# Quasi-Birth-and-Death Processes

Peter Taylor

Department of Mathematics and Statistics,
The University of Melbourne

Australian Research Council
Centre of Excellence for Mathematics
and Statistics of Complex Systems

# The M/M/1 queue

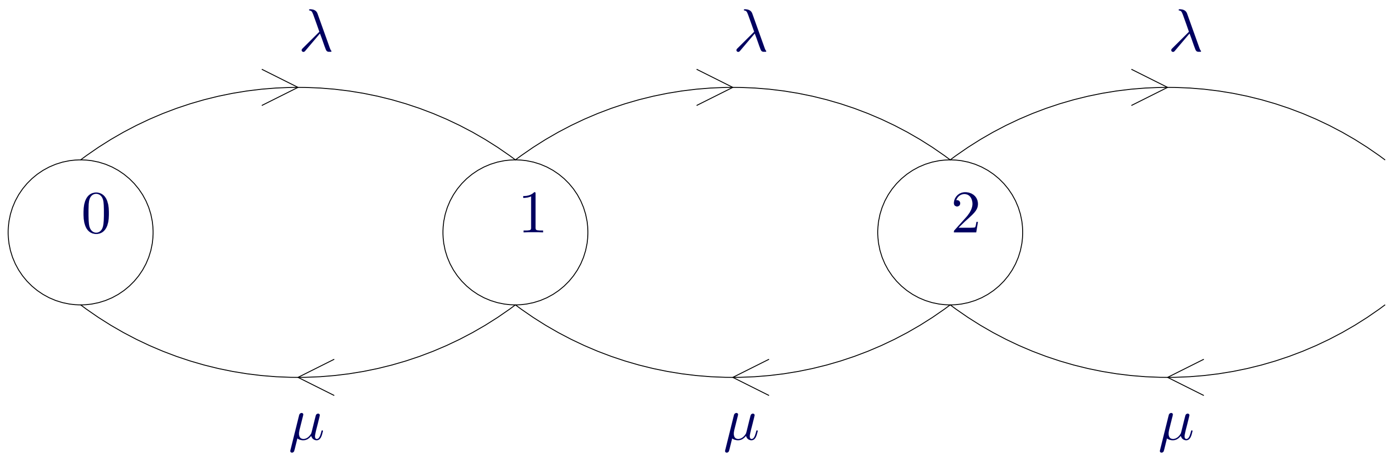



Figure 1: Transition structure for an M/M/1 queue

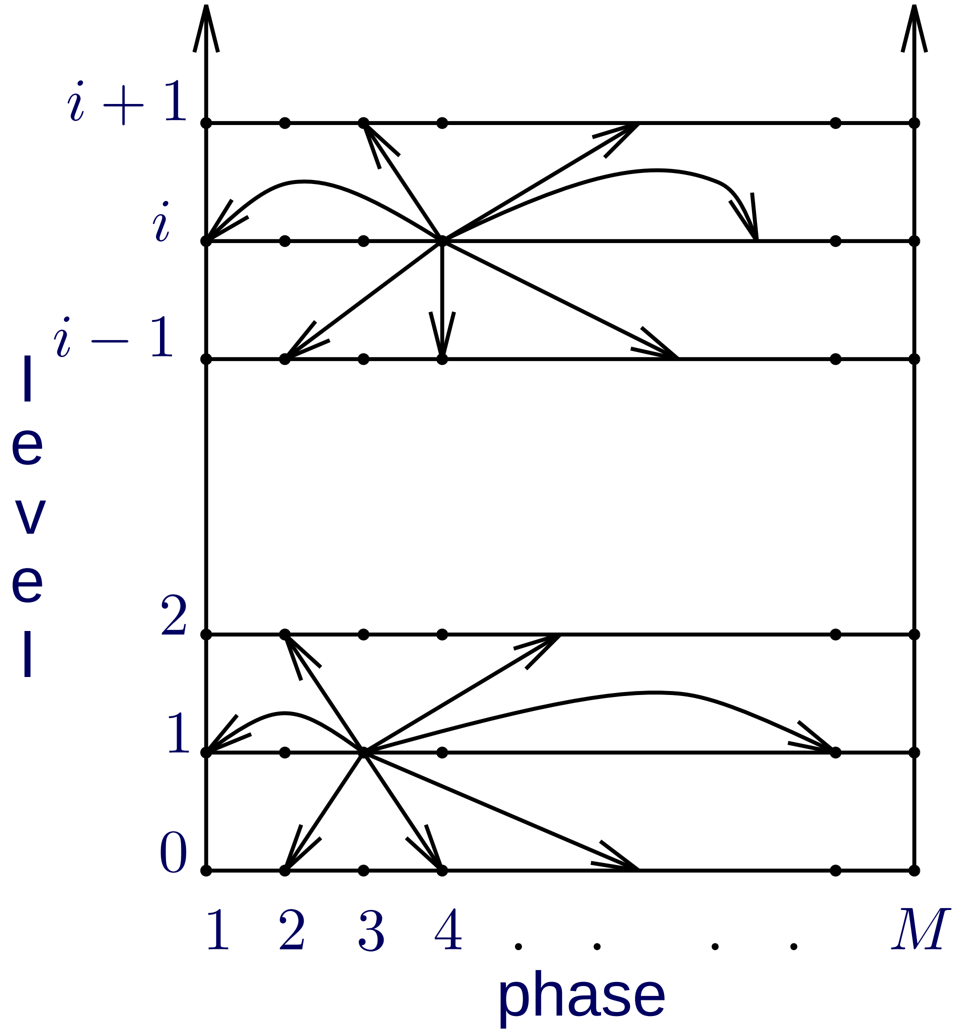
The M/M/1 queue can be modelled by a continuous-time Markov chain with state space $Z_+ = \{0, 1, \ldots\}$ and transition rates

