows water to flow out at rate ρ. Above the tank are some pipes, each of which can be open or closed and, when open, pours water into the tank at rate θ. If Z_t is the number of open taps at time t and if Y_t is the number of open pipes, then the content ξ_t of the tank obeys the differential equation

$$\frac{d\xi_t}{dt} = \theta Y_t - \rho Z_t, \tag{1.1}$$

at least while $0 < \xi_t < a$. The behaviour at $\xi = 0$ or $\xi = a$ is what you expect; when the tank is empty, the outflow ceases, when the tank is full, water flows over the top. The most interesting question is of course "What can one say about the equilibrium behaviour of this system?" In particular, what proportion of time will water be running over the top of the tank? Exactly this model is studied in [16] and similar ones in [5, 9 and 1].

There is a teletraffic interpretation in which "water running over the top of the tank" is interpreted as "data messages being lost," an undesirable situation in either language! See Mitra [16] for a description of the analogy. The water tank is called a *buffer* in the teletraffic language.

In the late 1970's and early 1980's, David Williams and a number of co-workers were studying the theory of Wiener–Hopf factorization of Markov

Received September 1992; revised March 1993.
[1]This work was supported by SERC Grant GR/H00444.
AMS 1991 *subject classifications*. 60J10, 60K25, 60K15.
Key words and phrases. Markov chain, fluid model, Wiener–Hopf factorization, invariant distribution, noisy Wiener–Hopf.

processes (see [3, 12, 13, 19 and 23]). The basic idea is to take a strong Markov process X, with state space E, and some function $v\colon E \to \mathbb{R} \setminus \{0\}$ and form the additive functional

$$\varphi_t \equiv \int_0^t v(X_s)\, ds. \tag{1.2}$$

In general, this additive functional may decrease as well as increase, so that the time change

$$\tau_t^+ \equiv \inf\{u\colon \varphi_u > t\}$$

will in general have jumps, and the time-changed process

$$Y_t^+ \equiv X(\tau_t^+),$$

which only takes values in the closure of $\{x\colon v(x) > 0\}$, will not be easy to describe, although it is clear that it will still be Markovian.

The fluid model described at the beginning can be related to this very simply, if we take as our Markov process (Y_t, Z_t)—we assume that the opening and closing of pipes and taps is controlled by a finite-state Markov chain—and the function

$$v(y, z) = \theta y - \rho z.$$

Now the buffer content process ξ is not quite the same as the additive functional φ, due to the effects when the buffer is full or empty. However, it turns out that *understanding the process Y^+ allows one easily to express the quantities of interest for ξ, in terms of the fundamental Wiener–Hopf factorization of* [3].

In Section 2 of this paper, we recall the Wiener–Hopf factorization of Markov chains from [3]. We give a proof which differs considerably from that of [3] and is more transparent.

In Section 3, we analyze the infinite-buffer invariant distribution, which is simply expressed in terms of the fundamental quantities of the Wiener–Hopf factorization. This has been obtained previously [2]. In Section 4, we similarly analyze the finite-buffer invariant law; once again, this is simply expressed in terms of the fundamental Wiener–Hopf factorization and appears here in greater generality and simplicity than any of the special cases previously analyzed. Indeed, the results are so *strikingly* simple that there has to be a probabilistic story to explain why the solutions to the differential equations reduce so far; Section 5 provides this explanation. Throughout the paper, we shall assume that the function $v\colon E \to \mathbb{R}$ defining the additive functional φ via (1.2) is nonvanishing; if it were allowed to vanish in places, the invariant law for ξ can still be computed as an extension of the nonvanishing case, and for completeness we include this in Section 6.

More recently, "noisy" Wiener–Hopf problems have been studied by Williams and Kennedy (see [8]). Such models for teletraffic queues had already appeared in [10] and we report on them in Section 7 of this paper. In particular, in terms of the fundamental quantities of the noisy Wiener–Hopf factorization, we once again provide simple expressions for the invariant laws of such a process. Again, the infinite-buffer solution has been obtained previously in [2].

The usefulness of all this hinges on an ability to compute the Wiener–Hopf factorization. Straightaway, it must be admitted that there is only a closed-form analytical solution in trivial cases. However, there are various numerical approaches, some of which are reasonably efficient. If there are n_+ states where $v > 0$, n_- where $v < 0$, it seems inevitable that the problem is $O(n_+ n_-(n_+ + n_-))$, but the constants do vary. A forthcoming paper explores a number of possible algorithms in more detail.

2. The basic Wiener–Hopf factorization. The theory presented here is essentially that of [3], with the same notation. However, the emphasis there was on the chain killed at a constant positive rate, which resulted in certain technical simplifications; here, we shall concentrate on the original chain.

So suppose that we are given an irreducible Markov chain $(X_t)_{t \geq 0}$ on a finite state space E, with Q-matrix Q, and some function v: $E \to R \setminus \{0\}$. We set $E_+ = \{i\colon v(i) > 0\}$, $E_- = \{i\colon v(i) < 0\}$, and we assume that both of these are nonempty. The additive functional φ is defined via

$$\varphi_t = \int_0^t v(X_s)\, ds,$$

and the time changes

$$\tau_t^{\pm} \equiv \inf\{u\colon \pm\varphi_u > t\}$$

lead to the time-changed processes

$$Y_t^{\pm} \equiv X(\tau_t^{\pm}).$$

Notice that φ is a *fluctuating* additive functional, and the times when φ crosses a new maximum (which are the only times which τ^+ can see) are always times when the underlying chain X is in E_+. Thus Y^+ (respectively, Y^-) is a Markov chain which lives in E_+ (respectively, E_-). *What are the generators of* $Y^{\pm}$? The answer is contained in the following fundamental result, but to state the result, we need some notation. Let $V \equiv \operatorname{diag}(v(i))$ be the diagonal matrix whose entries are the values of v, and, for any set S, let $\mathscr{Q}(S)$ be the set of irreducible $S \times S$ Q-matrices (nonnegative off-diagonal entries, nonpositive row sums). We shall say that $G \in \mathscr{Q}(S)$ is *recurrent* if its row sums are all zero; otherwise we shall call G *transient*.

DEFINITION. A Wiener–Hopf factorization of $V^{-1}Q$ is a quadruple $(Z_+, Q_+; Z_-, Q_-)$, where Z_+ is $|E_-| \times |E_+|$, Z_- is $|E_+| \times |E_-|$, $Q_{\pm} \in \mathscr{Q}(E_{\pm})$, such that

$$V^{-1}Q\begin{pmatrix} I & Z_- \\ Z_+ & I \end{pmatrix} = \begin{pmatrix} I & Z_- \\ Z_+ & I \end{pmatrix}\begin{pmatrix} Q_+ & 0 \\ 0 & -Q_- \end{pmatrix}. \tag{2.1}$$

THEOREM 1 (Barlow, Rogers and Williams [3]).

(i) *The quadruple* $(\Pi_+, G_+; \Pi_-, G_-)$ *is always a Wiener–Hopf factorization of* $V^{-1}Q$, *where the following hold*:

$$\Pi_+(i,j) = P[X(\tau_0^+) = j, \tau_0^+ < \infty | X_0 = i], \qquad i \in E_-, j \in E_+; \tag{2.2}$$

$$G_+ \text{ is the generator of } Y_+; \tag{2.3}$$

and Π_- *and* G_- *are defined analogously by interchanging* $+$ *and* $-$.

(ii) *If the matrix* $Q \in \mathscr{Q}(E)$ *is transient, then the Wiener–Hopf factorization is unique* [*and is therefore given by* (2.2) *and* (2.3)].

REMARKS.

(i) The original statement of Theorem 1 as it appeared in [3] concerned the Wiener–Hopf factorization of $V^{-1}(Q - cI)$, where $c > 0$ was introduced to guarantee the uniqueness of the factorization by making $Q - cI$ transient.

(ii) Observe that (2.1) is equivalent to the pair of statements

$$V^{-1}Q\begin{pmatrix} I \\ Z_+ \end{pmatrix} = \begin{pmatrix} I \\ Z_+ \end{pmatrix} Q^+, \tag{2.4)(i}$$

$$V^{-1}Q\begin{pmatrix} Z_- \\ I \end{pmatrix} = -\begin{pmatrix} Z_- \\ I \end{pmatrix} Q_-. \tag{2.4)(ii}$$

(iii) The nonuniqueness of the Wiener-Hopf factorization for recurrent Q is apparent in the simplest example. Take

$$Q = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}, \qquad V = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where $\beta < \alpha$. The statement (2.4)(i) here becomes

$$\begin{pmatrix} -\alpha & \alpha \\ -\beta & \beta \end{pmatrix}\begin{pmatrix} 1 \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ z \end{pmatrix} q,$$

which has the two solutions $(z, q) = (1, 0)$ and $(z, q) = (\beta/\alpha, \beta - \alpha)$. Now since $\beta < \alpha$, the chain spends longer on average in E_- than E_+. Therefore $\varphi \to -\infty$, and the process Y^+ dies out; thus the solution (2.2)–(2.3) which we want will be the second. The reader may find it interesting to analyze this example replacing Q by $Q - \varepsilon I$, and letting $\varepsilon \downarrow 0$.

