Analysis of the shortest queue problem

I.J.B.F. Adan, J. Wessels and W.H.M. Zijm

Eindhoven, May 1989
The Netherlands
ANALYSIS OF THE SHORTEST QUEUE PROBLEM

I.J.B.F. Adan *
J. Wessels *
W.H.M. Zijm **

University of Technology, Eindhoven

Abstract. In this paper we study a system consisting of two identical servers, each with exponentially distributed service times. Jobs arrive according to a Poisson stream. On arrival a job joins the shortest queue and in case both queues have equal length, he joins either queue with probability \( \frac{1}{2} \). We show that the stationary queue length distribution can be represented by an infinite sum of product form solutions, which satisfy nice recurrence relations. Due to the recurrence relations, the successive terms of the infinite sum are easily calculated. Moreover, the convergence of the infinite sum is exponentially fast and we provide bounds for the error of each partial sum. Based on these properties, a numerically highly attractive algorithm is obtained.

Keywords: difference equation, product form, queues in parallel, stationary queue length distribution.

Introduction

Consider a system consisting of two identical servers, each with exponentially distributed service times. Jobs arrive according to a Poisson stream. On arrival a job joins the shortest queue and in case both queues have equal length, he joins either queue with probability \( \frac{1}{2} \). This problem is known as the symmetric shortest queue problem and has been addressed by many authors. Haight [10] originally introduced the problem. Kingman [13] and Flatto and McKean [6] treated the problem by using a generating function analysis. They showed that the

* Eindhoven University of Technology, Department of Mathematics and Computer Science, P.O. Box 513, 5600 MB - Eindhoven, The Netherlands.

** Nederlandse Philips Bedrijven B.V., Centre for Quantitative Methods, Building HCM-721, P.O. Box 218, 5600 MD - Eindhoven, The Netherlands.

and

Eindhoven University of Technology, Department of Mathematics and Computer Science, P.O. Box 513, 5600 MB - Eindhoven, The Netherlands.
generating function for the equilibrium distribution of the lengths of the two queues is a meromorphic function. Then by the decomposition of the generating function into partial fractions, it follows that the equilibrium probabilities can be represented by an infinite sum of product form solutions. However, the decomposition leads to cumbersome formulas for the equilibrium probabilities. Fayolle and Iasnogorodski [5] proposed a method to determine the generating function for the equilibrium distribution, which applies to very general two dimensional random walks.

The method presented in this paper, is not based on a generating function analysis. Instead we treat the detailed balance equations as difference equations and analyze the structure of the solution directly. The main result is that the equilibrium probabilities can be represented by an infinite sum of product form solutions, which was, as we already mentioned, also recognized by Kingman and Flatto and McKean. Moreover, we provide nice recurrence relations for the terms in the infinite sum and due to these recurrence relations, the successive terms are easily calculated. The speed of convergence of the infinite sum is exponentially fast and we provide bounds for the error of each partial sum. Based on these properties, a numerically highly attractive algorithm is obtained.

The shortest queue problem has been extensively studied from a numerical point of view, see e.g. Gertsbakh [8], Grassmann [9] and Rao and Posner [14]. Conolly [4] treated the shortest queue problem with a finite buffer capacity and showed that this system can be solved efficiently. Using linear programming, Halfin [11] obtained bounds for the equilibrium distribution of the lengths of the two queues. These studies are all restricted to systems with two parallel queues. Hooghiemstra, Keane and Van de Ree [12] developed a power series method to calculate the stationary queue length distribution for fairly general multidimensional exponential queueing systems. Their method is not restricted to systems with two queues, but applies equally well to systems with more than two queues. So far as the shortest queue problem is concerned, Blanc [1,2] reported that the power series method works numerically satisfactory for the shortest queue system with 2 up to 25 parallel queues. However, a common disadvantage of these numerical methods is that in general no bounds can be given for the error of the numerical results.

The paper is organized as follows. In section 1 we will present the equilibrium equations. Then, in the next section, we will derive the main result, which states that the equilibrium probabilities can be represented by an infinite sum of product form solutions. In the following three sections we complete the proof of the main result. In section 6 we derive an explicit form for the normalizing constant. A summary of the results and some conclusions can be found in section 7. In the final two sections we will discuss some implications of the product form representation. In section 8 we show that the product from representation yields a complete asymptotic expansion for the equilibrium distribution, and in the final section we discuss the
numerical benefits of the product form representation and present some numerical results.

1. Equilibrium Equations

For simplicity of notation we suppose that the exponential servers have service times with unit mean and the Poisson arrival process has a rate $2\rho$ with $0 < \rho < 1$. The parallel queue system can be represented by a continuous time Markov process, whose state space consists of the pairs $(m, n)$, $m, n = 0, 1, \ldots$ where $m$ and $n$ are the lengths of the two queues. The transition rates in the upper triangle $n \geq m$ are illustrated in figure 1a, the rates in the lower triangle $n \leq m$ follow by reflection in the diagonal.

![Figure 1a: m-n transition rate diagram](image)

Let $\{p_{m,n}\}$ denote the equilibrium distribution of the lengths of the two queues. Then by symmetry

$$p_{m,n} = p_{n,m}$$

Therefore we can restrict the analysis to the probabilities $p_{m,n}$ in the triangle $n \geq m$. The equilibrium equations become for all $n > m$

$$p_{m,n} 2(\rho + 1) = p_{m-1,n} 2\rho + p_{m,n+1} + p_{m+1,n} \quad \text{if } m > 0, n > m+1$$

$$p_{m,m+1} 2(\rho + 1) = p_{m-1,m+1} 2\rho + p_{m,m+2} + p_{m+1,m+1} + p_{m,m} \rho \quad \text{if } m > 0, n = m+1$$

$$p_{0,n} (2\rho + 1) = p_{0,n+1} + p_{1,n} \quad \text{if } n > 1$$

$$p_{0,1} (2\rho + 1) = p_{0,2} + p_{1,1} + p_{0,0} \rho$$

and for all $n = m$,

$$p_{m,m} 2(\rho + 1) = p_{m-1,m} 2\rho + p_{m,m+1} + p_{m,m-1} 2\rho + p_{m+1,m} \quad \text{if } m > 0$$

$$p_{0,0} 2\rho = p_{0,1} + p_{1,0}$$
By symmetry, the equations for the diagonal can be written as

\[ p_{m,m} (\rho + 1) = p_{m-1,m} 2\rho + p_{m,m+1} \quad \text{if } m > 0 \]

\[ p_{0,0} \rho = p_{0,1} \]

Using the equations for the diagonal we can eliminate the probabilities \( p_{m,n} \) in the equations for the subdiagonal \( n = m+1 \). Then the following set of equations is obtained for the probabilities \( p_{m,n} \) in the upper triangle \( n > m \),

\[ p_{m,n} 2(\rho + 1) = p_{m-1,n} 2\rho + p_{m,n+1} + p_{m+1,n} \quad \text{if } m > 0, \quad n > m+1 \]

\[ p_{m,m+1} 2(\rho + 1) = p_{m-1,m+1} 2\rho + p_{m,m+2} + \]

\[ + (p_{m,m+1} 2\rho + p_{m+1,m+2}) \frac{1}{\rho + 1} + (p_{m-1,m} 2\rho + p_{m,m+1}) \frac{\rho}{\rho + 1} \]

\[ p_{0,n} (2\rho + 1) = p_{0,n+1} + p_{1,n} \quad \text{if } n > 1 \]

\[ p_{0,1}(2\rho + 1) = p_{0,2} + (p_{0,1} 2\rho + p_{1,2}) \frac{1}{\rho + 1} + p_{0,1} \]

Based on these equations the analysis can be further restricted to the probabilities \( p_{m,n} \) in the upper triangle \( n > m \). The equations for the diagonal can be used as definition for the probabilities \( p_{m,m} \).

We will prove that the equilibrium probabilities \( p_{m,n} \) in the upper triangle can be represented by an infinite sum of product form solutions. That is, for all \( n > m \) the probabilities \( p_{m,n} \) can be written as

\[ p_{m,n} = \sum_{i=0}^{\infty} c_{i} \xi_{i}^{m} \eta_{i}^{n} \]

This form is a decomposition on the horizontal and vertical axis. Since there is a drift along and to the diagonal in this system, we will use a decomposition on the diagonal and the vertical axis in the sequel of this paper. Thus instead of the coordinates \( m \) and \( n \), we will work with the coordinates \( m \) and \( r \), where \( r \) denotes the difference between the number of jobs in both queues, that is, \( r = n - m \). Then the upper triangle \( n \geq m \) in the \( m-n \) plane is transformed into the first quadrant in the \( m-r \) plane. In figure 1b we displayed the transition rate diagram for the new coordinates. Further, define \( q_{m,r} \) as the equilibrium probability that there are \( m \) jobs in the smaller queue and \( r \) jobs more in the larger one, then for all \( m \geq 0 \) and \( r \geq 0 \)

\[ q_{m,r} = p_{m,m+r} \]

For convenience, we will give the set of equilibrium equations in the new coordinates \( m \) and \( r \). For all \( m \geq 0 \) and \( r \geq 1 \),

\[ q_{m,r} 2(\rho + 1) = q_{m-1,r+1} 2\rho + q_{m,r+1} + q_{m+1,r-1} \quad \text{if } m > 0, \quad r > 1 \]
early the forms for $p_m$ and $q_m$ are equivalent, with $a_i = \zeta_i$ and $b_i = \eta_i$. Throughout the analysis we will use the trivial, but vital property that the equations on which the analysis is based, are linear, i.e. if two functions satisfy an equation, then any linear combination also satisfies the equation.

