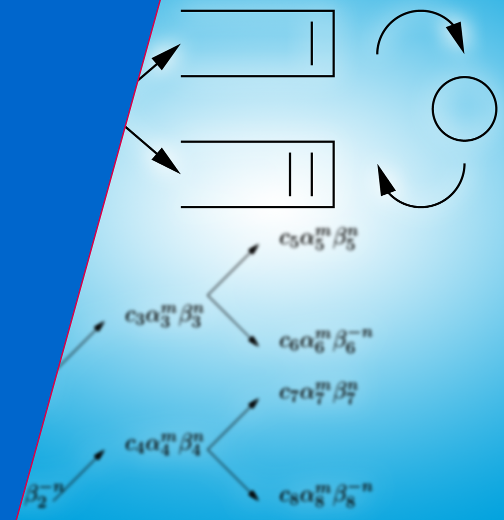
# Queueing models with multiple waiting lines: Direct methods

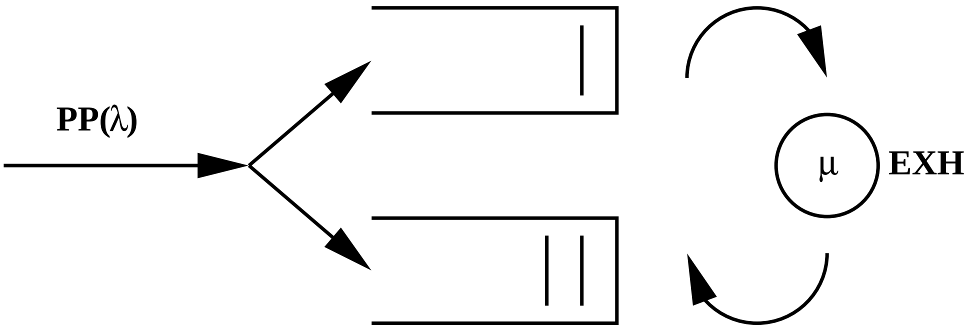


Technische Universiteit **Eindhoven** University of Technology

YEQT-VI 2012, Friday Nov 2

# Shortest queue and polling





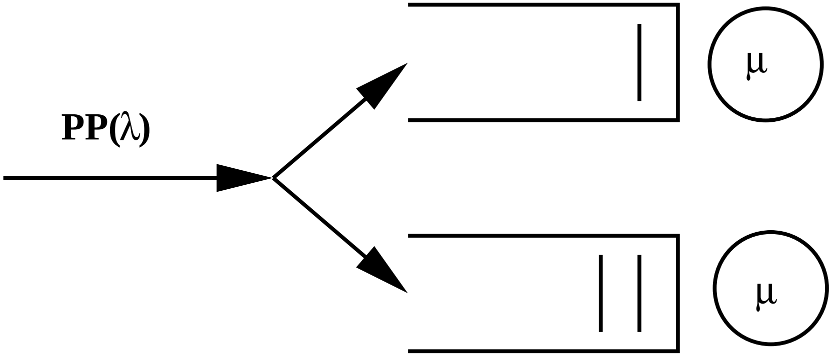

What is steady-state queue length distribution?

What is the waiting time distribution?

What is the mean waiting time?

# Shortest queue






What is steady-state queue length distribution?

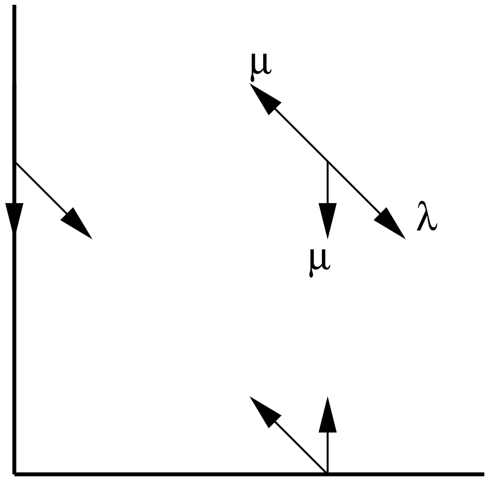
What is the waiting time distribution?

What is the mean waiting time?