(iv) We shall investigate the possible nonuniqueness more carefully below; it turns out that if there is more than one solution, *there can only be two solutions*. See Theorem 1.

(v) To simplify (and without any real loss of generality), we shall often make the assumption that

$$v: E \to \{-1, 1\}, \tag{2.5}$$

so that

$$V = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}.$$

This loses no generality because we could always time-change the original chain by $|v|$, making a new Q-matrix $\tilde{q}_{ij} \equiv |v_i|^{-1} q_{ij}$,, relative to which assumption (2.5) is valid.

(vi) As a piece of notation, we shall write m for the unique invariant distribution of the recurrent Q-matrix Q:

$$mQ = 0, \qquad \sum_i m_i = 1. \tag{2.6}$$

We partition m into its part on E_+ and its part on E_-, $m \equiv (m_+, m_-)$.

When $m(E_+) > m(E_-)$, the additive functional φ drifts to $+\infty$ and the process Y^- is transient (dies out). We may frequently find it helpful to imgine that a "dead" process has been sent to a graveyard state ∂, which has the property that $v_\partial = 0 = f(\partial)$ for any function f (apart from $I_{\{\partial\}}$). Although it may be helpful to *imagine* this, it is not helpful to incorporate this into the notation, and we shall not attempt to do so.

It is worth remarking that in the balanced case, $m(E_+) = m(E_-)$, both of the processes $Y^\pm$ live forever; the analysis of the balanced case is often anomalous and delicate.

PROOF OF THEOREM 1. For the sake of completeness, we give a proof of Theorem 1 which differs in significant respects from that of [3]. We shall concentrate on part (2.4)(i) of the Wiener–Hopf factorization, part (2.4)(ii) being wholly analogous. We shall also assume (2.5).

If we partition the Q-matrix as

$$Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

then (2.4)(i) can be rewritten as

$$A + BZ_+ = Q_+, \tag{2.7)(i}$$

$$-C - DZ_+ = Z_+ Q_+. \tag{2.7)(ii}$$

We shall now show that when $(Z_+, Q_+) = (\Pi_+, G_+)$ as in (2.2)–(2.3), then (2.7) is valid; thereafter we shall prove the uniqueness statement.

Equality (2.7)(i) is essentially obvious; A governs the jumps of the underlying chain X while it remains in E_+, and B gives the jump rates from E_+ into E_-. However, if the chain X jumps from $i \in E_+$ to $j \in E_-$, the clock φ starts to decrease and, when it eventually regains its current level, X will be in state k with probability $\Pi_+(j,k)$. Thus $B\Pi_+$ is the matrix of jump rates for Y^+ corresponding to jumps which "passed via E_- invisibly."

To understand (2.7)(ii), imagine the chain starts in some state $i \in E_-$ and wait until the first time τ that it enters E_+. Then

$$P^i(\tau \in dt,\, X_\tau = k) = (e^{tD}C)_{ik}\, dt, \qquad k \in E_+.$$

Now, by decomposing the path at τ, it is obvious that Π_+ and G_+ defined by (2.2) and (2.3) must satisfy

$$\Pi_+ = \int_0^\infty e^{tD} C e^{tG_+}\, dt. \tag{2.8}$$

From this,

$$\begin{aligned} D\Pi_+ + \Pi_+ G_+ &= \int_0^\infty \{D e^{tD} C e^{tG_+} + e^{tD} C e^{tG_+} G_+\}\, dt \\ &= [e^{tD} C e^{tG_+}]_0^\infty \\ &= -C, \end{aligned}$$

which is (2.7)(ii).

Now we turn to the uniqueness assertion of the theorem. If (Z_+, Q_+) satisfy

$$V^{-1} Q \begin{pmatrix} I \\ Z_+ \end{pmatrix} = \begin{pmatrix} I \\ Z_+ \end{pmatrix} Q^+,$$

then any eigenvector f of Q_+ with eigenvalue λ gives an eigenvector

$$\bar{f} \equiv \begin{pmatrix} I \\ Z_+ \end{pmatrix} f \tag{2.9}$$

of $V^{-1}Q$ with the same eigenvalue. Under the assumption that Q is transient, it is easy to show that $V^{-1}Q$ cannot have an eigenvalue on the imaginary axis ([3] has an elementary proof), so it must be that $\mathrm{Re}(\lambda) < 0$. Now the statement $V^{-1}Q\bar{f} = \lambda \bar{f}$ implies that

$$\bar{f}(X_t)\exp(-\lambda \varphi_t) \text{ is a martingale,} \tag{2.10}$$

which is bounded on $[0, \tau_t^+]$ since $\mathrm{Re}(\lambda) < 0$. Thus if ζ is the lifetime of the chain X, we have by the optional sampling theorem that

$$\begin{aligned} \bar{f}(i) &= E^i\big[\bar{f}(X(\tau_0^+)) : \tau_0^+ < \zeta\big] \\ &= \begin{cases} (\Pi_+ f)_i, & \text{if } i \in E_-, \\ f(i), & \text{if } i \in E_+. \end{cases} \end{aligned} \tag{2.11}$$

If we now assume that there is a basis $\{f_j\colon j = 1, \dots, n = |E_+|\}$ for $\mathbb{R}^{E_+}$ consisting of eigenvectors of Q_+, we deduce that, for each $j = 1, \dots, n$ and for each $i \in E_-$,

$$\bar{f}_j(i) = (Z_+ f_j)_i = (\Pi_+ f_j)_i,$$

and hence $Z_+ = \Pi_+$.

The assumption of the existence of a basis of eigenvectors is not necessary, because there is always a basis of Jordan vectors [vectors such that, for some k and λ, $(Q_+ - \lambda I)^k g = 0$] and [3] shows how this can be used similarly to conclude that $Z_+ = \Pi_+$. The fundamental idea is that if

$$\bar{g}_j \equiv (V^{-1}Q - \lambda I)^{k-j} \bar{g}, \qquad j = 1, \dots, k,$$

then, for any $m = 1, \dots, k$,

$$\sum_{j=1}^{m} \frac{(-\varphi_t)^{m-j}}{(m-j)!} \exp(-\lambda\varphi_t)\bar{g}_j(X_t) \text{ is a martingale}$$

bounded on $[0, \tau_t^+]$, for any $t \geq 0$ [for this, it is *essential* that $\mathrm{Re}(\lambda) < 0$]. We refer the reader to [3] for further discussion. □

REMARKS. Provided $m(E_+) \neq m(E_-)$, one of the processes $Y^{\pm}$ dies out and one of the matrices $\Pi_{\pm}$ is strictly substochastic, so that

$$S \equiv \begin{pmatrix} I & \Pi_- \\ \Pi_+ & I \end{pmatrix} \text{ is invertible,}$$

and the Wiener–Hopf factorization (2.1) is actually the statement that $V^{-1}Q$ is similar to

$$U \equiv \begin{pmatrix} G_+ & 0 \\ 0 & -G_- \end{pmatrix} = S^{-1}V^{-1}QS.$$

Thus the eigenstructure of $V^{-1}Q$ is that of the block-diagonal matrix U, and diagonalising $V^{-1}Q$ would tell us what G_+ and G_- were; for example, if $m(E_+) > m(E_-)$, then the eigenvalues of G_- are in the open left half-plane, so an eigenvalue λ of $V^{-1}Q$ with positive real part is (the negative of) an eigenvalue of G_-, and the remainder are eigenvalues of G_+.

However, the balanced case $m(E_+) = m(E_-)$ is more delicate. To begin with, S is not now invertible. Both of G_+ and G_- have a zero eigenvalue, so that U has a kernel of dimension 2, whereas $V^{-1}Q$ has a kernel of dimension only 1. In this case, G_+ and G_- have exactly one eigenvalue, 0, on the imaginary axis, and the remainder in the open left half-plane; $V^{-1}Q$ has a Jordan block

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{2.12}$$

in its Jordan decomposition. One way to see this latter is that if there was *not* such a Jordan block, we could represent $V^{-1}Q$ in Jordan form:

$$V^{-1}Q = T\begin{pmatrix} 0 & 0 \\ 0 & J \end{pmatrix}T^{-1},$$

where J is an $(|E| - 1) \times (|E| - 1)$ Jordan matrix. Thus the first column of T would have to be in the kernel of $V^{-1}Q$, therefore a multiple of $\mathbf{1}$, the vector of 1's; the first row of T^{-1} would have to be a multiple of the vector mV^{-1}. However, the inner product of the first column of T with the first row of T^{-1} is, of course, 1; but in the balanced case, $mV^{-1}\mathbf{1} = 0$. This contradiction establishes the claim.

Let us now record a trivial corollary of Theorem 1, which we label as a theorem because of its fundamental importance in any attempt to compute the Wiener–Hopf factorization.