$$
\begin{aligned}
q(n, n+1) &= \lambda, \quad n \geq 0 \\
q(n, n-1) &= \mu, \quad n \geq 1.
\end{aligned}
$$

So the transition matrix is

$$
\hat{Q} = \begin{bmatrix}
-\lambda & \lambda & 0 & 0 & \cdots \\
\mu & -(\mu+\lambda) & \lambda & 0 & \cdots \\
0 & \mu & -(\mu+\lambda) & \lambda & \cdots \\
0 & 0 & \mu & -(\mu+\lambda) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}.
$$

# A Quasi-Birth-and-Death Process



With a suitable ordering of the states, the transition matrix of a QBD has the block partitioned form

$$Q = \begin{bmatrix} \tilde{Q}_1 & Q_0 & 0 & 0 & \cdots \\ Q_2 & Q_1 & Q_0 & 0 & \cdots \\ 0 & Q_2 & Q_1 & Q_0 & \cdots \\ 0 & 0 & Q_2 & Q_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

When the QBD is in state $(k, i)$ we say that it is in level $k$ and phase $i$.

# Examples

- An M/M/1 queue in a random environment.
- A Ph/M/1 queue.
- An M/Ph/1 queue.
- MAP/Ph/1 queues.
- A model for a single cell in a PCS network with dynamic channel assignment (Lipper and Rumsewicz).
- Retrial models.
- Etc....

## The transition matrix

$$Q = \begin{bmatrix} \tilde{Q}_1 & Q_0 & 0 & 0 & \cdots \\ Q_2 & Q_1 & Q_0 & 0 & \cdots \\ 0 & Q_2 & Q_1 & Q_0 & \cdots \\ 0 & 0 & Q_2 & Q_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

looks like a block version of the transition matrix

$$\hat{Q} = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ \mu & -(\mu+\lambda) & \lambda & 0 & \cdots \\ 0 & \mu & -(\mu+\lambda) & \lambda & \cdots \\ 0 & 0 & \mu & -(\mu+\lambda) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

for the M/M/1 queue, so it is reasonable to ask what properties carry over from the M/M/1 queue in a “block sense" to a QBD.

