Faculteit der Economische Wetenschappen en Econometrie

Serie Research Memoranda (FEWEC)

A Dual Process Associated with the Evolution of the State of a Queueing Network at its Jumps

Richard J. Boucherie

Research Memorandum 1991-39
juni 1991

vrije Universiteit *amsterdam*

A dual process associated with the evolution of the state of a queueing network at its jumps

Richard J. Boucherie*

Abstract

For a product form queueing network, based on the potential interpretation of the transition rates, a dual process, associated with the queueing network at its jumps, is introduced. It is shown that this dual process can be chosen such that it describes the evolution of the states observed by a customer in transit. This gives a new interpretation and generalization of the arrival theorem. As is shown by various examples, the dual process gives insight into the behaviour of product form queueing networks.

Keywords: arrival theorem, dual process, Palm probabilities, product form, queueing network

1 Introduction

Many authors have considered queueing networks that can be modeled as a continuous-time Markov chain with product form equilibrium distribution due to a notion of local balance (cf. [8], [9], [11], [13], [14]). In state $\overline{n}$, describing the number of customers in the queues, a system dependent service rate for leaving queue i, $\frac{\psi(\overline{n}-e_i)}{\phi(\overline{n})}$, is assumed. Upon release customers are routed among the queues according to the routing function, $p_{ij}(\overline{n}-e_i)$. If a solution $\{c_i\}_{i=1}^N$ exists to the traffic equations

$$\sum_j \{\gamma_i p_{ij}(\overline{n}-e_i) - \gamma_j p_{ji}(\overline{n}-e_i)\} = 0, \tag{1.1}$$

expressing local balance, the equilibrium distribution is of product form:

$$\pi(\overline{n}) = B\phi(\overline{n}) \prod_k c_k^{n_k}.$$

The physical model for a transition assumes that in a transition $\overline{n} \to \overline{n} - e_i + e_j$ first a customer leaves queue i and subsequently this customer is routed to queue j. Thus, the transition virtually passes through $\overline{n} - e_i$ and consists of two parts: a service part and a routing part. This decomposition between service and routing is the basis of local balance as expressed in (1.1). In some cases, a queueing network

*Free University, Department of Econometrics, de Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

can also be modeled considering vacancies at the queues, i.e. positions at the queues where no customers are present. If a customer leaves queue i a vacancy is created at queue i. In a transition $\overline{n} \to \overline{n} - e_i + e_j$ first a vacancy is created at queue i and subsequently a vacancy is routed from queue j. Thus, in contrast to the customer process first a unit is created at queue i and the transition $\overline{m} \to \overline{m} + e_i - e_j$, where $\overline{m}$ describes the number of vacancies, virtually passes through $\overline{m} + e_i$. A second process with transitions $\overline{m} \to \overline{m} + e_j - e_k$ that virtually pass through $\overline{m} + e_j$ is the process describing the evolution of states as observed by a customer in transit. To this end, consider the evolution of the states of a queueing network. In a transition $\overline{n} \to \overline{n} - e_i + e_j$ a customer leaves queue i. In transit, this customer observes state $\overline{n} - e_i$. Then the customer is absorbed in queue j. Subsequently a customer may leave queue k. In transit this customer observes state $\overline{n} - e_i + e_j - e_k$. Thus a transition $\overline{n} - e_i \to \overline{n} - e_i + e_j - e_k$ is induced in which first a unit is created at queue j. The stochastic process describing transitions in which first a unit is created is called the dual process. The transition rates for this dual process are obtained from the transition rates of the primal process, the stochastic process describing the original queueing network, based on a potential interpretation of the transition rates partly due to [8], [14]. This approach leads to a generalization of the arrival theorem. In this generalization the process describing the evolution of states observed by customers in transit is given. Moreover, it is argued that a moving customer may influence the equilibrium distribution of the queueing network. Thus, a moving customer will not see the network as if the customer is not present, but this moving customer sees the network that is influenced by its own presence. The equilibrium distribution observed by a moving customer is shown to be the equilibrium distribution of the dual process.

In the examples, the symmetrical relation between the primal an the dual process is shown to be similar to the relation between the primal and dual problem in linear programming thus justifying the name dual process. Furthermore the relation between the transition rates for the primal process and the transition rates for the dual process is discussed. Finally, some comments on batch servicing and routing queueing networks are given. These networks are then used to explain the discrepancy between the equilibrium distribution in a discrete-time and corresponding continuous-time formulation of a queueing network.

This paper is organized as follows. In Section 2 the primal process is discussed. Some well-known results are reformulated in the setting of this paper and the potential interpretation of the transition rates is discussed. Section 3 presents the dual process. The transition rates of the dual process are defined based on the potential interpretation of the transition rates. Furthermore, a sufficient condition for the dual equilibrium distribution to be of product form is given. In Section 4, for a general class of primal processes, the dual routing function is given. Section 5 relates the transition rates and equilibrium distribution of the primal process to those of the dual process. In this section the Palm probabilities associated with the jumps of the primal process are studied. It is shown that moving customers see the dual equilibrium distribution thus generalizing the arrival theorem. Section 6 gives some examples and general remarks on the relation between the primal and the dual process and

Section 7 concludes the paper.

2 Preliminaries

Consider a continuous-time queueing network consisting of N queues, labeled $1,\ldots,N$. Assume that the queueing network can be represented by a stable, regular, continuous-time Markov chain $X = \{X(t),\ t \geq 0\}$ at state space $S \subseteq N_0^N = \{\overline{n} : \overline{n} = (n_1,\ldots,n_N),\ n_i \in N_0 = \{0,1,2,\ldots\},\ i = 1,\ldots,N\}$, where n_i denotes the number of customers at queue i, $i = 1,\ldots,N$. The transition rate from state $\overline{n} \in S$ to state $\overline{n}' \in S$ is denoted by $q(\overline{n},\overline{n}')$. As the Markov chain X describes a queueing network, it is natural to assume that transitions of X correspond to routing of customers, i.e. it is natural to assume that, for $\overline{n},\overline{n}' \in S$

$$q(\overline{n},\overline{n}') = 0 \ \text{ unless } \ \overline{n}' = \overline{n} - e_i + e_j,\ i,j \in \mathcal{N},\ i \neq j, \tag{2.1}$$

where $\mathcal{N} = \{0,\ldots,N\}$ if the queueing network is open and $\mathcal{N} = \{1,\ldots,N\}$ if the queueing network is closed, e_i, $i = 1,\ldots,N$, denotes the i-th unit vector and $e_0 = 0$. In a transition $\overline{n} \to \overline{n} - e_i + e_j$ first a customer leaves queue i and subsequently this customer is routed among the queues. In this transition the customer configuration $\overline{n} - e_i$ remains unchanged. Therefore, the transition can be written as $(\overline{n} - e_i) + e_i \to (\overline{n} - e_i) + e_j$ and can thus be regarded as virtually passing through $\overline{n} - e_i \in N_0^N$, a so-called dual state. The set of dual states plays an important role in the dual process, therefore this set is explicitly defined here.

Definition 2.1 (Dual state space) *For a stochastic process X at state space S with transition rates satisfying (2.1), the dual state space, S^d, is the set*

$$S^d = \{\overline{m} \in N_0^N | \exists i,j \in \mathcal{N},\ i \neq j : \overline{m} + e_i, \overline{m} + e_j \in S,\ q(\overline{m} + e_i, \overline{m} + e_j) > 0\}. \tag{2.2}$$

With this definition, the transition rates for X can now be defined. In agreement with the literature, assume that, for $\psi : S^d \to R^+$, $\phi : S \to R^+$, $p : S \times S \to R^+ \cup \{0\}$, $R^+ = (0,\infty)$, $p_{ij}(\overline{n} - e_i) \equiv p(\overline{n}, \overline{n} - e_i + e_j)$, $i,j \in \mathcal{N}$, the transition rates have the form

$$q(\overline{n}, \overline{n} - e_i + e_j) = \frac{\psi(\overline{n} - e_i)}{\phi(\overline{n})} p_{ij}(\overline{n} - e_i),\ \overline{n}, \overline{n} - e_i + e_j \in S,\ i,j \in \mathcal{N}. \tag{2.3}$$

Remark 2.2 (Transition rates) In the general formulation given above, (2.3) does not give rise to further restrictions on the transition rates. The form (2.3) is chosen in agreement with the literature (cf. [2], [3], [5], [6], [11]) and expresses a decomposition of the transition rates into a service part, ψ/ϕ, and a routing part, p_{ij}. At this moment, $p_{ij}(\overline{m})$ is an arbitrary function such that $p_{ii}(\overline{m}) = 0$ for all $\overline{m} \in S^d$, $i \in \mathcal{N}$. Note that the argument, $\overline{m}$, of $p_{ij}(\overline{m})$ explicitly states that a transition $\overline{m}+e_i \to \overline{m}+e_j$ occurs via the dual state $\overline{m}$. In the sequel, strong restrictions will be imposed on p_{ij} (cf. Remark 2.4). The functions ψ and ϕ are not subject to any further restrictions. □

The following lemma provides a sufficient condition for X to have a product form invariant measure, that is a non-negative solution, $\Phi = (\Phi(\overline{n}),\ \overline{n} \in S)$, to the global balance equations

$$\sum_{\overline{n}' \neq \overline{n}} \{\Phi(\overline{n})q(\overline{n},\overline{n}') - \Phi(\overline{n}')q(\overline{n}',\overline{n})\} = 0,\ \overline{n} \in S. \tag{2.4}$$

As transitions $\overline{n} \to \overline{n} - e_i + e_j$ are allowed only, the global balance equations can be written

$$\sum_{i,j \in \mathcal{N}} \{\Phi(\overline{n})q(\overline{n},\overline{n} - e_i + e_j) - \Phi(\overline{n} - e_i + e_j)q(\overline{n} - e_i + e_j,\overline{n})\} = 0,\ \overline{n} \in S. \tag{2.5}$$

If a solution, Φ, exists to the local balance equations

$$\sum_{j \in \mathcal{N}} \{\Phi(\overline{n})q(\overline{n},\overline{n} - e_i + e_j) - \Phi(\overline{n} - e_i + e_j)q(\overline{n} - e_i + e_j,\overline{n})\} = 0,\ i \in \mathcal{N},\ \overline{n} \in S, \tag{2.6}$$

then Φ is an invariant measure as can easily be seen by summing (2.6) over i. In Lemma 2.3 below, local balance of this form is expressed by (2.9). However, local balance can also be expressed as

$$\sum_{i \in \mathcal{N}} \{\Phi(\overline{n})q(\overline{n},\overline{n} - e_i + e_j) - \Phi(\overline{n} - e_i + e_j)q(\overline{n} - e_i + e_j,\overline{n})\} = 0,\ j \in \mathcal{N},\ \overline{n} \in S. \tag{2.7}$$

If a solution Φ exists to (2.7), then Φ is an invariant measure as can be seen by summing (2.7) over j. Note that a solution to (2.6) does not necessarily correspond to a solution to (2.7). In Section 3, (2.7) expresses local balance for the dual process, whereas in this section (2.6) expresses local balance for the primal process. The result of Lemma 2.3 is well-known. For example, for slightly different forms of the transition rates this result is obtained [3], [5], [8], [9], [14], however, as terms appearing in the proof of this lemma give a first glance at the dual process and the proof shows the symmetry between X and the dual process introduced in Section 3, both Lemma 2.3 and its proof are given here explicitly.

Lemma 2.3 *X allows an invariant measure Φ at S, given by*

$$\Phi(\overline{n}) = \phi(\overline{n}) \prod_{k=1}^{N} c_k^{n_k}, \quad \overline{n} \in S, \tag{2.8}$$

if for all $\overline{n} \in S$ the coefficients $\{c_i\}_{i=1}^{N}$ are a non-negative ($c_i \geq 0$ for all i) solution of

$$\sum_{j \in \mathcal{N}} \{\gamma_i p_{ij}(\overline{n} - e_i) - \gamma_j p_{ji}(\overline{n} - e_i)\} = 0,\ i \in \mathcal{N},\ \gamma_0 = 1. \tag{2.9}$$

Proof Insertion of (2.3) and (2.8) into the global balance equations at S gives for $\overline{n} \in S$

$$\begin{aligned}
&\sum_{\overline{n}' \neq \overline{n}} \{\Phi(\overline{n})q(\overline{n},\overline{n}') - \Phi(\overline{n}')q(\overline{n}',\overline{n})\} \\
&= \sum_{i,j \in \mathcal{N}} \left\{ \phi(\overline{n}) \prod_{k=1}^{N} c_k^{n_k} \frac{\psi(\overline{n}-e_i)}{\phi(\overline{n})} p_{ij}(\overline{n}-e_i) \right. \\
&\qquad \left. - \phi(\overline{n}-e_i+e_j) \prod_{k=1}^{N} c_k^{n_k-\delta_{ki}+\delta_{kj}} \frac{\psi(\overline{n}-e_i)}{\phi(\overline{n}-e_i+e_j)} p_{ji}(\overline{n}-e_i) \right\} \\
&= \sum_{i \in \mathcal{N}} \psi(\overline{n}-e_i) \prod_{k=1}^{N} c_k^{n_k-\delta_{ki}} \sum_{j \in \mathcal{N}} \{c_i p_{ij}(\overline{n}-e_i) - c_j p_{ji}(\overline{n}-e_i)\} \qquad (2.10) \\
&= 0,
\end{aligned}$$

where $\delta_{ij} = 1$ if $i = j$, 0 otherwise. The last equality is obtained from (2.9). □

Remark 2.4 (Blocking) At first glance, (2.9) does not give rise to major restrictions on the routing function p_{ij}. However, note that $p_{ij}(\overline{n}-e_i) \equiv p(\overline{n}, \overline{n}-e_i+e_j)$ and $p : S \times S \to R^+ \cup \{0\}$. Therefore, (2.9) implicitly assumes that, for all j in the summation, $\overline{n}-e_i+e_j \in S$. Moreover, the coefficients $\{c_i\}_{i=1}^N$ are state independent. This restricts blocking to the following cases, where $p_{ij}(\overline{m}) = \lambda_{ij} b_{ij}(\overline{m})$. (i) Reversible blocking: the solution $\{c_i\}_{i=1}^N$ to (2.9) satisfies $c_i\lambda_{ij} = c_j\lambda_{ji}$, $i,j \in \mathcal{N}$ and $b_{ij}(\overline{m}) = b_{ji}(\overline{m})$ for all $\overline{m} \in N_0^N$ such that both $\overline{m}+e_i$ and $\overline{m}+e_j \in S$, where S is an arbitrary subset of N_0^N. (ii) Indicator blocking: the solution $\{c_i\}_{i=1}^N$ to (2.9) satisfies the traffic equations $\sum_{j\in\mathcal{N}} \{c_i\lambda_{ij} - c_j\lambda_{ji}\} = 0$, $i \in \mathcal{N}$ and b_{ij} or S satisfies some very special conditions such as S is coordinate convex, i.e. $S = \{\overline{n} \in N_0^N | \sum_{j=1}^N n_j \leq M\}$ or b_{ij} is such that the stop protocol is satisfied (cf. Example 6.3), i.e. if $S = \{\overline{n} \in N_0^N | n_j \leq N_j,\ j = 1,\ldots,N\}$ and $m_k = N_k$ then $b_{ij}(\overline{m}) = 0$ for all i. Note that these blocking conditions cover a fairly wide class of blocking examples and contain practical blocking examples (cf. [11]). □