THEOREM 2. *Assuming the convention* (2.5), *the matrix* Π_+ *solves*

$$DZ + ZA + C + ZBZ = 0. \tag{2.13}$$

If Q *is transient, there is a unique substochastic solution* Z *to this equation.*

PROOF. If Z solves (2.13) and Z is substochastic, then (Z, Q_+) solves (2.4)(i), where $Q_+ \equiv A + BZ \in \mathscr{Q}(E_+)$. □

We conclude this section with the promised characterisation of all possible solutions in the case where Q is not transient.

THEOREM 3. *Suppose that* $Q \in \mathscr{Q}(E)$ *is recurrent and that, for some* $Q_+ \in \mathscr{Q}(E_+)$ *and* $|E_-| \times |E_+|$ *matrix* Z_+,

$$V^{-1}Q\begin{pmatrix} I \\ Z_+ \end{pmatrix} = \begin{pmatrix} I \\ Z_+ \end{pmatrix}Q_+.$$

(i) *If* Q_+ *is transient, then* $Q_+ = G_+$ *and* $Z_+ = \Pi_+$.
(ii) *If* Q_+ *is recurrent and* G_+ *is recurrent, then* $Q_+ = G_+$ *and* $Z_+ = \Pi_+$.
(iii) *If* Q_+ *is recurrent and* G_+ *is transient, then*

$$Q_+ = G_+ - (G_+\mathbf{1})\mu, \tag{2.14}$$

where μ *is the left eigenvector of* G_+ *whose eigenvalue has largest real part, and* μ *is normalised by the condition* $\mu\mathbf{1} = 1$.

The proof is deferred to the Appendix.

3. The invariant law for an infinite buffer. Let us now see how we may apply the theory of the previous section to obtain the invariant law of an infinite-buffer fluid model. We now think of the fluctuating additive functional φ as the difference between inflow and outflow, except that no fluid flows out when the buffer is empty. Therefore the buffer content at time t is

$$\xi_t \equiv \varphi_t - \underline{\varphi}_t, \qquad \underline{\varphi}_t \equiv \inf_{s \le t} \varphi_s.$$

In order to get an interesting limit distribution, we make the assumption (2.5) and also assume

$$m(E_-) > m(E_+), \quad \text{so that } \varphi_t \to -\infty \text{ a.s.} \tag{3.1}$$

One easy consequence of this is that

$$m_+ = m_-\Pi_+, \tag{3.2}$$

as we see easily from the Wiener–Hopf factorization (2.1) when we left-multiply by $(m_+, -m_-)$ and use the invertibility of G_+. We are now going to introduce the reversed-time quantities

$$\hat{Q} \equiv M^{-1}Q^TM, \qquad \hat{V} = -V, \qquad \hat{E}_+ = E_-, \qquad \hat{E}_- = E_+, \tag{3.3}$$

together with the time-reversed process $\hat{X}_t \equiv X_{-t}$ and its additive functional

$$\hat{\varphi}_t = \int_0^t \hat{v}(\hat{X}_s)\,ds = \varphi(-t).$$

Here, $M = \mathrm{diag}(m_i)$. The process $\hat{Y}^{\pm}$ and Q-matrices $\hat{G}_{\pm}$ are defined from $\hat{X}$ and $\hat{\varphi}$ in the same way that $Y^{\pm}$ and $G_{\pm}$ are defined from X and φ.

THEOREM 4. *For $x \geq 0$ and $j \in E$,*

$$\lim_{t\to\infty} P(X_t = j,\ \xi_t > x) = \begin{cases} m_j\big(\exp(x\hat{G}_-)\mathbf{1}\big)_j, & j \in E_+ \equiv \hat{E}_-, \\ m_j\big(\hat{\Pi}_- \exp(x\hat{G}_-)\mathbf{1}\big)_j, & j \in E_- \equiv \hat{E}_+. \end{cases} \tag{3.4}$$

PROOF. Assume that the process $(X_t)_{t\in\mathbb{R}}$ is in equilibrium, and $\varphi_0 = 0$. From Figure 1 it is easy to see that, for $x > 0$ and $j \in E$,

$$\{X_0 = j,\ \xi_0 > x\} = \Big\{\hat{X}_0 = j,\ \inf_{u\geq 0} \hat{\phi}_u < -x\Big\},$$

so that

$$\begin{aligned} \lim_{t\to\infty} P(X_t = j,\ \xi_t > x) &= P(X_0 = j,\ \xi_0 > x) \\ &= P\big(\hat{X}_0 = j\big)P\Big(\inf_{u\geq 0}\hat{\varphi}_u < -x|\hat{X}_0 = j\Big) \\ &= m_j P\big(\hat{Y}^- \text{ lives for at least } x|\hat{X}_0 = j\big). \end{aligned}$$

However, the generator of $\hat{Y}^-$ is $\hat{G}_-$, and so

$$P(X_0 = j,\ \xi_0 > x) = \begin{cases} m_j\big(\exp(x\hat{G}_-)\mathbf{1}\big)_j, & \text{if } j \in E_+ \equiv \hat{E}_-, \\ m_j\big(\hat{\Pi}_- \exp(x\hat{G}_-)\mathbf{1}\big)_j, & \text{if } j \in E_- \equiv \hat{E}_+, \end{cases}$$

as required. □

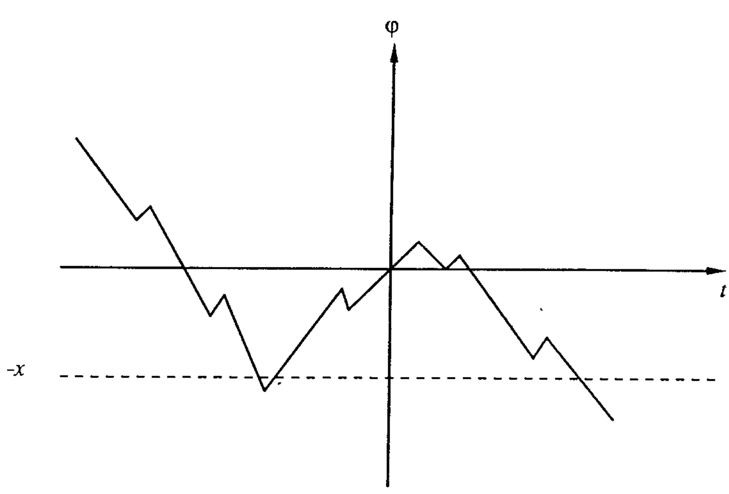

FIG. 1.

REMARKS. One could of course compute $\hat{G}_{\pm}$ by any algorithm which might be used to compute $G_{\pm}$. However, as a simple exercise in matrix manipulation, London, McKean, Rogers and Williams [12] prove that

$$\hat{\Pi}_{\pm} = \Pi^*_{\pm}, \tag{3.5)(i}$$

$$\hat{G}_{\pm} = \left(K_{\mp}^{-1} G_{\mp} K_{\mp}\right)^*, \tag{3.5)(ii}$$

where, for any matrix Z, Z^* is defined by $(Z^*)_{ij} \equiv m_j z_{ji}/m_i$, and

$$K_{+} = (I - \Pi_{-}\Pi_{+})^{-1}, \tag{3.6)(i}$$

$$K_{-} = (I - \Pi_{+}\Pi_{-})^{-1}. \tag{3.6)(ii}$$

(Notice that [12] took the reversal with the function v, rather than $-v$, as we have here; this accounts for the apparent interchange of plus and minus.) The conclusion is that *if* we have obtained $\Pi_{\pm}$ and $G_{\pm}$, then it is a trivial matter to deduce $\hat{\Pi}_{\pm}$ and $\hat{G}_{\pm}$.

The laws of other functionals of interest for this queueing model can be easily expressed in terms of the fundamental quantities of the Wiener–Hopf factorization. For example, the equilibrium probability of finding the buffer empty is given, from (3.4), by

$$\begin{aligned} P(\xi_0 = 0, X_0 = j) &= m_j\left(1 - (\hat{\Pi}_{-}1)_j\right) \\ &= m_j - (m_{+}\Pi_{-})_j \\ &= m_{-}(I - \Pi_{+}\Pi_{-})_j, \qquad j \in \hat{E}_{+} \equiv E_{-}, \end{aligned} \tag{3.7}$$

using (3.5)(i) and (3.2), respectively. This simple formula has a simple interpretation; m_j is the long-run proportion of time spent in state $j \in E_{-}$, and $(m_{-}\Pi_{+}\Pi_{-})_j$ is the long-run proportion of time spent in state j *when* φ *is at a level visited earlier*. We explain this more fully in Section 5. Summing (3.7) over j reveals that

$$P(\xi_0 = 0) = m(E_{-}) - m(E_{+}), \tag{3.8}$$

which is probabilistically obvious; the process $\underline{\varphi}$ is decreasing at rate 1 when $\xi = 0$, and at rate 0 otherwise, so

$$\begin{aligned} \lim_{t\to\infty} t^{-1} \int_0^t I_{\{\xi_s = 0\}}\, ds &= \lim_{t\to\infty} - t^{-1}\, \underline{\varphi}(t) \\ &= \lim_{t\to\infty} - t^{-1}\, \varphi(t) \\ &= -\sum m_i v_i \\ &= m(E_{-}) - m(E_{+}). \end{aligned}$$

We can obtain the limit distribution of the state of the chain X when a busy period starts; it is given simply by

$$\{m(E_{-}) - m(E_{+})\}^{-1} m_{-}(I - \Pi_{+}\Pi_{-})(-D^{-1}C),$$

since $(-D^{-1}C)_{ij} = P(X \text{ first enters } E_{+} \text{ at } j | X_0 = i)$, for $i \in E_{-}$.