2. Derivation of the main result

The objective in this section is to study the structure of the equilibrium probabilities. Particularly we will investigate whether the equilibrium probabilities have some kind of separable structure. Obviously, the equations (1)-(4) do not allow a separable solution of the form $q_{m,r} = \alpha^m \beta^r$. However, numerical experiments indicate that there exist $\alpha$ and $\beta$ such that

$$q_{m,r} \sim K \alpha^m \beta^r \quad \text{as } m \to \infty \text{ and } r \geq 1,$$

for some $K$. This is illustrated in figure 2 for the special case $\rho = 0.5$. In figure 2a we displayed the ratio of $q_{m,r}$ in the $m$ direction, which yields, at least for large $m$, the parameter $\alpha$. In figure 2b we displayed the ratio of $q_{m,r}$ in the $r$ direction, which yields the parameter $\beta$. The probabilities $q_{m,r}$ were computed by solving a finite capacity shortest queue system exactly, i.e. by means of a Markov chain analysis. In the example we computed the equilibrium distribution for a system where each queue has a maximal capacity of 15 jobs, which approximates well the infinite capacity system in case $\rho = 0.5$. 

\[
q_{m,1} \frac{2(\rho + 1)}{(m+1)2\rho + q_{m,2}} + \frac{1}{2} (q_{m,1} 2\rho + q_{m+1,1}) - \frac{\rho}{\rho + 1} \quad \text{if } m > 0, r = 1 \\
q_{0,r} (2\rho + 1) = q_{0,r+1} + q_{1,r-1} \quad \text{if } r > 0 \quad (3)
\]

\[
q_{0,1} (2\rho + 1) = q_{0,2} + \frac{1}{\rho + 1} + q_{0,1} \quad (4)
\]

and for all $m \geq 0$ and $r = 0$,

\[
q_{m,0} (2\rho + 1) = q_{m-1,2\rho + q_{m,1}} \quad \text{if } m > 0 \quad (5)
\]

\[
q_{0,0} \rho = q_{0,1} \quad (6)
\]
Figure 2a: the ratios $q_{m+1,r}/q_{m,r}$

Figure 2b: the ratios $q_{m,r+1}/q_{m,r}$

Clearly, as $p = 0.5$ we have $q_{m,r} \sim K 0.25^m 0.1^r$ for some $K$, which holds even for moderate $m$. The question is, what are in general the parameters $\alpha$ and $\beta$? Intuitively, $\alpha$ stands for the ratio of the probability that there are $m+2$ and $m$ jobs in the system. So a reasonable choice seems $\alpha = p^2$, which is supported by the numerical example. The parameter $\beta$ follows by observing that the form $\alpha^m \beta^r$ has to satisfy equation (1) in the interior of the set $\{(m,r), m \geq 0, r \geq 1\}$. Inserting this form in (1) and dividing both sides by the common term $\alpha^{m-1} \beta^{r-1}$ we get a quadratic form for the unknown $\beta$. This is stated in the following lemma.

Lemma 1

The form $\alpha^m \beta^r$ is a solution of equation (1) if and only if $\alpha$ and $\beta$ satisfy the quadratic form

$$\alpha \beta 2(p + 1) = \beta^2 2p + \alpha \beta^2 + \alpha^2$$

(8)

Putting $\alpha = p^2$ in (8) we obtain two roots $\beta = p$ and $\beta = p^2/(2 + p)$. The root $\beta = p$ yields the asymptotic solution $q_{m,r} \sim K p^{2m} \rho^r$ for some $K$, which corresponds to the equilibrium distribution of two independent $M|M|1$ queues, each with a workload $p$. It is very unlikely that the equilibrium distribution of the shortest queue problem behaves asymptotically like this distribution. Therefore the only reasonable choice is $\beta = p^2/(2 + p)$, which is also supported by the numerical example. Hence we empirically found that

$$q_{m,r} \sim K p^{2m} \left(\frac{\rho^2}{2 + \rho}\right)^r$$

as $m \to \infty$ and $r \geq 1$.

(9)

for some $K$. Actually, Kingman ([13], Theorem 5) and Flatto and McKean ([6], section 3) provided a rigorous proof for this asymptotic result.
Let $\alpha_0 = \rho^2$ and $\beta_0 = \rho^2/(2 + \rho)$. As is illustrated in figure 2 for the special case $\rho = 0.5$, the asymptotic solution $\alpha_0 \beta_0$ perfectly describes the behaviour of the equilibrium probabilities in the interior of the set $\{(m, r), m \geq 0, r \geq 1\}$ as well as near the boundary $r = 1$, but it does not capture the behaviour near the boundary $m = 0$. One easily verifies that $\alpha_0 \beta_0$ indeed satisfies equation (2) on the boundary $r = 1$ and that it violates equation (3) on the boundary $m = 0$. Obviously we can further improve the asymptotic solution by adding a term to correct the error on the boundary $m = 0$. For large $m$ this correction term should be small compared to the term $\alpha_0 \beta_0$ in order to avoid that it spoils the behaviour for large $m$.

Form the linear combination $\alpha_0 \beta_0 + c_0 \alpha \beta'$. We will try to choose $c_0$, $\alpha$ and $\beta$ such that this linear combination satisfies equation (3) and (1). Inserting it in equation (3) gives for all $r \geq 1$

$$(\beta_0 + c_0 \beta') (2\rho + 1) = (\beta_0 \beta^{r+1} + c_0 \beta^{r+1}) + (\alpha_0 \beta_0^{-1} + c_0 \alpha \beta^{-1}).$$

Since this must hold for all $r \geq 1$, we have to put $\beta = \beta_0$. Further we want $\alpha \beta_0$ to satisfy equation (1) in the interior of the set $\{(m, r), m \geq 0, r \geq 1\}$. By virtue of lemma 1 there are two $\alpha$'s such that $\alpha \beta_0$ satisfies equation (1), namely $\alpha_0 = \rho^2$ and $\alpha_1 = 2\rho^2/(2 + \rho)^2$. So we have to put $\alpha = \alpha_1$. Then for any $c_0$, the linear combination $\alpha_0 \beta_0 + c_0 \alpha \beta_0$ satisfies equation (1), because equation (1) is linear. Finally, dividing the above equation by the common term $\beta_0^{-1}$ gives an equation for the unknown $c_0$. Hence we can choose the coefficient $c_0$ such that the linear combination also satisfies equation (3).

In general, the result of this procedure can be stated as

**Lemma 2**

Let $x_1$ and $x_2$ be the roots of the quadratic form (8) for fixed $\beta$. Then the linear combination $k_1 x_1 \beta' + k_2 x_2 \beta'$ satisfies the equations (1) and (3) if $k_1$ and $k_2$ satisfy

$$k_2 = \frac{x_2 - \beta}{x_1 - \beta} k_1. \quad (10)$$

**Proof**

By virtue of lemma 1 the forms $x_1 \beta'$ and $x_2 \beta'$ both satisfy equation (1). Since equation (1) is linear, any linear combination also satisfies (1). Inserting the linear combination $k_1 x_1 \beta' + k_2 x_2 \beta'$ in (3) and dividing by the common term $\beta'^{-1}$ gives

$$(k_1 + k_2) \beta (2\rho + 1) = (k_1 + k_2) \beta^2 + k_1 x_1 + k_2 x_2,$$

which can be rewritten as

$$k_2 = \frac{\beta (2\rho + 1) - \beta^2 - x_1}{\beta (2\rho + 1) - \beta^2 - x_2} k_1. \quad (11)$$
Comparing figure 2 and 3, we see that this refinement also captures the behaviour of the equilibrium probabilities near the boundary \( m = 0 \). We conclude that

Then \( \alpha_0 \beta_0 + c_0 \alpha_1 \beta_0 \) satisfies the equations (1) and (3). For the special case of \( \rho = 0.5 \), we displayed in the figure below the same ratios as in figure 2 for the asymptotic solution \( q_{m,r} = \alpha_0 \beta_0 + c_0 \alpha_1 \beta_0 \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>( 0 )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>0.19</td>
<td>0.24</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>0.19</td>
<td>0.24</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

**Figure 3a:** the ratios \( q_{m+1,r} / q_{m,r} \)

**Figure 3b:** the ratios \( q_{m+1,r} / q_{m,r} \)

Comparing figure 2 and 3, we see that this refinement also captures the behaviour of the equilibrium probabilities near the boundary \( m = 0 \). We conclude that

\[
q_{m,r} \sim K (\alpha_0 \beta_0 + c_0 \alpha_1 \beta_0) \quad \text{as} \quad m+r \to \infty \quad \text{and} \quad r \geq 1,
\]

for some \( K \). We note that Flatto and McKean ([6], section 3) proved this statement, which is stronger than the asymptotic result (9).