Technische Universiteit
Eindhoven
University of Technology

# Shortest queue



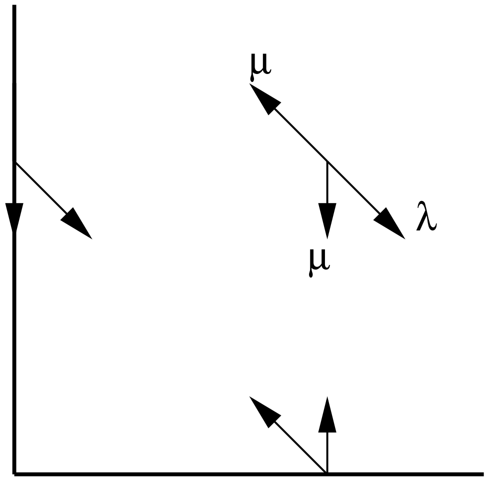



States are grid points $(m, n)$ where

- $m$ is length of shortest queue
- $n$ is difference between longest and shortest queue

What is steady-state queue length distribution $p_{m,n}$?

# Shortest queue

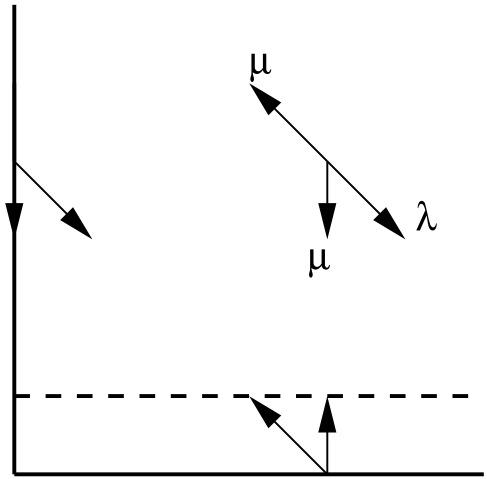



Balance equations ($\lambda + 2\mu = 1$)

$$
\begin{aligned}
p_{m,n} &= p_{m-1,n+1}\lambda + p_{m,n+1}\mu + p_{m+1,n-1}\mu \\
p_{m,1} &= p_{m-1,2}\lambda + p_{m,2}\mu + p_{m,0}2\mu + p_{m+1,0}\lambda \\
p_{m,0} &= p_{m,1}\lambda + p_{m+1,1}\mu \\
p_{0,n} &= p_{0,n}\mu + p_{0,n+1}\mu + p_{1,n-1}\mu
\end{aligned}
$$

# Shortest queue



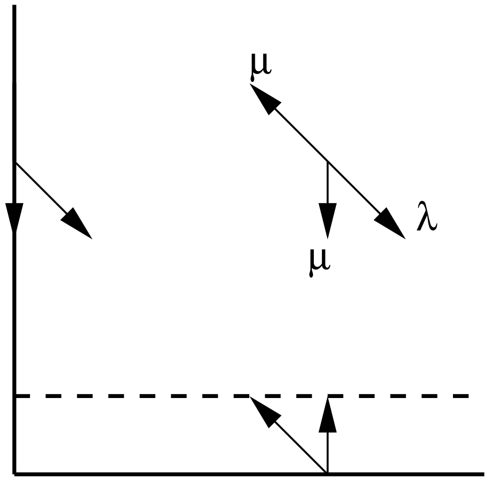



Balance equations ($\lambda + 2\mu = 1$)

$$
\begin{aligned}
p_{m,n} &= p_{m-1,n+1}\lambda + p_{m,n+1}\mu + p_{m+1,n-1}\mu \\
p_{m,1} &= p_{m-1,2}\lambda + p_{m,2}\mu + \\
&\quad (p_{m,1}\lambda + p_{m+1,1}\mu)2\mu + (p_{m-1,1}\lambda + p_{m,1}\mu)\lambda \\
p_{0,n} &= p_{0,n}\mu + p_{0,n+1}\mu + p_{1,n-1}\mu
\end{aligned}
$$

