NETWORKS OF QUASI-REVERSIBLE NODES

F. P. Kelly

Summary

Many analytical results are available for a network of queues when the nodes in the network have a simplifying property. This property, called here quasi-reversibility, was first identified by Muntz and has since been investigated by a number of authors. A closely related concept, partial balance, has been central to the investigation of insensitivity begun by Matthes.

Here we describe the concept of quasi-reversibility, provide new examples of quasi-reversible nodes, discuss the range of arrival rates for which a node remains quasi-reversible, and analyse a model of a communication network insensitive to patterns of dependence more general than have previously been considered.

1. Introduction

A great number of analytical results are available concerning the equilibrium behaviour of a queueing network when the nodes in the network have a certain simplifying property. This property, called here quasi-reversibility, was first identified by Muntz [17] and various examples of quasi-reversible nodes have been presented by Baskett, Chandy, Muntz and Palacios [2] and Kelly ([7], [8]).

Important aspects of the equilibrium behaviour of some quasi-reversible nodes are insensitive to the precise specification of the nodes. For example the equilibrium distribution of the number of customers in an M/G/K loss system depends on the service requirement distribution G only through its mean. This form of insensitivity have been extensively investigated by Matthes, Koenig, Nawrotski, Jansen and Schassberger ([11], [12], [15], [18], [19]; see also [5]). Central to this work is the feature of partial balance, a feature first observed in queueing networks by Whittle [24].

Quasi-reversibility and partial balance are closely related concepts, but it is perhaps worth pointing out that they direct attention to rather different aspects of complex systems. To illustrate their distinct emphases in a system amenable to analysis by either approach consider a closed network of J symmetric queues containing a total of N customers ([8], [20]). Quasi-reversibility focuses attention on the J queues and partial balance on the N customers. Both approaches lead to product form expressions for the equilibrium distribution; the first [7] stresses J factors, one for each queue, the second [11] stresses N factors, one for each customer. There are also technical distinctions: the first approach lends itself most easily to a treatment in terms of countable state space Markov processes, with results for systems involving arbitrary distributions having to be obtained by weak convergence arguments [1]; the second approach elegantly models various aspects of dependence using the theory of stationary point processes, but requires limiting arguments to cope with open systems. For further remarks on the relationship between the two approaches see [8] and [20].

In Section 2 of this paper quasi-reversibility is defined and the fundamental theorem for a network of quasi-reversible nodes is proved. The development presented is an extension of that to be found in Muntz [17] and Kelly [8], and has been much influenced by the ideas of

Melamed [16] and Walrand and Varaiya [23]. Section 3 describes a simple example of a quasi-reversible node. In Section 4 conditions are provided under which the arrival rates at a quasi-reversible node can be varied independently without destroying the property of quasi-reversibility. The linearity of the traffic equations arising in [2], [3] and [6] is a consequence of the ability to manipulate nodes in this manner. Not all quasi-reversible nodes can be so manipulated: Section 5 describes an example adapted from the work of Whittle on clustering phenomena [25] and touches on some of the unusual features, including non-linearity of the traffic equations, which may arise when such nodes form part of a network. Section 6 presents a clustering model exhibiting a form of insensitivity. The model allows patterns of dependence more general than have previously been considered and provides an interesting example with which to investigate the framework laid down in [11], [12], [15], [18] and [19]. A potential application to the modelling of communication networks is discussed.

2. Networks of Quasi-reversible Nodes

Let S be a countable set and let $Q = (q(x,x'), x,x' \in S)$ be a collection of non-negative real numbers. Set $q(x,x) = 0$ for all $x \in S$, and assume that

$$q(x) \triangleq \sum_{x' \in S} q(x,x')$$

is finite and positive for all $x \in S$. In the usual manner we can use Q to construct a continuous time Markov process with countable state space S: starting from state x allow the process to stay there for a period exponentially distributed with parameter $q(x)$, then move the process to state x' with probability $q(x,x')/q(x)$; allow the process to remain in state x' for a further period independent and exponentially distributed with parameter $q(x')$, and so on. The collection Q may be such that for some initial states this procedure generates an infinite number of jumps in a finite time with positive probability. When this is not so, i.e., in the non-explosive case, the procedure can be used to construct a Markov process defined on $[0,\infty)$ with an arbitrary initial distribution. In any event a discrete time Markov process defined on $(0,1,2,\ldots)$ can be constructed by using the same procedure but with the holding periods in each state set to one: call this process the jump chain associated with Q.

Suppose now that the transition rates Q admit a positive invariant measure $(\pi(x), x \in S)$, taken here to mean that $(\pi(x), x \in S)$ is a collection of positive numbers satisfying

$$\pi(x)q(x) = \sum_{x' \in S} \pi(x')q(x',x) \quad x \in S . \tag{1}$$

Define $q'(x,x')$ by

$$\pi(x)\, q'(x,x') = \pi(x')\, q(x',x), \tag{2}$$

let

$$Q' = (q'(x,x'),\ x,x' \in S)$$

and let

$$q'(x) = \sum_{x' \in S} q'(x,x').$$

Observe that equation (1) implies

$$q(x) = q'(x). \tag{3}$$

If $(x(t), t \in \mathbb{R})$ is a non-explosive stationary Markov process with transition rates Q and stationary distribution π then the transition rates Q' are precisely those of the reversed process $(x(-t), t \in \mathbb{R})$.