We added an extra term to compensate the error on the boundary \( m = 0 \). On the other hand we introduced a new error on the boundary \( r = 1 \), since the extra term violates equation (2). Because \( \alpha_1 < \alpha_0 \) the term \( \alpha_1 \beta_0 \) is very small compared to \( \alpha_0 \beta_0 \) even for moderate \( m \). Therefore its disturbing effect near the boundary \( r = 1 \) is practically negligible. However we can compensate this second order error on the boundary \( r = 1 \) in the same way as we did on the boundary \( m = 0 \), by again adding a correction term.
Lemma 3

Let $Y_1$ and $Y_2$ be the roots of the quadratic form (8) for fixed $\alpha$. Then the linear combination $k_1 Y_1 + k_2 Y_2$ satisfies the equations (1) and (2) if $k_1$ and $k_2$ satisfy

$$k_2 = -\frac{(\alpha + \rho)/y_2 - (\rho + 1)}{(\alpha + \rho)/y_1 - (\rho + 1)} k_1.$$  

(13)

Proof

By virtue of lemma 1 both $\alpha' y_1'$ and $\alpha' y_2'$ satisfy (1) and by linearity, also any linear combination. Inserting the linear combination $k_1 \alpha' y_1' + k_2 \alpha' y_2'$ in (2) and dividing both sides by the common term $\alpha' -1$ yields

$$(k_1 \alpha y_1 + k_2 \alpha y_2) 2(\rho + 1) = (k_1 y_1^2 + k_2 y_2^2) 2\rho + (k_1 \alpha y_1^2 + k_2 \alpha y_2^2) +

+ ((k_1 \alpha y_1 + k_2 \alpha y_2) 2\rho + k_1 \alpha^2 y_1 + k_2 \alpha^2 y_2) \frac{1}{\rho + 1} +

+ ((k_1 y_1 + k_2 y_2) 2\rho + k_1 \alpha y_1 + k_2 \alpha y_2) \frac{\rho}{\rho + 1}$$

By inserting equation (8) this reduces to

$$k_1 \alpha^2 + k_2 \alpha^2 = ((k_1 \alpha y_1 + k_2 \alpha y_2) 2\rho + k_1 \alpha^2 y_1 + k_2 \alpha^2 y_2) \frac{1}{\rho + 1} +$$
Applying lemma 3 with $Y_1 = f_{30}$, $Y_2 = f_{31}$, $a = a_1$, $k_1 = c_0$ and $k_2 = d_1$, yields
\[
\frac{(a + p)}{f_{30}} - \frac{(p + 1)}{2}.
\]
\[
\text{(17)}
\]
\[
\text{(16)}
\]
\[
\text{(15)}
\]
\[
\text{(14)}
\]

Since $y_1$ and $y_2$ are the roots of the quadratic form (8),
\[
y_1 y_2 = \frac{\alpha^2}{\alpha + 2p}.
\]
\[
\text{(15)}
\]

Multiplying the numerator and denominator of (14) by $(p + 1)$ and using (15) to rewrite $y_1$ and $y_2$, yields relation (13).

Applying lemma 3 with $y_1 = \beta_0$, $y_2 = \beta_1$, $\alpha = \alpha_1$, $k_1 = c_0$ and $k_2 = d_1$, yields
\[
d_1 = -\frac{(\alpha_1 + p)\beta_1 - (p + 1)}{(\alpha_1 + p)\beta_0 - (p + 1)} c_0.
\]

Then the linear combination $\alpha_0^\prime \beta_0^\prime + c_0 \alpha_1^\prime \beta_0^\prime + d_1 \alpha_1^\prime \beta_1^\prime$ satisfies both equations (1) and (2). Now we compensated the error on the boundary $r = 1$, but we introduced a new one on the boundary $m = 0$, since the compensating term $\alpha_1^\prime \beta_1^\prime$ violates equation (3). But it is clear how to continue this compensating procedure.

For the initial values $\alpha_0 = \rho^2$ and $\beta_0 = \rho^2 / (2 + \rho)$, generate the sequence
\[
\alpha_0 \rightarrow \beta_0 \rightarrow \alpha_1 \rightarrow \beta_1 \rightarrow \alpha_2 \rightarrow \beta_2 \rightarrow \ldots
\]
such that $\alpha_i$ and $\alpha_{i+1}$ are the roots of
\[
\alpha_i \beta_i (2(p + 1) = \beta_i^2 2p + \alpha \beta_i^2 + \alpha^2
\]
and $\beta_i$ and $\beta_{i+1}$ are the roots of
\[
\alpha_{i+1} \beta_{i+1} (2(p + 1) = \beta_{i+1}^2 2p + \alpha_{i+1} \beta_{i+1}^2 + \alpha_{i+1}^2
\]
In the appendix we show that the recursion relations for the numbers $\alpha_i$ and $\beta_i$ can be explicitly solved. A property that we will use later, is

Lemma 4

All the numbers $\alpha_i$ and $\beta_i$ are positive.
We note that the relations (18) and (19) immediately lead to
\[ a_{i+1} = \frac{2p \alpha_i}{\alpha_{i+1}}, \]
with initial values \(x_0 = p^2\) and \(\beta_0 = p^2 / (2 + p)\).

Proof

By induction. For \(\alpha_0\) and \(\beta_0\) it's obvious.
Assume both \(\alpha_i\) and \(\beta_i\) are positive. Since \(\alpha_i\) and \(\alpha_{i+1}\) are the roots of (16),
\[ \alpha_i \alpha_{i+1} = 2p \beta_i^2. \] (18)
Hence \(\alpha_{i+1} > 0\). Similarly, \(\beta_i\) and \(\beta_{i+1}\) are the roots of (17) and \(\alpha_i + 2p > 0\), thus
\[ \beta_i \beta_{i+1} = \frac{\alpha_i^2}{2p + \alpha_{i+1}}. \] (19)
Since both \(\alpha_{i+1} > 0\) and \(\beta_i > 0\), this yields \(\beta_{i+1} > 0\).

\[ \square \]

We note that the relations (18) and (19) immediately lead to

Corollary

For \(i = 0, 1, 2, \ldots\)
\[ a_{i+1} = \frac{2p \beta_i^2}{\alpha_i}, \]
\[ \beta_{i+1} = \frac{\alpha_i^2}{2p + \alpha_{i+1}} \beta_i, \]
with initial values \(\alpha_0 = p^2\) and \(\beta_0 = p^2 / (2 + p)\).

The corollary provides us a simple recursive scheme to produce the numbers \(\alpha_i\) and \(\beta_i\).

By virtue of lemma 1 all the solutions \(\alpha_m \beta_i^m\) and \(\alpha_{i+1} \beta_i^m\) satisfy equation (1) in the interior of the set \(\{(m, r), m \geq 0, r \geq 1\}\). Because equation (1) is linear, any linear combination also satisfies (1). Now form the solution \(x_{m,r}\), for all \(m \geq 0\) and \(r \geq 1\) defined as
\[ x_{m,r} = \sum_{i=0}^{\infty} d_i (\alpha_i^m + c_i \alpha_{i+1}^m) \beta_i^m \]
\[ = d_0 \alpha_0^m \beta_0^m + \sum_{i=0}^{\infty} (d_i c_i \beta_i^m + d_{i+1} \beta_i^m) \alpha_i^m, \] (20)
where in the first sum we formed pairs with a common factor \(\beta_i\) and in the second one with a common factor \(\alpha_{i+1}\). Put \(d_0 = 1\) and successively generate the coefficients \(c_i\) and \(d_{i+1}\) such that \((\alpha_i^m + c_i \alpha_{i+1}^m) \beta_i^m\) satisfies equation (3) on the boundary \(m = 0\) and \((d_i c_i \beta_i^m + d_{i+1} \beta_i^m) \alpha_i^m\) satisfies equation (2) on the boundary \(r = 1\). By virtue of lemma 2 and lemma 3, this yields for all \(i = 0, 1, \ldots\)
\[ c_i = -\frac{\alpha_{i+1} - \beta_i}{\alpha_i - \beta_i}, \] (21)
\[ d_{i+1} = \frac{(\alpha_{i+1} + 1)/\beta_{i+1} - (p + 1)}{(\alpha_{i+1} + 1)/\beta_{i} - (p + 1)} c_i d_i. \]  

(22)

For \( m \geq 0 \) and \( r = 0 \) the numbers \( x_{m,r} \) are not given by the product form representation (20), but defined by the equilibrium equations (5) and (6), yielding

\[ x_{m,0} = \frac{1}{p + 1} (x_{m-1,1} 2p + x_{m,1}), \quad \text{if } m > 0 \]

\[ x_{0,0} = \frac{x_{0,1}}{p}. \]

The following theorem establishes our main result: the solution \( \{x_{m,r}\} \) equals the equilibrium distribution \( \{q_{m,r}\} \) apart from a normalizing constant.