The following lemma considers the expected rate of transitions between a pair of queues. A similar result is obtained in [14] for the case of an open Jackson network and in [2] for the case of a reversible process with general transition rates. Here, a non-reversible process is considered with transition rates more general than in [14]. As in the previous lemma, both Lemma 2.5 and its proof contain terms appearing in the dual process. In order to give a proper formulation of this expected value, define $q_{ij} : S \to R^+ \cup \{0\}$, $i,j \in \mathcal{N}$ as

$$q_{ij}(\overline{n}) = q(\overline{n}, \overline{n}-e_i+e_j),\ \overline{n} \in S.$$

Lemma 2.5 *If Φ as given in (2.8) satisfies*

$$\sum_{\overline{n}\in S} \Phi(\overline{n}) = B^{-1} < \infty,$$

then under the distribution

$$\pi(\overline{n}) = B\Phi(\overline{n}),\ \overline{n} \in S,$$

for any $i, j \in \mathcal{N}$ the expected rate of transitions from queue i to queue j is given by

$$\boldsymbol{E}q_{ij} = B \sum_{\overline{m} \in S^d} \Psi(\overline{m}) c_i p_{ij}(\overline{m}), \tag{2.11}$$

where Ψ is defined as

$$\Psi(\overline{m}) = \psi(\overline{m}) \prod_{k=1}^{N} c_k^{m_k},\ \overline{m} \in S^d. \tag{2.12}$$

Proof Direct computation of the expectation value gives for any $i, j \in \mathcal{N}$

$$\begin{aligned}
\boldsymbol{E}q_{ij} &= \sum_{\overline{n} \in S} q(\overline{n}, \overline{n} - e_i + e_j)\pi(\overline{n}) \\
&= \sum_{\overline{n} \in S} \frac{\psi(\overline{n} - e_i)}{\phi(\overline{n})} p_{ij}(\overline{n} - e_i) B \phi(\overline{n}) \prod_{k=1}^{N} c_k^{n_k} \\
&= B \sum_{\overline{n} \in S} \psi(\overline{n} - e_i) \prod_{k=1}^{N} c_k^{n_k - \delta_{ki}} c_i p_{ij}(\overline{n} - e_i) \\
&= B \sum_{\overline{m} \in S_{ij}} \Psi(\overline{m}) c_i p_{ij}(\overline{m}),
\end{aligned}$$

where, for $i, j \in \mathcal{N}$,

$$S_{ij} = \{\overline{m} \in N_0^N | \overline{m} + e_i, \overline{m} + e_j \in S,\ p_{ij}(\overline{m}) > 0\}.$$

Note that, since $q(\overline{m} + e_i, \overline{m} + e_j) > 0$ if and only if $p_{ij}(\overline{m}) > 0$,

$$S^d = \bigcup_{i,j \in \mathcal{N}} S_{ij},$$

which completes the proof. □

Remark 2.6 (Dual process) The function $\Psi : S^d \to R^+ \cup \{0\}$, as defined in (2.12), appears both in the proof of Lemma 2.3 and in Lemma 2.5. The right-hand side of (2.11) gives rise to the following interpretation of Ψ. If Ψ satisfies

$$\sum_{\overline{m} \in S^d} \Psi(\overline{m}) = \left[B^d\right]^{-1} < \infty,$$

then under the distribution π^d at S^d defined by

$$\pi^d(\overline{m}) = B^d \Psi(\overline{m}),\ \overline{m} \in S^d,$$

the right-hand side of (2.11) can be written

$$\frac{B}{B^d} \sum_{\overline{m} \in S^d} c_i p_{ij}(\overline{m}) \pi^d(\overline{m}).$$

(2.11) expresses the probability flow from queue i to queue j and $p_{ij}(\overline{m})$ is the probability that a customer routes from queue i to queue j through $\overline{m}$. π^d can thus be interpreted as the probability that a transition passes through the dual state $\overline{m} \in S^d$, i.e. that a customer in transit observes state $\overline{m} \in S^d$. This will be formalized in Section 5.2. □

The remaining part of this section gives some interpretations of the process X with transition rates (2.3) and invariant measure (2.8). These interpretations will be used when the dual process is introduced. The definition of the dual process in Section 3 is based on the potential interpretation of the transition rates (2.3).

Interpretation 2.7 (Potentials (cf. [8], [14])) In accordance with [8], [14], the transition rates (2.3) can be interpreted as reflecting the potential difference between states $\overline{n}$ and $\overline{n}' = \overline{n} - e_i + e_j$. To this end, define the global or configuration potentials U representing the potential of configurations at S and V representing the potential of configurations at S^d. Note that, in the configuration $\overline{m} \in S^d$ one customer is released into the queueing network, which, in general, may influence the potential of the configuration. The intensity at which a customer is released from queue i in state $\overline{n} \in S$ is a function of the potential difference of the configurations $\overline{n}$ and $\overline{n} - e_i$, i.e. this intensity is a function of $V(\overline{n} - e_i) - U(\overline{n})$. Once a customer is released, the routing of the customer is determined by the potentials of the queues and not by the potential difference of the configurations. To this end, define $D_{ij}(\overline{m})$, the potential difference between queue i and queue j in configuration $\overline{m}$. The potential difference between states $\overline{n}$ and $\overline{n} - e_i + e_j$ is then given by $V(\overline{n} - e_i) - U(\overline{n}) - D_{ij}(\overline{n} - e_i)$. Define

$$\phi(\overline{n}) = \exp\left[-U(\overline{n})\right],\ \overline{n} \in S,$$

$$\psi(\overline{m}) = \exp\left[-V(\overline{m})\right],\ \overline{m} \in S^d,$$

$$p_{ij}(\overline{m}) = \exp\left[D_{ij}(\overline{m})\right],\ \overline{m} \in S^d,\ i,j \in \mathcal{N}, \tag{2.13}$$

then the transition rates have the form (2.3). Furthermore, Φ can now be written

$$\Phi(\overline{n}) = \exp\left[-U(\overline{n}) - \sum_{k=1}^{N} \zeta_k n_k\right],\ \overline{n} \in S,$$

where $\zeta_i = -\log c_i$, $i \in \mathcal{N}$, can be regarded as local or site potentials. However, as is stated in the literature (cf. [14]), this interpretation is valid only if X is reversible since the site potentials must satisfy

$$-\zeta_i + D_{ij}(\overline{m}) = -\zeta_j + D_{ji}(\overline{m}),\ \overline{m} \in S^d,\ i,j \in \mathcal{N}.$$

In Sections 3 and 4 this potential interpretation is used to construct the transition rates for the dual process. □

Interpretation 2.8 (Probabalistic) The transition rates (2.3) can also be interpreted probabilistically. To this end, define for $\overline{m} \in S^d$, $i,j \in \mathcal{N}$

$$\widehat{p}_{ij}(\overline{m}) = \frac{p_{ij}(\overline{m})}{p_i(\overline{m})}, \quad p_i(\overline{m}) = \sum_{j\in\mathcal{N}} p_{ij}(\overline{m}).$$

For $\overline{n}, \overline{n} - e_i + e_j \in S$, (2.3) can be written

$$q(\overline{n}, \overline{n}-e_i+e_j) = \frac{\sum_{i\in\mathcal{N}} \psi(\overline{n} - e_i)p_i(\overline{n} - e_i)}{\phi(\overline{n})} \frac{\psi(\overline{n} - e_i)p_i(\overline{n} - e_i)}{\sum_{i\in\mathcal{N}} \psi(\overline{n} - e_i)p_i(\overline{n} - e_i)} \widehat{p}_{ij}(\overline{n}-e_i). \quad (2.14)$$

As the total rate out of state $\overline{n}$ is given by

$$\sum_{\overline{n}'\neq\overline{n}} q(\overline{n}, \overline{n}') = \sum_{i,j\in\mathcal{N}} \frac{\psi(\overline{n} - e_i)}{\phi(\overline{n})} p_{ij}(\overline{n} - e_i) = \frac{\sum_{i\in\mathcal{N}} \psi(\overline{n} - e_i)p_i(\overline{n} - e_i)}{\phi(\overline{n})},$$

the first term in (2.14) can be interpreted as the total state-dependent service rate in state $\overline{n}$. The second term may now be interpreted as the probability that a customer is released from queue i and $\widehat{p}_{ij}$ is the state-dependent routing probability.

The form (2.3) for the transition rates has been introduced in the literature only recently (cf. [2], [3], [5], [6]). The first three references use the form (2.3), whereas the last reference uses the form

$$q(\overline{n}, \overline{n} - e_i + e_j) = \frac{\psi(\overline{n} - e_i)p_i(\overline{n} - e_i)}{\phi(\overline{n})} \widehat{p}_{ij}(\overline{n} - e_i).$$

The discussion above shows that these forms are equivalent. A similar observation is made in [3]. □

3 The dual process

For the stochastic process at S with transition rates as given in (2.3), this section defines a stochastic process at the dual state space S^d, as defined in Definition 2.1, with invariant measure Ψ, as defined in (2.12), such that the transition rates for this process can immediately be obtained from the transition rates (2.3). To this end, observe that a transition $\overline{n} \to \overline{n} - e_i + e_j$ for X, called the primal process, passes through the state $\overline{n} - e_i \in S^d$. Consider two successive transitions, say $\overline{n} \to \overline{n} - e_i + e_j \to \overline{n} - e_i + e_j - e_k + e_l$. As depicted in the upper half of Figure 1, this sequence of transitions passes through $\overline{n} - e_i$ and $\overline{n} - e_i + e_j - e_k$ and therefore induces a transition from state $\overline{n} - e_i \in S^d$ to state $\overline{n} - e_i + e_j - e_k \in S^d$, a transition for the dual process. The transition rates for the dual process (3.1) are defined by analogy with the transition rates for the primal process (2.3). In a transition $\overline{n} \to \overline{n} - e_i + e_j$ for the primal process X, first the potential difference $V(\overline{n} - e_i) - U(\overline{n})$ must be overcome. Then, in state $\overline{n} - e_i$, a customer is routed among the queues according to the site potential difference $D_{ij}(\overline{n} - e_i)$. A transition $\overline{n} - e_i \to \overline{n} - e_i + e_j - e_k$ for the dual process passes through $\overline{n} - e_i + e_j$. In this transition, first the global potential difference $U(\overline{n} - e_i + e_j) - V(\overline{n} - e_i)$ must be overcome, then, in state $\overline{n} - e_i + e_j$, according to the site potential difference $D^d_{kj}(\overline{n} - e_i + e_j)$ the process reaches state $\overline{n} - e_i + e_j - e_k$. From this construction, the transition rates for the dual process

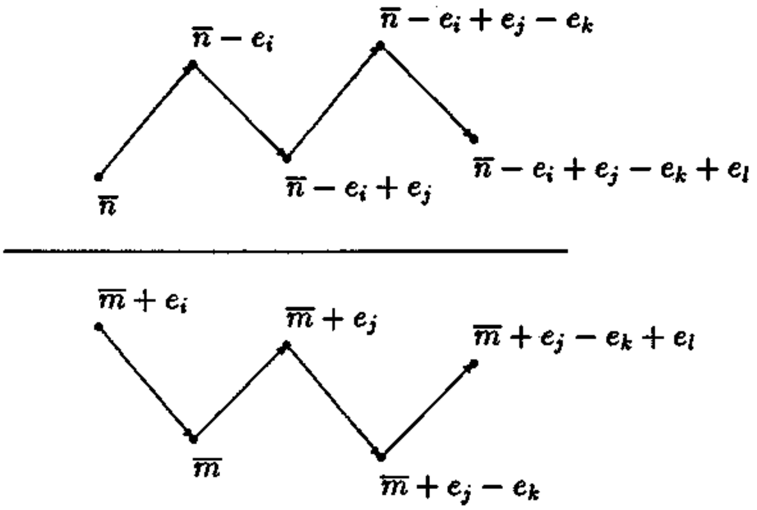

Figure 1. Sequence of states for the primal and the dual process ($\overline{m}=\overline{n}-e_i$).

are given by (3.1) below. The symmetry between the primal process and the dual process is symbolized in Figure 1. The lower half of this figure is a mirror-image of the upper half. Upwards transitions in both the lower and upper half correspond to the global potential difference and downwards transitions correspond to the site potential difference.

Definition 3.1 (Dual process) *For a stochastic process at S with transition rates given in (2.3), the dual process is the process $X^d = \{X^d(t),\ t \geq 0\}$ at S^d with transition rates q^d given by*

$$q^d(\overline{m},\overline{m}') = \begin{cases} \dfrac{\phi(\overline{m}+e_j)}{\psi(\overline{m})} p^d_{ij}(\overline{m}+e_j) & \overline{m},\overline{m}' \in S^d,\ \overline{m}' = \overline{m}+e_j-e_i,\ i,j \in \mathcal{N} \\ 0 & \textit{otherwise,} \end{cases} \tag{3.1}$$

where $p^d_{ij}(\overline{m}+e_j) \equiv p^d(\overline{m},\overline{m}+e_j-e_i)$, $\overline{m} \in S^d$, $i,j \in \mathcal{N}$ and $p^d : S^d \times S^d \to R^+ \cup \{0\}$.

Remark 3.2 (Transitions) In contrast to the primal process, a transition $\overline{m} \to \overline{m}+e_j-e_i$ for the dual process passes through state $\overline{m}+e_j$. This is explicitly visualized in the notation by writing $\overline{m}+e_j-e_i$ instead of $\overline{m}-e_i+e_j$ which is the usual notation for the primal process.

In Definition 3.1, the dual routing function p^d_{ij} is not related to the primal routing function p_{ij}. The relation between p^d_{ij} and p_{ij} depends on the interpretation of the dual process as is shown in Section 4. A few remarks on p^d_{ij} can be made here. Firstly, note that p^d_{ij} is not required to be a probability distribution. Secondly, the dual routing function, in general, does not satisfy $p^d_{jj}(\overline{n}) = 0$. This can easily be seen by observing Figure 1. If for the primal process a customer routes from queue i to queue j and next a customer routes from queue j to queue l, i.e. $k=j$ in Figure 1, then for the dual process a transition $\overline{m} \to \overline{m}$ occurs. Therefore, $q^d(\overline{m},\overline{m}) > 0$ and thus $p_{jj}(\overline{n}) > 0$. □

Similar to Lemma 2.3, a sufficient condition for X^d to possess a product form invariant measure can now be given. The proof of this result is a direct analog to the proof of Lemma 2.3 and shows the symmetry between the primal and dual process.