The distribution of the duration of a busy period is also easy to find; if the chain were in state $j \in E_+$ when the busy period began, then we are simply asking for the law of τ_0^-, which can in principle be found from the Wiener–Hopf factorization of $V^{-1}(Q - cI)$, as was observed in [3]. Indeed, in this case,

$$\Pi_-(i,j) = E\left(e^{-c\tau_0^-}; X(\tau_0^-) = j | X_0 = i\right),$$

so one finds (in principle) the Laplace transform of τ_0^-.

It is not the present purpose to exhaust all the possible calculations one could do on this model; the aim is rather to show the fundamental importance of the Wiener–Hopf factorization $(\Pi_+, G_+; \Pi_-, G_-)$. Perhaps more interesting is the computation of the invariant law for a finite buffer, to which we turn next.

4. The invariant law for a finite buffer. Once again in this section we make the assumption (2.5) that $|v| = 1$, and now we suppose we have a buffer of finite capacity $a > 0$. Once again, φ represents the difference between inflow and outflow, but, again, φ has to be modified in the obvious way to obtain the buffer content process ξ; we can express it as

$$\frac{d\xi_t}{dt} = \left[I_{(0,a)}(\xi_t) + I_{\{X_t \in E_+,\, \xi_t = 0\}} + I_{\{X_t \in E_-,\, \xi_t = a\}}\right] v(X_t), \tag{4.1}$$

for example.

How do we characterize the invariant distribution of the process (X_t, ξ_t)? The simple time-reversal argument of the previous section does not now work, so we need another method. An obvious approach is to solve the adjoint equations (or Kolmogorov forward equations) for the invariant law. Observe that the generator of (X_t, ξ_t) is given by

$$\begin{aligned} \mathscr{G}f(j,\xi) &= (Qf + Vf')(j,\xi) \\ &= \sum_k q_{jk} f(k,\xi) + v(j)\frac{df}{d\xi}(j,\xi), \end{aligned} \tag{4.2}$$

at least on functions f for which $f'(j,0) = 0$, for $j \in E_-$, and $f'(j,a) = 0$, for $j \in E_+$. We shall seek an invariant distribution of the form

(4.3)(i) $\quad P(X = j, \xi \in dx) = \pi_j(x)\,dx, \qquad 0 < x < a,\ j \in E,$

(4.3)(ii) $\quad P(X = j, \xi = 0) = p_-(j), \qquad j \in E_-,$

(4.3)(iii) $\quad P(X = j, \xi = a) = p_+(j), \qquad j \in E_+,$

where π is a C^1 function on $[0, a]$. We construct such an invariant distribution by solving the adjoint equations and, finally, argue that there can be no other invariant distribution and that this distribution is the limit as t tends to infinity of the law of (X_t, ξ_t).

We arrive finally at the main result of this section.

THEOREM 5. *The process (X, ξ) has a unique equilibrium distribution, which is of the form* (4.3)(i)–(iii). *We have explicitly, in the case $m(E_-) \neq m(E_+)$,*

$$p_- = m_-(I - \Pi_+\Pi_-)\exp(aG_-)K_-(a) \tag{4.4(i)}$$

$$p_+ = m_+(I - \Pi_-\Pi_+)\exp(aG_+)K_+(a) \tag{4.4(ii)}$$

where

$$K_\pm(a) \equiv \left(I - \Pi_\mp \exp(aG_\mp)\Pi_\pm \exp(aG_\pm)\right)^{-1}, \tag{4.5}$$

and in the balanced case $m(E_+) = m(E_-)$ we have that, for some θ,

$$p_\pm = \theta\nu_\pm(a), \tag{4.6}$$

where $\nu_\pm(a)$ is the unique probability on $E_\pm$ such that

$$\nu_\pm(a)K_\pm(a) = 0.$$

In either case, the function π is given in terms of $p_\pm$ by

$$\pi(x) = p_- Q\exp(xV^{-1}Q)V^{-1} = -p_+ Q\exp((x-a)V^{-1}Q)V^{-1}. \tag{4.7}$$

[*Here we write p_- instead of the clumsier notation* $(0\ \ p_-)$.] *In the balanced case, θ is determined by the normalization condition*

$$1 = \theta\left(1 + \nu_-(a)\exp(aQV^{-1})\mathbf{1}\right). \tag{4.8}$$

Finally, the distribution of (X_t, ξ_t) converges as $t \to \infty$ in variation norm to the equilibrium distribution.

If an invariant distribution has the form (4.3) then, for all test functions in the domain of $\mathscr{G}$, thinking of $p_\pm$ and π as row vectors, we have (e.g., as in [4], page 238) that

$$\begin{aligned} 0 &= p_- Qf(0) + p_+ Qf(a) + \int_0^a \pi(x)\mathscr{G}f(x)\,dx \\ &= p_- Qf(0) + p_+ Qf(a) + \pi(a)Vf(a) - \pi(0)Vf(0) \\ &\quad + \int_0^a (\pi(x)Q - \pi'(x)V)f(x)\,dx. \end{aligned}$$

Since the function f is arbitrary, we find that

$$\begin{aligned} \pi'(x)V &= \pi(x)Q, \\ p_- Q &= \pi(0)V, \\ p_+ Q &= -\pi(a)V, \end{aligned}$$

which has a solution of the form (4.7):

$$\begin{aligned} \pi(x) &= p_- Q\exp(xV^{-1}Q)V^{-1} \\ &= -p_+ Q\exp((x-a)V^{-1}Q)V^{-1} \end{aligned} \tag{4.9}$$

and it only remains to identify $p_{\pm}$. From (4.9) we see that p_+ and p_- are related by

(4.10) $$p_-Q = -p_+Q\exp(-aV^{-1}Q);$$

if we now right-multiply both sides of this equation by $\begin{pmatrix} I \\ \Pi_+ \end{pmatrix}$, we obtain

$$p_-Q\begin{pmatrix} I \\ \Pi_+ \end{pmatrix} = -p_+Q\begin{pmatrix} I \\ \Pi_+ \end{pmatrix}\exp(-aG_+),$$

using (2.1). However, once again using (2.1) and the fact that $p_{\pm}$ is concentrated on $E_{\pm}$, this reduces to

(4.11)(i) $$p_-\Pi_+G_+ = p_+G_+e^{-aG_+},$$

and similarly right-multiplying (4.6) by $\begin{pmatrix} \Pi_- \\ I \end{pmatrix}$ yields

(4.11)(ii) $$p_+\Pi_-e^{aG_-}G_- = p_-G_-.$$

The analysis now splits into three cases.

CASE 1 [$m(E_-) > m(E_+)$]. In this case, G_+ is invertible, so that (4.11)(i) yields

$$p_+ = p_-\Pi_+e^{aG_+};$$

substituting this into (4.11)(ii), we learn that

(4.12) $$p_-(I - \Pi_+e^{aG_+}\Pi_-e^{aG_-})G_- = 0.$$

If we write ν_- for the invariant distribution of G_-, (4.12) is reexpressed as

(4.13) $$p_- \propto \nu_-K_-(a);$$

since (as is easily verified) in this case $\nu_- \propto m_-(I - \Pi_+\Pi_-)$, we have that

(4.14)(i) $$p_- = cm_-(I - \Pi_+\Pi_-)K_-(a),$$

(4.14)(ii) $$p_+ = cm_-(I - \Pi_+\Pi_-)K_-(a)\Pi_+e^{aG_+},$$

for some constant $c > 0$. The constant c is fixed by the normalization condition that

$$\int_0^a \pi(x)\mathbf{1}\,dx + p_+\mathbf{1} + p_-\mathbf{1} = 1,$$

and a few calculations yield the conclusion that $c = 1$.

We may use the facts that $K_-(a)\Pi_+e^{aG_+} = \Pi_+e^{aG_+}K_+(a)$, $m_-(I - \Pi_+\Pi_-)$ is invariant for G_- and $m_+ = m_-\Pi_+$ to reexpress (4.14) in the symmetric form

$$p_- = m_-(I - \Pi_+\Pi_-)\exp(aG_-)K_-(a),$$
$$p_+ = m_+(I - \Pi_-\Pi_+)\exp(aG_+)K_+(a),$$

which is (4.4).

CASE 2 [$m(E_+) > m(E_-)$]. An exactly analogous argument will yield (4.4) as in Case 1.

CASE 3 [$m(E_+) = m(E_-)$]. As one expects, the balanced case is the most delicate.