Assume now that certain transitions are identified with arrivals or departures of customers (or items) of class $c \in C$, for C a finite or countable collection of possible customer classes. Specifically suppose that for each $(c,x) \in C \times S$ there are subsets $S^a(c,x)$, $S^d(c,x) \subset S$

satisfying

$$S^a(c_1,x) \cap S^a(c_2,x) = \phi \quad S^d(c_1,x) \cap S^d(c_2,x) = \phi \quad \forall c_1 \neq c_2 \tag{4}$$

$$\{(x,x'):\ x' \in S^a(c_1,x),\ x \in S^d(c_2,x')\} = \phi \quad \forall c_1,c_2 \tag{5}$$

with the interpretation that a transition from x to $x' \in S^a(c,x)$ signals the arrival of a customer of class c, and a transition from $x' \in S^d(c,x)$ to x signals the departure of a customer of class c. Condition (4) ensures that customers arrive singly and that they depart singly. This condition is essential in what follows. Condition (5) rules out the possibility that a single transition may signal both an arrival and a departure, and is made for convenience of exposition (cf. [8]). Since customer classes are essentially defined in terms of subsets of the state space S it will be natural in what follows to suppose that the symbol C fixes the family

$$((S^a(c,x), S^d(c,x)):\ (c,x) \in C \times S).$$

Call the node (Q,π,C) <u>quasi-reversible</u> if there exists a collection $(\alpha(c),\ \beta(c),\ c \in C)$ such that

$$\sum_{x' \in S^a(c,x)} q(x,x') = \alpha(c) \tag{6}$$

$$\sum_{x' \in S^d(c,x)} q'(x,x') = \beta(c) \tag{7}$$

for all $(c,x) \in C \times S$. If $(x(t), t \in \mathbb{R})$ is a non-explosive stationary Markov process with transition rates Q and equilibrium distribution π then quasi-reversibility reduces to the property that the state of the process at time $t, x(t)$, is independent of:

(i) the arrival times of class c customers, $c \in C$, subsequent to time t;

(ii) the departure times of class c customers, $c \in C$, prior to time t.

This property in turn implies that:

(i) arrival times of class c customers, for $c \in C$, form independent Poisson processes;

(ii) departure times of class c customers, for $c \in C$, form independent Poisson processes.

An especially simple example of a quasi-reversible node is the following system:

$$S = \{0,1,2\} \quad C = \{1\}$$

$$q(x,x') = \alpha \qquad x' = x + 1 \pmod 3$$

$$= \beta \qquad x' = x - 1 \pmod 3$$

$$\pi(x) = 1 \qquad x \in S$$

$$S^a(1,x) = x+1 \pmod 3 \qquad S^d(1,x) = x-1 \pmod 3.$$

This node can be viewed as acting as a source of rate α and a sink of rate β.

We shall now discuss how a number of quasi-reversible nodes can be linked together to form a network. Let $((Q_j, \pi_j, C), j = 1,2,\ldots,J)$ be a finite collection of quasi-reversible nodes, and use the subscript j generally to identify entities associated with the j^{th} node. Thus arrivals and departures are defined for node j in terms of the family

$$((S_j^a(c,x_j), S_j^d(c,x_j)): (c,x_j) \in C \times S_j)$$

and the collection of customer classes C is the same for each node. Let

$$(\xi,k): C \times \{1,2,\ldots,J\} \to C \times \{1,2,\ldots,J\}$$

be a bijection, with the interpretation that when a customer of class c departs from node j he transmutes into a customer of class $\xi(c,j)$ who then arrives at node $k(c,j)$. Assume that

$$\alpha_{k(c,j)}(\xi(c,j)) = \beta_j(c), \tag{8}$$

a requirement that will emerge as necessary to match departures of class c customers from node j with arrivals of class $\xi(c,j)$ customers at node $k(c,j)$.

Now define a collection of transition rates $Q = (q(x,x'), x,x' \in S)$ on the state space $S = S_1 \times S_2 \times \cdots \times S_J$ as follows. If

$$x = (x_1, x_2, \ldots, x_j, \ldots, x_k, \ldots, x_J) \tag{9}$$

and

$$x' = (x_1, x_2, \ldots, x_j', \ldots, x_k', \ldots, x_J) \tag{10}$$

where

$$x_j' \in S_j^d(c,x_j),\ k = k(c,j),\ x_k' \in S_k^a(\xi(c,j), x_k) \tag{11}$$

put

$$q(x,x') = q_j(x_j', x_j) \frac{q_k(x_k, x_k')}{\alpha_k(\xi(c,j))};$$

if

$$x = (x_1, x_2, \ldots, x_j, \ldots, x_J) \tag{12}$$

and

$$x' = (x_1, x_2, \ldots, x_j', \ldots, x_J) \tag{13}$$

where

$$x_j' \notin \bigcup_{c \in C} S_j^a(c,x_j),\ x_j \notin \bigcup_{c \in C} S_j^d(c,x_j') \tag{14}$$

put

$$q(x,x') = q_j(x_j, x_j');$$

otherwise put $q(x,x') = 0$. The transition rates Q are thus defined in the obvious way: a node behaves as it would in isolation except that arrivals are triggered exogenously, by departures from other nodes, rather than by an endogenous mechanism.

<u>Theorem</u>. The transition rates Q admit a positive invariant measure

$$\pi(x) = \prod_{j=1}^{J} \pi_j(x_j) \qquad x \in S$$

Proof. Define a collection of transition rates Q' in terms of $(Q'_j,\ j = 1,2,\ldots,J)$ as follows. If relations (9), (10) and (11) hold put

$$q'(x',x) = q'_k(x'_k,x_k)\,\frac{q'_j(x_j,x'_j)}{\beta_j(c)}\ ;$$

if relations (12), (13) and (14) hold put

$$q'(x',x) = q'_j(x'_j,x_j);$$

otherwise put $q'(x',x) = 0$. The definition (2) of Q'_j, $j = 1,2,\ldots,J$, and the equalities (8) imply that

$$\pi(x)q(x,x') = \pi(x')q'(x',x) \qquad x,x' \in S\ , \tag{15}$$

Since

$$q(x) = \sum_{j=1}^{J} [q_j(x_j) - \sum_{c\in C} \alpha_j(c)]$$

and

$$q'(x) = \sum_{j=1}^{J} [q'_j(x_j) - \sum_{c\in C} \beta_j(c)]$$

it follows from equations (3) and (8) that

$$q(x) = q'(x) \qquad x \in S$$

This and equation (15) establish the result:

$$\sum_{x\in S} \pi(x)q(x,x') = \pi(x')q(x') \qquad x \in S \qquad \square$$

The generality of our approach has resulted in some simplicity in the statement and proof of the Theorem, since we have been able to postpone asking whether Q is explosive, reducible or positive recurrent. In applications, however, we usually seek a unique stationary distribution rather than just a particular invariant measure. This final step must be justified by appeal to specific properties of the network under

consideration, but the general line of argument usually proceeds as follows ([9], [10]). If $S^* \subset S$ is a closed communicating class then $(\pi(x),\ x \in S^*)$ is a positive invariant measure for the transition rates $Q^* = (q(x,x'),\ x,x' \in S^*)$. If the Markov process constructed from Q^* is non-explosive then it has a stationary distribution if and only if

$$B^{-1} \triangleq \sum_{x\in S^*} \pi(x) < \infty\ , \tag{16}$$

and when this condition is satisfied $(B\pi(x),\ x \in S^*)$ is the unique stationary distribution. Observe that condition (16) may be satisfied even if some or all of the measures $(\pi_j,\ j = 1,2,\ldots,J)$ are not summable.

Sometimes interest is focused not directly on the Markov process constructed from Q or Q^*, but on chains embedded in this process - for example we may be interested in the state observed at times immediately following a particular sort of transition. To obtain results for such chains it is useful to consider the transition probabilities

$P = (p((x,y),\ (y,z)),\ x,y,z \in S)$ where

$$p((x,y),\ (y,z)) = q(y,z)/q(y).$$

These are just the transition probabilities of the Markov chain $((x(n),\ x(n+1)),\ n = 0,1,\ldots)$ formed by taking each successive pair of states of the jump chain $(x(n),\ n = 0,1,\ldots)$. It is immediately verified that an invariant measure for P is $(\pi(x)q(x,y),\ (x,y) \in S^2)$, and from this invariant measures for chains embedded in the sequence $((x(n),\ x(n+1)),\ n = 0,1,\ldots)$ can be readily deduced. If $(x(t),\ t \in \mathbb{R})$ is a non-explosive, irreducible stationary Markov process with transition rates Q^* and stationary distribution $(\pi(x),\ x \in S^*)$ then $(\pi(x)q(x,y),\ x,y \in S^*)$ is the essentially unique ~~stationary distribution~~ invariant measure for the chain $((x(n),\ x(n+1)),\ n = 0,1,\ldots)$ with state space $S^* \times S^*$. In this case $\pi(x)q(x,y)$ has an interpretation as the probability flux from state x to state y [8].

Often the embedded chain of interest has itself an invariant measure of product form but the appropriate closed communicating class to which the measure should be restricted differs in some respect from the state space S* of the original process. For example, results contained in [4], [8], [14] and [21] are concerned with a chain embedded at certain arrival times in a closed network and the appropriate class is isomorphic to the state space of a closed network with one less customer. As another example [4] if a closed network with homogenous customers is observed at just the times when the number in a particular queue increases from $n-1$ to n and if the state of the particular queue is deleted from the observation then the appropriate class for the resulting chain is isomorphic to the state space of a closed network with n less customers and one less queue.

3. A Many Server Queue

The examples of quasi-reversible nodes presented by Baskett, Chandy, Muntz and Palacios [2] and Kelly [7] are widely known. Here and in Section 5 we describe two simple examples not covered in those papers.

The first example is a queue with s servers at which customers of a single class arrive in a Poisson stream of rate α. The servers may differ in efficiency: specifically, a customer's service time at server i is exponentially distributed with parameter μ_i, for $i = 1,2,\ldots,s$. Define the state of the queue to be the vector $x = (n, i_1, i_2, \ldots, i_{s-n})$, read as (n) when $n \geq s$, where n is the number of customers at the queue and $i_1, i_2, \ldots, i_{s-n}$ is a list of the free servers arranged in order according to the length of time they have been free. Suppose that if a customer arrives to find the queue in state $x = (n, i_1, i_2, \ldots, i_{s-n})$ with $n < s$ he is allocated to server i_r with probability $p(r, s-n)$,

$r = 1,2,\ldots,s-n$. For example if $p(1,m) = 1$ for $m = 1,2,\ldots,s$ then a customer is always allocated to the server who has been idle for the longest time. If the customer arrives to find $n \geq s$ he waits in line. It is elementary to check that an invariant measure for the resulting transition rates Q is

$$\pi(x) = \prod_{r=1}^{s-n} \frac{\mu_{i_r}}{\alpha} \qquad n < s$$

$$= \left(\frac{\alpha}{\sum_{i=1}^{s} \mu_i} \right)^{n-s} \qquad n \geq s$$

The transition rates Q' defined by equation (2) are easily calculated and can be regarded as describing a similar s-server queue with a slightly different method of handling idle servers. It then follows that, with the obvious transitions signalling arrivals and departures of customers of the single class, the queue is quasi-reversible, with $\alpha(1) = \beta(1) = \alpha$. If $\Sigma\mu_i < \alpha$ the invariant measure π can be normalized to give the unique stationary distribution. In equilibrium the service time of a customer is distributed as a mixture of exponential distributions. The convex combination defining the mixture depends on the arrival rate α as well as on $\mu_1, \mu_2, \ldots, \mu_n$, and the service times of successive customers are dependent.

Note that if customers leaving the queue who have been served by server i are assigned class i then the queue is not quasi-reversible. In constrast if customers of class c, $c \in C$, arrive in independent Poisson streams of rate $\alpha(c)$, where $\alpha = \Sigma\alpha(c)$, and if a customer's class neither changes nor affects his progress as he passes through the queue then the queue is quasi-reversible. To show this the state of the process must be expanded from x to $(x, \underline{c})$ where $\underline{c} = (c_1, c_2, \ldots, c_n)$ determines the class of each customer in each possible position in the system: an invariant measure is then

$$\pi(x,\underline{c}) = \pi(x) \prod_{r=1}^{n} \frac{\alpha(c_i)}{\alpha}$$

and the conditions for quasi-reversibility are readily verified with $\beta(c) = \alpha(c)$ for all $c \in C$.