Lemma 3.3 *X^d allows an invariant measure Ψ at S^d, given by*

$$\Psi(\overline{m}) = \psi(\overline{m}) \prod_{k=1}^{N} d_k^{m_k}, \quad \overline{m} \in S^d, \tag{3.2}$$

if for all $\overline{m} \in S^d$ the coefficients $\{d_i\}_{i=1}^N$ are a positive ($d_i > 0$ for all i) solution of

$$\sum_{i \in \mathcal{N}} \left\{ \frac{1}{\gamma_j} p_{ij}^d(\overline{m} + e_j) - \frac{1}{\gamma_i} p_{ji}^d(\overline{m} + e_j) \right\} = 0, \quad i \in \mathcal{N}, \ \gamma_0 = 1. \tag{3.3}$$

Proof Insertion of (3.1) and (3.2) into the global balance equations at S^d gives for $\overline{m} \in S^d$

$$\begin{aligned}
&\sum_{\overline{m}' \neq \overline{m}} \left\{ \Psi(\overline{m}) q^d(\overline{m}, \overline{m}') - \Psi(\overline{m}') q^d(\overline{m}', \overline{m}) \right\} \\
&= \sum_{i,j \in \mathcal{N}} \Bigg\{ \psi(\overline{m}) \prod_{k=1}^{N} d_k^{m_k} \frac{\phi(\overline{m} + e_j)}{\psi(\overline{m})} p_{ij}^d(\overline{m} + e_j) \\
&\qquad - \psi(\overline{m} + e_j - e_i) \prod_{k=1}^{N} d_k^{m_k + \delta_{kj} - \delta_{ki}} \frac{\phi(\overline{m} + e_j)}{\psi(\overline{m} + e_j - e_i)} p_{ji}^d(\overline{m} + e_j) \Bigg\} \\
&= \sum_{j \in \mathcal{N}} \phi(\overline{m} + e_j) \prod_{k=1}^{N} d_k^{m_k + \delta_{kj}} \sum_{i \in \mathcal{N}} \left\{ \frac{1}{d_j} p_{ij}^d(\overline{m} + e_j) - \frac{1}{d_i} p_{ji}^d(\overline{m} + e_j) \right\} \\
&= 0,
\end{aligned}$$

where the last equality is obtained from (3.3). □

Remark 3.4 (Coefficients $\{d_i\}_{i=1}^N$) The dual routing function p_{ij}^d is not explicitly related to the primal routing function. Therefore, the coefficients $\{d_i\}_{i=1}^N$ for the dual process are not related to the coefficients $\{c_i\}_{i=1}^N$ for the primal process. In Section 4, p_{ij}^d will be chosen such that $d_i = c_i$, $i = 1, \ldots, N$ or $d_i = \frac{1}{c_i}$, $i = 1, \ldots, N$. In the first case the dual process is said to describe customers and in the second case the dual process describes vacancies or holes. □

Similar to Lemma 2.5, the expected rate of transition from queue i to queue j can now be computed for the dual process. To this end, define $q_{ij}^d : S^d \to R^+ \cup \{0\}$, $i, j \in \mathcal{N}$, as

$$q_{ij}^d(\overline{m}) = q^d(\overline{m}, \overline{m} + e_j - e_i), \ \overline{m} \in S^d.$$

Again, the symmetry between the dual process and the primal process is illustrated by Lemma 3.5.

Lemma 3.5 *If Ψ as given in (3.2) satisfies*

$$\sum_{\overline{m}\in S^d} \Psi(\overline{m}) = \left[B^d\right]^{-1} < \infty,$$

then under the distribution

$$\pi^d(\overline{m}) = B^d\Psi(\overline{m}),\ \overline{m} \in S^d,$$

for any $i,j \in \mathcal{N}$, $i \neq j$,

$$\boldsymbol{E}q_{ij}^d = B^d \sum_{\overline{n}\in S^-} \Phi(\overline{n})\frac{1}{d_j}p_{ij}^d(\overline{n}), \tag{3.4}$$

where S^- *is defined as*

$$S^- = \{\overline{n} \in N_0^N | \exists i,j \in \mathcal{N},\ i \neq j : \overline{n} - e_i, \overline{n} - e_j \in S^d,\ q^d(\overline{n} - e_j, \overline{n} - e_i) > 0\}. \tag{3.5}$$

Proof Direct computation of the expectation value gives for any $i,j \in \mathcal{N}$

$$\begin{aligned} \boldsymbol{E}q_{ij}^d &= \sum_{\overline{m}\in S^d} q^d(\overline{m}, \overline{m} + e_j - e_i)\pi^d(\overline{m}) \\ &= \sum_{\overline{m}\in S^d} \frac{\phi(\overline{m} + e_j)}{\psi(\overline{m})} p_{ij}^d(\overline{m} + e_j) B^d \psi(\overline{m}) \prod_{k=1}^N d_k^{m_k} \\ &= B^d \sum_{\overline{m}\in S^d} \phi(\overline{m} + e_j) \prod_{k=1}^N d_k^{m_k+\delta_{kj}} \frac{1}{d_j} p_{ij}^d(\overline{m} + e_j) \\ &= B^d \sum_{\overline{n}\in S_{ij}^d} \Phi(\overline{n}) \frac{1}{d_j} p_{ij}^d(\overline{n}), \end{aligned}$$

where, for $i,j \in \mathcal{N}$, $i \neq j$,

$$S_{ij}^d = \{\overline{n} \in N_0^N | \overline{n} - e_i, \overline{n} - e_j \in S^d,\ p_{ij}^d(\overline{n}) > 0\}.$$

Note that

$$S^- = \bigcup_{i,j\in\mathcal{N}} S_{ij}^d$$

which completes the proof. □

Remark 3.6 (Backward dualizing method) The set S^-, defined in (3.5), is the set of intermediate states for the transitions of the dual process. From the definition of the dual process, $q^d(\overline{n} - e_j, \overline{n} - e_i)$ is defined only if $\overline{n} \in S$. Therefore, $S^- \subseteq S$. In general, S^- is a strict subset of S, however, when the queueing network is open and unbounded, it may be the case that $S^- = S$.

The *forward dualizing method* is used to construct the dual process X^d from the primal process X. The state space S^d of the dual process consists of the intermediate states of the transitions of X and the transition rates for X^d are obtained via the potential interpretation. It is possible to define a *backward dualizing method* in various ways. The most direct method is the following *first backward dualizing method.*

The backward dual process of X^d is the process at S^- with transition rates given by (2.3). In this case, the backward dual process X^{dd} is not identical to the primal process unless $S^- = S$. However, without any knowledge of the primal process X, the backward dual process can be constructed from X^d. This backward dualizing method is discussed in Example 6.3. A different backward dualizing method is the following *second backward dualizing method.* A major drawback of this method is that it requires specific knowledge of the primal process. However, as is shown below, in this case the backward dual process is identical to the primal process, i.e. $X^{dd} = X$. For a state $\overline{m} \in S^d$, consider all states $\overline{m} + e_i$, $i \in \mathcal{N}$, such that $\overline{m} + e_i \in S$. The backward dual process X^{dd} is the process with state space S^{dd} given by

$$S^{dd} = \{\overline{n} \in N_0^N | \exists \overline{m} \in S^d,\ i \in \mathcal{N} \text{ such that } \overline{n} = \overline{m} + e_i \in S\},$$

and transition rates at S^{dd} given in (2.3). From the definition of the dual state space it is immediately clear that $S^{dd} = S$ and therefore that $X^{dd} = X$, unless S contains singletons ($\overline{n} \in S$ is called a singleton if for all $\overline{n}' \in S$: $q(\overline{n}, \overline{n}') = q(\overline{n}', \overline{n}) = 0$). □

4 The dual routing function

The dual process, with transition rates (3.1), is related to the primal process, with transition rates (2.3), through the functions ψ and ϕ only. The dual routing function p^d_{ij} is not related to the primal routing function p_{ij} in the definition of the dual process. Therefore, Section 3 defines a collection of dual processes for a collection of primal processes. In this section, based on the potential interpretation of the transition rates, the dual routing function p^d_{ij} is related to the primal routing function p_{ij} for a specific choice of this primal routing function. In Example 4.1 and Example 4.2, a state of the dual process is chosen to represent a customer configuration, whereas in Example 4.3 and Example 4.4, a state of the dual process represents the configuration of vacancies or holes.

Throughout this section, assume that for all $\overline{m} \in S^d$

$$p_{ij}(\overline{m}) = p_{ij}\mathbf{1}(\overline{m} + e_i, \overline{m} + e_j \in S).$$

Furthermore, when the queueing network is open assume that, for $M \leq \infty$, $M_i \in N_0$, $i = 1, \ldots, N$,

$$S = \{\overline{n} \in N_0^N | \sum_{i=1}^{N} n_i \leq M,\ n_i \geq M_i,\ i = 1, \ldots, N\},$$

with dual state space given by

$$S^d = \{\overline{n} \in N_0^N | \sum_{i=1}^{N} n_i \leq M - 1,\ n_i \geq M_i,\ i = 1, \ldots, N\}.$$

If $q^d(\overline{m}, \overline{m}') > 0$ for all $\overline{m}, \overline{m}' = \overline{m} + e_j - e_i \in S^d$, $i, j \in \mathcal{N}$, then

$$S^- = S \setminus \left\{ \overline{n} \in N_0^N | \sum_{i=1}^{N} n_i = M \text{ and } n_i = M_i \text{ for some } i \in \{1, \ldots, N\} \right\}. \quad (4.1)$$

If $M = \infty$ then $S^- = S = S^d$. However, as will be shown below, for the dual process to satisfy the dual traffic equations (3.3) at the boundary of S^d, in most examples, it must be assumed that $q^d(\overline{m}, \overline{m}') = 0$. Thus, in most cases, $S^- \subset S$. When the queueing network is closed assume that, for $M < \infty$, $M_i \in N_0$, $i = 1, \ldots, N$,

$$S = \{\overline{n} \in N_0^N | \sum_{i=1}^{N} n_i = M, \ n_i \geq M_i, \ i = 1, \ldots, N\},$$

in which case the dual state space is given by

$$S^d = \{\overline{n} \in N_0^N | \sum_{i=1}^{N} n_i = M - 1, \ n_i \geq M_i, \ i = 1, \ldots, N\}.$$

If $q^d(\overline{m}, \overline{m}') > 0$ for all $\overline{m}, \overline{m}' \in S^d$, $i, j \in \mathcal{N}$ then

$$S^- = S \setminus \{\overline{n} \in S | n_i = M_i \text{ for some } i \in \{1, \ldots, N\}\} \subset S. \quad (4.2)$$

Assume that a positive solution $\{c_i\}_{i=1}^N$ exists to the traffic equations. In the open network case they read

$$\sum_{j=0}^{N} \{\gamma_i p_{ij} - \gamma_j p_{ji}\} = 0, \ i = 1, \ldots, N, \ \gamma_0 = 1, \quad (4.3)$$

and in the closed network case

$$\sum_{j=1}^{N} \{\gamma_i p_{ij} - \gamma_j p_{ji}\} = 0, \ i = 1, \ldots, N. \quad (4.4)$$

Then, from Lemma 2.3, X allows an invariant measure Φ at S given by

$$\Phi(\overline{n}) = \phi(\overline{n}) \prod_{k=1}^{N} c_k^{n_k}, \ \overline{n} \in S.$$

Under these assumptions, the transition rates for the dual process will be given for four different cases. In all of these cases, due to a refinement of the site potential difference, the site potential difference D_{ij}^d for the dual process is related to the site potential difference D_{ij} for the primal process. This relation is then used to express p_{ij}^d as function of p_{ij}. To this end, recall (2.13) and the definition $\zeta_i = -\log c_i$, $i = 1, \ldots, N$, as given in Interpretation 2.7.

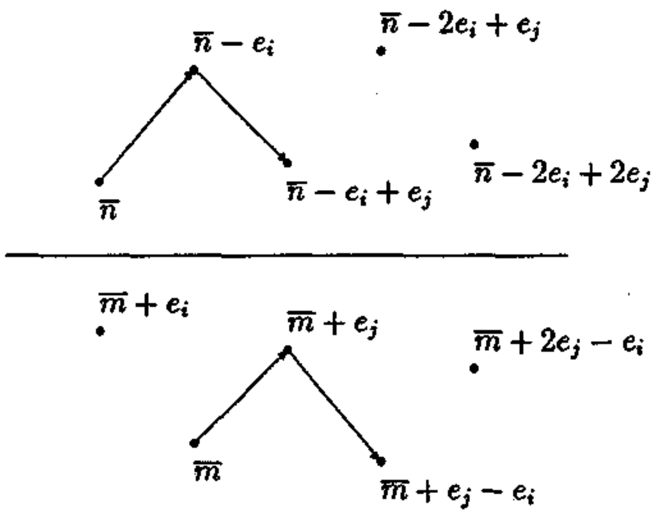

Figure 2. Similar transitions for the primal and the dual process ($\overline{m} = \overline{n} - e_i$).

4.1 Similar transitions

This example compares a transition $\overline{n} \to \overline{n} - e_i \to \overline{n} - e_i + e_j$ for the primal process to a transition $\overline{m} \to \overline{m} + e_j \to \overline{m} + e_j - e_i$ for the dual process, as depicted in Figure 2, where in both cases a unit is transferred from queue i to queue j. Assume the following transition mechanism for the primal process. Firstly, a customer is released from queue i according to the global potential difference, but this customer remains in the vicinity of queue i. The global state is changed into $\overline{n} - e_i$, representing the customer configuration of the customers remaining at the queues, with one customer released into the queueing network. Secondly, the customer is released from the vicinity of queue i and routed to queue j according to the site potential difference $D_{ij}(\overline{n} - e_i)$. From the transition mechanism above, it is natural for the dual states to represent customer configurations. A state for the dual process represents the customer configuration at the queues of the network, but with one customer that is not attached to a specific queue, i.e. a customer in transit from one queue to another. This gives the following transition mechanism for the dual process. Firstly, according to the global potential difference, the free moving customer is captured in the vicinity of queue j. Secondly, according to the site potential difference $D^d_{ij}(\overline{m}+e_j)$ a customer is released from queue i and the customer near queue j actually enters the queue. In the new state, $\overline{m} + e_j - e_i$, one customer is moving freely among the queues. Note that, for both the primal and the dual process, a customer is routed from queue i to queue j. However, for the primal process a customer must overcome the site potential ζ_i for leaving the vicinity of queue i to become a free moving customer that can be routed to queue j, whereas for the dual process a customer must overcome the site potential $-\zeta_j$ for entering queue j and thus allowing a customer to be released from queue i. Therefore, the following relation for the site potential difference for the primal and dual process can be deduced:

$$D^d_{ij}(\overline{m} + e_j) = -\zeta_i - \zeta_j + D_{ij}(\overline{m}), \; i, j \in \mathcal{N}. \tag{4.5}$$

For the routing function this implies

$$p_{ij}^d(\overline{n}) = c_i c_j p_{ij}, \ \overline{n} \in S^-, \ i,j \in \mathcal{N}. \tag{4.6}$$

If the primal process is reversible, i.e. the solution $\{c_i\}_{i=1}^N$ to the traffic equations (4.3), (4.4) satisfies

$$c_i p_{ij} = c_j p_{ji}, \ i,j \in \mathcal{N},$$

then for the dual routing function defined in (4.6), where S^- is given in (4.1), (4.2), the coefficients $\{c_i\}_{i=1}^N$ solve (3.3). Therefore, Lemma 3.3 implies that the dual process allows an invariant measure Ψ at S^d given by

$$\Psi(\overline{m}) = \psi(\overline{m}) \prod_{k=1}^N c_k^{m_k}, \ \overline{m} \in S^d. \tag{4.7}$$

However, if the primal process is non-reversible, for the dual process to possess a positive solution to (3.3) the dual routing function (4.6) must be modified to

$$p_{ij}^d(\overline{n}) = \begin{cases} c_i c_j p_{ij} & \text{if } n_k > M_k, \ k = 1,\dots,N, \ i,j \in \mathcal{N} \\ 0 & \text{otherwise.} \end{cases} \tag{4.8}$$

When the primal process represents a closed queueing network (4.8) is identical to (4.6) and S^- is given in (4.2), however, when the primal process represents an open queueing network (4.8) corresponds to (4.6) with S^- as given in (4.1) replaced by

$$S^- = \{\overline{n} \in N_0^N | \sum_{i=1}^N n_i \le M, \ n_k > M_k, \ k = 1,\dots,N\}. \tag{4.9}$$

For p_{ij}^d as given in (4.8) the coefficients $\{c_i\}_{i=1}^N$ solve (3.3) and the dual process allows an invariant measure Ψ as given in (4.7) above. Note that (4.7) is in agreement with the interpretation of the states of the dual process given above since (4.7) expresses the potential of the customer configuration $\overline{m}$.

