Research Article

Two Coupled Queues with Vastly Different Arrival Rates: Critical Loading Case

Charles Knessl[1] and John A. Morrison[2]

[1] *Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago,*
  851 South Morgan Street Chicago, IL 60607-7045, USA
[2] *Consultant Bell Laboratories, Alcatel-Lucent, 600 Mountain Avenue, Murray Hill, NJ 07974, USA*

Correspondence should be addressed to Charles Knessl, knessl@uic.edu

We consider two coupled queues with a generalized processor sharing service discipline. The second queue has a much smaller Poisson arrival rate than the first queue, while the customer service times are of comparable magnitude. The processor sharing server devotes most of its resources to the first queue, except when it is empty. The fraction of resources devoted to the second queue is small, of the same order as the ratio of the arrival rates. We assume that the primary queue is heavily loaded and that the secondary queue is critically loaded. If we let the small arrival rate to the secondary queue be $O(\varepsilon)$, where $0 \leq \varepsilon \ll 1$, then in this asymptotic limit the number of customers in the first queue will be large, of order $O(\varepsilon^{-1})$, while that in the second queue will be somewhat smaller, of order $O(\varepsilon^{-1/2})$. We obtain a two-dimensional diffusion approximation for this model and explicitly solve for the joint steady state probability distribution of the numbers of customers in the two queues. This work complements that in (Morrison, 2010), which the second queue was assumed to be heavily or lightly loaded, leading to mean queue lengths that were $O(\varepsilon^{-1})$ or $O(1)$, respectively.

1. Introduction

The study of two coupled queues is a fundamental problem in queueing theory and applied probability. Classic examples include the shortest queue problem [1–3], the longer queue problem [4], fork-join models [5–7] and two coupled queues with generalized processor sharing [8–10], which is the subject of the present investigation. Computing the joint probability distribution for these models typically leads to functional equations that may sometimes be recast as boundary value problems [11], such as Dirichlet and Riemann-Hilbert problems.

Generalized processor sharing (GPS) models have become quite popular in recent years, as they provide scheduling algorithms that yield both service differentiation among

different customer classes and also gains from statistical multiplexing. Some recent investigations and applications of such models appear in [12–15], where they are used, for example, for flow control in integrated service networks.

We consider here two parallel queues with respective Poisson arrival rates λ and $\varepsilon\sigma$, where $0 < \varepsilon \ll 1$. Thus, the arrivals to the second queue are much less frequent than those to the first, and we immediately scale the second arrival rate by ε, thus introducing σ. The service times are assumed to be exponentially distributed in both queues, with respective means $1/\mu$ and 1. Thus, we are taking the unit of time as the mean service time in the second queue. There is a single processor sharing server that works at unit rate and devotes $1 - \varepsilon\kappa = 1 - O(\varepsilon)$ of its capacity to the first queue and the remaining $\varepsilon\kappa$ to the second queue, provided both queues are nonempty. If one queue is empty, the processor devotes all of its capacity to the other queue. The total load is given by $\lambda/\mu + \varepsilon\sigma$, and we assume that the system is in heavy traffic so that this quantity will be close to 1. Hence, we define ω from the relation $\lambda/\mu + \varepsilon\sigma = 1 - \varepsilon\omega$ and assume that $\omega > 0$, so that the system is stable. This also means that the first queue is heavily loaded. The second queue has traffic intensity $\varepsilon\sigma/(\varepsilon\kappa) = \sigma/\kappa$, and we say it is underloaded if $\sigma/\kappa < 1$, overloaded if $\sigma/\kappa > 1$ and critically loaded if $\sigma/\kappa \approx 1$ (more precisely $\sigma/\kappa = 1 + O(\sqrt{\varepsilon})$). The underloaded and overloaded cases were analyzed in [16]. We denote by N_1 (resp., N_2) the number of customers in the first (resp., second) queue and the joint steady state probability distribution will be denoted by $p(m, n) = \mathrm{Prob}[N_1 = m, N_2 = n]$.

For the underloaded case most of the mass occurs on the scale $m = O(\varepsilon^{-1})$ and $n = O(1)$, so there will tend to be only a few customers in the second queue. Asymptotically, $p(m, n)$ has a product form behavior, with an exponential distribution in εN_1 and a geometric distribution in N_2 (see [16]). This analysis was recently extended to an arbitrary number of parallel queues by Morrison and Borst [17], as long as one queue is heavily loaded and all the others are underloaded (with similar assumptions about arrival rates and processor-sharing factors as above). For the overloaded case most of the probability mass occurs for both $m, n = O(\varepsilon^{-1})$, and in [16] a diffusion limit of the form $p(m, n) \sim \varepsilon^2 p_*(\zeta, \tau) = \varepsilon^2 p_*(\varepsilon m, \varepsilon n)$ is obtained. Here p_* may be characterized as the solution to a parabolic PDE, in the variables ζ and τ. Here, we will analyze the critically loaded case, which also leads to a diffusion limit with now $p(m, n) \sim \varepsilon^{3/2}\phi_0(\zeta, w) = \varepsilon^{3/2}\phi_0(\varepsilon m, \sqrt{\varepsilon}n)$, where ϕ_0 will satisfy an elliptic PDE. Thus the critically loaded case has $N_2 = O(\varepsilon^{-1/2})$ and leads to a somewhat more difficult problem than either the underloaded or overloaded cases.

We will obtain the PDE for $\phi_0(\zeta, w)$ as a limiting case of the difference equation(s) satisfied by $p(m, n)$, and explicitly solve the PDE by transform methods. We will obtain detailed results for the marginal distributions $\mathrm{Prob}[N_1 = m] = \sum_{n=0}^{\infty} p(m, n)$ and $\mathrm{Prob}[N_2 = n] = \sum_{m=0}^{\infty} p(m, n)$, as well as the mean queue lengths. We will also obtain other approximations to $p(m, n)$ that are valid on scales where $m = o(\varepsilon^{-1})$ and/or $n = o(\varepsilon^{-1/2})$. In particular, we shall show that $p(m, n)$ is $O(\varepsilon)$ on the scale $m = O(\varepsilon^{-1/2})$ and $n = O(1)$.

Previous work on this model includes Fayolle and Iasnogorodski [8] (see also [10]) and the more recent study of Guillemin and Pinchon [9]. There the authors consider the double generating function $F(x, y) = \sum_{m,n} x^m y^n p(m, n)$ and obtain a functional equation for the boundary values $F(x, 0)$ and $F(0, y)$. This is ultimately converted to a Dirichlet problem, which is solved to yield the boundary values of F in terms of elliptic integrals. One of the authors (J. A. Morrison) has verified that by analyzing the results in [8] and [9], in the asymptotic limit we consider and with the scaling $x = 1 - O(\varepsilon)$ and $y = 1 - O(\sqrt{\varepsilon})$, and then inverting the double transform, we also obtain our main approximation $p(m, n) \sim \varepsilon^{3/2}\phi_0(\zeta, w)$. However, this involves a lengthy calculation that takes far more work than our direct approach, which consists of deriving a limiting PDE and solving it, along with

appropriate boundary conditions. Also, this direct approach should work for other models of this type, including ones with finite capacities of customers, and with 3 or more coupled queues.

Other recent work on diffusion approximations for generalized processor sharing models includes Ramanan and Reiman [18] (see also references therein). This work, however, is more concerned with theoretical aspects of the diffusion approximations, such as convergence of the discrete problem to a certain diffusion process. Here, our focus is on obtaining the explicit solution to the limiting diffusion equation that arises from the balance equations. It is highly likely that this equation can be interpreted as the Kolmogorov forward equation for some appropriate diffusion process, but we do not consider such "process level" aspects here, as our approach is largely analytical.

Our approach has the merit that it can be used to compute correction terms to the diffusion approximations, and we do this in some cases here. Also, we treat scales other than the basic diffusion scale, where, for example, some of the variables remain discrete, which we then relate to the diffusion scale by asymptotic matching. This type of analysis is needed, for example, to accurately compute boundary probabilities.

From a mathematical viewpoint, the diffusion approximation we obtain (i.e., $\phi_0(\zeta, w)$) is somewhat nonstandard in that the density vanishes as $\zeta \to 0$ and the approximation breaks down for $\zeta = O(\sqrt{\varepsilon})$. Also, the corner behavior of the problem is much different than what is typical. In [19] we analyzed a more general version of this model in another heavy traffic limit, assuming that the arrival rates and processor-sharing factors were of comparable magnitude. That analysis led to an elliptic PDE that was more complicated than the one obtained here, but probably more representative of typical diffusion approximations to two coupled queues, such as those considered in [20–22].

The present scaling limit leads to a separable, elliptic PDE in the variables εm and $\sqrt{\varepsilon}n$. Since the boundary conditions are somewhat simpler than those for the diffusion model in [19], we are able to obtain a more explicit solution to this equation, using classical transform theory [23]. We then evaluate the solution in various limiting cases, to obtain even simpler results that yield more insight into model behavior.

Yet another analysis of the model considered here is done in [24], but there it was assumed that both the arrival and service rates of the secondary customers are small, while the server devotes comparable resources to each queue.

The paper is organized as follows. In Section 2, we state the problem more precisely and give the balance equations satisfied by $p(m, n)$. In Section 3, we summarize all of our main results. The derivations are given in Section 4 for the scale $m = O(\varepsilon^{-1})$, $n = O(\varepsilon^{-1/2})$ and in Section 5 for the other ranges of m, n.

Throughout the paper we will use the notation $f(x) \sim g(x)$ to mean $\lim_{x \to x_0}[f(x)/g(x)] = 1$, $f(x) = o(g(x))$ to mean $\lim_{x \to x_0}[f(x)/g(x)] = 0$, and $f(x) = O(g(x))$ to mean that $|f(x)/g(x)|$ is bounded for x sufficiently close to x_0.