**Theorem**

For all \( m \geq 0 \) and \( r \geq 0 \)

\[ q_{m,r} = C^{-1} x_{m,r}, \]

where \( C \) is a normalizing constant, defined as

\[ C = 2 \sum_{m=0}^{\infty} \sum_{r=1}^{\infty} x_{m,r} + \sum_{m=0}^{\infty} x_{m,0}. \]

In the following sections we will prove the theorem. First we will prove that \( \{x_{m,r}\} \) is well defined, that is, for all \( m \geq 0 \) and \( r \geq 1 \) the infinite sum (20) converges absolutely. Then we prove that \( \{x_{m,r}\} \) is a positive and convergent solution, that is, for all \( m \geq 0 \) and \( r \geq 0 \)

\[ x_{m,r} > 0 \]

and

\[ C = 2 \sum_{m=0}^{\infty} \sum_{r=1}^{\infty} x_{m,r} + \sum_{m=0}^{\infty} x_{m,0} < \infty. \]

The solution \( \{x_{m,r}\} \) formally satisfies the equations (1), (2) and (3), and by definition, the equations (5) and (6). In section 5 we prove that \( \{x_{m,r}\} \) also satisfies the balance equation in \((0, 1)\), where both boundaries meet each other. It will be clear that, by the uniqueness of the equilibrium distribution, this is sufficient to prove the theorem.
3. Asymptotics

In this section we will derive some preliminary results concerning the asymptotic behaviour of the numbers $\alpha_i$, $\beta_i$, $c_i$ and $d_i$. By virtue of lemma 4, we may define

Definition

For $i = 0, 1, 2, \ldots$

$$u_i = \frac{\alpha_i}{\beta_i}, \quad v_i = \frac{\alpha_{i+1}}{\beta_i}.\)$$

We will start to study the behaviour of the ratios $u_i$ and $v_i$, which can be easily analyzed. Then the behaviour of the constants $\alpha_i$, $\beta_i$, $c_i$ and $d_i$ is analyzed by expressing them in terms of the ratios $u_i$ and $v_i$. In the appendix we show that the recursion relations for the numbers $\alpha_i$ and $\beta_i$ can be explicitly solved. Hence the asymptotic behaviour of the numbers $\alpha_i$, $\beta_i$, $c_i$ and $d_i$ can also be obtained from the explicit expressions for $\alpha_i$ and $\beta_i$.

Since $\beta_i$ and $\beta_{i+1}$ are the roots of the quadratic form

$$\alpha_{i+1} \beta 2(p + 1) = \beta^2 2p + \alpha_{i+1} \beta^2 + \alpha_{i+1}^2,$$

they satisfy the relations

$$\beta_i \beta_{i+1} = \frac{\alpha_{i+1}^2}{(2p + \alpha_{i+1})}, \quad \beta_i + \beta_{i+1} = 2(p + 1) \frac{\alpha_{i+1}}{(2p + \alpha_{i+1})}.$$

Combining these relations gives

$$\frac{\alpha_{i+1}}{\beta_i} + \frac{\alpha_{i+1}}{\beta_{i+1}} = 2(p + 1),$$

or in terms of the ratios $u_i$ and $v_i$,

$$v_i + u_{i+1} = 2(p + 1). \quad (23)$$

Further, $\alpha_i$ and $\alpha_{i+1}$ are the roots of the quadratic form

$$\alpha \beta_i 2(p + 1) = \beta_i^2 2p + \alpha \beta_i^2 + \alpha^2,$$

so

$$\alpha_i \alpha_{i+1} = 2p \beta_i^2.$$

or dividing both sides by $\beta_i^2$,

$$u_i v_i = 2p. \quad (24)$$

Relations (23) and (24) immediately give us
Lemma 5

For \( i = 0, 1, 2, \ldots \)

\[
    u_{i+1} = 2(p+1) - \frac{2p}{u_i},
\]

\[
    v_i = 2(p+1) - \frac{2p}{v_{i+1}},
\]

with initial values \( u_0 = 2 + p \) and \( v_0 = 2p/(2 + p) \).

These iteration schemes are graphically illustrated in figure 4. The numbers \( A_1 \) and \( A_2 \) are the roots of the equation \( A = 2(p+1) - 2p/A \), that is, \( A_1 = p + 1 - \sqrt{p^2 + 1} \) and \( A_2 = p + 1 + \sqrt{p^2 + 1} \).

![Figure 4: the iteration schemes for \( u_i \) and \( v_i \)](image)

Then, by induction, lemma 5 yields

Corollary

\[ u_i \uparrow A_2 \quad \text{and} \quad v_i \downarrow A_1 \quad \text{as} \quad i \to \infty. \]

To analyze the behaviour of \( \alpha_i, \beta_i, c_i \) and \( d_i \) we have to express them in terms of \( u_i \) and \( v_i \). This is the contents of
Lemma 6

For all $i = 0, 1, 2, ...$

\[
\frac{\alpha_{i+1}}{\alpha_i} = \frac{v_i}{u_i}, \quad \frac{\beta_{i+1}}{\beta_i} = \frac{v_i}{u_{i+1}}, \quad c_i = \frac{1 - v_i}{u_i - 1}, \quad \frac{d_{i+1}}{d_i} = -\frac{v_i (u_{i+1} - \rho) (u_{i+1} - 1)}{u_{i+1} (v_i - \rho) (v_i - 1)} c_i.
\]

Proof

By the definition of $u_i$ and $v_i$, the first two relations for the ratios $\alpha_{i+1} / \alpha_i$ and $\beta_{i+1} / \beta_i$ are trivial and the third one for $c_i$ follows from relation (21) by dividing the numerator and denominator by $\beta_i$. The derivation of the relation for the ratio $d_{i+1} / d_i$ is a little more complicated. Multiplying the numerator and denominator in (22) by $\alpha_{i+1}$ and using the definition of the ratios $u_i$ and $v_i$, leads to

\[
\frac{d_{i+1}}{d_i} = -\frac{(\alpha_{i+1} + \rho) u_{i+1} - \alpha_{i+1} (p + 1)}{(\alpha_{i+1} + \rho) v_i - \alpha_{i+1} (p + 1)} c_i. \tag{25}
\]

and rewriting relation (19),

\[
\alpha_{i+1} = v_i u_{i+1} - 2\rho. \tag{26}
\]

Inserting (26) in (25) yields

\[
\frac{d_{i+1}}{d_i} = -\frac{(v_i u_{i+1} - \rho) u_{i+1} - (v_i u_{i+1} - 2\rho) (p + 1)}{(v_i u_{i+1} - \rho) v_i - (v_i u_{i+1} - 2\rho) (p + 1)} c_i
\]
\[
= -\frac{(u_{i+1}^2 - u_{i+1} (p + 1)) v_i + \rho (2(p + 1) - u_{i+1})}{(v_i^2 - v_i (p + 1)) u_{i+1} + \rho (2(p + 1) - v_i)} c_i.
\]

Finally, inserting relation (23) gives the expression in the lemma. \qed

Together with the corollary of lemma 5, this yields

Corollary

As $i \to \infty$, then

\[
\frac{\alpha_{i+1}}{\alpha_i} \downarrow \frac{A_1}{A_2}, \quad \frac{\beta_{i+1}}{\beta_i} \downarrow \frac{A_1}{A_2} < 1,
\]

\[
c_i \to \frac{1 - A_1}{A_2 - 1} < 1,
\]

\[
\frac{d_{i+1}}{d_i} \to -\frac{A_2}{A_1} \frac{1 - A_1}{A_2 - 1} < -1.
\]
Proof

The limiting behaviour of the ratios \( \alpha_{i+1} / \alpha_i \) and \( \beta_{i+1} / \beta_i \) and the coefficients \( c_i \) follow immediately from lemma 6 and the corollary of lemma 5, and further

\[
\frac{d_{i+1}}{d_i} \to - \frac{A_1 (A_2 - \rho) (A_2 - 1)}{A_2 (A_1 - \rho) (A_1 - 1)} c_{\infty} \quad \text{as } i \to \infty,
\]

where \( c_{\infty} = \lim_{i \to \infty} c_i \).

Since \( A_1 \) and \( A_2 \) are the roots of the quadratic from \( A^2 = 2(\rho + 1) A - 2\rho \), we have

\[
A_1 + A_2 = 2(\rho + 1), \quad A_1 A_2 = 2\rho.
\]

Hence

\[
\rho (A_1 + A_2) = \rho 2(\rho + 1) = A_1 A_2 (\rho + 1),
\]

and reordering the terms gives

\[
A_1 (A_2 - \rho) = A_2 \rho (1 - A_1), \quad A_2 (A_1 - \rho) = A_1 \rho (1 - A_2).
\]

Inserting in (27) yields

\[
\frac{d_{i+1}}{d_i} \to - \frac{A_2}{A_1} c_{\infty}.
\]

Further, by the corollary of lemma 5, it follows that for all \( i \),

\[
u_i > u_0 \quad \text{and} \quad v_i > v_0 > 0.
\]

In particular,

\[
u_i > 1 \quad \text{and} \quad 1 > \rho > v_i > 0.
\]

As a consequence,

**Corollary**

\[
1 > \alpha_0 > \beta_0 > \alpha_1 > \beta_1 > ... > 0
\]

and for all \( i \),

\[
c_i > 0 \quad \text{and} \quad \frac{d_{i+1}}{d_i} < 0.
\]
Proof

By inequality (28), the definition of \( u_i \) and \( v_i \) and lemma 4, it follows that for all \( i \),
\[
\alpha_i > \beta_i > \alpha_{i+1}. 
\]
Since \( \alpha_0 = \rho^2 < 1 \), this yields the first inequality.
The other ones follow from (28) and lemma 6.

Thus the terms in expression (20) for \( x_{m,r} \) are alternating.