The modification (4.8) of the dual routing function p_{ij}^d as given in (4.6) for a non-reversible primal process is not in conflict with the derivation of (4.6) since the potential interpretation of the transition rates is valid if the process is reversible only. The blocking protocol introduced in (4.8) is the protocol dual to the *stop protocol* for primal processes (cf. Remark 2.4) and will be discussed in Example 6.3.

An interesting observation for routing functions related through (4.6), i.e. for a reversible primal process, is stated in the following lemma, where the probability flow from queue i to queue j for the primal and the dual process is compared. As both the primal and the dual process describe the evolution of customer configurations in a queueing network, it seems natural to assume that the probability flow between queues for the primal process and the dual process must match. In many examples, however, this may not be the case.

Lemma 4.1 *If $S^- = S$ and (4.6) holds, then*

$$Eq_{ij}^d = Eq_{ij},\ i,j \in \mathcal{N},$$

if and only if

$$\sum_{\overline{n}\in S} \phi(\overline{n}) \prod_{k=1}^{N} c_k^{n_k} = \sum_{\overline{m}\in S^d} \psi(\overline{m}) \prod_{k=1}^{N} c_k^{m_k},$$

i.e. $B = B^d$ *for the appropriate definition of the normalizing constants for the primal and the dual process.*

Proof If (4.6) holds, then comparison of (2.11) and (3.4) gives

$$Eq_{ij} = Eq_{ij}^d,\ i,j \in \mathcal{N}$$

if and only if

$$B \sum_{\overline{m}\in S^d} \Psi(\overline{m}) c_i p_{ij}(\overline{m}) = B^d \sum_{\overline{n}\in S^-} \Phi(\overline{n}) \frac{1}{c_j} p_{ij}^d(\overline{n})$$

with B and B^d as defined in Lemma 2.5 and Lemma 3.2, respectively. □

4.2 Similar subtransitions

In the previous example, a transition $\overline{m} \to \overline{m}+e_j-e_i$ for the dual process is compared to a transition $\overline{n} \to \overline{n}-e_i+e_j$ for the primal process. Although this seems to be a natural way to obtain the routing function for the dual process, the subtransitions corresponding to D_{ij} and D_{ij}^d do not match. In fact, in the previous example, the subtransitions $\overline{n}-e_i \to \overline{n}-e_i+e_j$ and $\overline{m}+e_j \to \overline{m}+e_j-e_i$ are compared. As can be seen from Figure 3, the subtransition corresponding to $\overline{m}+e_j \to \overline{m}+e_j-e_i$ for the dual process is the subtransition $\overline{n}-e_i+e_j \to \overline{n}-2e_i+e_j$ for the primal process. However, this subtransition cannot occur as a consequence of a site potential difference for the primal process, since upwards subtransitions correspond to a global potential difference. Therefore, the subtransition $\overline{m}+e_j \to \overline{m}+e_j-e_i$ as a part of the transition $\overline{m} \to \overline{m}+e_j \to \overline{m}+e_j-e_i$ for the dual process is compared to the subtransition $\overline{n}-2e_i+e_j \to \overline{n}-e_i+e_j$ as a part of the transition $\overline{n}-2e_i+2e_j \to \overline{n}-2e_i+e_j \to \overline{n}-e_i+e_j$ for the primal process. Similar to Example 4.1, in the primal process a customer must overcome the site potential ζ_j for leaving the vicinity of queue j, whereas in the dual process a customers must overcome the site potential $-\zeta_j$ for entering queue j. This implies for the site potential differences, that

$$D_{ij}^d(\overline{m}+e_j) = -2\zeta_j + D_{ji}(\overline{n}-2e_i+e_j),\ i,j \in \mathcal{N},$$

which results in the following relation for the routing functions:

$$p_{ij}^d(\overline{n}) = c_j^2 p_{ji},\ \overline{n} \in S^-,\ i,j \in \mathcal{N}. \tag{4.10}$$

If the primal process is reversible (4.10) is equivalent to (4.6), however, if the primal process is non-reversible the modified version of (4.10)

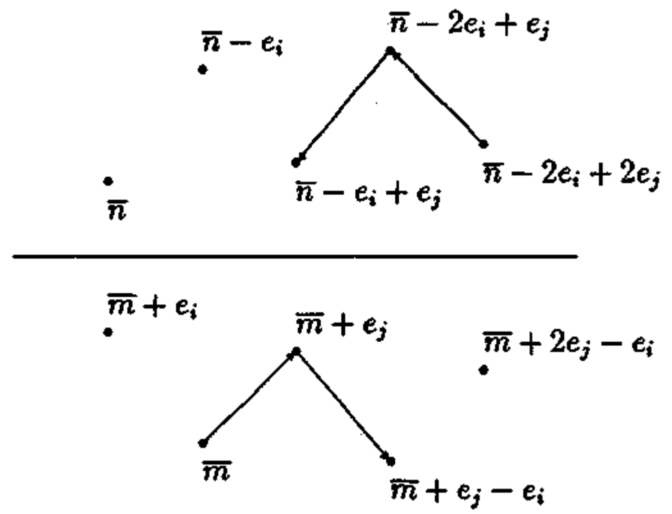

Figure 3. Similar subtransitions for the primal and the dual process ($\overline{m}=\overline{n}-e_i$).

$$p^d_{ij}(\overline{n}) = \begin{cases} c_j^2 p_{ji} & \text{if } n_k > M_k,\ k=1,\ldots,N,\ i,j\in\mathcal{N} \\ 0 & \text{otherwise.} \end{cases} \tag{4.11}$$

cannot be identified to (4.8). Therefore, in my opinion, the derivation of both (4.8) and (4.11) via reversible processes is justified. Similar to the previous example, $\{c_i\}_{i=1}^N$ solves (3.3), thus Ψ as given in (4.7) is an invariant measure for the dual process.

4.3 Similar transitions; reversed potential

As a direct consequence of the transition mechanism for the primal process, the previous examples assume that a state for the dual process represents a customer configuration. However, a state for the dual process may also be interpreted as representing the configuration of vacancies. To this end, reconsider the transition mechanism for the primal process as described in Example 4.1 and Figure 2. In a transition $\overline{n}\to\overline{n}-e_i\to\overline{n}-e_i+e_j$ first a customer leaves queue i, or equivalently a vacancy is created at queue i. For the dual process, in a transition $\overline{m}\to\overline{m}+e_j\to\overline{m}+e_j-e_i$ first a unit is created at queue j. Therefore, it seems natural to identify a unit in the dual process with a vacancy or hole in the primal process. As the site potential of a vacancy must be the reversed of the site potential of a customer, in the site potential interpretation the identification of a unit for the dual process to a vacancy corresponds to an additional site potential ζ_j instead of the $-\zeta_j$ in the derivation of (4.5) in Example 4.1. If, when routing, a unit for the dual process behaves similar to a unit for the primal process, the site potential relation (4.5) must be replaced by

$$D^d_{ij}(\overline{m}+e_j) = -\zeta_i+\zeta_j+D_{ij}(\overline{m}),\ i,j\in\mathcal{N}.$$

For the routing function this implies

$$p^d_{ij}(\overline{n}) = \frac{c_i}{c_j}p_{ij},\ \overline{n}\in S^-,\ i,j\in\mathcal{N}.$$

Similar to Example 4.1 for a non-reversible primal process the dual routing function must be modified to

$$p^d_{ij}(\overline{n}) = \begin{cases} \dfrac{c_i}{c_j} p_{ij} & \text{if } n_k > M_k,\ k = 1,\dots,N,\ i,j \in \mathcal{N} \\ 0 & \text{otherwise.} \end{cases}$$

By insertion of the dual routing function p^d_{ij} into (3.3), $d_i = \frac{1}{c_i}$, $i = 1,\dots,N$, is a solution to (3.3). Therefore, Lemma 3.3 implies that the dual process allows an invariant measure Ψ at S^d given by

$$\Psi(\overline{m}) = \psi(\overline{m}) \prod_{k=1}^{N} \left(\frac{1}{c_k}\right)^{m_k},\ \overline{m} \in S^d. \tag{4.12}$$

4.4 Similar subtransitions; reversed potential

By analogy with Example 4.2, the subtransitions depicted in Figure 3 may be compared in the case described in Example 4.3. If the primal process is reversible, this implies for the dual routing function

$$p^d_{ij}(\overline{n}) = p_{ji},\ \overline{n} \in S^-, \tag{4.13}$$

where S^- is defined in (4.1), (4.2). If the primal process is non-reversible, S^- as specified above must, in the open network case, be modified to (4.9). For p^d_{ij} thus defined the dual process allows an invariant measure Ψ at S^d given in (4.12). Note that (4.12) is in agreement with the interpretation of the states of the dual process since (4.12) expresses the potential of a vacancy configuration $\overline{m}$.

The relation between the primal routing function and the dual routing function (4.13) may also be interpreted directly as follows. If a customer routes from queue j to queue i in the primal process then an additional vacancy is created at queue j and a vacancy is deleted at queue i, i.e. a vacancy routes from queue i to queue j. Therefore, the dual routing function must be related to the primal routing function as given in (4.13).

A natural analog of Lemma 4.1 is the following lemma where the probability flow for a reversible primal process is related to the probability flow for the dual process. It seems natural to assume that the probability flow for customers in the primal process and the probability flow for vacancies in the dual process must match, i.e. to assume that $Eq^d_{ij} = Eq_{ji}$. In many examples, however, this may not be the case.

Lemma 4.2 *If $S^- = S$ and (4.13) holds, then*

$$Eq^d_{ij} = Eq_{ji},\ i,j \in \mathcal{N},$$

if and only if

$$\sum_{\overline{n}\in S} \phi(\overline{n}) \prod_{k=1}^{N} c_k^{n_k} = \sum_{\overline{m}\in S^d} \psi(\overline{m}) \prod_{k=1}^{N} \left(\frac{1}{c_k}\right)^{m_k},$$

i.e. $B = B^d$ for the appropriate definition of the normalizing constants for the primal and the dual process.

Proof Similar to the proof of Lemma 4.1. □

5 Relation between primal and dual process

For the primal process introduced in Section 2, a collection of dual processes is introduced in Section 3. Based on a refinement of the potential interpretation of the transition rates given in Interpretation 2.7, Section 4 gives some examples of dual processes with transition rates completely determined by the transition rates of the primal process. Although the dual process is completely determined by the primal process, in general, the primal process and corresponding dual process do not describe the same physical system (queueing network). The following definition of primal and dual processes may therefore be considered.

> A primal process is a process for which a transition $\overline{n} \to \overline{n} - e_i + e_j$ passes through $\overline{n} - e_i$, i.e. first a unit is deleted at queue i and then a unit is created at queue j.
>
> A dual process is a process for which a transition $\overline{n} \to \overline{n} - e_i + e_j$ passes through $\overline{n} + e_j$, i.e. first a unit is created at queue j and then a unit is deleted at queue i.
>
> A dual process corresponding to a given primal process is a dual process with transition rates completely determined by the primal process and a primal process corresponding to a given dual process is a primal process with transition rates completely determined by the dual process.

The dual processes given in Section 4 are examples of dual processes corresponding to the primal process given in Section 2. In practical cases, it may be desirable for the dual process corresponding to a given primal process to describe the same physical system. The problem of probability flow between queues being identical is considered in Lemmata 4.1 and 4.2. For a dual process to correspond to a primal process the most obvious condition, however, is that the probability flow between states is identical. This problem is considered in Section 5.1 below.

Remark 2.6 mentions that the equilibrium distribution of the dual process represents the probability that a transition for the primal process passes through a dual state. This will be formalized in Section 5.2, where the Palm probabilities associated with the jumps of the primal process are studied. Finally, the results of Section 5.1 and Section 5.2 are combined in the last part of Section 5.2, where a theorem similar to the arrival theorem is proven.

Throughout this section, assume that the primal process X is stationary and irreducible with unique stationary distribution π at S. Furthermore, assume that a positive solution $\{c_i\}_{i=1}^N$ to (2.9) exists, where $\sum_{j \in \mathcal{N}} p_{ij}(\overline{m}) = 1$, for all $i \in \mathcal{N}$, $\overline{m} \in S^d$.

5.1 Equal probability flow

For a dual process to correspond to a primal process, the most obvious condition on the transition rates of the dual process is that the probability flow between states of the dual process is induced by the probability flow between states of the primal process. To this end, reconsider Figure 1. If the dual process possesses an equilibrium distribution π^d, then the probability flow from state $\overline{m}$ to state $\overline{m}+e_j-e_k$ for the dual process is given by

$$\pi^d(\overline{m})q^d(\overline{m},\overline{m}+e_j-e_k).$$

For the primal process, this probability flow is induced by the probability flow from state $\overline{m}+e_i$ to state $\overline{m}+e_j-e_k+e_l$. Thus, the probability flow from state $\overline{m}$ to state $\overline{m}+e_j-e_k$ in the primal process is given by

$$\sum_{i,l\in\mathcal{N}} \pi(\overline{m}+e_i)q(\overline{m}+e_i,\overline{m}+e_j)q(\overline{m}+e_j,\overline{m}+e_j-e_k+e_l).$$

As the global part of the transition rates for the dual process is determined in the definition, (3.1), the dual routing function p^d_{ij} must be determined such that these flows match.