# Shortest queue






Basis solutions are products $\alpha^m \beta^n$ satisfying

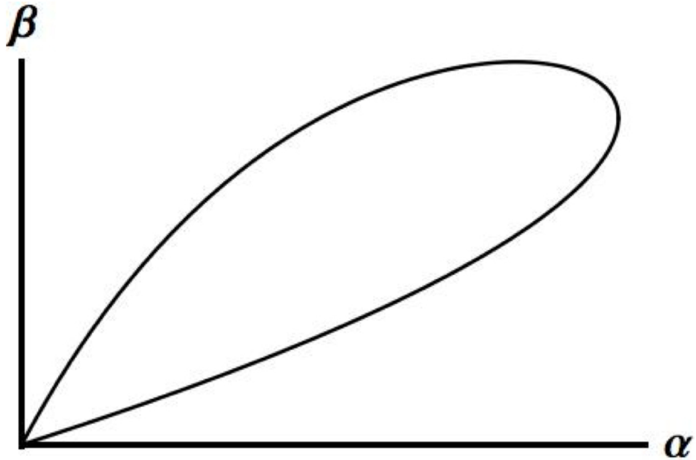
$$p_{m,n} \;=\; p_{m-1,n+1}\lambda + p_{m,n+1}\mu + p_{m+1,n-1}\mu$$

so $\alpha$ and $\beta$ are on the curve

$$\alpha\beta = \beta^2\lambda + \alpha\beta^2\mu + \alpha^2\mu$$

# Shortest queue



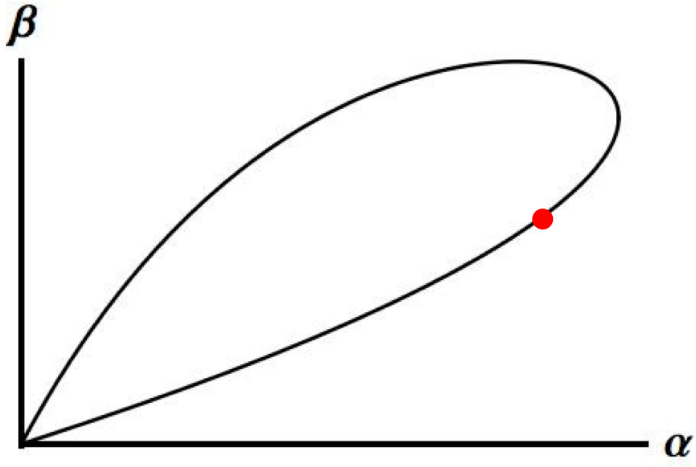



The curve of basis solutions

$$\alpha\beta = \beta^2\lambda + \alpha\beta^2\mu + \alpha^2\mu$$

# Shortest queue





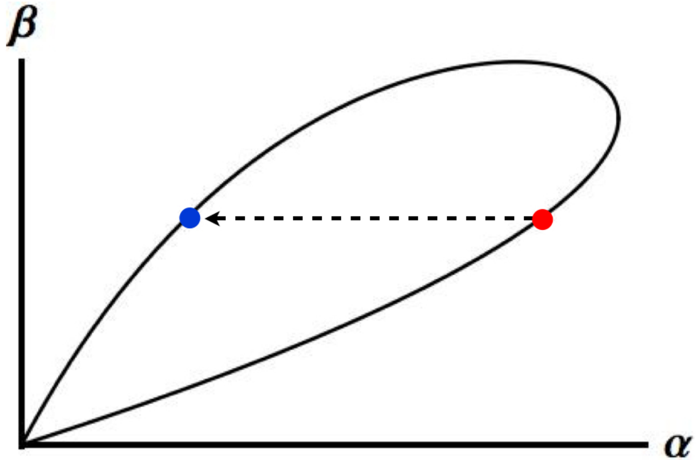

Initial basis solution satisfies balance equations for $n = 1$:

$$p_{m,n} \approx c_0 \alpha_0^m \beta_0^n$$

with $\alpha_0 = \rho^2$, $\beta_0 = \rho^2/(2+\rho)$ where $\rho = \frac{\lambda}{2\mu}$

# Shortest queue






Initial term $c_0 \alpha_0^n \beta_0^m$ violates balance equations for $m = 0$

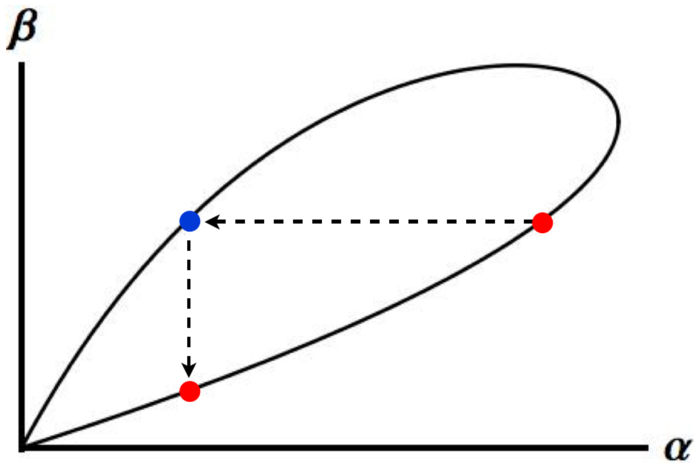
Add new term to compensate for this error:

$$p_{m,n} \approx c_0 \alpha_0^m \beta_0^n + c_1 \alpha_1^m \beta_1^n$$

where $\beta_1 = \beta_0$ and $c_1 = -c_0 \, \frac{\alpha_1 - \beta_1}{\alpha_0 - \beta_0}$

# Shortest queue






However, $c_1\alpha_1^n\beta_1^m$ violates balance equations for $n = 1$

Add again term:

$$p_{m,n} \approx c_0\alpha_0^m\beta_0^n + c_1\alpha_1^m\beta_1^n + c_2\alpha_2^m\beta_2^n$$

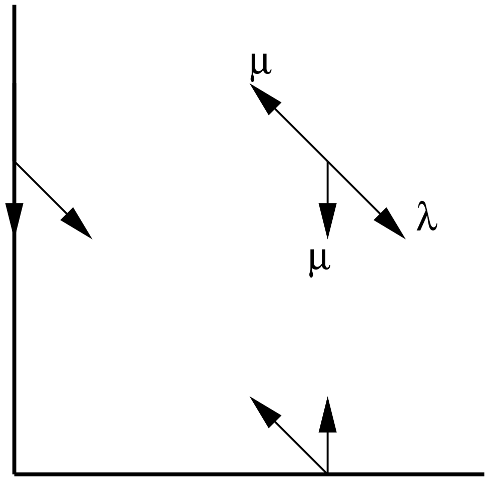
where $\alpha_2 = \alpha_1$ and $c_2 = -c_1 \frac{(\alpha_2+\rho)/\beta_2-(1+\rho)}{(\alpha_1+\rho)/\beta_1-(1+\rho)}$



TU/e Technische Universiteit Eindhoven University of Technology

# Shortest queue






Steady-state queue length distribution is given by:

$$p_{m,n} = c_0\alpha_0^m\beta_0^n + c_1\alpha_1^m\beta_1^n + c_2\alpha_2^m\beta_2^n + c_3\alpha_3^m\beta_3^n + \cdots$$



Technische Universiteit Eindhoven University of Technology

# Shortest queue



$$p_{m,n} = \sum_{i=0}^{\infty} c_i \alpha_i^m \beta_i^n$$

Observations:

- $\frac{1}{\alpha_i}, \frac{1}{\beta_i}$ are of form $A + B\left(ab^i + \frac{1}{ab^i}\right)$, so $\alpha_i, \beta_i \downarrow 0$ as $i \to \infty$
- series converges absolutely, faster further away from origin
- distribution of shortest queue

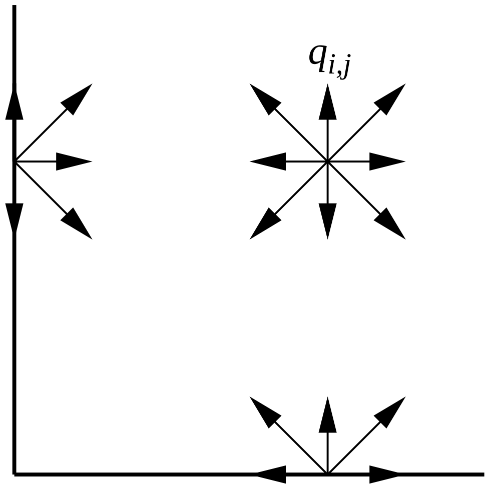
$$p_m = \sum_{n=0}^{\infty} p_{m,n} = \sum_{i=0}^{\infty} d_i (1 - \alpha_i) \alpha_i^m$$

- distribution and mean of waiting time

$$P(W > t) = \sum_{i=0}^{\infty} d_i \alpha_i e^{-\mu(1-\alpha_i)t}, \quad E(W) = \sum_{i=0}^{\infty} d_i \frac{\alpha_i}{\mu(1-\alpha_i)}$$