4. On the convergence

Now we are in the position to prove that for all \( m \geq 0 \) and \( r \geq 1 \) the numbers \( x_{m,r} \) are well defined. Consider a fixed \( m \geq 0 \) and \( r \geq 1 \). Then from the first corollary of lemma 6, both
\[
\frac{|d_{i+1} \alpha_m \beta_i|}{|d_i \alpha_m \beta_i|} \quad \text{and} \quad \frac{|d_{i+1} c_{i+1} \alpha_m \beta_{i+1}|}{|d_i \alpha_m \beta_i|} \rightarrow \frac{1 - A_1}{A_2 - 1} \left[ \frac{A_1}{A_2} \right]^{m+r-1} < 1 \quad \text{as} \quad i \rightarrow \infty.
\]
Hence, there exist positive constants \( M \) and \( R \), with \( R \) strictly less than unity, and both depending on \( m \) and \( r \), such that for all \( i \),
\[
|d_i \alpha_m \beta_i| \quad \text{and} \quad |d_i c_i \alpha_{m+1} \beta_i| \leq M R^i.
\]
This proves

Lemma 7

For all \( m \geq 0 \) and \( r \geq 1 \), both
\[
\sum_{i=0}^{\infty} d_i \alpha_m \beta_i \quad \text{and} \quad \sum_{i=0}^{\infty} d_i c_i \alpha_m \beta_i
\]
converge absolutely.

By virtue of this lemma, for all \( m \geq 0 \) and \( r \geq 1 \) the numbers \( x_{m,r} \) are well defined by the infinite sum (20), and it is allowed to change the order of summation, i.e.
\[
\sum_{i=0}^{\infty} d_i (\alpha_m + c_i \alpha_{m+1}) \beta_i = d_0 \alpha_m \beta_0 + \sum_{i=0}^{\infty} (d_i c_i \beta_i + d_{i+1} \beta_{i+1}) \alpha_{m+1}.
\]
As noted at the end of the previous section, the terms in the infinite sum (20) are alternating. So it is not immediately obvious whether the infinite sum is positive or negative. The following lemma will enable us to prove that \( \{x_{m,r}\} \) is a positive solution. It states that the terms in (20) are monotonously decreasing in modulus, at least with an uniform rate \( R = 4/(4 + 2\rho^2 + \rho^2) \).
Lemma 8

For all \( m \geq 0, r \geq 1 \) and \( i \geq 0 \),
\[
| d_{i+1} (\alpha_{i+1}^m + c_{i+1} \alpha_{i+2}^m) \beta_{i+1}^r | < R | d_i (\alpha_i^m + c_i \alpha_{i+1}^m) \beta_i^r |
\]
where \( R = 4 / (4 + 2p + p^2) < 1 \)

Proof

We will first prove the lemma for \( m = 0 \) and \( r = 1 \). Consider the ratio of both terms.

By lemma 6 it follows that
\[
\frac{|d_{i+1} (1 + c_{i+1}) \beta_{i+1}^r|}{|d_i (1 + c_i) \beta_i^r|} = \frac{v_i^2 (u_{i+1} - p) (u_{i+1} - 1)}{u_{i+1}^2 (\rho - v_i) (1 - v_i)} \frac{1 + c_{i+1}}{1 + c_i}.
\]

Further, again by lemma 6,
\[
c_i \frac{1 + c_{i+1}}{1 + c_i} = \frac{1 - v_i}{u_{i+1} - 1} \frac{u_{i+1} - v_{i+1}}{u_i - v_i}.
\]

Inserting in (29) yields
\[
\frac{|d_{i+1} (1 + c_{i+1}) \beta_{i+1}^r|}{|d_i (1 + c_i) \beta_i^r|} < \frac{v_i^2}{(\rho - v_i) (u_i - v_i)} \frac{u_{i+1} - v_{i+1}}{u_i - v_i} \frac{u_{i+1} - 1}{(u_{i+1} - p) (1 - v_i)} = R < 1,
\]
where in the second inequality we used that the numbers \( v_i \) are positive and monotonously decreasing and the numbers \( u_i - v_i \) are positive and monotonously increasing. This proves the lemma for \( m = 0 \) and \( r = 1 \).

Now consider a fixed \( m \geq 0 \) and \( r \geq 1 \).

Since the sequences \( (\alpha_i) \) and \( (\beta_i) \) are monotonously decreasing, it follows that for all \( i \),
\[
| d_{i+1} (\alpha_{i+1}^m + c_{i+1} \alpha_{i+2}^m) \beta_{i+1}^r | < | d_{i+1} (1 + c_{i+1}) \beta_{i+1}^r | \alpha_{i+1}^m \beta_{i+1}^r | < R | d_i (1 + c_i) \beta_i^r | \alpha_i^m \beta_i^r | < R | d_i (\alpha_i^m + c_i \alpha_{i+1}^m) \beta_i^r |.
\]

This completes the proof of the lemma.

This lemma gives us
Proof

The terms in expression (20) are alternating and, by lemma 8, strictly decreasing in modulus. Furthermore, the first term is positive. Hence $x_{m,r}$ is positive for $m \geq 0$ and $r \geq 1$, and by the definition of the numbers $x_{m,0}$, this implies that also $x_{m,0} > 0$ for all $m \geq 0$.

We end this section by proving that the normalizing constant $C$ is finite.

The normalizing constant is defined as

$$C = 2 \sum_{m=0}^{\infty} \sum_{r=1}^{\infty} x_{m,r} + \sum_{m=0}^{\infty} x_{m,0},$$

and by the definition of the numbers $x_{m,0}$, we may write

$$C = 2 \sum_{m=0}^{\infty} \sum_{r=1}^{\infty} x_{m,r} + \frac{1}{\rho + 1} \sum_{m=1}^{\infty} (x_{m-1,1} + 2\rho + x_{m,1}) + \frac{1}{\rho} x_{0,1}.$$  \hspace{1cm} (30)

Hence, for the finiteness of $C$, it suffices to prove

Lemma 9

$$\sum_{m=0}^{\infty} \sum_{r=1}^{\infty} x_{m,r} < \infty.$$

Proof

By (20),

$$\sum_{m=0}^{\infty} \sum_{r=1}^{\infty} x_{m,r} = \sum_{m=0}^{\infty} \sum_{r=1}^{\infty} \sum_{i=0}^{\infty} d_i (\alpha_i^m + c_i \alpha_i^{m+1}) \beta_i^r,$$

and we will show that the latter sum converges absolutely. Changing the order of summation and using the second corollary of lemma 6, yields

$$\sum_{m=0}^{\infty} \sum_{r=1}^{\infty} \sum_{i=0}^{\infty} |d_i (\alpha_i^m + c_i \alpha_i^{m+1}) \beta_i^r| = \sum_{i=0}^{\infty} |d_i| \left( \frac{1}{1 - \alpha_i} + \frac{c_i}{1 - \alpha_{i+1}} \right) \frac{\beta_i}{1 - \beta_i}. \hspace{1cm} (31)$$

By the second corollary of lemma 6 and the convergence of $c_i$, it follows that the numbers

$$\left( \frac{1}{1 - \alpha_i} + \frac{c_i}{1 - \alpha_{i+1}} \right) \frac{1}{1 - \beta_i}$$

Corollary

For all $m \geq 0$ and $r \geq 0$,

$x_{m,r} > 0$. 

are bounded and, by lemma 6, as $i \to \infty$, then the quotient $|d_{i+1} / d_i|$ tends to $(1 - A_1) / (A_2 - 1)$, which is strictly less than unity. Hence, there exist positive constants $M$ and $R$, with $R$ strictly less than unity, such that for all $i$,

$$
|d_i| \left( \frac{1}{1 - \alpha_i} + \frac{c_i}{1 - \alpha_i + 1} \right) \frac{\beta_i}{1 - \beta_i} \leq M R^i.
$$

Thus the sum (31) converges. \hfill \Box

5. Verification in $(0, 1)$

Up to now, we proved that for all $m = 0$ and $r \geq 0$ the numbers $x_{m,r}$ are well defined and strictly positive, and the normalizing constant $C$ is finite. Furthermore, we constructed these numbers such that they satisfy the equilibrium equations (1), (2) and (3), and by definition, the equations (5) and (6). In this section we will complete the proof of the theorem by showing that the numbers $x_{m,r}$ also satisfy the remaining equation (4). We may write equation (4) as

$$
x_{0,2} - x_{0,1} (2 \rho + 1) + (x_{0,1} 2 \rho + x_{1,1}) \frac{1}{\rho + 1} + x_{0,1} = 0,
$$

and we have to verify whether this identity holds.