Theorem 5.1 *The dual routing function is for $\overline{m}$, $\overline{m}+e_j-e_i\in S^d$, $i,j\in\mathcal{N}$ given by*

$$p^d_{ij}(\overline{m}+e_j)=\begin{cases}\dfrac{B}{B^d}\dfrac{\psi(\overline{m})\psi(\overline{m}+e_j-e_i)}{\phi(\overline{m}+e_j)^2}\displaystyle\prod_{k=1}^{N}\left(\frac{c_k}{d_k}\right)^{m_k+\delta_{kj}}d_j & \\ \qquad \text{if } \exists r,s\in\mathcal{N} \text{ such that } p_{rj}(\overline{m})>0,\ p_{is}(\overline{m}+e_j-e_i)>0, \\ 0 \qquad \text{otherwise,}\end{cases} \tag{5.1}$$

where $d_i>0$, $i=1,\ldots,N$, and

$$0<\left[B^d\right]^{-1}=\sum_{\overline{m}\in S^d}\psi(\overline{m})\prod_{k=1}^{N}d_k^{m_k}<\infty, \tag{5.2}$$

if and only if the dual process is an irreducible Markov chain with unique equilibrium distribution π^d at S^d given by

$$\pi^d(\overline{m})=B^d\psi(\overline{m})\prod_{k=1}^{N}d_k^{m_k},\ \overline{m}\in S^d, \tag{5.3}$$

and the probability flow from state $\overline{m}\in S^d$ to state $\overline{m}+e_j-e_k\in S^d$ is the same for both the dual process and the primal process for all $\overline{m}\in S^d$, $j,k\in\mathcal{N}$, i.e.

$$\pi^d(\overline{m})q^d(\overline{m},\overline{m}+e_j-e_k)=\sum_{i,l\in\mathcal{N}}\pi(\overline{m}+e_i)q(\overline{m}+e_i,\overline{m}+e_j)q(\overline{m}+e_j,\overline{m}+e_j-e_k+e_l). \tag{5.4}$$

Proof For $\overline{m},\overline{m}'\in S^d$ there exist i_0,i_1,j_0,j_1 such that $\overline{m}+e_{i_0},\overline{m}+e_{i_1},\overline{m}'+e_{j_0},\overline{m}'+e_{j_1}\in S$ and such that $q(\overline{m}+e_{i_0},\overline{m}+e_{i_1})>0$, $q(\overline{m}'+e_{j_0},\overline{m}'+e_{j_1})>0$. As X is irreducible, there exists a sequence of states $\overline{n}_0,\overline{n}_1,\overline{n}_2,\ldots,\overline{n}_{k-2},\overline{n}_{k-1},\overline{n}_k\in S$ such that $q(\overline{n}_i,\overline{n}_{i+1})>0$, $i=0,\ldots,k-1$, and $\overline{n}_0=\overline{m}+e_{i_0}$, $\overline{n}_1=\overline{m}+e_{i_1}$, $\overline{n}_{k-1}=\overline{m}'+e_{j_0}$, $\overline{n}_k=\overline{m}'+e_{j_1}$. Define $\overline{m}_i=\overline{n}_{i-1}-(\overline{n}_{i-1}-\overline{n}_i)^+$, $i=1,\ldots,k-1$, where $\overline{v}^+$ denotes the vector $\overline{v}$ with all non-positive entries replaced by zero. Then $\overline{m}_i\in S^d$, $i=1,\ldots,k-1$.

From (5.1), $p^d_{ij}(\overline{m}+e_j) > 0$ for all $i,j \in \mathcal{N}$, $\overline{m}, \overline{m}+e_j-e_i \in S^d$. This implies for the sequence $\overline{m}_1, \ldots, \overline{m}_{k-1} \in S^d$ that $q^d(\overline{m}_i, \overline{m}_{i+1}) > 0$, $i = 1, \ldots, k-1$. Since $\overline{m} = \overline{m}_1$ and $\overline{m}' = \overline{m}_{k-1}$ this implies that X^d is irreducible. Insertion of (5.1) into (3.3) gives that $\{d_i\}_{i=1}^N$ is a positive solution of (3.3). Therefore, from Lemma 3.3 and (5.2), π^d as given in (5.3) is the unique equilibrium distribution of the dual process. From Lemma 2.3, $\pi(\overline{n}) = B\phi(\overline{n}) \prod_{k=1}^N c_k^{n_k}$. Insertion of π, π^d and the transition rates q, q^d into (5.4), immediately gives in the non-trivial case (i.e. $p^d_{ij}(\overline{m}+e_j) > 0$) for the right-hand side of (5.4)

$$\begin{aligned}
&\sum_{i,l \in \mathcal{N}} \pi(\overline{m}+e_i) q(\overline{m}+e_i, \overline{m}+e_j) q(\overline{m}+e_j, \overline{m}+e_j-e_k+e_l) \\
&\overset{(2.3),(2.8)}{=} \quad \sum_{i,l \in \mathcal{N}} B \prod_{k=1}^N c_k^{m_k+\delta_{ki}} \psi(\overline{m}) p_{ij}(\overline{m}) \frac{\psi(\overline{m}+e_j-e_k)}{\phi(\overline{m}+e_j)} p_{kl}(\overline{m}+e_j-e_k) \\
&\overset{\sum_{l\in\mathcal{N}} p_{kl}(\overline{m}')=1}{=} \quad \sum_{i \in \mathcal{N}} c_i p_{ij}(\overline{m}) B\psi(\overline{m}) \prod_{k=1}^N c_k^{m_k} \frac{\psi(\overline{m}+e_j-e_k)}{\phi(\overline{m}+e_j)} \\
&\overset{(2.9)}{=} \quad B\psi(\overline{m}) \prod_{k=1}^N c_k^{m_k+\delta_{kj}} \frac{\psi(\overline{m}+e_j-e_k)}{\phi(\overline{m}+e_j)},
\end{aligned}$$

and for the left-hand side of (5.4)

$$\begin{aligned}
&\pi^d(\overline{m}) q^d(\overline{m}, \overline{m}+e_j-e_k) \\
&= \quad B^d \psi(\overline{m}) \prod_{k=1}^N d_k^{m_k} \frac{\phi(\overline{m}+e_j)}{\psi(\overline{m})} p^d_{kj}(\overline{m}+e_j) \\
&\overset{(5.1)}{=} \quad B\psi(\overline{m}) \prod_{k=1}^N d_k^{m_k} \frac{\psi(\overline{m}+e_j-e_k)}{\phi(\overline{m}+e_j)} \prod_{k=1}^N \left(\frac{c_k}{d_k}\right)^{m_k+\delta_{kj}} d_j,
\end{aligned}$$

which immediately establishes (5.4).

To prove the reversed statement, firstly note that (5.3) implies (5.2). Secondly, if for $\overline{m}, \overline{m}+e_j-e_k \in S^d$ it is not the case that $\exists r,s \in \mathcal{N}$ such that $p_{rj}(\overline{m}) > 0$ and $p_{ks}(\overline{m}+e_j-e_k) > 0$ then the right-hand side of (5.4) equals zero. Thus, since $\pi^d(\overline{m}) > 0$, $\psi, \phi > 0$ it must be the case that $p^d_{kj}(\overline{m}+e_j) = 0$. Thirdly, if such $r,s \in \mathcal{N}$ do exist then insertion of π, π^d, q as given in (2.3) and q^d as given in (3.1) into (5.4) gives, similarly to the derivation of (5.4) above, that the dual routing function has the form given in (5.1). □

If the dual routing function has the form (5.1), then the transition rates of the dual process have the form

$$q^d(\overline{m}, \overline{m}+e_j-e_i) = \frac{B}{B^d} \frac{\psi(\overline{m}+e_j-e_i)}{\phi(\overline{m}+e_j)} \prod_{k=1}^N \left(\frac{c_k}{d_k}\right)^{m_k+\delta_{kj}} d_j.$$

A similar observation is made in [2] for the case of a reversible process with $d_i = c_i$, where it is stated that $q(\overline{m}, \overline{m}+e_j-e_i) \propto \frac{\psi(\overline{m}+e_j-e_i)}{\phi(\overline{m}+e_j)}$ is dual to $q(\overline{n}, \overline{n}-e_i+e_j) \propto \frac{\psi(\overline{n}-e_i)}{\phi(\overline{n})}$. From the viewpoint of equal probability flow, it seems more logical to give as a definition for the dual transition rates

$$q^d(\overline{m}, \overline{m} + e_j - e_i) = \frac{\psi(\overline{m} + e_j - e_i)}{\phi(\overline{m} + e_j)} p^d_{ij}(\overline{m} + e_j). \tag{5.5}$$

However, in Section 3 the dual process is introduced via the potential interpretation of the transition rates. In this case, it is more natural to give the form (3.1) for the transition rates for the dual process as this form explicitly states that a dual transition from state $\overline{m}$ to state $\overline{m}+e_j-e_i$ passes through state $\overline{m}+e_j$. This makes the relation between the primal process and the dual process a symmetrical relation as can be seen when comparing Lemma 2.3 to Lemma 3.3 and Lemma 2.5 to Lemma 3.5. Then, the form (5.5) appears as a consequence of the specific assumption (5.4) guaranteeing equal probability flow between states. Note that a theorem similar to Theorem 5.1 above in which the primal routing function is related to the dual routing function for a given form of the dual routing function can easily be formulated.

5.2 Palm probabilities

Each type of transition $\overline{n} \to \overline{n}'$ ($\overline{n} \neq \overline{n}'$) for X can be associated with a subset H of $S \times S \setminus \text{diag}(S \times S)$. For example, a transition in which a customer moves from queue i to queue j corresponds to

$$H_{ij} = \bigcup_{\overline{m} \in S^d} \{(\overline{m} + e_i, \overline{m} + e_j),\ \overline{m} + e_i, \overline{m} + e_j \in S\},\ i,j \in \mathcal{N}, \tag{5.6}$$

Let N_H be the process counting the H-transitions of X. Assume that (5.2) holds for $d_i = c_i$, $i = 1, \ldots, N$, then for all $H \subseteq S \times S \setminus \text{diag}(S \times S)$

$$0 < \sum_{(\overline{n},\overline{n}') \in H} \pi(\overline{n}) q(\overline{n}, \overline{n}') < \infty$$

and the Palm probability P_H associated with N_H can be defined. The Palm probability of the event C given that an H-transition occurs is given by (cf. [1])

$$P_H(C) = \frac{\sum_{(\overline{n},\overline{n}') \in C} \pi(\overline{n}) q(\overline{n}, \overline{n}')}{\sum_{(\overline{n},\overline{n}') \in H} \pi(\overline{n}) q(\overline{n}, \overline{n}')},\ C \subseteq H.$$

For example, the probability that the unmoved customers are in state $\overline{m}$ when a customer moves from queue i to queue j is given by $P_{ij}(\overline{m}) \equiv P_{H_{ij}}(\{\overline{m}+e_i, \overline{m}+e_j\})$. Direct computation gives the following theorem.

Theorem 5.2 (Palm probabilities)

1. *The probability that the unmoved customers are in state $\overline{m} \in S^d$ when a customer moves from queue i to queue j, $i, j \in \mathcal{N}$ is given by*

$$P_{ij}(\overline{m}) = \frac{\Psi(\overline{m}) p_{ij}(\overline{m})}{\sum_{\overline{m} \in S^d} \Psi(\overline{m}) p_{ij}(\overline{m})},\ \overline{m} \in S^d,\ i,j \in \mathcal{N}. \tag{5.7}$$

2. *The probability that the unmoved customers are in state $\overline{m} \in S^d$ when a customer leaves queue i or enters queue i, $i \in \mathcal{N}$, is given by*

$$P_i(\overline{m}) = \frac{\Psi(\overline{m})}{\sum_{\overline{m}\in S^d}\Psi(\overline{m})},\ \overline{m}\in S^d,\ i\in\mathcal{N}. \tag{5.8}$$

For convenience, the exterior is regarded as queue 0, and

$$\Psi(\overline{m}) = \psi(\overline{m})\prod_{k=1}^{N} c_k^{m_k}.$$

Proof The proof of the statements of the Theorem consists of specifying the appropriate sets H and $C \subseteq H$.

1. H represents the transitions in which a customer moves from queue i to queue j, i.e. $H = H_{ij}$ as given in (5.6) and C represents the transition $\overline{m}+e_i \to \overline{m}+e_j$, i.e. $C = C_{ij}(\overline{m}) = \{(\overline{m}+e_i, \overline{m}+e_j)\}$. This gives

$$\begin{aligned} P_{ij}(\overline{m}) &= \frac{\pi(\overline{m}+e_i)q(\overline{m}+e_i,\overline{m}+e_j)}{\sum_{\overline{m}\in S^d}\pi(\overline{m}+e_i)q(\overline{m}+e_i,\overline{m}+e_j)} \\ &= \frac{B\psi(\overline{m})\prod_{k=1}^{N}c_k^{m_k}c_ip_{ij}(\overline{m})}{\sum_{\overline{m}\in S^d}B\psi(\overline{m})\prod_{k=1}^{N}c_k^{m_k}c_ip_{ij}(\overline{m})} \\ &= \frac{\Psi(\overline{m})p_{ij}(\overline{m})}{\sum_{\overline{m}\in S^d}\Psi(\overline{m})p_{ij}(\overline{m})}. \end{aligned}$$

2. If H represents the transitions in which a customer leaves queue i and C represents the transitions in which a customer leaves queue i in state $\overline{m}+e_i$, then $H = \bigcup_{j\in\mathcal{N}} H_{ij}$ and $C = \bigcup_{j\in\mathcal{N}} C_{ij}(\overline{m})$. This immediately gives $P_i(\overline{m})$. If H represents the transitions in which a customer enters queue i and C represents the transitions in which a customer enters queue i to state $\overline{m}+e_i$, then $H = \bigcup_{j\in\mathcal{N}} H_{ji}$ and $C = \bigcup_{j\in\mathcal{N}} C_{ji}(\overline{m})$. This immediately gives

$$\begin{aligned} P_i(\overline{m}) &= \frac{\sum_{j\in\mathcal{N}}\pi(\overline{m}+e_j)q(\overline{m}+e_j,\overline{m}+e_i)}{\sum_{j\in\mathcal{N}}\sum_{\overline{m}\in S^d}\pi(\overline{m}+e_j)q(\overline{m}+e_j,\overline{m}+e_i)} \\ &= \frac{\psi(\overline{m})\prod_{k=1}^{N}c_k^{m_k}\sum_{j\in\mathcal{N}}c_jp_{ji}(\overline{m})}{\sum_{\overline{m}\in S^d}\psi(\overline{m})\prod_{k=1}^{N}c_k^{m_k}\sum_{j\in\mathcal{N}}c_jp_{ji}(\overline{m})} \\ &\overset{(2.9)}{=} \frac{\Psi(\overline{m})\sum_{j\in\mathcal{N}}c_ip_{ij}(\overline{m})}{\sum_{\overline{m}\in S^d}\Psi(\overline{m})\sum_{j\in\mathcal{N}}c_ip_{ij}(\overline{m})} \\ &= \frac{\Psi(\overline{m})}{\sum_{\overline{m}\in S^d}\Psi(\overline{m})}. \end{aligned}$$

□

Intuitively, there seems to be a discrepancy between (5.7) and (5.8) as (5.7) and (5.8) should be such that $\sum_{j\in\mathcal{N}} P_{ij}(\overline{m}) = P_i(\overline{m})$. However, $P_{ij}(\overline{m})$ expresses the probability that the unmoved customers are in state $\overline{m}$ when a customer moves from queue i to queue j and not the probability that the unmoved customers are in state $\overline{m}$ when a customer moves from queue i to queue j given that a customer leaves queue i. Therefore, there is no discrepancy. For completeness, the following theorem considers the probabilities mentioned above.