First, if we define

$$L_t^x = \int_0^t I_{(\xi_u = x)}\, du,$$

for $x = 0, a$, then the content process ξ is related to the additive functional φ by

$$\xi_t = \varphi_t + L_t^0 - L_t^a.$$

Notice that

$$\frac{1}{t} L_t^0 \to p_- \mathbf{1}, \quad \text{a.s.},$$

$$\frac{1}{t} L_t^a \to p_+ \mathbf{1}, \quad \text{a.s.},$$

as $t \to \infty$, and the fact that $t^{-1}\ \varphi_t \to 0$ a.s. forces the conclusion

$$p_+ \mathbf{1} = p_- \mathbf{1}. \tag{4.15}$$

Returning to (4.11), we discover that

(4.16)(i) $\quad p_+ - p_- \Pi_+ \exp(aG_+)$ is invariant for G_+,

(4.16)(ii) $\quad p_- - p_+ \Pi_- \exp(aG_-)$ is invariant for G_-,

but $(p_+ - p_- \Pi_+ \exp(aG_+))1 = p_+ 1 - p_- 1 = 0$ by (4.13), so the only possibility is

$$p_+ = p_- \Pi_+ \exp(aG_+),$$

$$p_- = p_+ \Pi_- \exp(aG_-).$$

Thus p_+ is a multiple of the invariant law $\nu_+(a)$ of $\Pi_- \exp(aG_-)\Pi_+ \exp(aG_+)$; if we suppose $p_+ = \theta\nu_+(a)$, then we determine θ by the normalization condition

$$\begin{aligned} 1 &= p_+ \mathbf{1} + p_- \mathbf{1} + \int_0^a \pi(x) \mathbf{1}\, dx \\ &= 2\theta + \int_0^a p_- Q \exp(xV^{-1}) V^{-1} \mathbf{1}\, dx \\ &= 2\theta + \theta\nu_-(a) \{\exp(aQV^{-1}) - I\} \mathbf{1} \\ &= \theta + \theta\nu_-(a) \exp(aQV^{-1}) \mathbf{1}, \end{aligned}$$

which is (4.8).

Although this may not look very explicit, it is easily calculable in practice.

To complete the proof, we give a coupling argument (applicable equally to the infinite-buffer case) which proves uniqueness of the invariant law and convergence of the laws of (X_t, ξ_t) to it in variation norm. See [11] for more on the coupling method.

The uniqueness will follow immediately from the fact that, for any two starting distributions μ and μ' on $E \times [0, a]$, we can build on one sample space processes (X, ξ) and (X', ξ') with initial laws μ and μ', respectively, such that

$$P[\text{for some } T < \infty, (X_t, \xi_t) = (X'_t, \xi'_t) \text{ for all } t \geq T] = 1.$$

How do we build these two processes? Start by letting X and X' evolve independently, until the first time τ such that $X_\tau = X'_\tau$, at which time stick the particles together and make them follow a common trajectory thereafter. Of course in general $\xi_\tau \neq \xi'_\tau$ so if we suppose $\xi_\tau > \xi'_\tau$, we see that at all times $t \geq \tau$ the buffer content ξ_t must be at least as large as ξ'_t. So we just wait until ξ_t falls to 0, at which time $\xi_t = \xi'_t$ and from then on the two processes (X, ξ) and (X', ξ') coincide forever. □

REMARKS. Assuming that $m(E_-) > m(E_+)$, we can let $a \to \infty$ in the above results, and we should obtain the infinite-buffer answer of Section 3. This indeed happens, as the reader is invited to check using the reversal results (3.5) and (3.6). One can rephrase the results of this section in terms of the reversal, but to no advantage; in the finite-buffer situation, the reversal is a process of the same nature as the forward process, whereas in the infinite-buffer case, φ is transient to $+\infty$ rather than $-\infty$.

5. Sample-path explanation of the invariant law. We now give the promised elementary explanation of (3.7) and (4.14)(i), assuming that $m(E_-) > m(E_+)$ and assuming (2.5). To begin with, let us suppose that the underlying process is killed at an independent exponential time T of mean ε^{-1}, so that Q is replaced by $Q - \varepsilon I$. The Wiener–Hopf factorization shall, for this section only, depend implicitly on ε, but we shall omit ε from the notation and speak, for example, of Π_+, when we mean

$$\Pi_+(i, j) = E^i[\exp(-\varepsilon\tau_0^+); X(\tau_0^+) = j].$$

No confusion should arise. First, we explain the infinite-buffer result (3.7). If $N(y, j) = |\{t < T\colon \varphi_t = y, X_t = j\}|$, then the total time spent before T while $\varphi_t < 0$ and $X_t = j$ can be expressed in the two alternative forms

$$\int_0^T I_{\{X_t = j,\, \varphi_t < 0\}}\, dt = \int_{-\infty}^0 N(y, j)\, dy. \tag{5.1}$$

Now, for any $i, j \in E_-$, $y < 0$,

$$\begin{aligned} E^i[N(y, j)] &= \Big(\exp(-yG_-) \sum_{r \geq 0} (\Pi_+\Pi_-)^r\Big)(i, j) \\ &= \exp(-yG_-)(I - \Pi_+\Pi_-)^{-1}(i, j), \end{aligned} \tag{5.2}$$

because once φ reaches level y (when X is in state k), the probability that it will again be at level y in the future with $X = l$ is simply $\Pi_+\Pi_-(k, l)$. Thus

$$E^i \int_{-\infty}^0 N(y, j)\, dy = (-G_-)^{-1}(I - \Pi_+\Pi_-)^{-1}(i, j)$$

and

$$E^i \int_0^T I_{\{X_t = j,\, \varphi_t < 0\}}\, dt \sim E^i \int_0^T I_{\{X_t = j\}}\, dt$$
$$\sim \varepsilon^{-1} m_j \quad \text{as } \varepsilon \downarrow 0.$$

We conclude therefore that, for $i, j \in E_-$,

$$\begin{aligned} \varepsilon(-G_-)^{-1}(i,j) &= \varepsilon E^i \int_0^\infty I_{\{Y_t = j\}} \exp(-\varepsilon \tau_t^-)\, dt \\ &= \varepsilon E^i \int_0^\infty I_{\{X_t = j,\, \varphi_t = \underline{\varphi}_t\}} e^{-\varepsilon t}\, dt \\ &\sim m_-(I - \Pi_+\Pi_-)_j, \end{aligned} \tag{5.3}$$

so that the limit of $P(X_t = j,\ \varphi_t = \underline{\varphi}_t)$ must be $m_-(I - \Pi_+\Pi_-)_j$, as stated in (3.7). Of course, this argument does not prove existence of a limit, but does identify that limit if it exists. The existence of the limit follows from the coupling argument at the end of the previous section.

How do we now handle similarly the finite-buffer case? If the buffer has capacity a as before, and if $\xi_t = \varphi_t + L_t^0 - L_t^a$ is the buffer content at time t, it is obvious that $\xi_t \le \eta_t \equiv \varphi_t - \underline{\varphi}_t$, the content of an *infinite* buffer at time t. Thus at all times when $\varphi_t = \underline{\varphi}_t$, the finite buffer is certainly empty, although it may of course be empty at other times, too. The key observation is contained in the following result.

PROPOSITION. *Fix some $u > 0$, and define*

$$S = \inf\{t > \tau_u^- : \varphi_t = -u + a\}.$$

Then

$$\sigma_u \equiv \inf\{t > \tau_u^- : \varphi_t = -u,\ \xi_t = 0\} = \inf\{t > S : \varphi_t = -u\} \equiv \sigma_u'.$$

PROOF. From the definition of τ_u^-, it follows that

$$L^0(\sigma_u) - L^0(\tau_u^-) > 0,$$

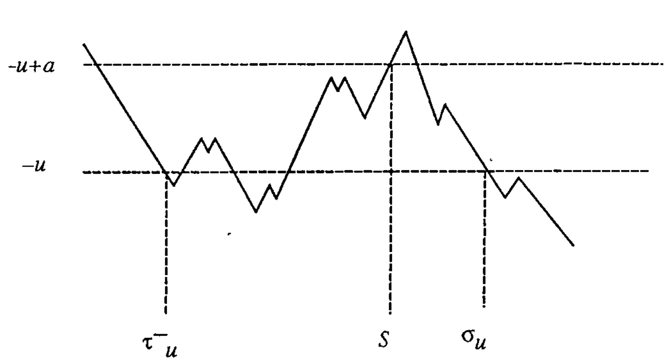

FIG. 2.

and therefore since $\xi(\tau_u^-) = \xi(\sigma_u) = 0$ and $\varphi(\tau_u^-) = \varphi(\sigma_u) = -u$, we must certainly have

$$L^a(\sigma_u) - L^a(\tau_u^-) = L^0(\sigma_u) - L^0(\tau_u^-) > 0;$$

so at some time in (τ_u^-, σ_u), the buffer must be full. Let the last such time be denoted by λ. Since L^a does not change in (λ, σ_u), the drop of $\varphi = \xi - L^0$ in this interval must be at least as great as the drop of ξ, which is a. Thus *at some time in* (τ_u^-, σ_u), φ *was at level* $-u + a$, and so $\tau_u^- < S < \sigma_u$.