2. Formulation

We consider two parallel infinite capacity queues for different traffic classes. The jobs arrive as Poisson processes with rates λ for the primary class, and $\varepsilon\sigma$ for the secondary class, where $0 < \varepsilon \ll 1$. Hence, the secondary jobs arrive much less frequently than the primary ones. Moreover, it is assumed that the primary and secondary jobs have exponentially distributed service requirements with mean service times $1/\mu$ and 1, respectively, where $\mu = O(1)$. The server works at unit rate, and if neither queue is empty devotes fractions $1 - \varepsilon\kappa$ and $\varepsilon\kappa$ of its

effort to the primary and secondary queues, respectively, where $\kappa = O(1)$. The corresponding service rates are $(1 - \varepsilon\kappa)\mu$ and $\varepsilon\kappa$. If one queue is empty the server works at unit rate on the other queue, so that this model is work conserving. It is assumed that the primary queue is heavily loaded, with

$$\frac{\lambda}{\mu} + \varepsilon\sigma = 1 - \varepsilon\omega, \quad 0 < \omega = O(1). \tag{2.1}$$

Moreover, we assume that the secondary queue is critically loaded, with

$$\sigma = \kappa + \delta\sqrt{\varepsilon}, \quad \delta = O(1), \tag{2.2}$$

where δ may have either sign. Thus, the asymptotic limit we consider has $\varepsilon \to 0^+$ with ω, μ, κ and δ fixed, and then σ and λ vary with ε so that (2.1) and (2.2) hold.

Since $\omega > 0$ the system is stable. Let $p(m, n)$ denote the stationary probability that there are m jobs in the primary queue and n jobs in the secondary queue. Then, the balance equations satisfied by $p(m, n)$ are

$$[\lambda + \varepsilon\sigma + (1 - \varepsilon\kappa)\mu + \varepsilon\kappa]p(m, n) = \lambda p(m - 1, n) + \varepsilon\sigma p(m, n - 1)$$

$$+ (1 - \varepsilon\kappa)\mu p(m + 1, n) + \varepsilon\kappa p(m, n + 1), \tag{2.3}$$

$$m \geqslant 1, \quad n \geqslant 1,$$

$$(\lambda + \varepsilon\sigma + \mu)p(m, 0) = \lambda p(m - 1, 0) + \mu p(m + 1, 0) + \varepsilon\kappa p(m, 1), \quad m \geqslant 1, \tag{2.4}$$

$$(\lambda + \varepsilon\sigma + 1)p(0, n) = \varepsilon\sigma p(0, n - 1) + (1 - \varepsilon\kappa)\mu p(1, n) + p(0, n + 1), \quad n \geqslant 1, \tag{2.5}$$

$$(\lambda + \varepsilon\sigma)p(0, 0) = \mu p(1, 0) + p(0, 1). \tag{2.6}$$

The normalization condition is

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p(m, n) = 1. \tag{2.7}$$

The mean number of jobs in the primary and secondary queues are

$$\mathbf{E}(N_1) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} mp(m, n),$$

$$\mathbf{E}(N_2) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} np(m, n), \tag{2.8}$$

From Little's Law, the corresponding mean waiting times are $\mathbf{E}(N_1)/\lambda$ and $\mathbf{E}(N_2)/(\varepsilon\sigma)$. From [8], the conservation law of Kleinrock implies, using (2.1) and (2.2), that

$$\frac{\mathbf{E}(N_1)}{\mu} + \mathbf{E}(N_2) = \frac{1}{\varepsilon\omega}\left(\frac{\lambda}{\mu^2} + \varepsilon\sigma\right)$$

$$= \frac{1}{\varepsilon\mu\omega}\left[1 + \varepsilon(\mu\kappa - \omega - \kappa) + \varepsilon^{3/2}(\mu - 1)\delta\right]. \tag{2.9}$$

3. Summary of Results

We consider $m = \zeta/\varepsilon = O(1/\varepsilon)$ and $n = w/\sqrt{\varepsilon} = O(1/\sqrt{\varepsilon})$, with

$$p\left(\frac{\zeta}{\varepsilon}, \frac{w}{\sqrt{\varepsilon}}\right) = \varepsilon^{3/2}\left[\phi_0(\zeta, w) + \sqrt{\varepsilon}\phi_1(\zeta, w) + O(\varepsilon)\right],$$

$$0 < \zeta = O(1), \quad 0 < w = O(1). \tag{3.1}$$

We then have.

Proposition 3.1.

$$\phi_0(\zeta, w) + I(\delta < 0)\frac{\delta\omega}{\kappa}\exp\left(-\omega\zeta + \frac{\delta w}{\kappa}\right)$$

$$= \frac{2\omega}{\pi}\exp\left(-\frac{\omega\zeta}{2} + \frac{\delta w}{2\kappa}\right)\int_0^\infty \frac{\beta[\beta\cos(w\beta) + (\delta/2k)\sin(w\beta)]}{(\beta^2 + \delta^2/(4\kappa^2))}\exp\left[-\sqrt{\omega^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa\beta^2}{\mu}}\frac{\zeta}{2}\right]d\beta. \tag{3.2}$$

This gives the limiting density of (N_1, N_2), which applies for fixed values of ζ and w. We next evaluate this density in various limiting cases, such as $w \to \infty$ and $\delta \to \pm\infty$, to gain more insight into its structure, and to verify consistency with results in [16].

Corollary 3.2. *If $\delta = O(1)$, $w \gg 1$ and $0 < \zeta/\sqrt{w} \ll 1$, then*

$$\phi_0(\zeta, w) \sim \frac{\omega(\delta^2 + \kappa\mu\omega^2)^{3/4}\zeta}{\sqrt{\pi\mu}\left(\delta + \sqrt{\delta^2 + \kappa\mu\omega^2}\right)w^{3/2}}$$

$$\times \exp\left(-\frac{\omega\zeta}{2}\right)\exp\left[-\frac{w}{2\kappa}\left(\sqrt{\delta^2 + \kappa\mu\omega^2} - \delta\right)\right]. \tag{3.3}$$

Corollary 3.3. *If $-\delta \gg 1, w/|\delta| > 0$ and $0 < \zeta = O(1)$, then*

$$
\begin{aligned}
\phi_0(\zeta, w) \quad \sim \quad & \frac{\omega|\delta|}{2\kappa} \exp\left(-\frac{\omega\zeta}{2}\right) \exp\left[-\frac{w}{2\kappa}\left(\sqrt{\delta^2 + \kappa\mu\omega^2} + |\delta|\right)\right] \exp\left(\frac{\mu\omega^2}{4|\delta|}w\right) \\
& \times \left\{\exp\left(-\frac{\omega\zeta}{2}\right) \mathrm{Erfc}\left[\frac{(\mu\omega w - |\delta|\zeta)}{2\sqrt{|\delta|\mu w}}\right] - \exp\left(\frac{\omega\zeta}{2}\right) \mathrm{Erfc}\left[\frac{(\mu\omega w + |\delta|\zeta)}{2\sqrt{|\delta|\mu w}}\right]\right\},
\end{aligned}
\tag{3.4}
$$

where the complementary error function is given by [25]

$$
\mathrm{Erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-u^2} du.
\tag{3.5}
$$

Corollary 3.4. *If $w \gg -\delta \gg 1$ and $0 < \zeta = O(1)$, then*

$$
\phi_0(\zeta, w) \sim \frac{2|\delta|^{5/2}\zeta}{\sqrt{\pi\mu}\kappa\mu\omega w^{3/2}} \exp\left(-\frac{\omega\zeta}{2}\right) \exp\left[-\frac{w}{2\kappa}\left(\sqrt{\delta^2 + \kappa\mu\omega^2} + |\delta|\right)\right].
\tag{3.6}
$$

Corollary 3.5. *If $-\delta \gg 1, 0 < w = O(1)$ and $0 < \zeta = O(1)$, then*

$$
\phi_0(\zeta, w) \sim \frac{\omega|\delta|}{\kappa} \exp\left(-\omega\zeta - \frac{|\delta|w}{\kappa}\right).
\tag{3.7}
$$

Remark 3.6. This matches with [16, Result 7], with $n = w/\sqrt{\varepsilon}, 0 < \varepsilon \ll 1$ and $|\delta|\sqrt{\varepsilon} \ll 1$.

Corollary 3.7. *If $\delta \gg 1, 0 < w/\delta = O(1)$ and $0 < \zeta = O(1)$, then*

$$
\phi_0(\zeta, w) \sim \frac{\omega\sqrt{\delta}\zeta}{2\sqrt{\pi\mu}w^{3/2}} \exp\left[-\frac{\delta}{4\mu w}\left(\zeta + \frac{\mu\omega w}{\delta}\right)^2\right].
\tag{3.8}
$$

Remark 3.8. This matches with [16, Result 4], with $\tau = w\sqrt{\varepsilon}, 0 < \varepsilon \ll 1$ and $\delta\sqrt{\varepsilon} \ll 1$.

In Propositions 3.9 and 3.16 below, we give the limiting marginals of N_2 and N_1, which apply for w and ζ fixed, respectively. Then, we simplify these marginal densities in various limiting cases.

Proposition 3.9. *The scaled lowest order asymptotic approximation to the stationary distribution of the number of jobs in the secondary queue is*

$$
\sum_{m=0}^\infty p\left(m, \frac{w}{\sqrt{\varepsilon}}\right) \sim \sqrt{\varepsilon} \int_0^\infty \phi_0(\zeta, w)d\zeta,
\tag{3.9}
$$

where

$$\int_0^\infty \phi_0(\zeta, w)d\zeta + I(\delta < 0)\frac{\delta}{\kappa}\exp\left(\frac{\delta w}{\kappa}\right) = \frac{4w}{\pi}\exp\left(\frac{\delta w}{2\kappa}\right)$$

$$\times \int_0^\infty \frac{\beta[\beta\cos(w\beta) + (\delta/2\kappa)\sin(w\beta)]}{(\beta^2 + \delta^2/4\kappa^2)\left[w + \sqrt{w^2 + \delta^2/(\kappa\mu) + 4\kappa\beta^2/\mu}\right]}d\beta. \tag{3.10}$$

Corollary 3.10. *If $\delta = O(1)$ and $w \gg 1$, then*

$$\int_0^\infty \phi_0(\zeta, w)d\zeta \sim \frac{4(\delta^2 + \kappa\mu w^2)^{3/4}}{\sqrt{\pi\mu}w\left(\delta + \sqrt{\delta^2 + \kappa\mu w^2}\right)w^{3/2}}\exp\left[-\frac{w}{2\kappa}\left(\sqrt{\delta^2 + \kappa\mu w^2} - \delta\right)\right], \tag{3.11}$$

which is consistent with (3.3).