Theorem 5.3 (Palm probabilities) *The probability that the unmoved customers are in state $\overline{m} \in S^d$ and a transition from queue i to queue j, $i,j \in \mathcal{N}$, occurs when a customer leaves queue i, $i \in \mathcal{N}$, is given by*

$$P_i(\overline{m}, i \to j) = \frac{\Psi(\overline{m})p_{ij}(\overline{m})}{\sum_{\overline{m} \in S^d} \Psi(\overline{m})}, \ \overline{m} \in S^d, \ i,j \in \mathcal{N},$$

and the probability that the unmoved customers are in state $\overline{m} \in S^d$ and a transition from queue i to queue j, $i,j \in \mathcal{N}$, occurs when a customer enters queue j, $j \in \mathcal{N}$, is given by

$$P_j(\overline{m}, i \to j) = \frac{\Psi(\overline{m})\frac{c_i}{c_j}p_{ij}(\overline{m})}{\sum_{\overline{m} \in S^d} \Psi(\overline{m})}, \ \overline{m} \in S^d, \ i,j \in \mathcal{N}.$$

Proof If H represents the transitions in which a customer leaves queue i and C represents the transition from state $\overline{m} + e_i$ to state $\overline{m} + e_j$ then $H = \bigcup_{j \in \mathcal{N}} H_{ij}$ and $C = C_{ij}(\overline{m})$. This immediately gives $P_i(\overline{m}, i \to j)$. If H represents the transitions in which a customer enters queue j and C represents the transition from state $\overline{m} + e_i$ to state $\overline{m} + e_j$ then $H = \bigcup_{i \in \mathcal{N}} H_{ij}$ and $C = C_{ij}(\overline{m})$. This gives

$$\begin{aligned} P_j(\overline{m}, i \to j) &= \frac{\pi(\overline{m} + e_i)q(\overline{m} + e_i, \overline{m} + e_j)}{\sum_{i \in \mathcal{N}} \sum_{\overline{m} \in S^d} \pi(\overline{m} + e_i)q(\overline{m} + e_i, \overline{m} + e_j)} \\ &= \frac{\psi(\overline{m}) \prod_{k=1}^N c_k^{m_k} c_i p_{ij}(\overline{m})}{\sum_{\overline{m} \in S^d} \psi(\overline{m}) \prod_{k=1}^N c_k^{m_k} \sum_{i \in \mathcal{N}} c_i p_{ij}(\overline{m})} \\ &\overset{(2.9)}{=} \frac{\Psi(\overline{m}) c_i p_{ij}(\overline{m})}{\sum_{\overline{m} \in S^d} \Psi(\overline{m}) c_j}. \end{aligned}$$

□

The Palm probability $P_H(\overline{m})$ is the probability that when an H-transition occurs a customer *sees* the other customers in state $\overline{m} \in S^d$. For example, $P_{ij}(\overline{m})$ is the probability that a customer in transit from queue i to queue j sees the non-moving customers in state $\overline{m} \in S^d$. The state observed by a customer in transit is a dual state representing the customer configuration at the queues with one customer in transit. The global potential of the state observed by a customer in transit may be influenced by this customer. Therefore, the distribution seen by a customer in transit does not necessarily correspond to the equilibrium distribution of the queueing network with this customer removed. From Theorem 5.2 above, in some cases $P_H(\overline{m}) = \pi^d(\overline{m})$, $\overline{m} \in S^d$, the equilibrium distribution of the dual process. This distribution is independent of the customer in transit. In those cases, arrivals see time averages (ASTA) as expressed by (5.8) and moving customers (units) see time averages (MUSTA) as expressed by (5.7). This is formalized in the following variant of the arrival theorem.

Theorem 5.4 (Arrival Theorem) *Consider a stationary Markov chain X with transition rates (2.3) such that a positive solution $\{c_i\}_{i=1}^N$ to (2.9) exists. The distribution observed by a customer arriving at a queue or departing from a queue is given by π^d as given in (5.3) with $d_i = c_i$, $i = 1, \ldots, N$. If, for $i,j \in \mathcal{N}$, $p_{ij}(\overline{m})$ is independent of $\overline{m}$ then the distribution observed by a customer in transit from queue i*

to queue j is given by π^d as specified above. Moreover, if moving customers (arriving at a queue, departing from a queue or in transit between queues) observe π^d, the evolution of the states observed by the moving customers is given by the dual process X^d with transition rates

$$q^d(\overline{m},\overline{m}+e_j-e_i) = \begin{cases} \dfrac{B}{B^d}\dfrac{\psi(\overline{m}+e_j-e_i)}{\phi(\overline{m}+e_j)}c_j, \\ \qquad \text{if } \exists r,s \in \mathcal{N} \text{ such that } p_{rj}(\overline{m}) > 0,\ p_{is}(\overline{m}+e_j-e_i) > 0, \\ 0 \qquad \text{otherwise,} \end{cases} \tag{5.9}$$

Proof From Theorem 5.2, π^d is the distribution observed by a moving customer in the cases specified in the theorem. The transition rates of the process describing the evolution of the states observed by the customers in transit must be such that the probability flow from state $\overline{m} \in S^d$ to state $\overline{m}' \in S^d$ is completely determined by the probability flow for X. As the equilibrium distribution of the process describing the evolution of the states observed by the customers in transit is given by π^d, Theorem 5.1 gives the last statement of the theorem. □

Remark 5.5 (Markov chain; dual process) Note that, by definition, the dual process is a Markov chain. As is shown in Theorem 5.4, the evolution of states as observed by a moving customer can be described by the dual process. However, as can immediately be seen by observing a sequence of transitions $\overline{n} \to \overline{n} - e_i \to \overline{n} - e_i + e_j \to \overline{n} - e_i + e_j - e_k \to \overline{n} - e_i + e_j - e_k + e_l$, as depicted in Figure 1, the transition probability from state $\overline{n} - e_i$ to state $\overline{n} - e_i + e_j - e_k$ must depend on $\overline{n}$ as the routing probability $p_{ij}(\overline{n} - e_i)$ for the first part of the transition sequence depends on $\overline{n}$. Therefore, for the primal process, the evolution of states as observed by a moving customer is not a Markov chain. Thus, the dual process is a Markov chain *constructed* such that it describes the evolution of the states as observed by a moving customer. □

Remark 5.6 (Literature) The arrival theorem is discussed by various authors (cf. [8], [13] and the references therein). Recently, for queueing networks with transition rates similar to (2.3), the arrival theorem is discussed in [6] for batch routing networks with transition rates discussed in Remark 2.8 and in [9] MUSTA is defined and proven. However, all authors assume $\psi = \phi$ when considering the arrival theorem, also when this is not necessary in the proof (cf. [6]). The motivation for this specific choice, $\psi = \phi$, is that the distribution seen by a moving customer must be identical to the equilibrium distribution of the queueing network with one customer removed. In my opinion, as I discussed in the preamble of Theorem 5.4, this assumption is not necessary. This is a direct consequence of the potential interpretation of the primal and dual process. A moving customer may influence the configuration potential of the state it observes. This influence is taken into account in ψ. If $\psi = \phi$, a moving customer does not influence the configuration potential of the state it observes. In this case, the first part of Theorem 5.4 reduces to the well-known version of the arrival theorem. To the best of my knowledge, the second part of Theorem 5.4 is new, in that the process describing the evolution of the states observed by the moving customers is not described in the literature. The appearance of this process is a direct consequence of the introduction of the dual process in Section 3. □

6 Examples and general remarks

This section gives some theoretical examples and some general remarks on the relation between the primal and the dual process. Firstly, in Example 6.1, the symmetrical relation between the primal and the dual process as discussed in Section 3, where Lemmata 2.3 and 2.5 are compared to Lemmata 3.3 and 3.5, is shown to be related to the relation between the primal and dual problem in linear programming. This justifies the name dual process given to the process introduced in Definition 3.1. Secondly, in Example 6.2, in the special case where the routing is state independent, it is shown that the traffic equations are a consequence of the transition rates and product form equilibrium distribution. Thirdly, in Examples 6.3 – 6.6, the transition rates and state spaces for the primal and dual process are discussed. Fourth, in Section 6.7, some comments on batch servicing and batch routing networks are given. These comments are used in Section 6.8 to explain the discrepancy between continuous-time and discrete-time results. Finally, Section 6.9 discusses customer-vacancy duality as reported in the literature.

6.1 Complementary slackness relations

This section gives a motivation for the name *dual process* for the process introduced in Definition 3.1. To this end, reconsider Definition 2.1 and Lemma 2.3. In the definition of the dual state space specific knowledge of the transition rates of the primal process is used ($\overline{m} \in S^d$ only if $\exists i, j \in \mathcal{N}$ such that $q(\overline{m}+e_i, \overline{m}+e_j) > 0$). This knowledge is not strictly necessary, it merely makes the introduction of the dual process more elegant. Therefore, the dual state space may be replaced by S^* defined as

$$S^* = \{\overline{m} \in N_0^N | \exists i, j \in \mathcal{N} : \overline{m}+e_i, \overline{m}+e_j \in S\}.$$

Moreover, the assumption $\psi(\overline{m}) > 0$ is not necessary. ψ may be replaced by $\psi : S^* \to R^+ \cup \{0\}$. Furthermore, p_{ij} may be defined as $p_{ij} : S^* \to R^+ \cup \{0\}$ (also see Remark 2.4). With these relaxed assumptions, Lemma 2.3 must be replaced by the following lemma.

Lemma 6.1 *X allows an invariant measure* Φ *at* S*, given by*

$$\Phi(\overline{n}) = \phi(\overline{n}) \prod_{k=1}^{N} c_k^{n_k}, \quad \overline{n} \in S,$$

if for all $\overline{n} \in S$ *the coefficients* $\{c_i\}_{i=1}^N$ *are a non-negative* ($c_i \geq 0$ *for all* i) *solution of*

$$\sum_{j\in\mathcal{N}} \{\gamma_i p_{ij}(\overline{n}-e_i) - \gamma_j p_{ji}(\overline{n}-e_i)\} = 0, \; i \in \mathcal{N}, \; \gamma_0 = 1, \tag{6.1}$$

for all $\overline{n}-e_i \in S^*$ such that $\psi(\overline{n}-e_i) > 0$.

The proof of Lemma 6.1 is identical to the proof of Lemma 2.3. The motivation for the last equality in the proof of Lemma 2.3 is changed. This equality is obtained since for fixed $\overline{n} \in S$, $\mathcal{N}$ can be divided into $\mathcal{N}_+ = \{i \in \mathcal{N} | \psi(\overline{n}-e_i) > 0\}$ and $\mathcal{N}_0 = \{i \in \mathcal{N} | \psi(\overline{n}-e_i) = 0\}$. At $\mathcal{N}_+$ (6.1) holds and thus the last equality in the proof of Lemma 2.3 is obtained.

For $\overline{b}$, the zero-vector, $\overline{c}$ defined as $[\overline{c}]_i = c_i$, $P(\overline{n})$ defined as $[P(\overline{n})]_{ij} = p_{ij}(\overline{n}-e_i)$, $\overline{\Psi}(\overline{n})$ defined as $[\overline{\Psi}(\overline{n})]_i = \psi(\overline{n}-e_i)\prod_{k=1}^N c_k^{n_k-\delta_{ki}}$, $i,j \in \mathcal{N}$, (2.10) can be written

$$\overline{\Psi}(\overline{n}) \cdot \left(\overline{b} - \overline{c}P(\overline{n})\right) = 0, \tag{6.2}$$

where $\cdot$ denotes inner product. (6.2) is valid since either $[\overline{\Psi}(\overline{n})]_i = 0$ or $[(\overline{b}-\overline{c}P(\overline{n}))]_i = 0$. Therefore, (6.2) resembles a complementary slackness relation from linear programming (cf. [10]), where $\overline{\Psi}(\overline{n})$ is the optimal solution of a dual problem and $\overline{c}P(\overline{n}) \leq \overline{b}$ is the feasibility condition for the primal problem. Therefore, the process with invariant measure $\psi(\overline{m})\prod_{k=1}^N c_k^{m_k}$ is called a dual process.

The relaxation of the assumptions stated in Section 2 to the form stated above creates a lot of difficulties when introducing the dual process. Although it is possible to use the relaxed assumptions, I have chosen to include *all* the assumptions of Section 2 on the transition rates, i.e. on S^d, ψ, ϕ and p_{ij} thus making the theory and the introduction of the dual process very transparent.

6.2 Traffic equations

In general, the reversed statement of Lemma 2.3, i.e. if Φ as given in (2.8) is an invariant measure at S then the coefficients $\{c_i\}_{i=1}^N$ are a positive solution to (2.9), is not true. However, if the routing function does not depend on the state, i.e. if

$$p_{ij}(\overline{m}) = p_{ij}, \; i,j \in \mathcal{N}, \; \overline{m} \in S^d, \tag{6.3}$$

then, from Lemma 2.5, the expected equilibrium rate from queue i to queue j is given by

$$Eq_{ij} = \frac{B}{B^d} c_i p_{ij}, \; i,j \in \mathcal{N}.$$

Clearly, in equilibrium, the probability flow for departures from queue i must equal the probability flow for arrivals to queue i, i.e.

$$\sum_{j\in\mathcal{N}} Eq_{ij} = \sum_{j\in\mathcal{N}} Eq_{ji}, \quad i \in \mathcal{N}.$$

As the expected equilibrium rate from queue i to queue j must be a non-negative number, with this specific choice for the routing function, the reversed statement of Lemma 2.3 is proven. If the process is reversible, assumption (6.3) is not necessary. In this case, for a general routing function $p_{ij}(\overline{m})$, the reversed statement of Lemma 2.3 can be proven (cf. [2]).

6.3 Blocking examples

This example shows that the dual state space depends on the blocking protocol. To this end, consider an open two station queueing network with finite capacity constraints on the number of customers in the queues and in the system. If M_1, M_2 customers are allowed at queue 1, 2, respectively and $M < M_1 + M_2$ customers are allowed in the system, the state space S for this process is given by

$$S = \{\overline{n} \in N_0^2 | n_1 \leq M_1,\ n_2 \leq M_2,\ n_1 + n_2 \leq M\}.$$

If all transitions $\overline{n} \to \overline{n} - e_i + e_j$ within S are allowed, the dual state space S^d is given by

$$S^d = \{\overline{m} \in N_0^2 | m_1 \leq M_1,\ m_2 \leq M_2,\ m_1 + m_2 \leq M - 1\}.$$

Dualizing back using the first backward dualizing method as described in Remark 3.6 gives under the assumption that all transitions $\overline{m} \to \overline{m} + e_j - e_i$ within S^d are allowed

$$S^- = S,$$

as can easily be seen from the definition (3.5) of S^-.