# Random walk in quarter plane



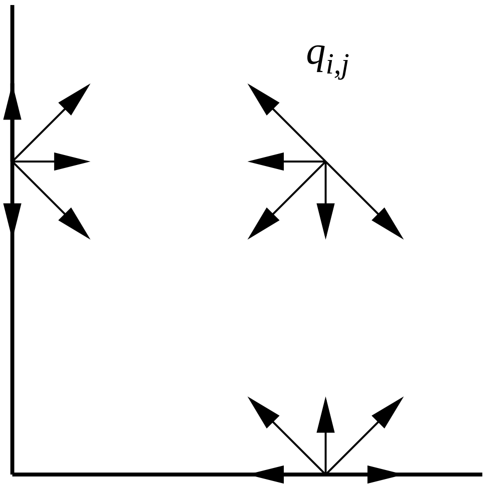



Then

$$p_{m,n} = \sum_{i=0}^{\infty} c_i \alpha_i^m \beta_i^n$$

# Random walk in quarter plane






Then

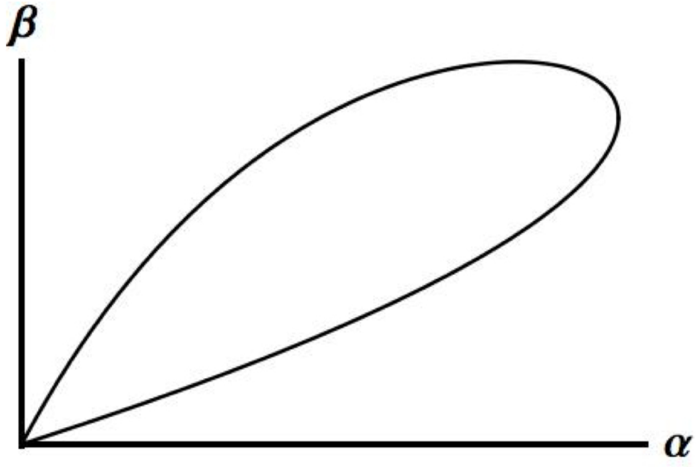
$$p_{m,n} = \sum_{i=0}^{\infty} c_i \alpha_i^m \beta_i^n$$

if no transitions to North, North-East and East:

$$q_{0,1} = q_{1,1} = q_{1,0} = 0$$

# Random walk in quarter plane



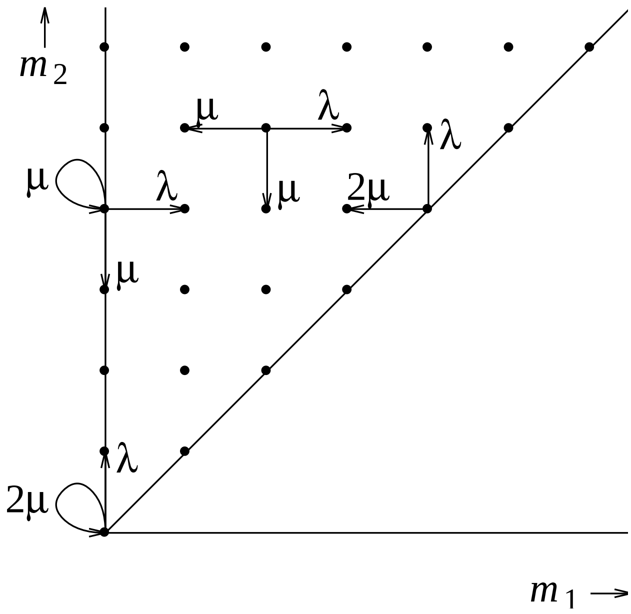



$q_{0,1} = q_{1,1} = q_{1,0} = 0$ implies that $\alpha$, $\beta$-curve is of the above form