Inserting expression (20) for the numbers $x_{m,r}$ gives

$$
\sum_{i=0}^{\infty} d_i \left[ (1 + c_i) \beta_i^2 - (1 + c_i) \beta_i (2 \rho + 1) \right] + \sum_{i=0}^{\infty} \frac{\beta_i (2 \rho + \alpha_i)}{\rho + 1} + c_i \frac{\beta_i (2 \rho + \alpha_i+1)}{\rho + 1} + \sum_{i=0}^{\infty} d_i (1 + c_i) \beta_i = 0.
$$

We have chosen the coefficients $c_i$ such that the terms $(\alpha_i^m + c_i \alpha_i^m + 1) \beta_i^2$ satisfy equation (3) on the boundary $m = 0$. Inserting $(\alpha_i^m + c_i \alpha_i^m + 1) \beta_i^2$ in (3) and dividing both sides by $\beta_i^{-1}$, yields

$$
(1 + c_i) \beta_i (2 \rho + 1) = (1 + c_i) \beta_i^2 + (\alpha_i + c_i \alpha_{i+1}).
$$

Hence we can rewrite the first sum in equation (32) as

$$
- \sum_{i=0}^{\infty} d_i (\alpha_i + c_i \alpha_{i+1}).
$$

Using relation (26), the second sum can be written as

$$
\sum_{i=0}^{\infty} d_i \left[ \frac{\nu_{i-1}}{\rho + 1} \alpha_i + c_i \frac{\mu_{i+1}}{\rho + 1} \alpha_{i+1} \right],
$$

where $\nu_{-1} = \rho$, and finally, using that
1 + c_i = \frac{u_i}{u_i - 1} + c_i \frac{v_i}{v_i - 1},

which follows from lemma 6, the third sum can be written as

\[ \sum_{i=0}^{\infty} d_i \left( \frac{\alpha_i}{u_i - 1} + c_i \frac{\alpha_{i+1}}{v_i - 1} \right). \]

Hence, for the left hand side of equation (32) we obtain the form

\[ \sum_{i=0}^{\infty} d_i \left[ \left( -1 + \frac{v_{i-1}}{\rho + 1} + \frac{1}{u_i - 1} \right) \alpha_i + c_i \left( -1 + \frac{u_{i+1}}{\rho + 1} + \frac{1}{v_i - 1} \right) \alpha_{i+1} \right], \]

or taking terms with a common factor \( \alpha_i \),

\[ d_0 \left[ -1 + \frac{v_{i-1}}{\rho + 1} + \frac{1}{u_0 - 1} \right] \alpha_0 + \]

\[ + \sum_{i=0}^{\infty} \left[ d_i c_i \left( -1 + \frac{u_{i+1}}{\rho + 1} + \frac{1}{v_i - 1} \right) + d_{i+1} \left( -1 + \frac{v_i}{\rho + 1} + \frac{1}{u_{i+1} - 1} \right) \right] \alpha_{i+1}. \]

One easily verifies that the first term vanishes and using relation (23) yields

\[ d_i c_i \left[ -1 + \frac{u_{i+1}}{\rho + 1} + \frac{1}{v_i - 1} \right] + d_{i+1} \left[ -1 + \frac{v_i}{\rho + 1} + \frac{1}{u_{i+1} - 1} \right] = \]

\[ = d_i c_i \frac{v_i (u_{i+1} - \rho)}{(p + 1) (v_i - 1)} + d_{i+1} \frac{u_{i+1} (v_i - \rho)}{(p + 1) (u_{i+1} - 1)} \]

\[ = d_i c_i \frac{v_i (u_{i+1} - \rho) (u_{i+1} - 1) + d_{i+1} u_{i+1} (v_i - \rho) (v_i - 1)}{(p + 1) (v_i - 1) (u_{i+1} - 1)} \]

and by virtue of lemma 6, we have chosen the coefficients \( d_i \) such that the numerator vanishes.

6. Explicit form for \( C \)

In this section we derive an explicit form for the normalizing constant \( C \) by using the generating function. The derivation is based on the same ideas used by Kingman [13] and Flatto and McKean [6]. Define

\[ F(x, y) = \sum_{m=0}^{\infty} \sum_{r=0}^{\infty} q_{m,r} x^m y^r. \]

The probabilities \( q_{m,0}, m > 0 \), can be eliminated by inserting (5),

\[ F(x, y) = \sum_{m=0}^{\infty} \sum_{r=1}^{\infty} q_{m,r} x^m y^r + \frac{1}{\rho + 1} \sum_{m=1}^{\infty} (q_{m-1,1} 2\rho + q_{m,1}) + q_{0,0}, \]
and inserting expression (20) leads to

\[
F(x, y) = C^{-1} \left\{ \sum_{i=0}^{\infty} d_i \left[ \frac{1}{1 - \alpha_i x} + \frac{c_i}{1 - \alpha_{i+1} x} \right] \frac{\beta_i y}{1 - \beta_i y} + \frac{1}{\rho + 1} \sum_{i=0}^{\infty} d_i \left[ \frac{(2p + \alpha_i) x}{1 - \alpha_i x} + \frac{(2p + \alpha_i) x}{1 - \alpha_{i+1} x} \right] \beta_i + x_{0,0} \right\}.
\]  

(33)

Hence, \( F(x, y) \) is defined for all \( x \) and \( y \), except for the planes \( x = 1/\alpha_i \) and \( y = 1/\beta_i \), and explicitly determined up to the constant \( C \). The equations (1)-(6) can be reduced to a single equation for \( F(x, y) \). From the equations (1)-(6) it follows that

\[
q_{m,r} \left( 2(p + 1) = q_{m-1,r+1} 2p + q_{m,r+1} + q_{m+1,r-1} \right) \quad \text{if } m > 0, r > 1
\]

\[
q_{0,r} \left( 2(p + 1) = q_{0,r+1} + q_{1,r-1} + q_{0,r} \right) \quad \text{if } m = 0, r > 1
\]

\[
q_{m,1} \left( 2(p + 1) = q_{m-1,2} 2p + q_{m,2} + q_{m+1,0} + q_{m,0} \right) \quad \text{if } m > 0, r = 1
\]

\[
q_{0,1} \left( 2(p + 1) = q_{0,2} + q_{1,0} + q_{0,0} \rho + q_{0,1} \right) \quad \text{if } m = 0, r = 1
\]

\[
q_{m,0} \left( 2(p + 1) = q_{m-1,1} 2p + q_{m,1} + q_{m,0} (2p + 1) \right) \quad \text{if } m > 0, r = 0
\]

\[
q_{0,0} \left( 2(p + 1) = q_{0,0} (2p + 1) + q_{0,0} \right) \quad \text{if } m = 0, r = 0
\]

Multiplying the equations with \( x^{m+1} y^{r+1} \) and summing over all \( m \geq 0 \) and \( r \geq 0 \) leads to

\[
F(x, y) g(x, y) = F(x, 0) h(x, y) + F(0, y) k(x, y),
\]

where

\[
g(x, y) = y^2 + x (2p + 1) - 2p y \quad \text{if } m = 0, r = 1
\]

\[
h(x, y) = x (2p + 1) - (p + 1) x y - \rho x y^2,
\]

\[
k(x, y) = y (y - x).
\]

If \( x \) and \( y \) satisfy the quadratic form \( g(x, y) = 0 \) and \( F(x, y) \) has no singularity in \( (x, y) \), that is, \( x \neq 1/\alpha_i \) and \( y \neq 1/\beta_i \) for all \( i \), then

\[
F(x, 0) h(x, y) + F(0, y) k(x, y) = 0.
\]

(34)

Thus equation (34) relates \( F(x, 0) \) and \( F(0, y) \) whenever \( x \) and \( y \) satisfy \( g(x, y) = 0 \) and \( F(x, y) \) has no singularity in \( (x, y) \). The value of \( F(0, 1) \) can be obtained by a balance argument: the average number of jobs arriving per unit time = the average number of jobs departing per unit time. The average number of jobs arriving per unit time equals the arrival rate \( 2p \). On the other hand, jobs depart at rate \( 2 \), except when at least one of both queues is empty. The departure rate equals \( 1 \) in case one queue is empty, and \( 0 \) in case both queues are empty. Hence, we obtain

\[
2p = 2 - P \text{ (only one queue is empty) } - 2P \text{ (both queues are empty)}
\]
Immediate from (35) and Lemma 10, we have
\[
    \text{Coronary} \quad c = \frac{P (2+p)}{2 (1+p) (1-p) (2-p)}
\]

so
\[
    F(0, 1) = 1 - p.
\]

Applying relation (34) to \( x = 1/2p \) and \( y = 1 \) and then to \( x = 1/2p \) and \( y = 1/p \), yields
\[
    F(0, 1/p) = \frac{h(1/2p, 1/p) k(1/2p, 1)}{k(1/2p, 1/p) k(1/2p, 1)} F(0, 1) = (1 - p) (2 - p).
\]

(35)

Now, the constant \( C \) can be expressed in terms of \( F(0, 1/p) \).

Lemma 10
\[
    C = \frac{\rho (2+p)}{2 (1+p) F(0, 1/p)}
\]

Proof

Since \( F(x, 0) \) has a singularity in \( x = 1/p^2 \), we cannot apply relation (34) to \( x = 1/p^2 \) and \( y = 1/p \). Instead we will take limits \( x \to 1/p^2 \) and \( y \to 1/p \), such that \( x \) and \( y \) satisfy \( g(x, y) = 0 \). For \( x > 1/(p^2 + 1) \), the quadratic form \( g(x, y) = 0 \) has two real roots \( y_- \) and \( y_+ \).

\[
    y_\pm = (p + 1) x \pm \sqrt{x (x - 1)}.
\]

For \( x = 1/p^2 \), it follows that \( y_- = 1/p \). Further, in a neighbourhood of \( x = 1/p^2 \),
\[
    h(x, y_-) = h(1/p^2, 1/p) + \frac{d}{dx} h(x, y_-) \bigg|_{x=1/p^2} (x - 1/p^2) + O\left[ (x - 1/p^2)^2 \right]
\]
\[
    = (2 + p) (p - 1) (x - 1/p^2) / 2p + O\left[ (x - 1/p^2)^2 \right].
\]

(36)

Inserting \( y_- \) in equation (34) and letting \( x \to 1/p^2 \), then by (33) and (36), this yields the lemma.