The above discussion shows that the queue is quasi-reversible for all values of the arrival rates $\alpha(c)$, $c \in C$, satisfying $\Sigma\alpha(c) < \infty$. Chandy, Howard and Towsley [5] have observed that symmetric queues also have this property. The next Section derives the property in a more general setting.

4. Varied Arrival Rates

Consider a quasi-reversible node (Q,π,C) with arrival and departure rates $\alpha(c)$, $\beta(c)$, $c \in C$. Define a new collection of transition rates $Q^+ = (q^+(x,x'),\ x,x' \in S)$ by

$$q^+(x,x') = \frac{\alpha^+(c)}{\alpha(c)}\, q(x,x') \text{ if } x' \in S^a(c,x)$$

$$= q(x,x') \quad \text{otherwise}$$

where $\alpha^+(c) = 0$ if and only if $\alpha(c) = 0$, and $\Sigma\alpha^+(c) < \infty$. The interpretation here is that the arrival rate of class c customers has been altered from $\alpha(c)$ to $\alpha^+(c)$, for $c \in C$. The next result gives sufficient conditions for the altered node to be quasi-reversible.

<u>Proposition</u>. If

(a) there exists a function $n: C \times S \to \mathbb{Z}$ such that
if either $q(x,x') > 0$ or $q(x',x) > 0$ then

$$x' \in S^a(c,x) \cup S^d(c,x) \iff n(c,x') = n(c,x) + 1$$

$$x' \notin S^a(c,x) \cup S^d(c,x) \iff n(c,x') = n(c,x)$$

(b) $\alpha(c) = \beta(c) \quad c \in C$

then the node (Q^+,π^+,C) is quasi-reversible, where

$$\pi^+(x) = \pi(x) \prod_{c\in C} \left[\frac{\alpha^+(c)}{\alpha(c)}\right]^{n(c,x)} .$$

Remark. The integer $n(c,x)$ can be regarded as the number of class c customers in the node when its state is x. Conditions (a) and (b) can then be interpreted as a requirement that the node be customer conserving for each class $c \in C$. Condition (b) can be deduced from condition (a) when the Markov process constructed from Q is positive recurrent. All the nodes considered in [2] and [7] and the networks formed from these nodes can be formulated so that they satisfy conditions (a) and (b).

Proof. Let $\underline{n} = (n(c),\ c \in C) \in \mathbb{Z}^C$. Consider a collection of transition rates Q_Ψ defined on the state space $\{\underline{n}: \Sigma n(c) < \infty\}$ by

$$q(\underline{n},\ T_c\underline{n}) = \alpha(c)\, \frac{\Psi(T_c\underline{n})}{\Psi(\underline{n})}$$

$$q(T_c\underline{n},\underline{n}) = \alpha(c)$$

where

$$T_{c'}\underline{n} = (n(c) + I[c=c'],\ c \in C)$$

with all other transition rates zero. An invariant measure for Q_Ψ is clearly Ψ. If we identify a transition from $\underline{n}$ to $T_c\underline{n}$ as a <u>departure</u> of a customer of class c and a transition from $T_c\underline{n}$ to $\underline{n}$ as an <u>arrival</u> of a customer of class c then the node (Q_Ψ,Ψ,C) is quasi-reversible, with $\beta(c) = \alpha(c)$ for all $c \in C$.

Now form a network from the nodes (Q,π,C) and (Q_Ψ,Ψ,C) by having a departure of a class c customer from one node trigger the arrival of a class c customer at the other node, for each $c \in C$. Observe that those network states in which the state of the second node $(n(c),\ c \in C)$ corresponds precisely to the list $(n(c,x),\ c \in C)$ derived from the state x of the first node form a closed class and so an invariant measure over this class is

$$\Psi(n(c,x),\ c \in C)\,\pi(x)$$

The choice

$$\Psi(\underline{n}) = \prod_{c\in C} \left(\frac{\alpha^+(c)}{\alpha(c)}\right)^{n(c)} \tag{17}$$

and the bijection

$$((n(c,x),\ c \in C),\ x) \leftrightarrow x$$

establish that π^+ is a positive invariant measure for Q^+. Substitution into equations (6) and (7) then shows that (Q^+,π^+,C) is quasi-reversible, with arrival and departure rates $\alpha^+(c)$ for customers of class c, $c \in C$. □

Choices more general than expression (17) can be made for the function Ψ, and some of these produce nodes quasi-reversible with respect to customer classifications less fine than C ([8], see also [22]). Indeed any quasi-reversible node with state space $\{\underline{n},\ \Sigma' n(c) < \infty\}$ and arrival and departures rates $\alpha(c)$, $c \in C$, can be joined with the node (Q,π,C) to form a network, and the outcome viewed as a variation of the arrival rates at the node (Q,π,C). The result of Lam [13] can be interpreted in this way.

Examples of quasi-reversible nodes which do not satisfy condition (a) can be constructed from the reversible migration processes introduced by Kingman [10] (see [8]). While these nodes do not conserve customers of each class c, $c \in C$, they can be regarded as conserving customers unidentified by class. It is possible to show that if the arrival rates $\alpha(c)$, $c \in C$, at such a node are all multiplied by the same factor the resulting node is quasi-reversible. In the next Section we shall discuss a node which does not conserve even unclassified customers, but first we shall describe an example at the other extreme where the arrival rates can depend on more than the information contained in $(n(c,x),\ c \in C)$.