Corollary 3.11. *If $-\delta \gg 1$ and $0 < w/|\delta| = O(1)$, then*

$$\int_0^\infty \phi_0(\zeta, w)d\zeta \sim \frac{|\delta|}{\kappa}\exp\left[-\frac{w}{2\kappa}\left(\sqrt{\delta^2 + \kappa\mu w^2} + |\delta|\right)\right]$$

$$\times \left[\left(1 + \frac{\mu w^2 w}{2|\delta|}\right)\exp\left(\frac{\mu w^2 w}{4|\delta|}\right)\text{Erfc}\left(\frac{w}{2}\sqrt{\frac{\mu w}{|\delta|}}\right) - w\sqrt{\frac{\mu w}{\pi|\delta|}}\right]. \tag{3.12}$$

Corollary 3.12. *If $w \gg -\delta \gg 1$, then*

$$\int_0^\infty \phi_0(\zeta, w)d\zeta \sim \frac{8|\delta|^{5/2}}{\sqrt{\pi\mu}\kappa\mu w^3 w^{3/2}}\exp\left[-\frac{w}{2\kappa}\left(\sqrt{\delta^2 + \kappa\mu w^2} + |\delta|\right)\right], \tag{3.13}$$

which is consistent with (3.6).

Corollary 3.13. *If $-\delta \gg 1$ and $0 < w = O(1)$, then*

$$\int_0^\infty \phi_0(\zeta, w)d\zeta \sim \frac{|\delta|}{\kappa}\exp\left(-\frac{|\delta|w}{\kappa}\right), \tag{3.14}$$

which is consistent with (3.7).

Corollary 3.14. *If $\delta \gg 1$ and $0 < w/\delta = O(1)$, then*

$$\int_0^\infty \phi_0(\zeta, w)d\zeta \sim w\sqrt{\frac{\mu}{\delta}}\left[\frac{1}{\sqrt{\pi w}}\exp\left(-\frac{\mu w^2}{4\delta}w\right) - \frac{w}{2}\sqrt{\frac{\mu}{\delta}}\text{Erfc}\left(\frac{w}{2}\sqrt{\frac{\mu w}{\delta}}\right)\right], \tag{3.15}$$

which is consistent with (3.8).

Remark 3.15. This matches with [16, Result 3], with $\tau = w\sqrt{\varepsilon}$, $0 < \varepsilon \ll 1$ and $\delta\sqrt{\varepsilon} \ll 1$.

Proposition 3.16. *The scaled asymptotic approximation to the stationary distribution of the number of jobs in the primary queue is*

$$\sum_{n=0}^{\infty} p\left(\frac{\zeta}{\varepsilon}, n\right) = \varepsilon w \exp(-w\zeta) + \varepsilon^{3/2} Q_1(\zeta) + O\left(\varepsilon^2\right), \tag{3.16}$$

where

$$Q_1(\zeta) - I(\delta > 0)\delta(1 - w\zeta)\exp(-w\zeta)$$

$$= \frac{w\mu}{\pi} \exp\left(-\frac{w\zeta}{2}\right) \int_0^{\infty} \frac{\beta^2}{\left(\beta^2 + \delta^2/4\kappa^2\right)^2}$$

$$\times \left\{ \left[w + \sqrt{w^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa\beta^2}{\mu}} \right] \exp\left[-\sqrt{w^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa\beta^2}{\mu}} \frac{\zeta}{2} \right] - 2w \exp\left(-\frac{w\zeta}{2}\right) \right\} d\beta. \tag{3.17}$$

Corollary 3.17. *If $|\delta| \gg 1$, then*

$$\sum_{n=0}^{\infty} p\left(\frac{\zeta}{\varepsilon}, n\right) = \varepsilon \exp(-w\zeta)\left[w + \sqrt{\varepsilon}I(\delta > 0)\delta(1 - w\zeta)\right] + O\left(\varepsilon^2\right). \tag{3.18}$$

Remark 3.18. This matches with [16, Result 7] for $\delta < 0$, and with [16, Result 2] for $\delta > 0$ since, for $\delta\sqrt{\varepsilon} \ll 1$,

$$(w + \sigma - \kappa)\exp[-(w + \sigma - \kappa)\zeta] = \exp(-w\zeta)\left[w + \sqrt{\varepsilon}\delta(1 - w\zeta)\right] + O(\varepsilon). \tag{3.19}$$

We next give expansions for the mean queue lengths, for $\varepsilon \to 0$ (with w, κ, μ, δ fixed).

Proposition 3.19. *The lowest order asymptotic approximation to the stationary mean secondary queue length is*

$$\mathbf{E}(N_2) \sim \frac{s}{\sqrt{\varepsilon}},$$

$$s = \frac{\delta}{\mu w^2} \left\{ 1 + \frac{1}{2\pi c^4} \left[\left(1 - c^4\right)\cos^{-1}(-c) - \left(1 + c^4\right)\cos^{-1}c - 2c\sqrt{1 - c^2}\left(1 - 2c^2\right) \right] \right\}, \tag{3.20}$$

where

$$-1 < c = \frac{\delta}{\sqrt{\delta^2 + \kappa\mu w^2}} < 1, \quad 0 < \cos^{-1}(\pm c) < \pi. \tag{3.21}$$

Remark 3.20. We have verified from (3.17) that $\int_0^\infty \zeta Q_1(\zeta)d\zeta = -\mu s$ so that from (3.16),

$$\mathbf{E}(N_1) = \frac{1}{\varepsilon \omega} - \frac{\mu s}{\sqrt{\varepsilon}} + O(1),$$
(3.22)

which is consistent with (2.9).

Corollary 3.21. *If $|\delta|\sqrt{\varepsilon} \ll 1$, then*

$$s \sim \frac{\delta}{\mu\omega^2}, \quad if \ \delta \gg 1; \qquad s \sim -\frac{\kappa}{\delta}, \quad if \ -\delta \gg 1.$$
(3.23)

These match with [16, Result 1 and Corollary 10], respectively.

We next give some asymptotic results for $\phi_0(\zeta, w)$ that apply for δ fixed and ζ and/or $w \to \infty$. We also give the "corner" behavior as $(\zeta, w) \to (0, 0)$.

Proposition 3.22. (i) $\zeta, w \to \infty$ *with $\delta < 0$ and $0 \leqslant w/\zeta < |\delta|/(\mu\omega)$,*

$$\phi_0(\zeta, w) \sim \frac{\omega|\delta|}{\kappa} \exp\left(-\omega\zeta - \frac{|\delta|}{\kappa}w \right),$$
(3.24)

(ii) $\zeta, w \to \infty$ *with $0 < w/\zeta < \infty$ for $\delta > 0$, or $|\delta|/(\mu\omega) < w/\zeta < \infty$ for $\delta \leqslant 0$,*

$$\phi_0(\zeta, w) \sim K(\zeta, w) \exp\left(-\frac{\omega}{2}\zeta + \frac{\delta}{2\kappa}w - \frac{1}{2}\sqrt{\omega^2 + \frac{\delta^2}{\mu\kappa}}\sqrt{\frac{\mu}{\kappa}w^2 + \zeta^2} \right),$$
(3.25)

$$K(\zeta, w) = \sqrt{\frac{\mu}{\pi\kappa}} \frac{\omega b_s}{b_s + \delta/(2\kappa)} \left(\omega^2 + \frac{\delta^2}{\kappa\mu} \right)^{1/4} \zeta\left(\zeta^2 + \frac{\mu}{\kappa}w^2 \right)^{-3/4},$$
(3.26)

$$b_s = b_s\left(\frac{w}{\zeta} \right) = \frac{\mu}{2\kappa}\left(\omega^2 + \frac{\delta^2}{\kappa\mu} \right)^{1/2} \frac{w}{\sqrt{\zeta^2 + (\mu/\kappa)w^2}},$$

(iii) $\zeta \to \infty$, $w = O(1)$, $\delta > 0$,

$$\phi_0(\zeta, w) \sim \frac{\omega}{\delta}\frac{\mu^{3/2}}{\sqrt{\kappa}}\left(\omega^2 + \frac{\delta^2}{\kappa\mu} \right)^{3/4}\frac{1}{\sqrt{\pi}\zeta^{3/2}}\left(w + \frac{2\kappa}{\delta} \right)\exp\left[\frac{\delta w}{2\kappa} \right]\exp\left[-\frac{1}{2}\left(\omega + \sqrt{\omega^2 + \frac{\delta^2}{\kappa\mu}} \right)\zeta \right],$$
(3.27)

(iv) $w, \zeta \to \infty$, $\zeta = O(\sqrt{w})$,

$$\phi_0(\zeta, w) \sim \sqrt{\frac{2b_+\kappa}{\mu}} \frac{b_+ w}{b_+ + \delta/(2\kappa)} \frac{\zeta}{\sqrt{\pi} w^{3/2}} \exp\left[\left(\frac{\delta}{2\kappa} - b_+\right)w - \frac{w}{2}\zeta - \frac{b_+\kappa}{2\mu} \frac{\zeta^2}{w}\right], \tag{3.28}$$

$$b_+ = \frac{\sqrt{\mu}}{2\sqrt{\kappa}} \sqrt{w^2 + \frac{\delta^2}{\kappa\mu}}, \tag{3.29}$$

(v) $\zeta, w \to 0$ with $0 < w/\zeta < \infty$,

$$\phi_0(\zeta, w) \sim \frac{2w}{\pi} \sqrt{\frac{\mu}{\kappa}} \frac{\zeta}{\zeta^2 + (\mu/\kappa)w^2}, \tag{3.30}$$

(vi) $\zeta \to 0$ with $0 < w < \infty$,

$$\phi_0(\zeta, w) \sim \sqrt{\frac{\kappa}{\mu}} \frac{w}{\pi} \zeta a(w) \exp\left(\frac{\delta w}{2\kappa}\right), \tag{3.31}$$

$$a(w) = \frac{\delta}{2\kappa} \frac{1}{w} \int_{-\infty}^{\infty} e^{iw\beta} \frac{(\mu/4\kappa)(w^2 + \delta^2/\mu\kappa) - i(\beta\delta/2\kappa)}{(\beta + i\delta/2\kappa)^2 [\beta^2 + (\mu/4\kappa)(w^2 + \delta^2/\kappa\mu)]^{1/2}} d\beta$$

$$\tag{3.32}$$

$$+ \frac{\mu}{4\kappa}\left(w^2 + \frac{\delta^2}{\mu\kappa}\right) \frac{1}{w^2} \int_{-\infty}^{\infty} e^{iw\beta} \left[\beta^2 + \frac{\mu}{4\kappa}\left(w^2 + \frac{\delta^2}{\kappa\mu}\right)\right]^{-3/2} d\beta.$$

Remark 3.23. The results show that for $\delta < 0$ the density ϕ_0 is asymptotically of product form in a sector, and distinctly nonproduct form in the complimentary sector. For $\delta > 0$ the product form behavior is absent. Item (v) shows that ϕ_0 has an integrable singularity near the corner, while item (vi) shows that $\phi_0 = O(\zeta)$ as $\zeta \to 0$. The second integral in (3.32) may be expressed in terms of a modified Bessel function, using the identity

$$\int_{-\infty}^{\infty} e^{iZU} \left(Z^2 + 1\right)^{-3/2} dZ = 2U K_1(U). \tag{3.33}$$

The approximation $p(m, n) \approx \varepsilon^{3/2} \phi_0(\zeta, w)$ is only valid for $m = O(\varepsilon^{-1})$ and $n = O(\varepsilon^{-1/2})$. For other ranges of (m, n), other expansions must be constructed, and these we summarize below.