However, now we observe that the buffer is certainly *full* at time S (just as the buffer is empty at any time where $\varphi_t = \underline{\varphi}_t$) and likewise must be *empty* at $\sigma_u' = \inf\{t > S: \varphi_t = -u\}$. □

So now if we define

$$N_a(y, j) = |\{t: \varphi_t = y, \xi_t = 0, X_t = j\}|,$$

we have as in (5.2) that, for $i, j \in E_-$ and $y < 0$,

$$\begin{aligned} E^i N_a(y,j) &= \Big(\exp(-yG_-) \sum_{r \geq 0} (\Pi_+ \exp(aG_+)\Pi_- \exp(aG_-))^r\Big)(i,j) \\ &= \big(\exp(-yG_-)(I - \Pi_+ \exp(aG_+)\Pi_- \exp(aG_-))^{-1}\big)(i,j) \\ &= (\exp(-yG_-)K_-(a))(i,j). \end{aligned}$$

Thus

$$\begin{aligned} &\varepsilon E^i \int_0^T I_{\{\varphi_t < 0,\, \xi_t = 0,\, X_t = j\}}\, dt \\ &\quad = \varepsilon \int_{-\infty}^0 (\exp(-yG_-)K_-(a))(i,j)\, dy \\ &\quad = \varepsilon(-G_-)^{-1} K_-(a)(i,j) \\ &\quad \to m_-(I - \Pi_+\Pi_-)K_-(a)_j; \qquad \varepsilon \downarrow 0, \end{aligned} \tag{5.4}$$

by the argument in (5.3). However, the limit of (5.4) as $\varepsilon \downarrow 0$ is the limiting probability that $\xi = 0$, $X = j$, yielding (4.14)(i), as promised.

6. Invariant law when V may vanish. In this section, we allow the possibility that v takes the value 0; as before, we may without loss of generality assume that v takes values in $\{-1, 0, 1\}$ and define $E_0 = \{i: v(i) = 0\}$ and $E_\pm$ as before. Writing $E. = E_+ \cup E_-$, we partition Q as

$$Q = \begin{pmatrix} Q_{\cdot\cdot} & Q_{\cdot 0} \\ Q_{0\cdot} & Q_{00} \end{pmatrix} \begin{pmatrix} E. \\ E_0 \end{pmatrix}.$$

If we time-change out the time spent in E_0, we see a chain $\tilde{X}$ which lives in $E.$ and has generator

$$\tilde{Q} \equiv Q_{\cdot\cdot} - Q_{\cdot 0} Q_{00}^{-1} Q_{0\cdot}.$$

We can carry out the Wiener–Hopf factorization of this chain exactly as before and obtain the invariant law for the buffer of finite capacity a in terms of the density $\tilde{\pi}$ and the two row vectors $\tilde{p}_+$ and $\tilde{p}_-$. The defining properties are that

(6.1)(i) $$0 = \int_0^a \tilde{\pi}(x)\left[\tilde{Q}f(x) + Vf'(x)\right] dx$$
$$+\tilde{p}_-\tilde{Q}f(0) + \tilde{p}_+\tilde{Q}f(a) \quad \text{for all } f \in \mathscr{D}(\tilde{\mathscr{G}});$$

(6.1)(ii) $$\tilde{p}_+\mathbf{1} + \tilde{p}_-\mathbf{1} + \int_0^a \tilde{\pi}(x)\mathbf{1}\, dx = 1.$$

Here, as before

$$\mathscr{D}(\tilde{\mathscr{G}}) = \{f\colon E.\times[0,a] \to \mathbb{R} | f(j,\cdot) \text{ is } C^1$$
$$\text{for all } j,\ f'(i,0) = 0 = f'(k,a) \text{ for } i \in E_-,\ k \in E_+\}.$$

If we now define the E-vectors p_+, p_-, $\pi(x)$ by

$$p_\pm = \tilde{p}_\pm(I, -Q_{\cdot 0}Q_{00}^{-1}),$$
$$\pi(x) = \tilde{\pi}(x)(I, -Q_{\cdot 0}Q_{00}^{-1}),$$

it is trivial to verify that, for all

$$f \in \mathscr{D}(\mathscr{G}) = \{f\colon E \times [0,a] \to \mathbb{R} | f(j,\cdot) \text{ is } C^1$$
$$\text{for all } j,\ f'(i,0) = 0 = f'(k,a) \text{ for } i \in E_-,\ k \in E_+\},$$

$$0 = \int_0^a \pi(x)\{Qf(x) + Vf'(x)\} + p_-Qf(0) + p_+Qf(a),$$

so that (π, p_+, p_-) is (a multiple of) the invariant law for the process including the states of E_0. We leave the reader the minor task of verifying that if m is the invariant law of Q, then we obtain the invariant law by multiplying (π, p_+, p_-) by $m(E.)$. Uniqueness of the invariant law follows again as in Section 4.

7. Noisy Wiener–Hopf problems. As before, we take an irreducible Markov chain X on a finite set E, a function $v\colon E \to \mathbb{R}$ and now we assume given an independent Brownian motion B, using which we form the (continuous fluctuating) additive function φ defined by

$$\varphi_t \equiv \varepsilon B_t + \int_0^t v(X_s)\, ds.$$

Exactly as before, we define the time changes $\tau_t^\pm = \inf\{u\colon \pm\varphi_u > t\}$ and the processes $Y_t^\pm \equiv X(\tau_t^\pm)$. The processes $Y^\pm$ are Markov chains, but straightaway notice that $Y^\pm$ take values in the whole of E, in contrast to the earlier construction. We shall write $\Gamma_\pm$ for the generators of $Y^\pm$, which are characterized by the following result.

THEOREM. *The generator Γ_+ (respectively, Γ_-) is the unique $Z \in \mathscr{Q}(E)$ such that*

(7.1)(i) $$\tfrac{1}{2}\varepsilon^2 Z^2 - VZ + Q = 0$$

(*respectively,*

(7.1)(ii) $$\tfrac{1}{2}\varepsilon^2 Z^2 + VZ + Q = 0).$$

REMARKS. This result is explicit in [2] and implicit in [8]. We give here a short proof.

PROOF OF THEOREM 6. Fix some positive a and some h: $E \to \mathbb{R}$, and define

(7.2) $$f(j,x) = E\left[h(Y_a^+) : \tau_a^+ < \infty | X_0 = j, \varphi_0 = x\right], \qquad j \in E,\ x < a.$$

Then $f(X_t, \varphi_t)$ is a martingale until τ_a^+; so, by Itô's formula,

(7.3) $$\tfrac{1}{2}\varepsilon^2 f'' + Vf' + Qf = 0.$$

However, on the other hand,

$$f(\cdot, x) = \exp[(a-x)\Gamma_+]h;$$

so substituting into (7.3) yields

$$\left(\tfrac{1}{2}\varepsilon^2\Gamma_+^2 - V\Gamma_+ + Q\right)f = 0.$$

Since h is arbitrary, we conclude that Γ_+ solves (7.1)(i). As for uniqueness, from any solution Z to (7.1)(i) we could build $\tilde{f}$ by

$$\tilde{f}(\cdot, x) = \exp[(a-x)Z]h.$$

Then $\tilde{f}(X_t, \phi_t)$ would be a bounded martingale on $[0, \tau_a^+]$, whence $\tilde{f} = f$ defined by (7.2). □

The aim now is to present briefly the analogues of the invariant law results of Sections 3 and 4. The methods are essentially the same, so we shall comment little.