Assume that the primal routing function $p_{ij}(\overline{m})$ is state independent, i.e. $p_{ij}(\overline{m}) = p_{ij}$ for all $\overline{m} \in S^d$. Then, unless X is reversible, X does not posses a product form equilibrium distribution. For a product form equilibrium distribution to exist, a special blocking protocol must be used for blocking certain transitions near the boundary of S. To this end, consider the *stop protocol*, also referred to as *communication blocking* (cf. [11]). In the stop protocol, if one queue reaches its limit all other queues are stopped and customers are not allowed to enter the queueing network. Graphically, this protocol can easily be explained. To this end, consider the primal traffic equations (2.9):

$$\sum_{j \in \mathcal{N}} \{\gamma_i p_{ij}(\overline{m}) - \gamma_j p_{ji}(\overline{m})\} = 0,\ i \in \mathcal{N}. \tag{6.4}$$

For each $i \in \mathcal{N}, \overline{m} \in S^d$, (6.4) states that the flow for transitions in which a customer leaves queue i ($\sum_j c_i p_{ij}$) balances with the flow for transitions in which a customer enters queue i ($\sum_j c_j p_{ji}$). For fixed $\overline{m}$ this is visualized in Figure 4, where for $i \in \mathcal{N}$ these flows are depicted. As $p_{ij}(\overline{m}) \equiv p(\overline{m} + e_i, \overline{m} + e_j)$ is independent of $\overline{m}$, for all $\overline{m}$ flows are balanced in exactly the same way. But if $\overline{m}, \overline{m} + e_1 \in S, \overline{m} + e_2 \notin S$, i.e. if $m_2 = M_2$, transitions $\overline{m} \leftrightarrow \overline{m} + e_2$, $\overline{m} + e_1 \leftrightarrow \overline{m} + e_2$ are prohibited. Therefore, for the flow balance not to be disturbed, also the transitions $\overline{m} \leftrightarrow \overline{m} + e_1$ must be blocked. A similar argument holds if $m_1 = M_1$. The blocking protocol prohibiting the transitions $\overline{m} \leftrightarrow \overline{m} + e_1$ if $m_2 = M_2$ and $\overline{m} \leftrightarrow \overline{m} + e_2$ if $m_1 = M_1$ is called the *stop protocol.* If the stop protocol is used the dual state space S^d is given by

$$S_s^d = \{\overline{m} \in N_0^2 | m_1 \leq M_1 - 1,\ m_2 \leq M_2 - 1,\ m_1 + m_2 \leq M - 1\}.$$

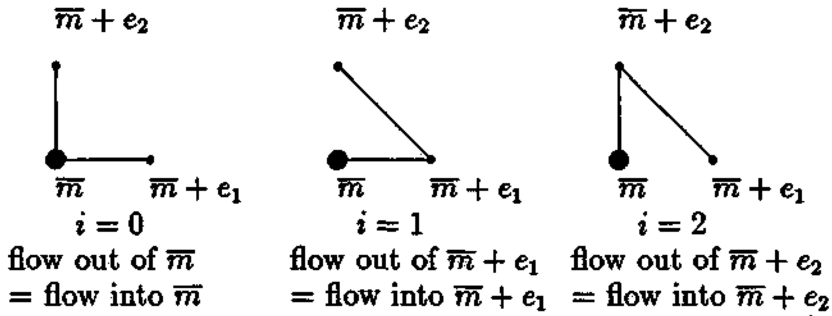

Figure 4. Balance of flows; primal process.

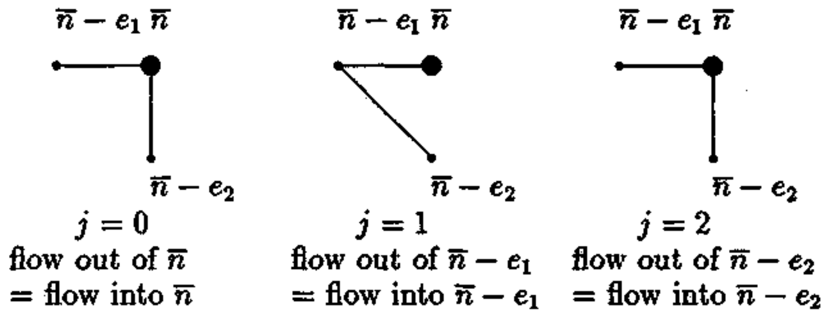

Figure 5. Balance of flows; dual process.

Similar to the assumption on the primal routing function, assume that the dual routing function is state independent, i.e. $p^d_{ij}(\overline{n}) = p^d_{ij}$ for all $\overline{n} \in S^-$. Then, for the dual process to posses a product form equilibrium distribution, transitions near the boundary of S^d have to be blocked. By contrast with the primal process, where transitions near the upper boundary of S have to be blocked, for the dual process to allow a product form equilibrium distribution transitions near the lower boundary of S^d must be blocked. To realize this, reconsider the dual traffic equations (3.3):

$$\sum_{i\in\mathcal{N}} \left\{ \frac{1}{\gamma_j} p^d_{ij}(\overline{n}) - \frac{1}{\gamma_i} p^d_{ji}(\overline{n}) \right\} = 0, \; j \in \mathcal{N}.$$

In Figure 5, for fixed $\overline{n}$, the balanced flows are depicted. As $p^d_{ij}(\overline{n})$ is independent of $\overline{n}$, for all $\overline{n}$ flows are balanced in exactly the same way. Therefore, if $\overline{n}, \overline{n}-e_1 \in S^d_s$, $\overline{n}-e_2 \notin S^d$, i.e. if $n_2 = 0$, transitions $\overline{n} \leftrightarrow \overline{n}-e_2$, $\overline{n}-e_1 \leftrightarrow \overline{n}-e_2$ are prohibited. Therefore, for the flow balance not to be disturbed, the transitions $\overline{n} \leftrightarrow \overline{n}-e_1$ must be blocked. A similar argument shows that if $n_1 = 0$ transitions $\overline{n} \leftrightarrow \overline{n}-e_2$ must be blocked. The protocol stopping arrivals to and departures from the non-starved queues if a queue is starved is called the *dual stop protocol.* Under the dual stop protocol, dualizing back according to the first backward dualizing method gives

$$S^-_s = \{\overline{n} \in N^2_0 | 0 < n_1 \le M_1 - 1, \; 0 < n_2 \le M_2 - 1, \; n_1 + n_2 \le M\}.$$

Clearly, $S_s^d \subset S^d$ and $S_s^- \subset S^-$. Thus, the dual state space depends on the specific blocking protocol.

6.4 Alternative transition rates; intermediate states

In many practical examples, the form of the transitions (e.g. $\overline{m} \to \overline{m} - e_i + e_j$) and the equilibrium distribution are observed only. In these cases the transition rates must be determined from these observations. This example gives two alternatives for these rates. Moreover, this example gives an alternative justification of Example 4.1.

Consider a dual process at state space S^d. Then a transition $\overline{m} \to \overline{m} + e_j - e_i$ for this process passes through $\overline{m} + e_j$ and the transition rates q^d have the form

$$q^d(\overline{m}, \overline{m} + e_j - e_i) = \frac{\phi(\overline{m} + e_j)}{\psi(\overline{m})} p_{ij}^d(\overline{m} + e_j). \tag{6.5}$$

In practical situations, however, the intermediate state, $\overline{m} + e_j$, is not observed. Therefore, a transition $\overline{m} \to \overline{m} + e_j - e_i$ for a process at S^d could pass through $\overline{m} - e_i$ resulting in transition rates

$$q^d(\overline{m}, \overline{m} - e_i + e_j) = \frac{\phi(\overline{m} - e_i)}{\psi(\overline{m})} p_{ij}(\overline{m} - e_i), \tag{6.6}$$

where $\phi : S^\star \to R^+$, $p_{ij} : S^\star \to R^+ \cup \{0\}$ and

$$S^\star = \{\overline{n} \in N_0^N | \exists i, j \in \mathcal{N},\ i \neq j : \overline{n} + e_i, \overline{n} + e_j \in S^d,\ q^d(\overline{n} + e_i, \overline{n} + e_j) > 0\}.$$

Assume that the process at S^d allows an equilibrium distribution π^d given by

$$\pi^d(\overline{m}) = B^d \psi(\overline{m}) \prod_{k=1}^N c_k^{m_k},\ \overline{m} \in S^d,$$

then the routing function p_{ij}^d or p_{ij} must be determined such that π^d satisfies the global balance equations (2.4) at S^d. As the routing functions reflect the site potential difference and state $\overline{m} + e_j$ contains one customer extra at both queue j and queue i when compared to $\overline{m} - e_i$, it is not unnatural to assume

$$p_{ij}^d(\overline{m} + e_j) = c_i c_j p_{ij}(\overline{m} - e_i).$$

Then, if the routing is state independent (e.g. $p_{ij}(\overline{m} - e_i) = p_{ij} \mathbf{1}(\overline{m} - e_i \in S^\star)$), from Example 4.1, $\{c_i\}_{i=1}^N$ is a solution to (2.9) if and only if it is a solution to (3.3).

This example shows that, unless specific information on the transition rates is available (e.g. speeds, blocking), (6.5) and (6.6) are indistinguishable descriptions of the process at S^d.

6.5 Alternative transition rates; traffic equations

Based on the intermediate states, the previous example gives a theoretical argument for the identification of a primal and a dual process. However, a more practical observation may be the basis of this identification. To this end, consider a process X at state space $S = \{\overline{n} \in N_0^N \mid \sum_{i=1}^N n_i = M\}$ with transition rates

$$q(\overline{n}, \overline{n} - e_i + e_j) = \mu_i p_{ij}, \ i, j \in \mathcal{N}. \tag{6.7}$$

Furthermore, assume a positive solution $\{c_i\}_{i=1}^N$ is known to

$$\sum_{i \in \mathcal{N}} \{\gamma_j \mu_i p_{ij} - \gamma_i \mu_j p_{ji}\} = 0, \ j \in \mathcal{N}.$$

Consider the dual formulation for X with dual routing function

$$p_{ij}^d = \mu_i p_{ij}, \ i, j \in \mathcal{N}$$

and $\phi \equiv 1$, $\psi \equiv 1$. Since (6.7) is independent of the intermediate state $\overline{n} - e_i$ or $\overline{n} + e_j$, the transition rates in the dual formulation are given by (6.7) also. From Lemma 3.3, X allows an invariant measure

$$\psi(\overline{n}) = \prod_{k=1}^N \left(\frac{1}{c_k}\right)^{n_k}, \ \overline{n} \in S.$$

6.6 Self dual process

A primal process is called *self dual* if a dual process corresponding to this primal process exists such that the primal process and the dual process are statistically indistinguishable.

Consider a primal process at state space S with transition rates (2.3) and equilibrium distribution π. Then for the dual process at state space S^d defined in (2.2) with transition rates (3.1) and equilibrium distribution π^d, the process is self dual if and only if

(i) $S^d = S$,
(ii) $q^d(\overline{n}, \overline{n} + e_j - e_i) = q(\overline{n}, \overline{n} - e_i + e_j)$, $\overline{n}, \overline{n} - e_i + e_j \in S$,
(iii) $\pi^d(\overline{n}) = \pi(\overline{n})$, $\overline{n} \in S$.

Note that the intermediate states for a dual transition and a primal transition are not the same. However, as these dual states are not observed during a transition this is not a problem.

As an illustration consider the following simple example. Consider an open queueing network consisting of N single server queues. The service rate at queue i is μ_i. Upon release from queue i a customer is routed to queue j according to the routing probability $\hat{p}_{ij}$. Assume that a positive solution $\{c_i\}_{i=1}^N$ exists to

$$\gamma_i \hat{p}_{ij} = \gamma_j \hat{p}_{ji}, \ i, j \in \mathcal{N}, \ \gamma_0 = \mu_0.$$

This queueing network can be modeled as a primal process at state space $S = N_0^N$ with transition rates $q(\overline{n}, \overline{n} - e_i + e_j) = \mu_i \hat{p}_{ij}$, where μ_0 is the rate of the Poisson arrival process to the queueing network. For $\phi \equiv 1$, $\psi \equiv 1$ and primal routing function $p_{ij} = \mu_i \hat{p}_{ij}$ the transition rates have the form (2.3). As $\{\frac{c_i}{\mu_i}\}_{i=1}^N$ is a positive solution to (2.9), for $\mu_i > c_i$, $i = 1, \ldots, N$, the primal process possesses an equilibrium distribution π given by

$$\pi(\overline{n}) = B \prod_{k=1}^{N} \left(\frac{c_i}{\mu_i}\right)^{n_k}. \tag{6.8}$$

The dual state space for this primal process is $S^d = S$. Consider the dual process with transition rates (3.1) with dual routing function

$$p_{ij}^d = p_{ij}, \; i, j \in \mathcal{N}.$$

Then $\{\frac{\mu_i}{c_i}\}_{i=1}^N$ is a positive solution to (3.3). Thus the dual process allows an invariant measure π^d given in (6.8) and the process is self dual.