Proposition 3.24. (i) $m = S/\sqrt{\varepsilon} = O(\varepsilon^{-1/2})$, $n = w/\sqrt{\varepsilon} = O(\varepsilon^{-1/2})$,

$$p(m, n) \sim \varepsilon^2 [S\mathcal{P}_+(w) + \mathcal{P}_-(w)],$$

$$\mathcal{P}_+(w) = \frac{\omega}{\pi}\sqrt{\frac{\kappa}{\mu}} \exp\left(\frac{\delta w}{2\kappa}\right) a(w),$$ (3.34)

$$\mathcal{P}_-(w) = \mu \int_w^\infty \mathcal{P}_+(u) du,$$

(ii) $m = S/\sqrt{\varepsilon} = O(\varepsilon^{-1/2})$, $n = O(1)$,

$$p(m, n) \sim \varepsilon \frac{2\omega}{\pi} \int_0^\pi \cos\left(\frac{\Omega}{2}\right) \cos\left[\left(n + \frac{1}{2}\right)\Omega\right] \exp\left[-2\sqrt{\frac{\kappa}{\mu}} S \sin\left(\frac{\Omega}{2}\right)\right] d\Omega$$ (3.35)

(iii) $m, n = O(1)$,

$$p(m, n) \sim \varepsilon^{3/2} \frac{2\omega}{\pi}\sqrt{\frac{\kappa}{\mu}} \left[\frac{4}{(2n+3)(2n-1)} m + \frac{4n}{4n^2 - 1}\mu\right], \quad n \geqslant 1,$$ (3.36)

$$p(m, 0) \sim \varepsilon\omega.$$

Remark 3.25. We comment that for $n = O(1)$ and $\zeta > 0$ $(m = O(\varepsilon^{-1}))$ $p(m, n) \sim \varepsilon^{3/2}\phi_0(\zeta, 0)$ so that the diffusion approximation still applies. For $m = O(1)$ and $w > 0$ item (i) still applies, and then $p(m, n) \sim \varepsilon^2 \mathcal{P}_-(w)$, which is independent of m and can be used to estimate the boundary probabilities $p(0, n)$, which are $O(\varepsilon^2)$ for $n = O(\varepsilon^{-1/2})$. Note that in item (ii) and for $n = 0$ in item (iii), $p(m, n)$ is $O(\varepsilon)$, which is larger than the order of magnitude $(O(\varepsilon^{3/2}))$ of the diffusion approximation on the (ζ, w) scale. But, the total mass in the range in (ii) is $O(\sqrt{\varepsilon})$, while that in ranges (i) and (iii) is $O(\varepsilon)$.

4. Analysis of the Main Diffusion Approximation

If we let $m = \zeta/\varepsilon$, $n = w/\sqrt{\varepsilon}$ and use (3.1), (2.1), and (2.2), in (2.3) and (2.4), we obtain to lowest order

$$\mu \frac{\partial^2 \phi_0}{\partial \zeta^2} + \mu\omega \frac{\partial \phi_0}{\partial \zeta} + \kappa \frac{\partial^2 \phi_0}{\partial w^2} - \delta \frac{\partial \phi_0}{\partial w} = 0, \quad \zeta > 0, \ w > 0,$$ (4.1)

and the boundary condition

$$\kappa \frac{\partial \phi_0}{\partial w}(\zeta, 0) = \delta\phi_0(\zeta, 0), \quad \zeta > 0.$$ (4.2)

We will discuss the second boundary condition, along $\zeta = 0$, after (4.16).

We let

$$\phi_0(\zeta, w) = \exp\left(-\frac{w\zeta}{2} + \frac{\delta w}{2\kappa}\right)\Phi_0(\zeta, w). \tag{4.3}$$

It follows that

$$\mu\frac{\partial^2 \Phi_0}{\partial \zeta^2} + \kappa\frac{\partial^2 \Phi_0}{\partial w^2} - \frac{1}{4}\left(\mu\omega^2 + \frac{\delta^2}{\kappa}\right)\Phi_0 = 0, \quad \zeta > 0, \; w > 0, \tag{4.4}$$

$$\frac{\partial \Phi_0}{\partial w}(\zeta, 0) = \frac{\delta}{2\kappa}\Phi_0(\zeta, 0). \tag{4.5}$$

To solve (4.4) and (4.5) we apply a transform in the w variable. Using the theory of distributions and Green's functions for ordinary differential equations (see [23, p. 294, exercise 4.24]), we have the following transform pair:

$$G(B) = \int_0^\infty [B\cos(Bx) + A\sin(Bx)]F(x)dx, \tag{4.6}$$

$$F(x) = -I(A < 0)CAe^{Ax} + \frac{2}{\pi}\int_0^\infty \frac{B\cos(Bx) + A\sin(Bx)}{A^2 + B^2}G(B)dB. \tag{4.7}$$

Here, the constant C appears only when $A < 0$ and the term Ae^{Ax} corresponds to a single discrete eigenvalue in the spectral theory. By multiplying (4.7) by e^{Ax} and integrating from $x = 0$ and $x = \infty$, we find that $C = 2\int_0^\infty F(x)e^{Ax}dx$. Applying (4.6) with $B = \beta$ and $x = w$, we let

$$\Omega_0(\zeta, \beta) = \int_0^\infty \left[\beta\cos(w\beta) + \frac{\delta}{2\kappa}\sin(\beta w)\right]\Phi_0(\zeta, w)dw, \tag{4.8}$$

then integration by parts and the boundary condition (4.5) leads to

$$\int_0^\infty \left[\beta\cos(\beta w) + \frac{\delta}{2\kappa}\sin(\beta w)\right]\frac{\partial^2 \Phi_0}{\partial w^2}(\zeta, w)dw = -\beta^2\Omega_0(\zeta, \beta). \tag{4.9}$$

Hence, from (4.4),

$$\frac{\partial^2 \Omega_0}{\partial \zeta^2} - \frac{1}{4}\left(\omega^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa\beta^2}{\mu}\right)\Omega_0 = 0, \tag{4.10}$$

so that

$$\Omega_0(\zeta, \beta) = \Omega_0(0, \beta)\exp\left[-\sqrt{\omega^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa\beta^2}{\mu}}\frac{\zeta}{2}\right]. \tag{4.11}$$

Applying the inversion formula (4.7)

$$\Phi_0(\zeta, u) + I(\delta < 0)\frac{\delta}{\kappa}\exp\left(\frac{\delta u}{2\kappa}\right)\int_0^\infty \exp\left(\frac{\delta w}{2\kappa}\right)\Phi_0(\zeta, w)dw$$

$$= \frac{2}{\pi}\int_0^\infty \frac{[\beta\cos(u\beta) + (\delta/2\kappa)\sin(u\beta)]}{(\beta^2 + \delta^2/(4\kappa^2))}\Omega_0(\zeta, \beta)d\beta. \tag{4.12}$$

Let

$$P_0(\zeta) = \int_0^\infty \phi_0(\zeta, w)dw. \tag{4.13}$$

Then, from (4.1) and (4.2),

$$\frac{d^2 P_0}{d\zeta^2} + \omega\frac{dP_0}{d\zeta} = 0, \quad \zeta > 0. \tag{4.14}$$

But, from the normalization condition (2.7), (3.1) and the Euler-Maclaurin summation formula [26], $\int_0^\infty P_0(\zeta)d\zeta = 1$, so that

$$P_0(\zeta) = \omega e^{-\omega\zeta}. \tag{4.15}$$

Hence, from (4.3) and (4.13),

$$\int_0^\infty \exp\left(\frac{\delta w}{2\kappa}\right)\Phi_0(\zeta, w)dw = \omega e^{-\omega\zeta/2}. \tag{4.16}$$

Also, if we use (3.1), (2.1), and (2.2), in (2.5), we obtain the lowest order boundary condition $\partial\phi_0/\partial w(0, w) = 0$, for $w > 0$. We conclude that $\phi_0(0, w)$ and $\Phi_0(0, w)$ are proportional to a delta function at $w = 0^+$ and hence, from (4.8) and (4.16), that $\Omega_0(0, \beta) = \omega\beta$. Proposition 3.1 follows from (4.3), (4.11), (4.12), and (4.15).

We may rewrite the integral in (3.2) as

$$\frac{1}{2}\int_{-\infty}^\infty \frac{\beta e^{iw\beta}}{(\beta + i\delta/(2\kappa))}\exp\left[-\sqrt{\omega^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa\beta^2}{\mu}}\frac{\zeta}{2}\right]d\beta. \tag{4.17}$$

We then deform the contour of integration to one around a cut in the β-plane from $(i/2\kappa)\sqrt{\delta^2 + \kappa\mu\omega^2}$ to $i\infty$, and let

$$\beta = i\left(\frac{1}{2\kappa}\sqrt{\delta^2 + \kappa\mu\omega^2} + y\right). \tag{4.18}$$

For $\delta < 0$ there is a contribution from the pole at $\beta = -i\delta/(2\kappa)$, and we obtain

$$
\phi_0(\zeta, w) = \frac{2w}{\pi} \exp\left(-\frac{w\zeta}{2} + \frac{\delta w}{2\kappa}\right) \exp\left(-\frac{w}{2\kappa}\sqrt{\delta^2 + \kappa\mu\omega^2}\right)
$$

$$
\times \int_0^\infty \frac{\left(\sqrt{\delta^2 + \kappa\mu\omega^2} + 2\kappa y\right)e^{-wy}}{\left(\delta + \sqrt{\delta^2 + \kappa\mu\omega^2} + 2\kappa y\right)} \sin\left(\sqrt{y}\sqrt{\kappa y + \sqrt{\delta^2 + \kappa\mu\omega^2}}\,\frac{\zeta}{\sqrt{\mu}}\right) dy. \tag{4.19}
$$

For $\delta = O(1)$, $w \gg 1$ and $0 < \zeta/\sqrt{w} \ll 1$, the main contribution to the integral comes from $y = O(1/w)$, and Corollary 3.2 follows.