Let x be the state of a series of first come first served M/M/1

queues at which arrivals of customers of class c form a Poisson process of unit rate for $c \in C$, where C is a finite set. Suppose that a customer's class neither changes nor affects his progress as he moves through the series of queues. The resulting node (Q,π,C) is quasi-reversible. Let $\underline{c}(x) = (c_1,c_2,\ldots,c_n)$ be the classes of the n customers in the series of queues arranged in order of their arrival at the first queue in the series so that, for example, c_1 is the class of the customer who has been in the node the least time. Observe that if at some point in time $\underline{c}(x)$ is given, its future evolution can be tracked by a simple updating procedure applied whenever an arrival at or departure from the node occurs. By joining the node (Q,π,C) to another quasi-reversible node it is possible to show that if the arrival rate of class c customers is altered to

$$\frac{\Psi(c,\underline{c}(x))}{\Psi(\underline{c}(x))}$$

when the state of the node is x then an invariant measure for the resulting system is

$$\Psi(\underline{c}(x))\ \pi(x) \qquad x \in S$$

provided

$$\sum_{c\in C} \Psi(\underline{c},c) = \sum_{c\in C} \Psi(c,\underline{c}) .$$

For example if

$$\Psi(\underline{c}) = p^{M(\underline{c})}$$

where

$$M(\underline{c}) = \#\{i:\ c_i = c_{i+1},\ 1 \le i \le n-1\}$$

then an arriving customer of class c is lost with probability p when $c = c_1$.

5. A Clustering Node

We shall now discuss in detail a quasi-reversible node at which the arrival rates $\alpha(c)$, $c \in C$, cannot be varied independently without loosing quasi-reversibility. Let $S = \mathbb{N}^2$ and let the non-zero transition rates from the collection Q be given by

$$q((n_1,n_2),\ (n_1+1,n_2)) = \alpha(1)$$
$$q((n_1,n_2),\ (n_1,n_2+1)) = \alpha(2)$$
$$q((n_1,n_2),\ (n_1-1,n_2)) = \mu_1 n_1$$
$$q((n_1,n_2),\ (n_1,n_2-1)) = \mu_2 n_2$$
$$q((n_1,n_2),\ (n_1-2,n_2+1)) = \gamma_{12} n_1 (n_1-1)$$
$$q((n_1,n_2),\ (n_1+2,n_2-1)) = \gamma_{21} n_2$$

Provided that

$$\left.\begin{array}{l} \delta_1^2\,\gamma_{12} = \delta_2\,\gamma_{21} \\ \text{where} \\ \alpha(1) = \delta_1\mu_1 \qquad \alpha(2) = \delta_2\mu_2 \end{array}\right\} \qquad (18)$$

the rates Q admit an invariant measure

$$\pi(n_1,n_2) = \frac{\delta_1^{n_1}}{n_1!}\ \frac{\delta_2^{n_2}}{n_2!}$$

With $C = \{1,2\}$ and

$$S^a(1,\ (n_1,n_2)) = S^d(1,\ (n_1,n_2)) = (n_1+1,n_2)$$
$$S^a(2,\ (n_1,n_2)) = S^d(2,\ (n_1,n_2)) = (n_1,n_2+1)$$

the node (Q,π,C) is quasi-reversible. It is not difficult to show that if the arrival rates $\alpha(1),\alpha(2)$ are altered to $\eta_1\,\alpha(1)$, $\eta_2\,\alpha(2)$ respectively the resulting system is quasi-reversible if and only if $\eta_2 = \eta_1^2$. Condition (a) of the preceding Section does not hold: all that is conserved is the sum $n_1 + 2n_2$.

The presence in a network of a quasi-reversible node without property (a) can give rise to interesting phenomena not observed in the networks considered in [2] and [7]. For example, suppose we form a network from the above node labelled 1, and two single-class quasi-reversible nodes, labelled 2 and 3, satisfying conditions (a) and (b). Link the nodes as indicated in Figure 1: the important point to notice here is that an item leaving node 1 will eventually return as an item of a different class.

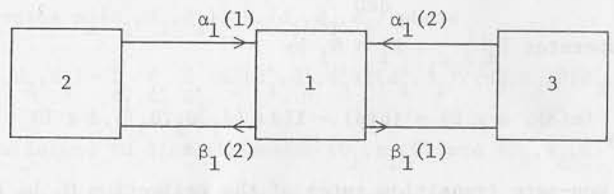


Figure 1. Network With Clustering Node

In addition to equations (8) we must now satisfy the non-linear constraint (18): a solution exists, given by

$$\alpha_1(1) = \alpha_1(2) = \beta_1(1) = \beta_1(2) = \frac{\mu_1^2\gamma_{21}}{\mu_2\gamma_{12}}\ .$$

The resulting network has, then, an invariant measure given by the fundamental theorem as a product of the invariant measures for each node. The invariant measure for node 1 will be summable if $\delta_1 < 1$ and $\delta_2 < 1$, that is if

$$\mu_1\gamma_{21} < \mu_2\gamma_{12} \quad \text{and} \quad \mu_1^2\gamma_{21} < \mu_2^2\gamma_{12}\ .$$

If in addition the invariant measures for nodes 2 and 3 are summable, then the product form is proportional to a stationary distribution for the Markov process, and in equilibrium the states of the three nodes are independent. Independence may thus be obtained in an irreducible network with no identifiable exogenous arrival streams. If the process is observed at, say, those instants when two class 1 items in node 1 are uniting to form a class 2 item and if the units so involved are left out of the description of the system, then the resulting Markov chain has a stationary distribution identical to that of the process.

6. An Insensitive Clustering Network

We begin this Section with another example of a clustering node.

Let D be a countable set of unit types, and let

$$S_1 = \{\underline{n} = (n(d),\ d \in D): \sum_{d\in D} n(d) < \infty\} .$$

Define the operator $T_{d_1d_2d_3}: S_1 \to S_1$ by

$$T_{d_1d_2d_3}(n(d),\ d \in D) = (n(d) - I[d \in \{d_1,d_2,d_3\}], d \in D) ,$$

and let the non-zero transition rates of the collection Q_1 be given by

$$q_1(\underline{n}, T_{d_1d_2d_3}\underline{n}) = \alpha_1(d_1,d_2,d_3) \prod_{i=1}^{3} \left[\frac{\lambda(d_i)}{\rho(d_i)}\right] n(d_i) \qquad (19)$$

$$q_1(T_{d_1d_2d_3}\underline{n}, \underline{n}) = \alpha_1(d_1,d_2,d_3) . \qquad (20)$$

An invariant measure for Q_1 is

$$\pi_1(\underline{n}) = \prod_{d\in D} \left[\frac{\rho(d)}{\lambda(d)}\right]^{n(d)} \frac{1}{n(d)!}$$

Set $C = D^3$ and identify transitions (19) and (20) as signalling respectively the departure or arrival of an item of class (d_1,d_2,d_3). The system (Q_1,π_1,C) is then quasi-reversible with $\alpha_1(d_1,d_2,d_3) = \beta_1(d_1,d_2,d_3)$.