# Stability

The single server queue is stable if and only if $\lambda < \mu$.
Let $\boldsymbol{x}$ be the solution to

$$\boldsymbol{x}[Q_0 + Q_1 + Q_2] = \mathbf{0}.$$

Then the QBD is positive recurrent if

$$\boldsymbol{x}Q_0\boldsymbol{e}' < \boldsymbol{x}Q_2\boldsymbol{e}',$$

null recurrent if

$$\boldsymbol{x}Q_0\boldsymbol{e}' = \boldsymbol{x}Q_2\boldsymbol{e}'$$

and transient if

$$\boldsymbol{x}Q_0\boldsymbol{e}' > \boldsymbol{x}Q_2\boldsymbol{e}'.$$

# The Stationary Distribution

The stationary distribution for a stable M/M/1 queue is derived by solving the system of second order difference equations

$$\pi(n)(\lambda + \mu) = \pi(n-1)\lambda + \pi(n+1)\mu,$$

for $n \geq 1$ and

$$\pi(0)\lambda = \pi(1)\mu.$$

The characteristic equation is

$$x^2\mu - x(\lambda + \mu) + \lambda,$$

which has roots $1$ and $\rho = \lambda/\mu$. Thus a summable solution for $\pi(n)$ exists only if $\rho < 1$, in which case

$$\pi(n) = (1-\rho)\rho^n.$$

Now let's think about the QBD. Assume it is positive recurrent and write the stationary distribution as $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \ldots)$. Then the equations for the stationary distribution are

$$\boldsymbol{\pi}_{n-1} Q_0 + \boldsymbol{\pi}_n Q_1 + \boldsymbol{\pi}_{n+1} Q_2 = \mathbf{0}$$

for $n \geq 1$ and

$$\boldsymbol{\pi}_0 \tilde{Q}_1 + \boldsymbol{\pi}_1 Q_2 = \mathbf{0}.$$

If we happen to be able to find a nonnegative matrix $R$ and a positive vector $\boldsymbol{x}_0$ such that

$$Q_0 + RQ_1 + R^2 Q_2 = 0,$$

$$\boldsymbol{x}_0 \left[ \tilde{Q}_1 + RQ_2 \right] = \mathbf{0},$$

and

$$\boldsymbol{x}_0 \sum_{k=0}^{\infty} R^k \boldsymbol{e} = \boldsymbol{x}_0 [I - R]^{-1} \boldsymbol{e} = 1,$$

then the stationary distribution would be given by

$$\boldsymbol{\pi}_n = \boldsymbol{x}_0 R^n.$$

We proceed by using the fact that a matrix with a physical interpretation has the required properties.

Let $\tau_i$ be the mean sojourn time in phase $i$ of level $k$. Define $R$ to be the matrix whose $(i, j)$th entry is $1/\tau_i$ times the expected sojourn time in phase $j$ of level $k+1$ before first return to level $k$ conditional on the process starting in phase $i$ of level $k$.

Then, some analysis gives us

$$Q_0 + RQ_1 + R^2Q_2 = 0.$$

The physical interpretation also gives us the fact that the $(i, j)$th entry of $R^\ell$ is $1/\tau_i$ times the expected sojourn time in phase $j$ of level $k+\ell$ before first return to level $k$ conditional on the process starting in phase $i$ of level $k$.

Under the assumption that the QBD is positive recurrent, we know that the total expected time spent in higher levels before the process returns to level $k$ must be finite, and so

$$\sum_{\ell=0}^{\infty} R^\ell \boldsymbol{e}$$

is finite.

Finally, some more analysis tells us that, under the assumption of positive recurrence, the matrix $\tilde{Q}_1 + RQ_2$ is an irreducible generator matrix and so there must be a positive solution $\boldsymbol{x}_0$ to

$$\boldsymbol{x}_0 \left[\tilde{Q}_1 + RQ_2\right] = \mathbf{0}.$$

Thus, the matrix $R$ that we defined above satisfies all the conditions we need. We have proved that the QBD has the matrix-geometric stationary distribution

$$\boldsymbol{\pi}_n = \boldsymbol{x}_0 R^n.$$

# How do we calculate $R$?

The matrix $R$ is, in fact, the minimal non-negative solution to the matrix quadratic equation

$$Q_0 + RQ_1 + R^2Q_2 = 0.$$

There are a number of good algorithms for deriving this solution.

I shall discuss these in the final part of this presentation.