If $-\delta \gg 1$, $w > 0$ and $0 < \zeta = O(1)$, then the integral in (4.19) is approximated by

$$
\frac{|\delta|}{2\kappa} \int_0^\infty \frac{e^{-wy}}{(y + \mu\omega^2/4|\delta|)} \sin\left(\zeta\sqrt{\frac{|\delta|y}{\mu}}\right) dy = \frac{|\delta|}{\kappa} \int_0^\infty \frac{x\exp(-(\mu w/|\delta|)x^2)}{(x^2 + \omega^2/4)} \sin(\zeta x) dx, \tag{4.20}
$$

which leads [27] to Corollary 3.3. Corollaries 3.4 and 3.5 follow from the asymptotic approximation [25] $\mathrm{Erfc}(z) \sim \exp(-z^2)/(\sqrt{\pi}z)$, $z \gg 1$, and the limiting value $\mathrm{Erfc}(-\infty) = 2$. If $\delta \gg 1$, $0 < w/\delta = O(1)$ and $0 < \zeta = O(1)$, then from (4.19), we obtain the approximation

$$
\phi_0(\zeta, w) \sim \frac{w}{\pi} \exp\left(-\frac{w\zeta}{2} - \frac{\mu\omega^2}{4\delta}w\right) \int_0^\infty e^{-wy} \sin\left(\sqrt{\frac{\delta y}{\mu}}\,\zeta\right) dy, \tag{4.21}
$$

and hence Corollary 3.7.

Proposition 3.9 follows from Proposition 3.1, (3.1) and the Euler-Maclaurin summation formula [26]. Next, from (4.19)

$$
\int_0^\infty \phi_0(\zeta, w) d\zeta = \frac{8w}{\pi}\sqrt{\mu} \exp\left(\frac{\delta w}{2k}\right) \exp\left(-\frac{w}{2\kappa}\sqrt{\delta^2 + \kappa\mu\omega^2}\right)
$$

$$
\times \int_0^\infty \frac{\left(\sqrt{\delta^2 + \kappa\mu\omega^2} + 2\kappa y\right)\sqrt{y}\sqrt{\kappa y + \sqrt{\delta^2 + \kappa\mu\omega^2}}e^{-wy}}{\left(\delta + \sqrt{\delta^2 + \kappa\mu\omega^2} + 2\kappa y\right)\left[\mu\omega^2 + 4y\left(\kappa y + \sqrt{\delta^2 + \kappa\mu\omega^2}\right)\right]} dy. \tag{4.22}
$$

For $\delta = O(1)$ and $w \gg 1$, the main contribution to the integral comes from $y = O(1/w)$, and Corollary 3.10 follows. If $-\delta \gg 1$ and $0 < w/|\delta| = O(1)$, then

$$
\int_0^\infty \phi_0(\zeta, w) d\zeta \sim \frac{w\sqrt{\mu|\delta|}}{\pi\kappa} \exp\left[-\frac{w}{2\kappa}\left(\sqrt{\delta^2 + \kappa\mu\omega^2} + |\delta|\right)\right] \int_0^\infty \frac{\sqrt{y}e^{-wy}}{(y + \mu\omega^2/(4|\delta|))^2} dy, \tag{4.23}
$$

which leads [27] to Corollary 3.11. Corollaries 3.12, 3.13, and 3.14 follow directly from Corollaries 3.4, 3.5, and 3.7, respectively.

We now consider the first order correction term $\phi_1(\zeta, w)$ in (3.1). If we use (3.1), (2.1), and (2.2), in (2.3) and (2.4), we obtain

$$\mu \frac{\partial^2 \phi_1}{\partial \zeta^2} + \mu \omega \frac{\partial \phi_1}{\partial \zeta} + \kappa \frac{\partial^2 \phi_1}{\partial w^2} - \delta \frac{\partial \phi_1}{\partial w} + \delta \left(\mu \frac{\partial \phi_0}{\partial \zeta} + \frac{1}{2} \frac{\partial^2 \phi_0}{\partial w^2} \right) = 0, \tag{4.24}$$

for $\zeta > 0$ and $w > 0$, and the boundary condition

$$\kappa \frac{\partial \phi_1}{\partial w}(\zeta, 0) - \delta \phi_1(\zeta, 0) + \mu \frac{\partial^2 \phi_0}{\partial \zeta^2}(\zeta, 0) + \mu(\omega + \kappa) \frac{\partial \phi_0}{\partial \zeta}(\zeta, 0) + \frac{\kappa}{2} \frac{\partial^2 \phi_0}{\partial w^2}(\zeta, 0) = 0. \tag{4.25}$$

We let

$$P_1(\zeta) = \int_0^\infty \phi_1(\zeta, w) dw. \tag{4.26}$$

It follows, from (4.15) and (4.24)–(4.26), that

$$\mu \frac{d^2 P_1}{d\zeta^2} + \mu \omega \frac{dP_1}{d\zeta} + \mu \frac{\partial^2 \phi_0}{\partial \zeta^2}(\zeta, 0) + \mu(\omega + \kappa) \frac{\partial \phi_0}{\partial \zeta}(\zeta, 0) + \frac{\kappa}{2} \frac{\partial^2 \phi_0}{\partial w^2}(\zeta, 0) - \frac{\delta}{2} \frac{\partial \phi_0}{\partial w}(\zeta, 0) = \delta \mu \omega^2 e^{-\omega \zeta}. \tag{4.27}$$

Hence, from (4.1) at $w = 0^+$,

$$P_1(\zeta) = Q_1(\zeta) - \frac{1}{2} \phi_0(\zeta, 0), \tag{4.28}$$

where

$$\frac{d^2 Q_1}{d\zeta^2} + \omega \frac{dQ_1}{d\zeta} = \delta \omega^2 e^{-\omega \zeta} - \kappa \frac{\partial \phi_0}{\partial \zeta}(\zeta, 0). \tag{4.29}$$

From Proposition 3.1,

$$\phi_0(\zeta, 0) + I(\delta < 0) \frac{\delta \omega}{\kappa} e^{-\omega \zeta}$$

$$= \frac{2\omega}{\pi} \exp \left(-\frac{\omega \zeta}{2} \right) \int_0^\infty \frac{\beta^2}{(\beta^2 + \delta^2/(4\kappa^2))} \exp \left[-\sqrt{\omega^2 + \frac{\delta^2}{\kappa \mu} + \frac{4\kappa \beta^2}{\mu}} \frac{\zeta}{2} \right] d\beta. \tag{4.30}$$

It follows that

$$
\frac{d^2 Q_1}{d\zeta^2} + \omega \frac{dQ_1}{d\zeta} - I(\delta > 0)\delta\omega^2 e^{-\omega\zeta}
$$
$$
= \frac{\omega\kappa}{\pi} e^{-\omega\zeta/2} \int_0^\infty \frac{\beta^2 \left[\omega + \sqrt{\omega^2 + \delta^2/(\kappa\mu) + 4\kappa\beta^2/\mu}\right]}{(\beta^2 + \delta^2/(4\kappa^2))} \exp\left[-\sqrt{\omega^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa\beta^2}{\mu}}\,\frac{\zeta}{2}\right] d\beta. \tag{4.31}
$$

From (3.1), (4.13), (4.26), (4.28), and the Euler-Maclaurin summation formula [26],

$$
\sum_{n=0}^\infty p\left(\frac{\zeta}{\varepsilon}, n\right) = \varepsilon P_0(\zeta) + \varepsilon^{3/2} Q_1(\zeta) + O\left(\varepsilon^2\right). \tag{4.32}
$$

Hence,

$$
\int_0^\infty Q_1(\zeta) d\zeta = 0. \tag{4.33}
$$

Proposition 3.16 follows in an elementary manner from (4.31) and (4.33).

If we let $2\kappa y = \sqrt{\delta^2 + \kappa\mu\omega^2}(\cosh u - 1)$ in (4.22), we obtain

$$
\int_0^\infty \phi_0(\zeta, w) d\zeta = \frac{2\omega}{\pi}\sqrt{\frac{\mu}{\kappa}} \int_0^\infty \frac{\sinh^2 u \cosh u \exp\left[-(w/2\kappa)\sqrt{\delta^2 + \kappa\mu\omega^2}(\cosh u - c)\right]}{(\cosh u + c)\left(\cosh^2 u - c^2\right)} du, \tag{4.34}
$$

where c is given by (3.21). Hence,

$$
\int_0^\infty \int_0^\infty w\phi_0(\zeta, w) d\zeta dw = \frac{8\delta\left(1 - c^2\right)^{3/2}}{\pi\mu\omega^2 c} \int_0^\infty \frac{\sinh^2 u \cosh u}{(\cosh u + c)^2 (\cosh u - c)^3} du. \tag{4.35}
$$

The evaluation of the integral in (4.35) is routine, but tedious, and Proposition 3.19 follows from (3.1), since

$$
\mathbf{E}(N_2) \sim \frac{1}{\sqrt{\varepsilon}} \int_0^\infty \int_0^\infty w\phi_0(\zeta, w) d\zeta dw. \tag{4.36}
$$

We next establish the various asymptotic formulas in Proposition 3.22. We first note that the integrand in Proposition 3.1 is an even function of β, and if β is viewed as complex it has a simple pole at $\beta = -i\delta/(2\kappa)$ and branch points at

$$
\beta = \pm i\frac{\sqrt{\mu}}{2\sqrt{\kappa}}\sqrt{\omega^2 + \frac{\delta^2}{\kappa\mu}} \equiv ib_\pm. \tag{4.37}
$$

Then, we can represent, for any δ, ϕ_0 as the contour integral

$$\phi_0(\zeta, w) = \exp\left[-\frac{w}{2}\zeta + \frac{\delta}{2\kappa}w\right]\frac{1}{\pi}\int_C \frac{\beta w}{\beta + i\delta/2\kappa}$$

$$\times \exp\left[iw\beta - \frac{1}{2}\zeta\sqrt{w^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa}{\mu}\beta^2}\right]d\beta. \tag{4.38}$$

Here, C is a horizontal contour in the β-plane, on which

$$\max\left\{-\frac{\delta}{2\kappa}, 0\right\} < \text{Im}(\beta) < b_+. \tag{4.39}$$

The condition in (4.39) insures that if $\delta < 0$ the pole at $\beta = i|\delta|/(2\kappa)$ lies below the contour C.