A few remarks on self duality are to be made. Firstly, note that, in general, for (i) and (iii) to be valid, it need not be the case that (ii) is satisfied. Moreover, (ii) will, in general, not be satisfied. Secondly, from the above example, it seems to be the case that the transition rates for the dual process depend on the specific choice for ϕ, ψ and p_{ij}. This is not the case. To this end, consider the obvious choice for these functions:

$$\phi(\overline{n}) = \psi(\overline{n}) = \prod_{k=1}^{N} \left(\frac{1}{\mu_k}\right)^{n_k}, \quad p_{ij} = \hat{p}_{ij}.$$

In this case, for the process to be self dual, the dual routing function has to be defined as

$$p_{ij}^d(\overline{n}) = \mu_j \mu_i p_{ij}, \; \overline{n} \in S. \tag{6.9}$$

6.7 Batch servicing and routing

This paper introduces the notion of duality for stochastic processes in the setting of queueing networks in which single changes are allowed only. This notion of duality can immediately be generalized to queueing networks in which batch service and routing is allowed, however, the analysis will become more complex. Therefore, in this paper I have chosen to introduce the notion of duality for stochastic processes describing single changes queueing networks only. This section gives some remarks on the generalization of the setting in the present paper to queueing networks allowing batch servicing and batch routing also. To this end, consider the batch service batch routing queueing network as defined in [3]. For $\overline{g} = (g_1, \ldots, g_N)$ the number of customers leaving the queues and $\overline{g}' = (g_1', \ldots, g_N')$ the number of customers entering the queues in a transition, for $\overline{n}, \overline{n}' = \overline{n} - \overline{g} + \overline{g}' \in S$ the transition rates are (cf. [2], [3], [5], [6])

$$q(\overline{n}, \overline{n} - \overline{g} + \overline{g}') = \frac{\psi(\overline{n} - \overline{g})}{\phi(\overline{n})} p(\overline{g}, \overline{g}'; \overline{n} - \overline{g}).$$

Assume a positive solution $\{c_i\}_{i=1}^N$ exists to

$$\sum_{\overline{g}'} \left\{ \prod_k \gamma_k^{g_k} p(\overline{g}, \overline{g}'; \overline{m}) - \prod_k \gamma_k^{g'_k} p(\overline{g}', \overline{g}; \overline{m}) \right\} = 0. \tag{6.10}$$

From [3], the process allows an invariant measure Φ given by

$$\Phi(\overline{n}) = \phi(\overline{n}) \prod_k c_k^{n_k}, \ \overline{n} \in S.$$

The dual state space S^d can immediately be generalized to batch servicing and routing queueing networks. Also, the dual process can be defined similar to Definition 3.1. The transition rates for the dual process are

$$q^d(\overline{m}, \overline{m} + \overline{g}' - \overline{g}) = \frac{\phi(\overline{m} + \overline{g}')}{\psi(\overline{m})} p^d(\overline{g}, \overline{g}'; \overline{m} + \overline{g}').$$

Furthermore, if a positive solution $\{d_i\}_{i=1}^N$ exists to

$$\sum_{\overline{g}} \left\{ \prod_k \left(\frac{1}{\gamma_k}\right)^{g'_k} p^d(\overline{g}, \overline{g}'; \overline{m} + \overline{g}') - \prod_k \left(\frac{1}{\gamma_k}\right)^{g_k} p^d(\overline{g}', \overline{g}; \overline{m} + \overline{g}') \right\} = 0, \tag{6.11}$$

the dual process allows an invariant measure Ψ given by

$$\Psi(\overline{m}) = \psi(\overline{m}) \prod_{k=1}^N d_k^{m_k}, \ m_k \in S^d.$$

Thus, for batch servicing and routing queueing networks the dual process can be defined. Furthermore, the results of Section 5 can immediately be generalized.

6.8 Discrete-time queueing networks

[12] considers a discrete-time queueing network with *early arrivals*, i.e. arrivals occur just after the beginning of a time slot, while departures take place just before the end of a time slot. The state of the network is the number of customers present at the queues at the beginning of a time slot. Thus, for $\overline{x}_n = (x_1, \ldots, x_N)$ the number of customer present at the beginning of a time slot n, $\overline{a}_n = (a_1, \ldots, a_N)$ the number of customers arriving at the queues in time slot n, $\overline{d}_{n+1} = (d_1, \ldots, d_N)$ the number of customers departing form the queues in time slot n, the state at the beginning of time slot $n+1$ is given by

$$\overline{x}_{n+1} = \overline{x}_n + \overline{a}_n - \overline{d}_{n+1}.$$

As arrivals occur before departures, the transition $\overline{x}_n \to \overline{x}_{n+1}$ passes through $\overline{x}_n + \overline{a}_n$. In [12] this observation is explicitly used since the probability that departures $\overline{d}_{n+1}$ occur depends on $\overline{x}_n + \overline{a}_n$. Thus, the process in [12] is a dual process in the setting of the present paper. The primal process corresponding to this dual process is the following. Let $\overline{y}_n$ denote the intermediate state of the queueing network in time slot n, i.e. the state of the queueing network between arrivals and departures in time slot n. For this process departures occur before arrivals. The evolution of this process is given by

$$\overline{y}_{n+1} = \overline{y}_n - \overline{d}_{n+1} + \overline{a}_{n+1}.$$

Thus the primal process corresponds to *late arrivals.*

As this paper considers continuous-time single changes queueing networks only, the discrete-time queueing network with batch arrivals described here cannot be modeled. However, as is argued in Section 6.7 above, the notion of duality can immediately be generalized to a continuous-time model with batch service and batch routing. As the notion of duality introduced in this paper seems to be able to explain the difference between the results obtained in a continuous-time setting and the results obtained in a discrete-time setting, it is illustrative to consider the discrete-time model here in more detail. To this end, consider the transition probabilities given in [12]. The probability of servicing customers $\overline{d}_{n+1}$ in time slot n is given by

$$P\{\overline{d}_{n+1}|\overline{x}_n, \overline{a}_n\} = \prod_{i=1}^{N} \frac{c_i(x_i + a_i)}{d_i!}\alpha_i(x_i + a_i)\alpha_i(x_i + a_i - 1)\cdots\alpha_i(x_i - d_i + 1).$$

Customers served at queue i are independently routed to queue j according to the routing probability p_{ij}. The probability that customers $\overline{a}_n$ arrive into the network in time slot n is given by

$$P\{\overline{a}_n\} = \prod_{i=1}^{N} \frac{\lambda_i^{a_i}}{a_i!} e^{-\lambda_i}.$$

In [3] this model is discussed. It is shown that this discrete-time model corresponds to a continuous-time model with transition rates (2.3) with

$$\phi(\overline{n}) = \prod_{i=1}^{N} \frac{1}{c_i(n_i)} \prod_{k=1}^{n_i} \frac{1}{\alpha_i(k)},$$

$$\psi(\overline{m}) = \prod_{i=1}^{N} \prod_{k=1}^{m_i} \frac{1}{\alpha_i(k)},$$

and independent routing of customers. Thus, the equilibrium distribution for the primal process is given by

$$\pi(\overline{n}) = B \prod_{i=1}^{N} \frac{1}{c_i(n_i)} \prod_{k=1}^{n_i} \frac{c_i}{\alpha_i(k)},$$

where c_i is a positive solution to the traffic equations. This distribution is obtained in [3], [5]. The dual equilibrium distribution reads

$$\pi^d(\overline{n}) = B^d \prod_{i=1}^{N} \prod_{k=1}^{n_i} \frac{c_i}{\alpha_i(k)},$$

as is given in [12], where the process is observed at the end of time slots. Note that the dual routing function has to be chosen very carefully to obtain the correct form for the dual equilibrium distribution (cf. Section 4). A similar observation is made in [6], where the dual states are called base states. There the equilibrium distribution on base states is obtained. However, [6] does not introduce a process on these base states.

6.9 Customer-vacancy duality

Consider a closed queueing network consisting of single server queues only. The server at queue i works at rate μ_i, $i = 1, \ldots, N$, and the number of customers at queue i is not to exceed a given capacity constraint M_i, $i = 1, \ldots, N$. The state space of the primal process describing this queueing network when M customers are present in the network is given by

$$S_{\overline{M},M} = \{\overline{n} \in N_0^N | \sum_{i=1}^{N} n_i = M,\ n_i \leq M_i,\ i = 1, \ldots, N\},$$

where $\overline{M} = (M_1, \ldots, M_N)$. The transition rates for the primal process are

$$q(\overline{n}, \overline{n} - e_i + e_j) = \mu_i \hat{p}_{ij},\ \overline{n}, \overline{n} - e_i + e_j \in S_{\overline{M},M},\ i, j \in \mathcal{N},$$

where $\hat{p}_{ij}$ is the routing probability. Assume a positive solution $\{c_i\}_{i=1}^{N}$ exists to the traffic equations

$$\sum_{j=1}^{N} \{c_i \hat{p}_{ij} - c_j \hat{p}_{ji}\} = 0,\ i \in \mathcal{N}.$$

Similar to Example 6.6, define the primal routing function p_{ij} as

$$p_{ij}(\overline{m}) = \mu_i \hat{p}_{ij} \mathbf{1}(\overline{m} + e_i, \overline{m} + e_j \in S_{\overline{M},M}),$$

then for $M \leq \min_{i \in \mathcal{N}} M_i$, from Lemma 2.3, the primal process allows an equilibrium distribution π at $S_{\overline{M},M}$ given by

$$\pi(\overline{n}) = B \prod_{k=1}^{N} \left(\frac{c_k}{\mu_k}\right)^{n_k},\ \overline{n} \in S_{\overline{M},M}.$$

If $M > \min_{i\in\mathcal{N}} M_i$ the equilibrium distribution cannot immediately be obtained in closed form. However, for $M \geq \sum_{i=1}^N M_i - \min_{i\in\mathcal{N}} M_i$ the equilibrium distribution can be obtained by considering the vacancies at the queues. To this end, note that when n_i customers are present at queue i, $m_i = M_i - n_i$ vacancies are present at queue i. Furthermore, when a customer routes from queue i to queue j a vacancy is routed from queue j to queue i. Moreover, in a transition $\overline{n} \to \overline{n} - e_i + e_j$ first a customer leaves queue i and is subsequently routed to queue j. This corresponds to the creation of a vacancy at queue i and subsequently the removal of a vacancy at queue j. Thus the queueing network can be described using vacancies instead of customers and the process describing the evolution of the states in the vacancy description is a dual process. The state space of this dual process is

$$S^d_{\overline{M},M} = \{\overline{m} \in N_0^N \mid \sum_{i=1}^N m_i = \sum_{i=1}^N M_i - M,\ m_i \leq M_i,\ i = 1,\ldots,N\}$$

and the transition rates are

$$q^d(\overline{m},\overline{m} + e_j - e_i) = \mu_j \hat{p}_{ji},\ \overline{m},\overline{m} + e_j - e_i \in S^d_{\overline{M},M},\ i,j \in \mathcal{N}.$$

For the dual process to posses an equilibrium distribution π^d based on the dual traffic equations the dual stop protocol must be used (cf. Example 6.3). However, as is argued in Example 6.5, the intermediate state for a transition $\overline{m} \to \overline{m} + e_j - e_i$ is not observed. Thus this transition can be considered as passing through $\overline{m} - e_i$. In this primal description blocking will not occur. Then, if $\{d_i\}_{i=1}^N$ solves

$$\sum_{j\in\mathcal{N}} \{d_i \mu_j p_{ji} - d_j \mu_i p_{ij}\} = 0,\ i \in \mathcal{N}, \tag{6.12}$$

the vacancy process possesses an equilibrium distribution π^d given by

$$\pi^d(\overline{m}) = B^d \prod_{k=1}^N d_k^{m_k},\ \overline{m} \in S^d_{\overline{M},M}. \tag{6.13}$$

From (6.13) the equilibrium distribution for customers can immediately be obtained by inserting $n_i = M_i - m_i$, $i = 1,\ldots,N$. This gives

$$\pi(\overline{n}) = B \prod_{k=1}^N \left(\frac{1}{d_k}\right)^{n_k},\ \overline{n} \in S_{\overline{M},M}.$$

In the special case that the queueing network is a cyclic network a method considering vacancies at the queues is used in [4]. If the queueing network is cyclic,

$$d_i = \frac{1}{\mu_{i-1}},\ i = 2,\ldots,N,\ d_1 = \frac{1}{\mu_N},$$

is a solution to (6.12) and the equilibrium distribution π is given by

$$\pi(\overline{n}) = B \prod_{k=1}^N \mu_k^{n_{k-1}},\ \overline{n} \in S_{\overline{M},M},$$

as obtained in [4]. This example generalizes the approach used in [4] to non cyclic queueing networks, however, this example is not a straightforward application of the dual process presented in this paper. Note that the dual process presented here gives a justification for the dual routing function presented in Example 4.4.

7 Conclusion

For a product form queueing network with general transition rates described in Section 2, in Section 3 a dual process is introduced, based on the potential interpretation of the transition rates. In contrast to the primal process where in a transition first a unit leaves a queue and is subsequently routed among the queues, in the dual process first a unit enters a queue and subsequently a unit is released from a queue in the network. It is shown that the dual process is closely related to the process describing the queueing network when one customer is routed among the queues thus describing the queueing network as observed by a customer in transit. This is formalized in Section 5 where the Palm probabilities associated with the jumps of the primal process are studied. As a direct consequence of the introduction of the dual process, in the arrival theorem, the equilibrium distribution of the state as observed by a moving customer is not necessarily equal to the equilibrium distribution of the state of the queueing network with this customer removed. Furthermore, the arrival theorem is generalized to also include the process describing the evolution of the state observed by a moving customer. The relation between primal and dual process is discussed. Firstly, it is shown that the symmetrical relation between the primal and dual process as discussed in Section 2 and 3 is similar to the relation between the primal and dual problem in linear programming. Secondly, blocking of transitions is discussed. It is shown that the dual process depends on the specific blocking protocol for the primal process. Thirdly, the transitions are discussed. In these examples transition rates for the primal and dual process are constructed such that the primal and dual process describe a queueing network at the same state space with equal transition rates in the primal and dual description. Fourth, batch service and routing is discussed. It is shown that the approach of this paper can easily be generalized to batch service and routing queueing networks with a product form equilibrium distribution. As an application a discrete-time queueing network is discussed. For this network the discrepancy between the equilibrium distribution obtained in a discrete-time formulation and the equilibrium distribution obtained in a continuous-time formulation is discussed.

The approach taken in this paper is restricted to stochastic processes allowing single changes only, i.e. to stochastic processes modelling queueing networks in which one customer is allowed to change queues in a transition only. The assumptions can be relaxed (cf. Sections 6.1, 6.7), but the analysis will become very untransparent. Therefore, the present setting (a continuous-time product form queueing network with single changes) is chosen to introduce the dual process and to highlight the relationship between the primal and dual processes.

References

[1] Baccelli, F. and Brémaud, P.(1987). *Palm probabilities and stationary queues.* Lecture Notes in Statistics 41, Springer-Verlag.

[2] Boucherie, R. J. and N. M. van Dijk (1990). "Spatial birth- death processes with multiple changes and applications to batch service networks and clustering processes." *Adv. Appl. Prob.* **22** 433-455.

[3] Boucherie, R. J. and N. M. van Dijk (1991). "Product forms for queueing networks with state dependent multiple job transitions." *To appear: Adv. Appl. Prob.*

[4] Gordon, W. J. and G. F. Newell (1967). "Cyclic queueing systems with restricted length queues." *Operations Research* **15** 266-277.

[5] Henderson, W., C. E. M. Pearce, P. G. Taylor and N. M. van Dijk (1990). "Closed queueing networks with batch services." *Queueing Systems* **6** 59-70.

[6] Henderson, W. and P. G. Taylor (1990). "Some new results on queueing networks with batch movement." *Research Report*

[7] Hordijk, A. and N. M. van Dijk (1983). "Networks of queues Part I: Job-local-balance and the adjoint process. Part II: General routing and service characteristics." *Lecture notes in Control and Informational Sciences.* Springer-Verlag. **Vol. 60** 158-205.

[8] Kelly, F. P.(1979). *Reversibility and stochastic networks.* Wiley.

[9] Serfozo, R. F.(1989). "Markovian network processes: congestion dependent routing and processing." *Queueing Systems* **5** 5-36.

[10] Schrijver, A.(1986). *Theory of linear and integer programming.* Wiley.

[11] Van Dijk, N. M.(1991). "Product forms for queueing networks with limited clusters." *Research Report*

[12] Walrand, J.(1983). "A discrete-time queueing network." *J. Appl. Prob.* **20** 903-909.

[13] Walrand, J.(1988). *An introduction to queueing networks.* Prentice Hall.

[14] Whittle, P.(1986). *Systems in stochastic equilibrium.* Wiley.