# A variation

Sometimes the QBD has the block-partitioned form

$$Q = \begin{bmatrix} \tilde{Q}_1 & \tilde{Q}_0 & 0 & 0 & \cdots \\ \tilde{Q}_2 & Q_1 & Q_0 & 0 & \cdots \\ 0 & Q_2 & Q_1 & Q_0 & \cdots \\ 0 & 0 & Q_2 & Q_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

In this case the stationary distribution has the form

$$\boldsymbol{\pi}_n = \boldsymbol{x}_0 \tilde{R} R^{n-1}$$

where

$$\tilde{R} = -\tilde{Q}_0 \left[Q_1 + RQ_2\right]^{-1}$$

and $\boldsymbol{x}_0$ satisfies

$$\boldsymbol{x}_0 \left[\tilde{Q}_1 + \tilde{R}\tilde{Q}_2\right] = 0.$$

# Hitting Probabilities

Let's assume that the QBD starts in phase $i$ of level $k$. We are often interested in calculating the probability that it hits level $k-1$ in finite time, and does so in phase $j$. That is we are interested in the hitting probabilities on lower levels.

Let us store these probabilities in a matrix $G$. Thus, the $(i,j)$th entry of the matrix $G$ is the probability that the QBD will first enter level $k-1$ in phase $j$ given that it starts in phase $i$ of level $k$.

The probabilities $G_{ij}$ are the same as for the discrete-time QBD observed only at jump times. This has a transition matrix of the form

$$P = \begin{bmatrix} \tilde{D}_1 & D_0 & 0 & 0 & \cdots \\ D_2 & D_1 & D_0 & 0 & \cdots \\ 0 & D_2 & D_1 & D_0 & \cdots \\ 0 & 0 & D_2 & D_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The $D_i$ can be written in terms of the $Q_i$ via the relations

$$\begin{aligned} D_0 &= \Psi^{-1} Q_0 \\ D_1 &= \Psi^{-1} Q_1 + I \\ D_2 &= \Psi^{-1} Q_2 \end{aligned}$$

where $\Psi$ is a diagonal matrix with entries $1/\tau(i)$. Also

$$\tilde{D}_1 = \tilde{\Psi}^{-1} \tilde{Q}_1 + I$$

where $\tilde{\Psi}$ is a diagonal matrix whose entries are the negative of the diagonal entries of $\tilde{Q}_1$.

Now if the process starts in state $(k, i)$, then there are three possibilities for a path that will first hit level $k - 1$ in state $(k - 1, j)$,

- it can go straight to state $(k - 1, j)$, with probability $[D_2]_{ij}$,
- it can go to a state of the form state $(k, \ell)$, and then eventually hit level $k - 1$ in state $(k - 1, j)$, with probability $[D_1]_{i\ell} G_{\ell j}$, or
- it can go to a state of the form state $(k + 1, m)$, subsequently hit level $k$ in state $(k, \ell)$ and then eventually hit level $k - 1$ in state $(k - 1, j)$, with probability $[D_0]_{im} G_{m,\ell} G_{\ell j}$.

Taking these possibilities into account, we derive the equation

$$\begin{aligned} G &= D_2 + D_1 G + D_0 G^2 \\ &= \Psi^{-1} Q_2 + \left[\Psi^{-1} Q_1 + I\right] G + \Psi^{-1} Q_0 G^2 \end{aligned}$$

and so

$$Q_2 + Q_1 G + Q_0 G^2 = 0.$$

As with $R$, we are looking for the minimal nonnegative solution to this equation.

# Some extensions

Processes with transition matrices of the form

$$Q = \begin{bmatrix} \tilde{Q}_1 & Q_0 & 0 & 0 & \cdots \\ \tilde{Q}_2 & Q_1 & Q_0 & 0 & \cdots \\ \tilde{Q}_3 & Q_2 & Q_1 & Q_0 & \cdots \\ \tilde{Q}_4 & Q_3 & Q_2 & Q_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

are called Markov chains of GI/M/1 type.

Processes with transition matrices of the form

$$Q = \begin{bmatrix} \tilde{Q}_1 & \tilde{Q}_2 & \tilde{Q}_3 & \tilde{Q}_4 & \cdots \\ Q_0 & Q_1 & Q_2 & Q_3 & \cdots \\ 0 & Q_0 & Q_1 & Q_2 & \cdots \\ 0 & 0 & Q_0 & Q_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

are called Markov chains of M/G/1 type.

Markov chains of GI/M/1 type and of M/G/1 type can be analysed using matrix-analytic methods.

Like QBDs, Markov chains of GI/M/1 type have matrix-geometric stationary distributions.

Analysis of truncated and level-dependent versions of these processes is also possible.