If $\delta > 0$ we can shift the contour to the real β-axis, and then (4.38) becomes the same as (3.2). If $\delta < 0$ the pole must be taken into account in making this shift, and the residue from this pole yields the exponential terms in the left hand side of (3.2).

To evaluate (4.38) for $\zeta, w \to \infty$ we employ the saddle point method. There is a saddle point where

$$\frac{d}{d\beta}\left[iw\beta - \frac{1}{2}\zeta\sqrt{w^2 + \frac{\delta^2}{\kappa\mu} + \frac{4\kappa}{\mu}\beta^2}\right] = 0, \tag{4.40}$$

so that

$$\beta = ib_s = ib_s\left(\frac{w}{\zeta}\right) \equiv i\frac{\mu}{2\kappa}\sqrt{w^2 + \frac{\delta^2}{\kappa\mu}}\frac{w}{\sqrt{(\mu/\kappa)w^2 + \zeta^2}}. \tag{4.41}$$

The saddle is on the imaginary axis and the directions of steepest descent are $\arg(\beta - ib_s) = 0, \pi$. Then shifting the contour C into another horizontal contour through ib_s leads to (3.24). Such a shift is always permissible if $\delta > 0$, but if $\delta < 0$ we need the saddle to lie above the pole, that is, $b_s > |\delta|/(2\kappa)$, and this occurs precisely when $w/\zeta > |\delta|/(\mu w)$. We thus obtain the condition in item (i) of Proposition 3.22. If $\delta < 0$ and $w/\zeta < |\delta|/(\mu w)$ the pole dominates the saddle point contribution, and we obtain (3.23).

For $\zeta \to \infty$ and $w = O(1)$ a different analysis is needed, as now $b_s \to 0$ so the saddle approaches the real axis, where the integrand in (4.38) has as simple zero. In this case (which applies only if $\delta > 0$), we shift C back to the real axis and expand the integrand for $\beta \to 0$. Using

$$\frac{\beta w}{\beta + i\delta/(2\kappa)}e^{iw\beta} = \frac{2\beta\kappa w}{i\delta}\left[1 + i\beta\left(w + \frac{2\kappa}{\delta}\right) + O\left(\beta^2\right)\right], \quad \beta \to 0, \tag{4.42}$$

we thus obtain

$$
\phi_0(\zeta, w) \sim \exp\left[-\left(\frac{w}{2} + \frac{1}{2}\sqrt{w^2 + \frac{\delta^2}{\kappa\mu}}\right)\zeta\right] \exp\left(\frac{\delta}{2\kappa}w\right)
$$

$$
\times \frac{2\kappa w}{\pi\delta} \int_{-\infty}^{\infty} \left[\frac{\beta}{i} + \beta^2\left(w + \frac{2\kappa}{\delta}\right) + O(\beta^3)\right] \exp\left[-\frac{\kappa\zeta}{\mu\Delta}\beta^2\right] d\beta,
$$

(4.43)

where $\Delta = \sqrt{w^2 + \delta^2/(\kappa\mu)}$. Evaluating the integral(s) in (4.43) leads to (3.25). If we consider the opposite limit, where $\zeta = O(1)$ and $w \to \infty$, then the saddle point approaches the upper branch point at ib_+. But by deforming $\mathcal{C}$ to an integral about the branch cut we can show that the final result coincides with the expansion of (3.25) for $w/\zeta \gg 1$, which is given by (3.28).

Next, we consider the corner behavior of ϕ_0 as $\zeta, w \to 0$. Now, the main contribution to the integral will come from where $|\beta|$ is large. From (3.2) for $\delta > 0$, we then obtain

$$
\phi_0(\zeta, w) \sim \frac{w}{\pi} \int_{-\infty}^{\infty} e^{iw\beta} e^{-\sqrt{(\kappa/\mu)}\zeta|\beta|} d\beta
$$

$$
= \frac{w}{\pi}\sqrt{\frac{\mu}{\kappa}}\frac{2}{\zeta}\left[1 + \frac{\mu}{\kappa}\frac{w^2}{\zeta^2}\right]^{-1},
$$

(4.44)

which yields (3.30), and this can be shown to remain valid for $\delta < 0$.

Finally, we fix w and let $\zeta \to 0$. Simply, setting $\zeta = 0$ in (3.2) leads to a divergent integral. However, an integration by parts leads to, for $\delta > 0$,

$$
\phi_0(\zeta, w) = \frac{w}{\pi} \exp\left(\frac{\delta}{2\kappa}w\right) \int_{-\infty}^{\infty} \exp\left[iw\beta - \sqrt{\frac{\kappa}{\mu}}\sqrt{\beta^2 + \frac{\mu}{4\kappa}\Delta^2}\zeta\right]\left[-g_0'(\beta)\right] d\beta,
$$

(4.45)

where $\Delta = \sqrt{w^2 + \delta^2/\kappa\mu}$ and

$$
g_0(\beta) = \frac{\beta}{\beta + i\delta/2\kappa}\left[iw - \frac{\beta}{\sqrt{\beta^2 + (\mu/4\kappa)\Delta^2}}\sqrt{\frac{\kappa}{\mu}}\zeta\right]^{-1}.
$$

(4.46)

Expanding (4.46) for $\zeta \to 0$ and noting that, by contour integration (if $\delta > 0$),

$$
\int_{-\infty}^{\infty} e^{iw\beta}\frac{d}{d\beta}\left[\frac{\beta}{\beta + i\delta/(2\kappa)}\right] d\beta = 0,
$$

(4.47)

we write the integrand as

$$e^{iw\beta}\left[1 - \sqrt{\frac{\kappa}{\mu}}\sqrt{\beta^2 + \frac{\mu}{4\kappa}\Delta^2}\zeta + O\left(\zeta^2\right)\right]$$

$$\times\left\{\frac{i}{w}\frac{d}{d\beta}\left(\frac{\beta}{\beta + i\delta/(2\kappa)}\right) + \frac{\zeta}{w^2}\frac{d}{d\beta}\left[\frac{\beta}{\beta + i\delta/(2\kappa)}\frac{\beta}{\sqrt{\beta^2 + (\mu/(4\kappa))\Delta^2}}\right] + O\left(\zeta^2\right)\right\}. \tag{4.48}$$

By using (4.47), identifying the $O(\zeta)$ terms in (4.48), and explicitly performing the differentiation with respect to β, we ultimately obtain (3.31) and (3.32). This completes the (sketched) derivation of Proposition 3.22.

5. Analysis of Boundary and Corner Regions

We analyze cases where $m = o(\varepsilon^{-1})$ and/or $n = o(\varepsilon^{-1/2})$. While these carry mass that is asymptotically small, they must be considered to insure that $p(m, n)$ is properly normalized to higher orders in ε, and to compute higher order approximations to the moments. Also, to determine $\phi_0(\zeta, w)$ we used the boundary condition $\phi_0(0, w) = w\delta(w)$, and analysis of cases where ζ and w are small will allow us to examine this condition more carefully.

First we observe from (3.31) that $\phi_0(\zeta, w)$ vanishes linearly as $\zeta \to 0$, which indicates a nonuniformity in the asymptotics. We first consider the scale $m = O(1)$ with $w > 0$ and set

$$p(m, n) = \varepsilon^{\nu_1}\tilde{p}(m, w; \varepsilon) \sim \varepsilon^{\nu_1}\tilde{p}(m, w). \tag{5.1}$$

Here, ν_1 is a constant that will be determined by asymptotic matching. From (2.3), in terms of the variables m and w, we have

$$\left[\lambda + \mu + \varepsilon(\sigma + \kappa - \kappa\mu)\right]\tilde{p}(m, w; \varepsilon) = \lambda\tilde{p}(m - 1, w; \varepsilon)$$

$$+ \varepsilon\sigma\tilde{p}(m, w - \sqrt{\varepsilon}; \varepsilon)$$

$$+ \mu(1 - \varepsilon\kappa)\tilde{p}(m + 1, w; \varepsilon)$$

$$+ \varepsilon\kappa\tilde{p}(m, w + \sqrt{\varepsilon}; \varepsilon). \tag{5.2}$$

Thus, the leading term must satisfy, since $\lambda = \mu + O(\varepsilon)$, $2\mu\tilde{p}(m, w) = \mu\tilde{p}(m - 1, w) + \mu\tilde{p}(m + 1, w)$, and we write $\tilde{p}$ as

$$\tilde{p}(m, w) = m\tilde{p}_+(w) + \tilde{p}_-(w), \tag{5.3}$$

which is a linear function of m. The expansion for $m = O(1)$ must satisfy the boundary condition (2.5), which is

$$(\lambda + \varepsilon\sigma + 1)\tilde{p}(0, w; \varepsilon) = \varepsilon\sigma\tilde{p}(0, w - \sqrt{\varepsilon}; \varepsilon) + \mu(1 - \varepsilon\kappa)\tilde{p}(1, w; \varepsilon) + \tilde{p}(0, w + \sqrt{\varepsilon}; \varepsilon). \tag{5.4}$$

To leading order, (5.4) implies that $\tilde{p}(0, w) = \tilde{p}(1, w)$ so that (5.3) becomes $\tilde{p}(m, w) = \tilde{p}_-(w)$. But, as $m \to \infty, \varepsilon^{\nu_1} \tilde{p}_-(w)$ cannot match to $\varepsilon^{3/2} \phi_0(\zeta, w)$, since ϕ_0 vanishes as $\zeta = m\varepsilon \to 0$. This indicates that we must analyze another scale, which has $m \gg 1$ and $m \ll 1/\varepsilon$ (with w fixed).