Consider now a second node which operates as follows. Items of class $(d_1,d_2,d_3) \in C$ arrive in a Poisson stream of rate $\alpha_2(d_1,d_2,d_3)$. They pass independently through the node, an item labelled (d_1,d_2,d_3) on arrival taking a random period of time whose distribution is

determined by (d_1,d_2,d_3) and which can be represented by a passage time in a countable state space Markov process. Let the mean of this random period be $m(d_1,d_2,d_3)$. Upon arrival at the node an item of class (d_1,d_2,d_3) is allocated a second label (d_1',d_2',d_3') with probability $P(d_1,d_1')P(d_2,d_2')P(d_3,d_3')$, where $P: D^2 \to [0,1]$ is a transition probability matrix, and on departure it leaves as an item of class (d_1',d_2',d_3'). Without difficulty (although not without tedium) it is possible to define formally the node (Q_2,π_2,C) corresponding to this description and to show that it is quasi-reversible with arrival and departure rates $\alpha_2(d_1,d_2,d_3)$, $\beta_2(d_1,d_2,d_3)$ where

$$\beta_2(d_1,d_2,d_3) = \sum_{d_1'} \sum_{d_2'} \sum_{d_3'} \alpha_2(d_1',d_2',d_3')P(d_1',d_1)P(d_2',d_2)P(d_3',d_3). \qquad (21)$$

We now intend to link the nodes (Q_1,π_1,C) and (Q_2,π_2,C) together. To satisfy condition (8) we must ensure that

$$\alpha_1(d_1,d_2,d_3) = \beta_1(d_1,d_2,d_3) = \alpha_2(d_1,d_2,d_3) = \beta_2(d_1,d_2,d_3) \qquad (22)$$

Suppose that the transition matrix P admits a positive invariant measure $(\rho(d),\ d \in D)$, and use the symbol $\leftrightarrows$ to indicate the communication relation induced by P. From the equality (21) it follows that condition (22) is met when

$$\alpha_1(d_1,d_2,d_3) = \alpha_2(d_1,d_2,d_3) = \rho(d_1)\rho(d_2)\rho(d_3)f(d_1,d_2,d_3)$$

for any function $f: D^3 \to [0,\infty)$ satisfying

$$d_1 \leftrightarrows d_1',\ d_2 \leftrightarrows d_2',\ d_3 \leftrightarrows d_3' \implies f(d_1,d_2,d_3) = f(d_1',d_2',d_3') .$$

The resulting network will then have invariant measure given by the fundamental theorem.

As an application of the above discussion, consider a communication network as illustrated in Figure 2. A collection of centres are connected by channels. Two centres may be in communication via a joining channel, in which case the triple so formed must be disjoint

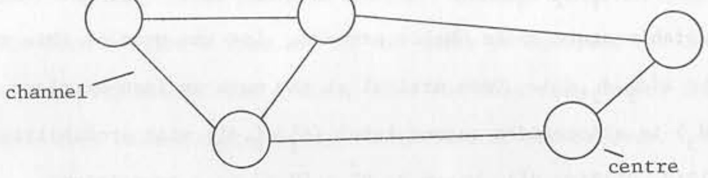


Figure 2. Communication Network

from all other such triples. If a centre or channel is not a member of a triple call it idle. To relate this communication network to the preceding discussion identify centres and channels as the basic units, so that each centre or channel is labelled with an element $d \in D$. The idle units then correspond to the occupants of node 1, and a linked triple corresponds to an item in node 2. Each time a centre or channel takes part in linked triple its state d changes in accordance with the transition probabilities P, but independently of the states of the units to which it is linked. To model the fact that centres and channels retain their geographical identify assume the transition matrix P has a number of closed communicating classes, one for each centre and one for each channel. To indicate which triples are geographically feasible let $f(d_1,d_2,d_3) = 1$ when the triple (d_1,d_2,d_3) identifies two centres and a channel which physically joins them, and let $f(d_1,d_2,d_3) = 0$ otherwise.

The requirements that there be exactly one unit associated with each of the closed communicating classes determined by P and that linked triples be geographically feasible identifies a closed communicating class S* for the overall network Q. If $\Sigma\rho(d) < \infty$ then the invariant measure for Q will be summable over S*, and so its restriction to S* will be the unique stationary distribution. Various consequences follow from the form of this distribution. For example, the equilibrium probability that $((d_1^i,d_2^i,d_3^i),\ i = 1,2,\dots,I)$ gives the

list of linked triples (with their current states) and that $(d_1^o,d_2^o,\dots,d_{N-3I}^o)$ gives the list of idle units (with their current states) is

$$B\left(\prod_{j=1}^{N-3I}\frac{\rho(d_j^o)}{\lambda(d_j^o)}\right)\prod_{i=1}^{I}\rho(d_1^i)\rho(d_2^i)\rho(d_3^i)m(d_1^i,d_2^i,d_3^i) \qquad (23)$$

when B is a normalizing constant, obtainable by summation. Observe the influence on this probability of $\lambda(d)$, the propensity of a node in state d to link, and $m(d_1,d_2,d_3)$, the mean link time of a triple (d_1,d_2,d_3). Various possibilities for the link time distributions are available. For example suppose that for each $d \in D$ we have a distribution F_d. If X_d is a random variable with distribution F_d the link time of the triple (d_1,d_2,d_3) could be distributed as, say,

$$\min(X_{d_1},X_{d_2},X_{d_3}) \qquad (24)$$

or

$$X_{d_1} + X_{d_2} + X_{d_3}\ . \qquad (25)$$

The technical restriction to passage times prevents the choice $X_{d_1}X_{d_2}X_{d_3}$, but observe that when we can write $m(d_1,d_2,d_3) = m(d_1)m(d_2)m(d_3)$ the product form (23) separates further.