We shall consider the buffer with capacity a, $0 < a \le \infty$, so that the generator of (X, ξ), where ξ is still the content process, is

(7.4) $$\mathscr{G}f = \tfrac{1}{2}\varepsilon^2 f'' + Vf' + Qf,$$

with boundary conditions

(7.5) $$f'(j,x) = 0 \quad \text{for } j \in E,\ x = 0, a.$$

The adjoint equation determining the invariant law is now quite a lot simpler, since there is no time spent in the empty or full states. We obtain

$$0 = \int \pi(x)\mathscr{G}f(x)\,dx \quad \text{for all } f \in \mathscr{D}(\mathscr{G}),$$

and integrating this by parts gives us

(7.6)(i) $$\tfrac{1}{2}\varepsilon^2\pi'' - \pi'V + \pi Q = 0 \qquad \text{in } (0,a)$$

(7.6)(ii) $$\tfrac{1}{2}\varepsilon^2\pi' = \pi V \quad \text{at } 0, a.$$

If we recall the reversed process introduced in (3.3), then (7.6)(i) can be reexpressed as

$$\tfrac{1}{2}\varepsilon^2\pi'' + \pi'\hat{V} + \pi M^{-1}\hat{Q}^T M = 0.$$

So if we define $f \equiv (\pi M^{-1})^T$, this is again

(7.7)(i) $$\tfrac{1}{2}\varepsilon^2 f'' + \hat{V}f' + \hat{Q}f = 0$$

and the boundary condition transforms to

(7.7)(ii) $$\tfrac{1}{2}\varepsilon^2 f' = -\hat{V}f \quad \text{at } 0, a.$$

CASE 1 (Infinite buffer). We now assume that $m(E_-) > m(E_+)$, without which there is no invariant law. Then $\hat{\Gamma}_-$ is transient and $\hat{\Gamma}_+$ is recurrent. The general solution to (7.7)(i) is of the form

$$f(x) = \exp(-x\hat{\Gamma}_+)g + \exp(x\hat{\Gamma}_-)h; \tag{7.8}$$

to give a solution which is integrable, we set $g = 0$ and now pick h so that the boundary condition is satisfied and the normalization condition holds. It is easy to calculate that

$$f(x) = -\exp(x\hat{\Gamma}_-)\hat{\Gamma}_-\mathbf{1}$$

and, in particular, that

$$\pi(x)^T = -M\exp(x\hat{\Gamma}_-)\hat{\Gamma}_-\mathbf{1}. \tag{7.9}$$

This can also be expressed as

$$P(X=j,\ \xi > x) = m_j\big(\exp(x\hat{\Gamma}_-)\mathbf{1}\big)_j, \tag{7.8}$$

which is analogous to (3.4) and admits the same very simple proof.

This result is also in [2].

CASE 2 (Finite buffer). Let us again assume the buffer has capacity $a \in (0,\infty)$, and first we take $m(E_-) > m(E_+)$.

Looking for a solution of the form

$$f(x) \equiv (\pi(x)M^{-1})^T = \exp((a-x)\hat{\Gamma}_+)g + \exp(x\hat{\Gamma}_-)h, \tag{7.9}$$

the boundary condition tells us

$$\left(\tfrac{1}{2}\varepsilon^2\hat{\Gamma}_+ - \hat{V}\right)\exp(a\hat{\Gamma}_+)g = \left(\tfrac{1}{2}\varepsilon^2\hat{\Gamma}_- + \hat{V}\right)h, \tag{7.10)(i}$$

$$\left(\tfrac{1}{2}\varepsilon^2\hat{\Gamma}_+ - \hat{V}\right)g = \left(\tfrac{1}{2}\varepsilon^2\hat{\Gamma}_- + \hat{V}\right)\exp(a\hat{\Gamma}_-)h. \tag{7.10)(ii}$$

So if we write $h = \hat{\Gamma}_- u$ and $g = \hat{\Gamma}_+ w + \alpha\mathbf{1}$, as we may, we conclude that

$$-Q\exp(a\hat{\Gamma}_+)w - \alpha\hat{V}\mathbf{1} = -Qu, \tag{7.11)(i}$$

$$-Qw - \alpha\hat{V}\mathbf{1} = Q\exp(a\hat{\Gamma}_-)u. \tag{7.11)(ii}$$

Left-multiplying by the invariant measure m of Q tells us that $\alpha = 0$, and

$$Q\big(u - \exp(a\hat{\Gamma}_+)w\big) = 0,$$

$$Q\big(w - \exp(a\hat{\Gamma}_-)u\big) = 0.$$

This implies that, for some α and β,

$$u - \exp(a\hat{\Gamma}_+)w = \alpha\mathbf{1}, \qquad w - \exp(a\hat{\Gamma}_-)u = \beta\mathbf{1}.$$

However, observe that we could add any multiple of $\mathbf{1}$ to w without changing g, so we assume without loss of generality that $\beta = 0$, and thus

$$u = \alpha\big(I - \exp(a\hat{\Gamma}_+)\exp(a\hat{\Gamma}_-)\big)^{-1}\mathbf{1}, \tag{7.12)(i}$$

$$w = \exp(a\hat{\Gamma}_-)u, \tag{7.12)(ii}$$

for some constant α fixed by the normalization requirement. It is not hard to confirm that $\alpha = 1$ for normalization, and hence

$$(\pi(x)M^{-1})^T = \Big\{\exp\big[(a-x)\hat{\Gamma}_+\big]\hat{\Gamma}_+\exp\big(a\hat{\Gamma}_-\big) + \exp\big(x\hat{\Gamma}_-\big)\hat{\Gamma}_-\Big\} \times\Big(I - \exp\big(a\hat{\Gamma}_+\big)\exp\big(a\hat{\Gamma}_-\big)\Big)^{-1}\mathbf{1}. \tag{7.13}$$

Once again, the balanced case $m(E_+) = m(E_-)$ requires a more delicate treatment and ends with a less satisfying result. The general form (7.9) of the solution and the boundary conditions (7.10) remain the same, but now the generic form of h and g must be

$$g = \hat{\Gamma}_+ w + \alpha\mathbf{1}, \qquad h = \hat{\Gamma}_- u + \beta\mathbf{1},$$

although we could add $c\mathbf{1}$ to g, and subtract $c\mathbf{1}$ from h without altering solution (7.9); so we lose no generality in assuming that $\beta = 0$, so that (7.11) still holds. Since $m\hat{V}\mathbf{1} = 0$, we can no longer left-multiply by m and deduce that $\alpha = 0$, but we *can* represent

$$\hat{V}\mathbf{1} = Qz,$$

for some z, and thus we obtain

$$Q\Big(u - \exp\big(a\hat{\Gamma}_+\big)w - \alpha z\Big) = 0,$$

$$Q\Big(\exp\big(a\hat{\Gamma}_-\big)u - w - \alpha z\Big) = 0.$$

This implies that, for some b and c,

$$u - \exp\big(a\hat{\Gamma}_+\big)w - \alpha z = b\mathbf{1}, \tag{7.14)(i}$$

$$\exp\big(a\hat{\Gamma}_-\big)u - w - \alpha z = c\mathbf{1}, \tag{7.14)(ii}$$

a pair of equations with a certain amount of redundancy; any multiple of $\mathbf{1}$ could be added to u, w or z without affecting anything. So let us assume that $\nu_+ w = 0 = \nu_- u = mz$, where $\nu_\pm$ are the invariant laws of $\Gamma_\pm$, and we write u', w', z' for the vectors u, w, z with the bottom row omitted. Thus

$$u = J_- u', \qquad w = J_+ w', \qquad z = Jz',$$

for certain matrices J and $J_\pm$ which are easy to specify explicitly. If we write R for the $(n-1)\times n$ matrix $(I_{n-1}\ \mathbf{0})$, $n = |E|$, then (7.14)(i) and (ii) become

$$u' - R\exp\big(a\hat{\Gamma}_+\big)J_+ w' - \alpha z' = b\mathbf{1},$$

$$R\exp\big(a\hat{\Gamma}_-\big)J_- u' - w' - \alpha z' = c\mathbf{1},$$

so that

$$\Big(I - R\exp\big(a\hat{\Gamma}_+\big)\exp\big(a\hat{\Gamma}_-\big)J_-\Big)u' = b\mathbf{1} - cR\exp\big(a\hat{\Gamma}_+\big)J_+\mathbf{1} + \alpha\Big(I - R\exp\big(a\hat{\Gamma}_+\big)J\Big)z. \tag{7.15}$$

Now the matrix $I - R\exp(a\hat{\Gamma}_+)\exp(a\hat{\Gamma}_-)J_-$ is invertible, because if x were annihilated by it, we would have

$$x = R\exp\big(a\hat{\Gamma}_+\big)\exp\big(a\hat{\Gamma}_-\big)J_- x = R\exp\big(na\hat{\Gamma}_+\big)\exp\big(na\hat{\Gamma}_-\big)J_- x,$$

for every n. Letting $n \to \infty$ and using the fact that $\nu_- J_- = 0$ shows that $x = 0$. Thus (7.15) actually does determine u' (in terms of the parameters b, c and α); similarly for w'. The parameters b and c can now be determined (in terms of α) from (7.14)(i) and (ii) and the conditions that $v_+ w = v_- u = 0$. Finally, α can be fixed by the normalization condition on the solution.

APPENDIX

PROOF OF THEOREM 3. First, we remark that the process φ contains an embedded random walk, by looking at φ at successive return times to some distinguished state i_0; from this, it is not hard to see that

$$\text{(A.1)} \qquad \text{either} \quad P\Big(\sup_t \varphi_t = +\infty\Big) = 1 \quad \text{or} \quad P(\varphi_t \to -\infty) = 1.$$

(i) If we return to the uniqueness argument of Theorem 1 and take an eigenvector f of Q_+, $Q_+ f = \lambda f$, we may apply the optional sampling theorem to the martingale $\bar{f}(X_t)e^{-\lambda t}$. We shall find that, by stopping at $n \wedge \tau_0^+$,

$$\text{(A.2)} \qquad \begin{aligned} \bar{f}(i) &= E^i\big[\bar{f}(X(\tau_0^+)) : \tau_0^+ < n\big] \\ &\quad + E^i\big[\exp(-\lambda\varphi_n)\bar{f}(X_n) : n \le \tau_0^+\big], \qquad i \in E_-. \end{aligned}$$

Now the last term tends to zero as $n \to \infty$; indeed, the random variable $\exp(-\lambda\varphi_n)\bar{f}(X_n)$ is bounded on the events $\{n \le \tau_0^+\}$, and, by (A.1), *either* τ_0^+ is finite a.s. *or else* $\varphi_t \to -\infty$. In the first case, the events $\{n \le \tau_0^+\}$ decrease to the null event $\{\tau_0^+ = +\infty\}$, in the second, the random variables $\exp(-\lambda\varphi_n)\bar{f}(X_n)$ tend to zero. This uses, of course, the fact that $\mathrm{Re}(\lambda) < 0$, a consequence of the transience of Q_+. The argument that $Z_+ = \Pi_+$ is now exactly as in Theorem 1.