We thus set $m = S/\sqrt{\varepsilon}$ and

$$p(m, n) = \varepsilon^{\nu_2} P(S, w; \varepsilon) \sim \varepsilon^{\nu_2} P(S, w). \tag{5.5}$$

By rewriting (2.3) on the (S, w) scale, we obtain

$$\mu \left[1 - \varepsilon(w + \kappa) - \varepsilon^{3/2} \delta \right] \left[P(S, w; \varepsilon) - P(S - \sqrt{\varepsilon}, w; \varepsilon) \right]$$

$$= \varepsilon \sigma \left[P(S, w - \sqrt{\varepsilon}; \varepsilon) - P(S, w; \varepsilon) \right]$$

$$+ \varepsilon \kappa \left[P(S + \sqrt{\varepsilon}, w; \varepsilon) - P(S, w; \varepsilon) \right] \tag{5.6}$$

$$+ \mu(1 - \varepsilon \kappa) \left[P(S + \sqrt{\varepsilon}, w; \varepsilon) - P(S, w; \varepsilon) \right].$$

The limiting form of (5.6) as $\varepsilon \to 0$ is $P_{SS} = 0$, so we write

$$P(S, w) = S P_+(w) + P_-(w). \tag{5.7}$$

On the (S, w) scale the boundary condition (2.5), using also $\lambda = \mu[1 - \varepsilon(w + \sigma)]$ and $\sigma = \kappa + \sqrt{\varepsilon} \delta$, becomes

$$[\mu + 1 + \varepsilon(\sigma - \sigma\mu - \omega\mu)] P(0, w; \varepsilon) = \varepsilon \sigma P(0, w - \sqrt{\varepsilon}; \varepsilon)$$

$$+ P(0, w + \sqrt{\varepsilon}; \varepsilon) \tag{5.8}$$

$$+ \mu(1 - \varepsilon\kappa) P(\sqrt{\varepsilon}, w; \varepsilon).$$

To leading order as $\varepsilon \to 0$, (5.8) implies that

$$P_w(0, w) + \mu P_S(0, w) = 0. \tag{5.9}$$

Combining (5.7) and (5.9), we thus write the approximation on the (S, w) scale as

$$p(m, n) \sim \varepsilon^{\nu_2} \left[S P_+(w) + \mu \int_w^\infty P_+(u) du \right]. \tag{5.10}$$

We then determine ν_2 and $P_+(w)$ by asymptotically matching (5.10), as $S \to \infty$, to $\varepsilon^{3/2} \phi_0(\zeta, w)$ as $\zeta \to 0$. Noting that $S = \zeta/\sqrt{\varepsilon}$, we find, in view of (3.31), that

$$\nu_2 = 2, \qquad P_+(w) = \frac{\omega}{\pi} \sqrt{\frac{\kappa}{\mu}} \exp\left(\frac{\delta w}{2\kappa} \right) a(w). \tag{5.11}$$

This establishes (3.34) in Proposition 3.24. Expression (5.10) remains valid for $m = O(1)$ as then $S = \sqrt{\varepsilon}m \to 0$ and we obtain $p(m,n) \sim \varepsilon^2 \mu \int_w^\infty P_+(u)du$, which is independent of m and consistent with our previous analysis for $m = O(1)$.

We consider some limiting cases of the (S, w) scale result in (3.34). As $\delta \to \infty$ from (3.28) we obtain $b_+ \sim \delta/(2\kappa)$ and then (3.32) yields, after some computations,

$$a(w) \sim \exp\left(-\frac{w\delta}{2\kappa}\right)\frac{\sqrt{2\pi}}{w^{3/2}}\frac{1}{2}\sqrt{\frac{\delta}{2\kappa}}; \quad \delta \longrightarrow +\infty, \ w = O(1). \tag{5.12}$$

It follows from (3.34) that

$$P_+(w) \sim \frac{w}{2\sqrt{\pi}}\sqrt{\frac{\delta}{\kappa}}\frac{1}{w^{3/2}}, \quad P_-(w) \sim \frac{w}{\sqrt{\pi}}\sqrt{\delta\mu}\frac{1}{\sqrt{w}}; \quad \delta \longrightarrow +\infty, \ w = O(1). \tag{5.13}$$

The above results are consistent with those of Morrison in [16] for the case $\sigma > \kappa$, where it was shown that

$$\sum_{m=0}^\infty p(m,n) \sim \frac{\varepsilon w\sqrt{\mu}}{\sqrt{\pi(\sigma - \kappa)\varepsilon n}} \sim \frac{p(m,n)}{\varepsilon(\sigma - \kappa)}; \quad \sigma > \kappa, \ n \gg 1. \tag{5.14}$$

Hence, for $\delta = (\sigma - \kappa)/\sqrt{\varepsilon} \to \infty$ our approximation $p(m,n) \sim \varepsilon^2 P_-(w)$ agrees with (5.14).

We will next consider the scale $m, n = O(1)$ and matching this to the (S, w) (or (m, w)) scale(s) will require the behavior of (3.34) as $w \to 0$. For $w \to 0$ the second integral in the definition of $a(w)$ in (3.32) dominates, and we obtain

$$a(w) \sim \frac{2}{w^2}, \quad w \longrightarrow 0^+, \tag{5.15}$$

so that

$$P_+(w) \sim \frac{w}{\pi}\sqrt{\frac{\kappa}{\mu}}\frac{2}{w^2}, \quad w \longrightarrow 0^+. \tag{5.16}$$

Also, $P_-(w)$ will be less singular ($O(w^{-1})$) than $P_+(w)$, and, since $w = n/\sqrt{\varepsilon}$, for $w \to 0^+$ the approximation $\varepsilon^2[\sqrt{\varepsilon}mP_+(w) + P_-(w)]$ becomes of order $O(\varepsilon^{3/2})$ on the (m, n) scale.

On the scale $m, n = O(1)$ we will show that $p(m, 0)$ is asymptotically larger than $p(m, n)$ for $n \geqslant 1$, which we just inferred to be $O(\varepsilon^{3/2})$. We thus set

$$p(m,n) = \begin{cases} \varepsilon^{3/2}Q(m,n) + O(\varepsilon^2), & n \geqslant 1, \\ \varepsilon Q_*(m) + \varepsilon^{3/2}Q_*^{(1)}(m) + O(\varepsilon^2), & n = 0. \end{cases} \tag{5.17}$$

On the (m, n) scale, the limiting form of (2.3), with (5.17), becomes

$$2\mu Q(m,n) = \mu Q(m-1,n) + \mu Q(m+1,n), \tag{5.18}$$

while the boundary condition (2.5) yields

$$(\mu + 1)Q(0, n) = \mu Q(1, n) + Q(0, n + 1), \quad n \geqslant 2. \tag{5.19}$$

Hence, $Q(m, n)$ is linear in m, and writing

$$Q(m, n) = Q_+(n) \cdot m + Q_-(n), \tag{5.20}$$

we then obtain from (2.5), $Q_-(n) - Q_-(n + 1) = \mu Q_+(n)$ which we sum to get

$$Q_-(n) = \mu \sum_{j=n}^{\infty} Q_+(j). \tag{5.21}$$

Next, we consider the boundary condition (2.4) and the corner condition (2.6). Since $\lambda = \mu + O(\varepsilon)$, we find that

$$2\mu Q_*(m) = \mu Q_*(m - 1) + \mu Q_*(m + 1), \tag{5.22}$$

and $Q_*^{(1)}(m)$ also satisfies (5.22). Hence, we have

$$Q_*(m) = \gamma + \gamma_* m, \tag{5.23}$$

and $Q_*^{(1)}(m) = \gamma^{(1)} + \gamma_*^{(1)} m$. The corner condition (2.6), with (5.17), yields to leading order $(O(\varepsilon))$

$$\mu Q_*(0) = \mu Q_*(1), \tag{5.24}$$

and at the next order $(O(\varepsilon^{3/2}))$

$$\mu Q_*^{(1)}(0) = \mu Q_*^{(1)}(1) + Q(0, 1). \tag{5.25}$$

It follows from (5.24) that $\gamma_* = 0$. Also, (2.5) with $n = 1$ becomes $(\mu + 1)p(0, 1) = \mu p(1, 1) + p(0, 2) + O(\varepsilon p(0, 0))$ so that (5.19) holds to leading order even if $n = 1$. To summarize, we have obtained

$$p(m, n) \sim \begin{cases} \varepsilon^{3/2} \left[m Q_+(n) + \mu \sum_{j=n}^{\infty} Q_+(j) \right], & n \geqslant 1, \\ \varepsilon \gamma + O(\varepsilon^{3/2}), & n = 0. \end{cases} \tag{5.26}$$

It remains to determine $Q_+(n)$ and the constant γ. By letting $m = S/\sqrt{\varepsilon}$ and $n \to \infty$ and matching to the (S, w) scale, we conclude that as $n \to \infty$

$$Q_+(n) \sim \frac{w}{\pi}\sqrt{\frac{\kappa}{\mu}}\frac{2}{n^2}, \qquad Q_-(n) \sim \frac{w}{\pi}\sqrt{\mu\kappa}\frac{2}{n}. \tag{5.27}$$

While these relations are consistent with (5.21), they do not determine $Q_+(n)$ for $n = O(1)$.

We thus examine another scale, that has $m = S/\sqrt{\varepsilon} = O(\varepsilon^{-1/2})$ and $n = O(1)$. By matching to (5.26) as $m \to \infty$, we expect that $p(m, n)$ will be $O(\varepsilon)$ on this scale, so we set

$$p(m, n) \sim \varepsilon\overline{Q}(S, n). \tag{5.28}$$

Then, from (2.3), we obtain to leading order

$$\mu\overline{Q}_{SS}(S, n) + \kappa\left[\overline{Q}(S, n+1) + \overline{Q}(S, n-1) - 2\overline{Q}(S, n)\right] = 0, \quad n \geqslant 1, \tag{5.29}$$

while the boundary condition (2.4) leads to

$$\mu\overline{Q}_{SS}(S, 0) + \kappa\left[\overline{Q}(S, 1) - \overline{Q}(S, 0)\right] = 0. \tag{5.30}$$

Thus (5.29) corresponds to driftless diffusion in the S variable and a driftless random walk in the discrete n variable. We may replace (5.30) by the "artificial" boundary condition $\overline{Q}(S, -1) = \overline{Q}(S, 0)$, by extending (5.29) to hold also at $n = 0$.