Immediate from (35) and Lemma 10, we have

Corollary
\[
    C = \frac{\rho (2+p)}{2 (1+p) (1-p) (2-p)}
\]
7. Conclusions

In the previous sections we proved that \( \{x_{m,r}\} \) is a positive and convergent solution of the equilibrium equations. By a result of Foster ([7], Theorem 1), this proves that the shortest queue system is ergodic. Since the equilibrium distribution for an ergodic system is unique, the collection \( \{x_{m,r}\} \) can be normalized to produce the equilibrium distribution \( \{q_{m,r}\} \).

Let us summarize the

**Main results**

For all \( m \geq 0 \) and \( r \geq 0 \)

\[
q_{m,r} = C^{-1} x_{m,r},
\]

where

\[
x_{m,r} = \sum_{i=0}^{\infty} d_i (\alpha_i^m + c_i \alpha_{i+1}^m) \beta_i^r
\]

if \( m \geq 0, r \geq 1 \)

\[
x_{m,0} = \frac{1}{\rho + 1} (x_{m-1,1} 2\rho + x_{m,1})
\]

if \( m > 0, r = 0 \)

\[
x_{0,0} = \frac{x_{0,1}}{\rho}
\]

and

\[
C = \frac{\rho (2 + \rho)}{2 (1 + \rho) (1 - \rho) (2 - \rho)}.
\]

The numbers \( \alpha_i, \beta_i, c_i \) and \( d_i \) satisfy for \( i = 0, 1, 2, \ldots \)

\[
\alpha_{i+1} = \frac{2\rho \beta_i^2}{\alpha_i},
\]

\[
\beta_{i+1} = \frac{\alpha_{i+1}^2}{(2\rho + \alpha_{i+1}) \beta_i},
\]

with initial values \( \alpha_0 = \rho^2 \) and \( \beta_0 = \rho^2 / (2 + \rho) \), and

\[
c_i = -\frac{\alpha_{i+1} - \beta_i}{\alpha_i - \beta_i},
\]

\[
d_{i+1} = -\frac{(\alpha_{i+1} + \rho) / \beta_{i+1} - (\rho + 1)}{(\alpha_{i+1} + \rho) / \beta_i - (\rho + 1)} c_i d_i,
\]

with initial value \( d_0 = 1 \).
Remark

The probabilities \( q_{m,0} \) are not given by the representation (20), but instead, for all \( m \geq 0 \)

\[
q_{m,0} = \sum_{i=0}^{\infty} d_i \left( \frac{\alpha_i}{\alpha_i + \rho} \alpha_{i+m}^m + c_i \frac{\alpha_{i+1}}{\alpha_{i+1} + \rho} \alpha_{i+1}^{m+1} \right).
\]

In the final two sections we will discuss some implications of the product form representation. First we will show that the product form representation yields a complete asymptotic expansion for large \( m \) and \( r \). Next we will discuss the numerical benefits of the product form representation.

8. Asymptotic expansion

Flatto and McKean [6] proved that

\[
q_{m,r} - K (\alpha_0^m + c_0 \alpha_0^m) \beta_0 = o \left( d_{j-1} (\alpha_{j-1}^{m-1} + c_{j-1} \alpha_{j-1}^{m-1}) \beta_{j-1}^{m-1} \right)
\]  

as \( m+r \to \infty \) and \( r \geq 1 \), \hspace{1cm} (37)

where \( K = 2 (1 + \rho) (1 - \rho) (2 - \rho) / \rho(2 + \rho) \). In view of this asymptotic result, we note that the product form representation (20) yields a complete asymptotic expansion as \( m + r \to \infty \) and \( r \geq 1 \). The proof is based on the following two lemmas. Since the numbers \( \alpha_j \) and \( \beta_j \) are monotonously decreasing, it immediately follows that

**Lemma 11**

For \( j = 1, 2, ... \)

\[
d_j (\alpha_j^m + c_j \alpha_j^{m+1}) \beta_j = o \left( d_{j-1} (\alpha_{j-1}^{m-1} + c_{j-1} \alpha_{j-1}^{m-1}) \beta_{j-1}^{m-1} \right)
\]  
as \( m+r \to \infty \) and \( r \geq 1 \).

Thus successive terms in the sum (20) are indeed refinements. Since the terms in (20) are alternating and monotonously decreasing in modulus, the error of each partial sum is bounded by the final term of the partial sum. Hence, we have the following set of \( O \)-formulas.

**Lemma 12**

For all \( j = 1, 2, ... \)

\[
q_{m,r} = C^{-1} \sum_{i=0}^{j-1} d_i (\alpha_i^m + c_i \alpha_i^{m+1}) \beta_i^j + O \left( d_j (\alpha_j^m + c_j \alpha_j^{m+1}) \beta_j^j \right)
\]  
as \( m+r \to \infty \) and \( r \geq 1 \).
The $O$-formula for $j = 1$ improves the asymptotic equivalence (37), as

$$C^{-1}d_0 \left( \alpha_0^n + c_0 \alpha_1^n \right) \beta_0^n + O \left[ d_1 \left( \alpha_1^n + c_1 \alpha_2^n \right) \beta_1^n \right] = C^{-1}d_0 \left( \alpha_0^n + c_0 \alpha_1^n \right) \beta_0^n (1 + o(1))$$

as $m+r \to \infty$ and $r \geq 1$. Accordingly, the formula for $j = 2$ improves the one for $j = 1$, and so on. The $O$-formulas of lemma 12 can be represented by the formula (see e.g. de Bruijn [3]),

$$q_{m,r} \approx \sum_{i=0}^{\infty} d_i \left( \alpha_i^m + c_i \alpha_{i+1}^m \right) \beta_i^r \quad \text{as } m+r \to \infty \text{ and } r \geq 1.$$

9. Numerical results

From a numerical point of view the product form representation is very attractive for several reasons. First of all, the terms in the representation for the numbers $x_{m,r}$ can be generated by a simple recursion scheme and decrease exponentially fast. Secondly, the terms are alternating and monotonously decreasing in modulus. This makes that the error in the partial sum can be bounded by the modulus of the final term in the partial sum.

In the table below we computed the numbers $q_{m,r}$ with a relative error of 0.1%. The numbers between brackets denote the number of terms in (20) needed to attain the accuracy.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$q_{0,1}$</th>
<th>$q_{0,2}$</th>
<th>$q_{1,1}$</th>
<th>$q_{1,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.1591 (14)</td>
<td>0.0100 (3)</td>
<td>0.0156 (3)</td>
<td>0.0007 (2)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1580 (10)</td>
<td>0.0233 (3)</td>
<td>0.0441 (4)</td>
<td>0.0047 (2)</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1100 (8)</td>
<td>0.0275 (4)</td>
<td>0.0606 (4)</td>
<td>0.0118 (3)</td>
</tr>
<tr>
<td>0.9</td>
<td>0.0380 (6)</td>
<td>0.0140 (4)</td>
<td>0.0350 (4)</td>
<td>0.0104 (3)</td>
</tr>
</tbody>
</table>

Table 1

Let us investigate the rate of convergence of the terms in the expression for $x_{m,r}$ as a function of $\rho$. From the corollary of lemma 6, it follows that for all $m \geq 0$ and $r \geq 1$,

$$\frac{d_{i+1} \left( \alpha_0^{m+1} + c_{i+1} \alpha_1^{m+1} \right) \beta_i^r}{d_i \left( \alpha_i^m + c_i \alpha_{i+1}^m \right) \beta_i^r} \to \frac{1 - A_1}{A_2 - 1} \left[ \frac{A_1}{A_2} \right]^{m+r-1} \quad \text{as } i \to \infty.$$

The factor $(1 - A_1)/(A_2 - 1)$ is monotonously decreasing for $0 < \rho < 1$, and

$$\lim_{\rho \to 0} \frac{1 - A_1}{A_2 - 1} = 1, \quad \lim_{\rho \to 1} \frac{1 - A_1}{A_2 - 1} = \frac{\sqrt{2} - 1}{\sqrt{2} + 1}.$$

Further, the factor $A_1/A_2$ is monotonously increasing for $0 < \rho < 1$, and
\[
\lim_{\rho \to 0} \frac{A_1}{A_2} = 0, \quad \lim_{\rho \to 1} \frac{A_1}{A_2} = \frac{2 - \sqrt{2}}{2 + \sqrt{2}}.
\]

Hence, if \( m > 0 \) or \( r > 1 \), then the convergence of the terms in the expression for \( x_{m,r} \) is very fast for all \( \rho \), namely at least with rate \( (2 - \sqrt{2})/(2 + \sqrt{2}) \approx 0.1715 \). But if \( m = 0 \) and \( r = 1 \), then the convergence is slow for small \( \rho \). This is illustrated in Table 1. For instance, for all values of \( \rho \) we need at most 4 terms to compute \( x_{0,2} \) with an accuracy of 0.1%. However, for \( \rho = 0.7 \) we need 8 terms to compute \( x_{0,1} \) with an accuracy of 0.1% and this number is nearly doubled for \( \rho = 0.3 \).