Now focus attention on a single unit. The sequence of states taken by the unit is easily described, forming a Markov chain with transition matrix P. However the sequence of link times associated with the unit has a much more complicated structure depending not only on the unit's own state but also on the states of the units to which it happens to be linked. The resulting pattern of dependence in the sequences of link times associated with the various units is markedly more complex than occurs, for example, in the dependent sequences of service requirements associated with the various customers in a closed network of symmetric queues.

We shall now attempt to formulate the model of this Section within the framework provided by Matthes, Koenig, Nawrotski, Jansen and Schassberger ([11], [12], [15], [18], [19]). Using the terminology of [19], units can be identified as the elements of a generalized semi-Markov scheme provided link times of distinct triples have a common distribution, and the partial balance conditions are then found to be satisfied. However the scheme is not disconnected, since more than one element can be activated at the same time. If link times are generated from the forms (24) or (25) units can again be identified as elements, but the resulting formulation violates conditions imposed by the framework of a generalized semi-Markov scheme. Of course the correct formulation arises when we identify the set of elements with the set of possible linked triples - the items of the network formulation.

Various generalizations of the model of this Section can be carried through without disturbing its tractability: for example units can link to form larger clusters, and the linear factor n(d) in transition rate (19) can be generalized to reflect, perhaps, duplicate channels responding passively to link demands from centres. The guiding principle in the exploration of such generalizations is that they must leave node 1 quasi-reversible.

7. References

[1] Barbour, A. D. (1976) Networks of queues and the method of stages. Adv. in Appl. Probab. 8, 584-591.

[2] Baskett, F., Chandy, K. M., Muntz, R. R., and Palacios, F. G. (1975) Open, closed and mixed networks of queues with different classes of customers. J. Assoc. Comput. Mach., 22, 248-260.

[3] Beutler, F. J., Melamed, B., and Zeigler, B. P. (1977) Equilibrium properties of arbitrarily interconnected queueing networks. In P. R. Krishnaiah (Ed.), Multivariate Analysis IV, North-Holland, Amsterdam. pp. 351-370.

[4] Chandy, K. M., Herzog, U. and Woo, L. (1975) Parametric analysis of queueing networks. IBM J. Res. Develop. 19, 36-42.

[5] Chandy, K. M., Howard, J. H., and Towsley, D. F. (1977) Product form and local balance in queueing networks. J. Assoc. Comput. Mach., 24, 250-263.

[6] Kelly, F. P. (1975) Networks of queues with customers of different types. J. Appl. Probab., 12, 542-554.

[7] Kelly, F. P. (1976) Networks of queues. Adv. in Appl. Probab., 8, 416-432.

[8] Kelly, F. P. (1979) Reversibility and Stochastic Networks, Wiley, New York.

[9] Kendall, D. G. and Reuter, G. E. H. (1957) The calculation of the ergodic projection for Markov chains and processes with a countable infinity of states. Acta Math., 97, 103-143.

[10] Kingman, J. F. C. (1969) Markov population processes. J. Appl. Probab., 6, 1-18.

[11] König, D., and Jansen, U. (1974) Stochastic processes and properties of invariance for queueing systems with speeds and temporary interruptions. Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes, and of the 1974 European Meeting of Statisticians, Czechoslovak Academy of Sciences. pp. 335-343.

[12] König, D., Matthes, K., and Nawrotzki, K. (1967) Verallgemeinerungen der Erlangschen und Engsetschen Formeln (Eine Methode in der Bedienungstheorie), Akademie-Verlag, Berlin.

[13] Lam, S. S. (1977) Queueing Networks with population size constraints. IBM J. Res. Develop., 21, 370-378.

[14] Lavenberg, S. S. and Reiser, M. (1980) Stationary state probabilities at arrival instants for closed queueing networks with multiple types of customers. J. Appl. Probab., 17, 1048-1061.

[15] Matthes, K. (1962) Zur Theorie der Bedienungsprozesse. Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes, Czechoslovak Academy of Sciences. pp. 513-528.

[16] Melamed, B. (1979) On Poisson traffic processes in discrete-state Markovian systems with application to queueing theory. Adv. in Appl. Probab., 11, 218-239.

[17] Muntz, R. R. (1972) Poisson Departure Processes and Queueing Networks, IBM Research Report RC4145. A shortened version appeared in Proceedings of the Seventh Annual Conference on Information Science and Systems, Princeton, 1973, pp. 435-440.

[18] Schassberger, R. (1977) Insensitivity of steady-state distributions of generalized semi-Markov processes, Part I. Ann. Probab., 5, 87-99.

[19] Schassberger, R. (1978) Insensitivity of steady-state distributions of generalized semi-Markov processes, Part II. Ann. Probab., 6, 85-93.

[20] Schassberger, R. (1978) The insensitivity of stationary probabilities in networks of queues. Adv. in Appl. Probab., 10, 906-912.

[21] Sevcik, K. C. and Mitrani, I. (1981) The distribution of queueing network states at input and output instants. J. Assoc. Comput. Mach., 28, 358-371.

[22] Towsley, D. (1980) Queueing network models with state-dependent routing. J. Assoc. Comput. Mach., 27, 323-337.

[23] Walrand, J. and Varaiya, P. (1980) Interconnections of Markov chains and quasi-reversible queueing networks. Stochastic Process. Appl., 10, 209-219.

[24] Whittle, P. (1967) Nonlinear migration processes. Bull. Inst. Internat. Statist. 42, 642-647.