To analyze (5.29) consider the discrete Green's function problem

$$G(n+1) + G(n-1) + (v-2)G(n) = -\delta(n, n_0),$$
$$G(-1) = G(0), \tag{5.31}$$

where v is a "spectral" parameter. This is easily solved to yield

$$G = G(n; n_0) = \frac{1}{\sqrt{v(v-4)}}\left[A^{|n-n_0|} + A^{n+n_0+1}\right],$$
$$A = A(v) = \frac{1}{2}\left[2 - v - \sqrt{v(v-4)}\right]. \tag{5.32}$$

By integrating (5.32) over a loop in the complex v-plane that encircles the branch cut $\mathrm{Re}(v) \in [0, 4]$, we obtain

$$\delta(n, n_0) = \frac{1}{2\pi i}\oint \frac{[A(v)]^{|n-n_0|} + [A(v)]^{n+n_0+1}}{\sqrt{v(v-4)}}dv$$
$$= \frac{1}{\pi}\int_0^\pi \{\cos[(n-n_0)\Omega] + \cos[(n+n_0+1)\Omega]\}d\Omega, \tag{5.33}$$

where we evaluated the contour integral explicitly. Here, $\delta(n, n_0)$ is the Kronecker delta symbol. From (5.33), we infer that for any sequence $\{f(n)\}$

$$f(n) = \frac{2}{\pi} \sum_{L=0}^{\infty} \int_0^{\pi} f(L) \cos\left[\left(n + \frac{1}{2}\right)\Omega\right] \cos\left[\left(L + \frac{1}{2}\right)\Omega\right] d\Omega, \tag{5.34}$$

which leads to the transform pair

$$g(\Omega) = \sum_{L=0}^{\infty} f(L) \cos\left[\left(L + \frac{1}{2}\right)\Omega\right]$$
$$f(L) = \frac{2}{\pi} \int_0^{\pi} g(\Omega) \cos\left[\left(L + \frac{1}{2}\right)\Omega\right] d\Omega. \tag{5.35}$$

Returning to (5.29) we set

$$\overline{Q}(S, n) = \frac{2}{\pi} \int_0^{\pi} q(s, \Omega) \cos\left[\left(n + \frac{1}{2}\right)\Omega\right] d\Omega, \tag{5.36}$$

and applying the transform in (5.35) to (5.29) and (5.30) leads to the ODE

$$q_{SS} + 2\frac{\kappa}{\mu}(\cos\Omega - 1)q = 0. \tag{5.37}$$

Using $\cos\Omega - 1 = 2\sin^2(\Omega/2)$ and requiring also q to decay as $S \to \infty$, we thus have

$$q = f(\Omega) \exp\left[-2\sqrt{\frac{\kappa}{\mu}} \sin\left(\frac{\Omega}{2}\right) s\right], \tag{5.38}$$

and hence

$$\overline{Q}(S, n) = \frac{2}{\pi} \int_0^{\pi} f(\Omega) \cos\left[\left(n + \frac{1}{2}\right)\Omega\right] \exp\left[-2\sqrt{\frac{\kappa}{\mu}} \sin\left(\frac{\Omega}{2}\right) S\right] d\Omega. \tag{5.39}$$

We determine the function $f(\Omega)$ by asymptotic matching.

First we consider the matching of $\varepsilon\overline{Q}(S, n)$ as $S \to \infty$ to $\varepsilon^{3/2}\phi_0(\zeta, 0)$ as $\zeta = \sqrt{\varepsilon}S \to 0$. For $S \to \infty$ the major contribution to the integral in (5.39) will come from near $\Omega = 0$, and the Laplace method yields

$$\overline{Q}(S, n) \sim \frac{2}{\pi} \sqrt{\frac{\mu}{\kappa}} \frac{f(0)}{S}. \tag{5.40}$$

Setting $w = 0$ in (3.2) and letting $\zeta \to 0$, or, equivalently, using (3.30), we conclude that $f(0) = w$. Now, we match the (S, n) and (m, n) scales.

For $n \geqslant 1$, $\varepsilon \overline{Q}(S, n)$ as $S \rightarrow 0$ should agree with $\varepsilon^{3/2} Q(m, n)$ as $m \rightarrow \infty$, which is possible only if $\overline{Q}(0, n) = 0$. Thus, $f(\Omega)$ in (5.39) must be orthogonal to $\cos[(n + 1/2)\Omega]$ for all $n \geqslant 1$ and hence $f(\Omega) = k \cos(\Omega/2)$. The constant k must equal ω, since $f(0) = \omega$.

Having determined $f(\Omega)$ we proceed to let $S \rightarrow 0$ in (5.39). If $n = 0$, $\overline{Q}(S, 0) \sim 2\omega\pi^{-1} \int_0^\pi \cos^2(\Omega/2) d\Omega = \omega$ so that $\gamma = \omega$ in (5.26). If $n \geqslant 1$

$$
\begin{aligned}
\overline{Q}(S, n) &\sim \frac{2\omega}{\pi}(-2S)\sqrt{\frac{\kappa}{\mu}} \int_0^\pi \cos\left(\frac{\Omega}{2}\right) \sin\left(\frac{\Omega}{2}\right) \cos\left[\left(n + \frac{1}{2}\right)\Omega\right] d\Omega \\
&= \frac{2\omega}{\pi}\sqrt{\frac{\kappa}{\mu}} \frac{S}{(n + 3/2)(n - 1/2)}.
\end{aligned}
\tag{5.41}
$$

Thus, by matching $\varepsilon^{3/2} Q(m, n)$ as $m \rightarrow \infty$ to $\varepsilon \overline{Q}(S, n)$ as $S \rightarrow 0$ we conclude that

$$
Q_+(n) = \frac{2\omega}{\pi}\sqrt{\frac{\kappa}{\mu}} \frac{4}{(2n + 3)(2n - 1)}.
\tag{5.42}
$$

Then, we easily obtain

$$
\sum_{j=n}^\infty Q_+(j) = \frac{2\omega}{\pi}\sqrt{\frac{\kappa}{\mu}} \frac{4n}{4n^2 - 1}.
\tag{5.43}
$$

Now, Q_+, Q_- in (5.20) are determined, as is γ in (5.26), and we have established (3.36) in Proposition 3.24. Also, using $f(\Omega) = \omega \cos(\Omega/2)$ in (5.39) leads to (3.35) in Proposition 3.24.

Thus, we have used a singular perturbation analysis to approximate $p(m, n)$ on scales other than the (ζ, w) scale. While these carry mass that is $o(1)$, the size of $p(m, n)$ may actually be larger than $O(\varepsilon^{3/2})$, as is evident from (3.35).

Acknowledgment

References

[1] I. J.-B. F. Adan, J. Wessels, and W. H. M. Zijm, "Analysis of the symmetric shortest queue problem," *Communications in Statistics. Stochastic Models*, vol. 6, no. 4, pp. 691–713, 1990.

[2] I. J.-B. F. Adan, J. Wessels, and W. H. M. Zijm, "Analysis of the asymmetric shortest queue problem," *Queueing Systems*, vol. 8, no. 1, pp. 1–58, 1991.

[3] L. Flatto and H. P. McKean, "Two queues in parallel," *Communications on Pure and Applied Mathematics*, vol. 30, no. 2, pp. 255–263, 1977.

[4] L. Flatto, "The longer queue problem," *Probability in the Engineering and Informational Sciences*, vol. 3, pp. 537–559, 1989.

[5] L. Flatto, "Two parallel queues created by arrivals with two demands. II," *SIAM Journal on Applied Mathematics*, vol. 45, no. 5, pp. 861–878, 1985.

[6] L. Flatto and S. Hahn, "Two parallel queues created by arrivals with two demands. I," *SIAM Journal on Applied Mathematics*, vol. 44, no. 5, pp. 1041–1054, 1984.

[7] X. Tan and C. Knessl, "A fork-join queueing model: diffusion approximation, integral representations and asymptotics," *Queueing Systems*, vol. 22, no. 3-4, pp. 287–322, 1996.

[8] G. Fayolle and R. Iasnogorodski, "Two coupled processors: the reduction to a Riemann-Hilbert problem," *Zeitschrift für scheinlichkeitstheorie und Verwandte Gebiete*, vol. 47, no. 3, pp. 325–351, 1979.

[9] F. Guillemin and D. Pinchon, "Analysis of generalized processor-sharing systems with two classes of customers and exponential services," *Journal of Applied Probability*, vol. 41, no. 3, pp. 832–858, 2004.

[10] A. G. Konheim, I. Meilijson, and A. Melkman, "Processor-sharing of two parallel lines," *Journal of Applied Probability*, vol. 18, no. 4, pp. 952–956, 1981.

[11] J. W. Cohen and O. J. Boxma, *Boundary Value Problems in Queueing System Analysis*, vol. 79 of *North-Holland Mathematics Studies*, North-Holland, Amsterdam, The Netherlands, 1983.

[12] A. Elwalid and D. Mitra, "Design of generalized processor sharing schedulers which statistically multiplex heterogeneous QoS classes," in *Proceedings of the 18th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '99)*, vol. 3, pp. 1220–1230, New York, NY, USA, 1999.

[13] K. Kumaran, G. Margrave, D. Mitra, and K. Stanley, "Novel techniques for the design and control of generalized processor sharing of multiple QoS classes," in *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '00)*, pp. 932–941, Tel Aviv, Israel, 2000.

[14] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the single-node case," *IEEE/ACM Transactions on Networking*, vol. 1, no. 3, pp. 344–357, 1993.

[15] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the multiple node case," *IEEE/ACM Transactions on Networking*, vol. 2, no. 2, pp. 137–150, 1994.

[16] J. A. Morrison, "Two queues with vastly different arrival rates and processor-sharing factors," *Queueing Systems*, vol. 64, no. 1, pp. 49–67, 2010.

[17] J. A. Morrison and S. C. Borst, "Interacting queues in heavy traffic," *Queueing Systems*, vol. 65, no. 2, pp. 135–156, 2010.

[18] K. Ramanan and M. I. Reiman, "The heavy traffic limit of an unbalanced generalized processor sharing model," *The Annals of Applied Probability*, vol. 18, no. 1, pp. 22–58, 2008.

[19] C. Knessl and J. A. Morrison, "Heavy traffic analysis of two coupled processors," *Queueing Systems*, vol. 43, no. 3, pp. 173–220, 2003.

[20] G. J. Foschini, "Equilibria for diffusion models of pairs of communicating computers—symmetric case," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 273–284, 1982.

[21] J. M. Harrison and M. I. Reiman, "On the distribution of multidimensional reflected Brownian motion," *SIAM Journal on Applied Mathematics*, vol. 41, no. 2, pp. 345–361, 1981.

[22] G. F. Newell, *Approximate Behavior of Tandem Queues*, vol. 171 of *Lecture Notes in Economics and Mathematical Systems*, Springer, Berlin, Germany, 1979.

[23] I. Stakgold, *Boundary Value Problems of Mathematical Physics. Vol. I*, Macmillan, New York, NY, USA, 1967.

[24] J. A. Morrison, "Processor sharing for two queues with vastly different rates," *Queueing Systems*, vol. 57, no. 1, pp. 19–28, 2007.

[25] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics*, vol. 32 of *Die Grundlehren der mathematischen Wissenschaften*, Springer, New York, NY, USA, 3rd edition, 1966.

[26] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Physics*, National Bureau of Standards, Washington, DC, USA, 1964.

[27] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, Boston, Mass, USA, Russian edition, 1994.