We can also obtain expressions for the moments of the waiting time. Let \( W_k \) denote the \( k \)-th moment of the waiting time, \( k = 1, 2, \ldots \), then
\[
W_k = 2 \sum_{m=1}^{\infty} \sum_{r=1}^{\infty} m^k q_{m,r} + \sum_{m=1}^{\infty} m^k q_{m,0}.
\]

By Equation (5), this can be written as
\[
W_k = 2 \sum_{m=1}^{\infty} \sum_{r=1}^{\infty} m^k q_{m,r} + \frac{1}{1 + \rho} \sum_{m=1}^{\infty} m^k \left[ q_{m-1,1} 2 \rho + q_{m,1} \right].
\]

Inserting the expression for \( q_{m,r} \) and then changing the order of summation yields
\[
W_k = C^{-1} \sum_{i=0}^{\infty} \left\{ 2 \, d_i \left[ f_k(\alpha_i) + c_i f_k(\alpha_{i+1}) \right] - \frac{\beta_i}{1 - \beta_i} + \frac{1}{1 + \rho} \, d_i \left[ f_k(\alpha_i) (\alpha_i^{-1} 2 \rho + 1) + c_i f_k(\alpha_{i+1}) (\alpha_{i+1}^{-1} 2 \rho + 1) \right] \beta_i \right\}, \tag{38}
\]
where
\[
f_k(x) = \sum_{m=1}^{\infty} m^k x^m.
\]

For example, if \( |x| < 1 \),
\[
\begin{align*}
f_1(x) &= x / (1 - x)^2, \\
f_2(x) &= x (1 + x) / (1 - x)^3, \\
f_3(x) &= x (1 + 4x + x^2) / (1 - x)^4.
\end{align*}
\]

In the table below we computed the first three moments \( W_1 \), \( W_2 \) and \( W_3 \). The sum in expression (38) was computed with a relative accuracy of 0.1%. The numbers between brackets denote the number of terms in (38) to attain the accuracy.
<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$W_1$</th>
<th>$W_2$</th>
<th>$W_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.1441 (13)</td>
<td>0.1740 (13)</td>
<td>0.2426 (12)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4262 (8)</td>
<td>0.7210 (8)</td>
<td>1.6068 (7)</td>
</tr>
<tr>
<td>0.7</td>
<td>1.1081 (6)</td>
<td>3.2757 (5)</td>
<td>13.868 (4)</td>
</tr>
<tr>
<td>0.9</td>
<td>4.4748 (4)</td>
<td>42.732 (3)</td>
<td>608.59 (2)</td>
</tr>
</tbody>
</table>

Table 2

References


Appendix: explicit forms for \(\alpha_i\) and \(\beta_i\)

Let us study the sequence of reciprocals of the numbers \(\alpha_i\) and \(\beta_i\).

**Definition**

For any \(x_0\), define the sequence

\[
\ldots \prec y_{-2} \prec y_{-1} \prec x_0 \prec y_0 \prec x_1 \prec y_1 \prec \ldots
\]

such that \(y_i\) and \(y_{i+1}\) are the roots of

\[
x_{i+1} y_{i+1} = x_{i+1}^2 + 2p + x_{i+1} + y^2. \tag{39}
\]

and \(x_i\) and \(x_{i+1}\) are the roots of

\[
x_{i+1} y_{i+1} = x_{i+1}^2 + 2p + x + y_i^2. \tag{40}
\]

By dividing both sides of equation (16) by \(\alpha^2 \beta^2\) and equation (17) by \(\alpha_i^2 \beta_i^2\), one easily verifies that the sequence of reciprocals \(1/\alpha_0, 1/\beta_0, 1/\alpha_1, 1/\beta_1, \ldots\) satisfies the quadratic forms (40) and (39). Hence, for the special choice of \(x_0 = 1/p^2\), it follows that one of the branches, originating in \(x_0\), generates the sequence of reciprocals. If we further specify that \(y_{-1} = 1/p\) and \(y_0 = (2 + p)/p^2\), then the right branch \(x_0, y_0, x_1, y_1, \ldots\) generates the reciprocals, thus

**Lemma 13**

If \(y_{-1} = 1/p\) and \(y_0 = (2 + p)/p^2\) (and thus \(x_0 = 1/p^2\)), then for \(i = 0, 1, 2, \ldots\)

\[
x_i = \frac{1}{\alpha_i}, \quad y_i = \frac{1}{\beta_i}.
\]

Kingman ([13], Lemma 3) proved

**Lemma 14**

Any sequence \(\ldots, y_{-2}, y_{-1}, y_0, y_1, y_2, \ldots\) has the form

\[
y_i = A + \mu (a \lambda^i + a^{-1} \lambda^{-i}),
\]

where

\[
A = \frac{1 + p}{2(1 + p^2)}, \quad \lambda = \frac{\beta + 1 - \sqrt{p^2 + 1}}{\beta + 1 + \sqrt{p^2 + 1}}, \quad \mu = \frac{2^{-3/2}}{1 + p^2},
\]

where
and $a$ is an arbitrary complex number.

From this representation one can easily obtain a representation for the numbers $x_i$. Since $y_{i-1}$ and $y_i$ are the roots of

$$x_i y 2(p + 1) = x^2 + 2p + x + y^2,$$

we have that

$$x_i = \frac{y_{i-1} + y_i}{2(p + 1)} = \frac{2A + \mu (\lambda^{1/2} + \lambda^{-1/2}) (a \lambda^{-1/2} + a^{-1} \lambda^{-i+1/2})}{2(p + 1)}.$$

Using that

$$(\lambda^{1/2} + \lambda^{-1/2})^2 = \lambda + \lambda^{-1} + 2 = \frac{2(1 + p)^2}{\rho} = \frac{A^2}{\mu^2},$$

so

$$\lambda^{1/2} + \lambda^{-1/2} = \frac{A}{\mu},$$

this expression can be simplified to

$$x_i = \frac{A (2 + a \lambda^{i-1/2} + a^{-1} \lambda^{-i+1/2})}{2(p + 1)}.$$

For the special choice of $y_{-1} = 1/\rho$ and $y_0 = (2 + p)/\rho^2$, the constant $a$ follows from the equations

$$y_{-1} = \frac{1}{\rho} = A + \mu (a \lambda^{-1} + a^{-1} \lambda),$$

$$y_0 = \frac{2 + p}{\rho^2} = A + \mu (a + a^{-1}).$$

The latter equation can be written as

$$a + a^{-1} = \mu^{-1} \left[ \frac{2 + p}{\rho^2} - A \right].$$

Since $(2 + p)/\rho^2 - A > 2$ and $\mu < 1$ for $0 < \rho < 1$, the left hand side of equation (42) is larger than 2. Hence, equation (42) has two positive roots. Let $v$ be the positive root, which is less than unity, then the other one is $1/v$. The appropriate root follows from equation (41). Assuming that $a = 1/v$, gives

$$\frac{1}{\rho} = A + \mu (v^{-1} \lambda^{-1} + v \lambda) > A + \mu (v^{-1} + v) = \frac{2 + p}{\rho^2},$$

which is clearly a contradiction. Hence $a = v$. By lemma 13, this proves
Lemma 15

For all $i = 0, 1, 2, ...$

\[
\frac{1}{\alpha_i} = \frac{A \left(2 + a \lambda^{i-1/2} + a^{-1} \lambda^{-i+1/2}\right)}{2(p + 1)}.
\]

\[
\frac{1}{\beta_i} = A + \mu \left(a \lambda^i + a^{-1} \lambda^{-i}\right),
\]

where $a$ is the positive root, less than unity, of

\[
\frac{2 + \rho}{p^2} = A + \mu (a + a^{-1}).
\]
List of COSOR-memoranda - 1989

<table>
<thead>
<tr>
<th>Number</th>
<th>Month</th>
<th>Author</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>M 89-01</td>
<td>January</td>
<td>D.A. Overdijk</td>
<td>Conjugate profiles on mating gear teeth</td>
</tr>
<tr>
<td>M 89-02</td>
<td>January</td>
<td>A.H.W. Geerts</td>
<td>A priori results in linear-quadratic optimal control theory</td>
</tr>
<tr>
<td>M 89-03</td>
<td>February</td>
<td>A.A. Stoorvogel, H.L. Trentelman</td>
<td>The quadratic matrix inequality in singular $H_\infty$ control with state feedback</td>
</tr>
<tr>
<td>M 89-04</td>
<td>February</td>
<td>E. Willekens, N. Veraverbeke</td>
<td>Estimation of convolution tail behaviour</td>
</tr>
<tr>
<td>M 89-05</td>
<td>March</td>
<td>H.L. Trentelman</td>
<td>The totally singular linear quadratic problem with indefinite cost</td>
</tr>
<tr>
<td>M 89-06</td>
<td>April</td>
<td>B.G. Hansen</td>
<td>Self-decomposable distributions and branching processes</td>
</tr>
<tr>
<td>M 89-07</td>
<td>April</td>
<td>B.G. Hansen</td>
<td>Note on Urbanik's class $L_\alpha$</td>
</tr>
<tr>
<td>M 89-08</td>
<td>April</td>
<td>B.G. Hansen</td>
<td>Reversed self-decomposability</td>
</tr>
<tr>
<td>M 89-09</td>
<td>April</td>
<td>A.A. Stoorvogel</td>
<td>The singular zero-sum differential game with stability using $H_\infty$ control theory</td>
</tr>
<tr>
<td>M 89-10</td>
<td>April</td>
<td>L.J.G. Langenheoff, W.H.M. Zijm</td>
<td>An analytical theory of multi-echelon production/distribution systems</td>
</tr>
</tbody>
</table>