[25] Whittle, P. (1972) Statistics and critical points of polymerization processes. Supplement Adv. in Appl. Probab., 199-220.

Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, ENGLAND

Discussant's Report on
"Networks of Quasi-reversible Nodes,"
by F. P. Kelly

This elegant presentation contains some important results. The proof of the product form theorem given in Section 2 is a nice illustration of the technique which consists in guessing Q' to verify some invariant measure. The idea of considering invariant measures instead of invariant probability measures pays off in Section 4.

It is probably useful to complement the algebraic aspects of the theory emphasized in that presentation with some comments on the probabilistic interpretation of the concepts and results.

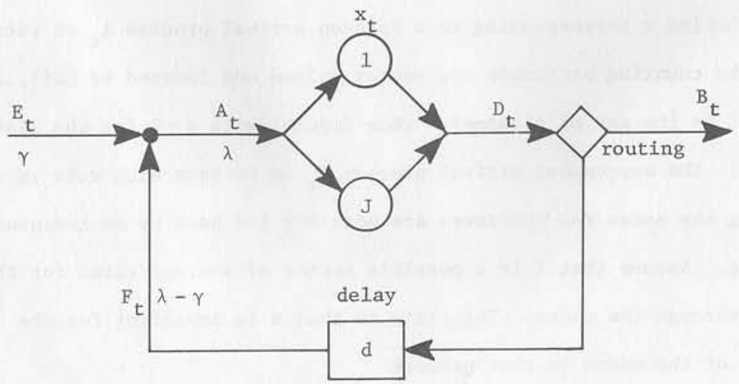
Notice that, in the stationary case, equations (6) [resp. (7)] say that the rates of the arrival processes [resp. the reversed departure processes] at time t are independent of x_t. Hence the equivalence with the conclusions (i), (ii).

The product form theorem relates independence properties: quasi-reversibility and product form. To explain why the calculations of Section 2 of [23] go through, I would like to sketch a probabilistic argument which isolates the role of quasi-reversibility and hopefully contributes to the intuitive understanding of those results.

Consider J nodes which are quasi-reversible under an invariant distribution π corresponding to a Poisson arrival process A_t of rate λ (all the counting processes are vector valued and indexed by $C\times\{1,\dots,J\}$, where C is the set of classes). (See figure, with $d = 0$ for the time being.) The exogeneous arrival process E_t is Poisson with rate γ; after leaving the nodes the customers are possibly fed back by an independent routing. Assume that λ is a possible vector of average rates for the flows through the nodes. The claim is that π is invariant for the states of the nodes in that network.

This is the argument. For $d > 0$, introduce a pure delay d in the links between nodes (see figure). Denote by x_t the resulting state process for the nodes, by A_t [resp. D_t] the total number of arrivals [resp. departures] at the nodes. Let F_t be the output of the delay box. Assume that x_0 has distribution π and that $F_{(0,d]} = \{F_t, 0 < t \leq d\}$ (the contents of the delay box at $t = 0$) is Poisson with rate $\lambda - \gamma$ and independent of x_0 and of $(E_t, t \geq 0)$. Then $A_{(0,d]} = F_{(0,d]} + E_{(0,d]}$ is Poisson with rate λ and independent of x_0. Thus (x_t, A_t, D_t) will behave for t in (0,d] as if the nodes were in isolation. By quasi-reversibility, it follows that $D_{(0,d]}$ is Poisson and independent of x_d. By independence of the routing, the same is true of $F_{(d,2d]}$. Also, x_t has distribution π for t in (0,d]. By induction, this proves that x_t has distribution π for all t. By letting d go to zero, one can then show that π must be an invariant distribution for the original network. (This is easy if the original network is a regular Markov chain.)

Notice also that the argument for $d > 0$ shows that the exit process $B_{(0,t]}$ is Poisson and independent of x_t and of $F_{(t,t+d]}$. This leads to the quasi-reversibility of the network. It is also clear that A_t is not Poisson in general: $A_{(t,t+d]}$ and $A_{(t+d,t+2d]}$ are generally not independent. The same argument shows that an invariant distribution for the open network remains invariant for the associated closed network.



Discussant: Dr. J. Walrand, Cornell University, 1434 Hanshaw Road, Ithaca, New York 14850.

Editors:

Ralph L. Disney
Department of Industrial Engineering
and Operations Research
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia 24061, USA

Teunis J. Ott
Bell Laboratories
Hölmdel, New Jersey 07733, USA

Applied probability-computer science.

(Progress in computer science ; no. 2-3)
1. Electronic data processing--Congresses.
2. Electronic digital computers--Programming--
Congresses. 3. Probabilities--Congresses.
4. Queuing theory--Congresses. I. Disney, Ralph L.,
1928- . II. Ott, Teunis J. III. Applied
Probability Technical Section-College of the
Operations Research Society of America, the Institute
of Management Sciences. IV. Series.
QA75.5.A66 1982 0001.64 82-18506

ISBN 3-7643-3067-8 (v. 1)
ISBN 3-7643-3093-7 (v. 2)

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Applied probability-computer science, the
interface / sponsored by Applied Probability
Techn. Sect., College of the Operations Research
Soc. of America, The Inst. of Management Sciences.
- Boston ; Basel ; Stuttgart : Birkhäuser

1981. January 5 - 7, 1981, Florida Atlantic
University, Boca Raton, Florida.
Vol. 1 (1982).
(Progress in computer science ; Vol. 2)
ISBN 3-7643-3067-8

NE: Florida Atlantic University Boca Raton, Fla.
GT

ISBN 3-7646-3067-8

Printed in USA

Applied Probability– Computer Science: The Interface Volume I

Sponsored by Applied Probability Technical Section
College of the Operations Research Society of America
The Institute of Management Sciences
January 5-7, 1981
Florida Atlantic University
Boca Raton, Florida

Ralph L. Disney,
Teunis J. Ott,
editors

TIMS

ORSA

1982

Birkhäuser
Boston • Basel • Stuttgart