(ii) First, note that Q_+ is irreducible because Q is. Also, it is easy to seee that there can be no vector f such that

$$Q_+^2 f = 0 \neq Q_+ f$$

(because this would imply that $Q_+ f = c\mathbf{1}$ and, for some finite Markov chain ζ, $f(\zeta_t) - ct$ is a martingale). Thus Q_+ has 0 as a simple eigenvalue, and its eigenspaces of eigenvalues with negative real part span the remaining $n - 1$ dimensions. So let us assume that

$$Q_+ f_j = \lambda_j f_j, \qquad j = 1, \dots, n,$$

and $\{f_1, \dots, f_n\}$ forms a basis for $\mathbb{R}^n$ (if this assumption fails, we can always take a basis of Jordan vectors, as explained in Theorem 1; the proof is made a little clearer by this simplifying assumption). Now, exactly as in (A.2), for each $j = 1, \dots, n$,

$$\text{(A.3)} \qquad \bar{f}_j(i) = (\Pi_+ f_j)(i), \qquad i \in E_-,$$

since the assumption that G_+ is recurrent implies that $\tau_0^+ < \infty$ a.s.

(iii) The analysis of this final case proceeds as for case (ii) as far as (A.3), *which now can only be guaranteed to hold for* $j = 2, \dots, n$.

What we know therefore is that

$$Z_+ f_j = \Pi_+ f_j \quad \text{for } j = 2, \dots, n$$

and hence

$$Z_+ Q_+ = \Pi_+ Q_+.$$

This implies that

$$Z_+ - \Pi_+ = w\mu,$$

where μ is the (everywhere positive) invariant distribution of Q_+, and w is some as yet unknown vector. Thus

$$\begin{aligned} V^{-1}Q\begin{pmatrix} I \\ Z_+ \end{pmatrix} &= V^{-1}Q\begin{pmatrix} I \\ \Pi_+ \end{pmatrix} + V^{-1}Q\begin{pmatrix} 0 \\ w \end{pmatrix}\mu \\ &= \begin{pmatrix} I \\ \Pi_+ \end{pmatrix} G_+ + V^{-1}Q\begin{pmatrix} 0 \\ w \end{pmatrix}\mu \\ &= \begin{pmatrix} I \\ Z_+ \end{pmatrix} Q_+ \\ &= \begin{pmatrix} I \\ \Pi_+ \end{pmatrix} Q_+. \end{aligned}$$

Thus

$$V^{-1}Q\begin{pmatrix} 0 \\ w \end{pmatrix}\mu = \begin{pmatrix} I \\ \Pi_+ \end{pmatrix}(Q_+ - G_+),$$

which says that

(A.4)(i) $$Bw\mu = Q_+ - G_+,$$

(A.4)(ii) $$-Dw\mu = \Pi_+(Q_+ - G_+).$$

So right-multiplying (A.4)(i) by $\mathbf{1}$ tells us that $Bw\mathbf{1} = -G_+\mathbf{1}$, and so

(A.5) $$Q_+ = G_+ - G_+\mathbf{1}\mu.$$

Left-multiplying (A.5) by μ shows that μ is a left eigenvector of G_+ with eigenvalue $\alpha \equiv \mu G_+\mathbf{1}$. Because μ is a probability distribution and is everywhere positive, we may define

$$\tilde{G}_+(i,j) \equiv \mu_j G_+(j,i)/\mu_i - \alpha\delta_{ij}$$

and observe that $\tilde{G}_+$ is an irreducible recurrent Q-matrix, all of whose eigenvalues apart from 0 therefore have a strictly negative real part. Since $\mathrm{sp}(\tilde{G}_+) = \mathrm{sp}(G_+) - \alpha$, everything is now proved. □

Acknowledgments. I should like to thank Soren Asmussen, through whom I first became aware of the interest in fluid models, for stimulating discussion and George O'Brien for inviting me to speak at the 21st Conference on Stochastic Processes and their Applications, Toronto, June 1992, which gave me the incentive to put this material together.

REFERENCES

[1] ANICK, D., MITRA, D. and SONDHI, M. M. (1982). Stochastic theory of a data-handling system with multiple sources. *Bell System Tech. J.* **61** 1871–1894.

[2] ASMUSSEN, S. (1992). Markov-modulated reflected Brownian motion and fluid flow. Unpublished manuscript.

[3] BARLOW, M. T., ROGERS, L. C. G. and WILLIAMS, D. (1980). Wiener–Hopf factorization for matrices. *Séminaire de Probabilités XIV. Lecture Notes in Math.* **784** 324–331. Springer, Berlin.

[4] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterisation and Convergence.* Wiley, New York.

[5] GAVER, D. P. and LEHOCZKY, J. P. (1982). Channels that cooperatively service a data stream and voice messages. *IEEE Trans. Comm.* **30** 1153–1161.

[6] HALFIN, S. (1972). Steady-state distribution for the buffer content of an M/G/1 queue with a varying service rate. *SIAM J. App. Math.* **23** 356–363.

[7] HALFIN, S. and SEGAL, M. (1972). A priority queueing model for a mixture of two types of customers. *SIAM J. Appl. Math.* **23** 369–379.

[8] KENNEDY, J. E. and WILLIAMS, D. (1990). Probabilistic factorization of a quadratic matrix polynomial. *Math. Proc. Cambridge Philos. Soc.* **107** 591–600.

[9] KOSTEN, L. (1974). Stochastic theory of a multi-entry buffer. Delft Progress Report. 10–18, 44–50, 103–115.

[10] LEHOCZKY, J. P. and GAVER, D. P. (1982). Performance evaluation of voice/data queueing systems. *Applied Probability—Computer Science: The Interface* (R. L. Disney and T. J. Ott) **1** 329–346.

[11] LINDVALL, T. (1992). *Lectures on the Coupling Method.* Wiley, New York.

[12] LONDON, R. R., MCKEAN, H. P., ROGERS, L. C. G. and WILLIAMS, D. (1982). A martingale approach to some Wiener–Hopf problems I. *Séminaire de Probabilités XVI. Lecture Notes in Math.* **920** 41–67. Springer, Berlin.

[13] LONDON, R. R., MCKEAN, H. P., ROGERS, L. C. G. and WILLIAMS, D. (1982). A martingale approach to some Wiener–Hopf problems II. *Séminaire de Probabilités XVI. Lecture Notes in Math.* **920** 68–90. Springer, Berlin.

[14] LOYNES, R. M. (1962). A continuous-time treatment of certain queues and infinite dams. *J. Austral. Math. Soc.* **2** 484–498.

[15] MCGILL, P. (1989). Some eigenvalue identities for Brownian motion. *Math. Proc. Cambridge Philos. Soc.* **105** 587–596.

[16] MITRA, D. (1988). Stochastic theory of a fluid model of producers and customers coupled by a buffer. *Adv. in Appl. Probab.* **20** 646–676.

[17] NEUTS, M. F. (1981). *Matrix-Geometric Solutions in Stochastic Models.* Johns Hopkins Univ. Press.

[18] NEUTS, M. F. (1989). *Structured Stochastic Matrices of the M/G/1 Type and Their Applications.* Dekker, New York.

[19] ROGERS, L. C. G. (1983). Wiener–Hopf factorization of diffusions and Lévy processes. *Proc. London Math. Soc.* **47** 177–191.

[20] ROGERS, L. C. G. and SHI, Z. (1992). Computing the invariant law of a fluid model. *J. Appl. Probab.* To appear.

[21] SEVASTYANOV, B. A. (1962). Influence of storage bin capacity on the average standstill time of a production line. *Theory Probab. Appl.* **7** 429–438.

[22] STERN, T. E. and ELWALID, A. I. (1991). Analysis of separable Markov-modulated rate models for information-handling systems. *Adv. Appl. Probab.* **23** 105–139.

[23] WILLIAMS, D. (1982). A potential-theoretic note on the quadratic Wiener–Hopf equation for Q-matrices. *Séminaire de Probabilités XVI. Lecture Notes in Math.* **920** 91–94. Springer, Berlin.

SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF LONDON
QUEEN MARY AND WESTFIELD COLLEGE
MILE END ROAD
LONDON E1 4NS
ENGLAND