which guarantees that $\alpha_i$, $\beta_i \to 0$ as $i \to \infty$

# Shortest queue: Bounds





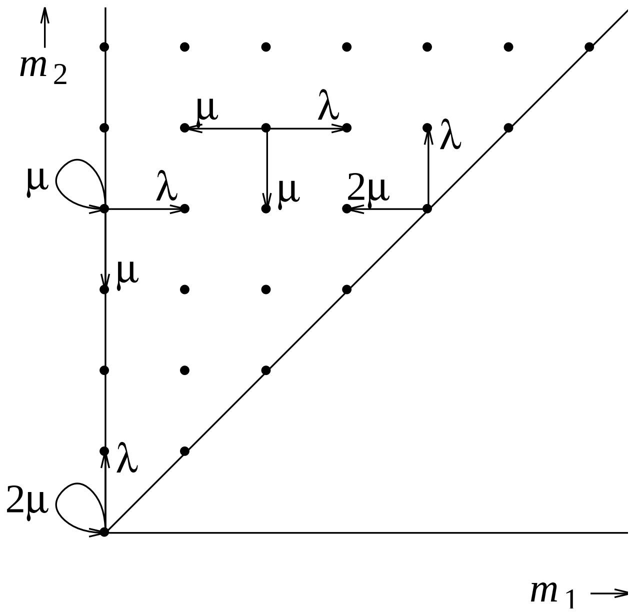

States are grid points $(m_1, m_2)$ where

- $m_1$ is length of shortest queue
- $m_2$ is length of longest queue



Technische Universiteit Eindhoven University of Technology

# Shortest queue: Bounds

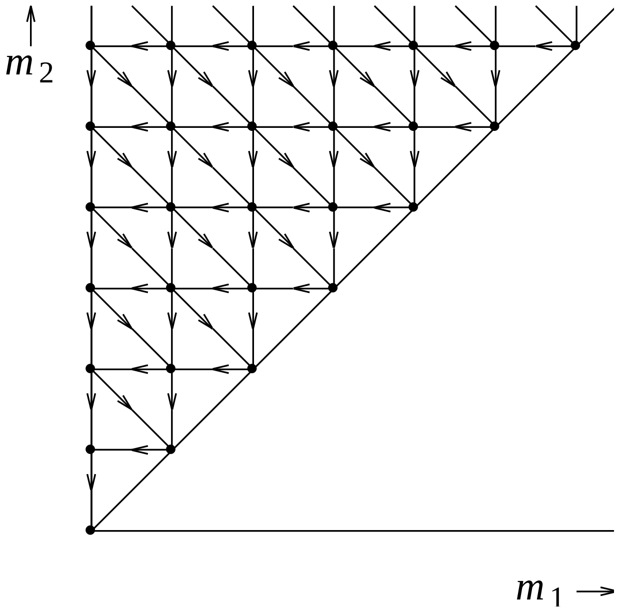



Average cost ($\lambda + 2\mu = 1$):

- $c(m_1, m_2) = m_1 + m_2$ cost per period
- $g$ is long-run average cost (mean number in system): Bounds for $g$?


Technische Universiteit Eindhoven University of Technology

# Shortest queue: Bounds




Precedence relations: state $(m_1, m_2)$ is more attractive than $(n_1, n_2)$ if

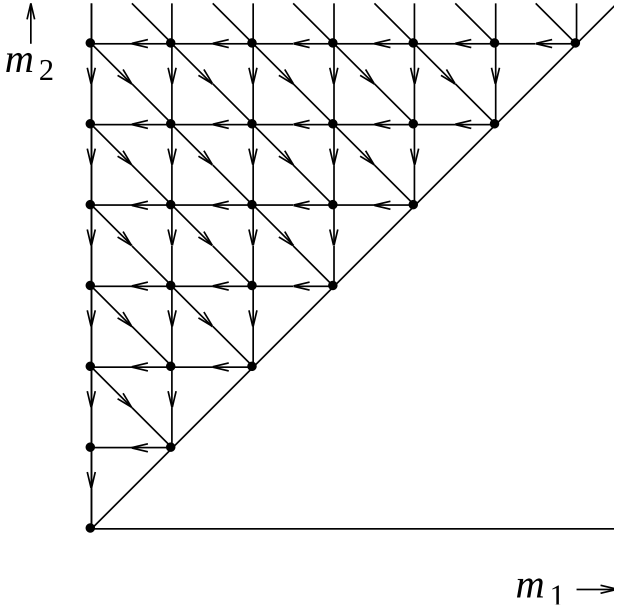
$$v_t(m_1, m_2) \leq v_t(n_1, n_2), \quad t = 0, 1, 2, \ldots$$

where $v_t(m_1, m_2)$ is the $t$-period expected cost starting in $(m_1, m_2)$


Technische Universiteit Eindhoven University of Technology

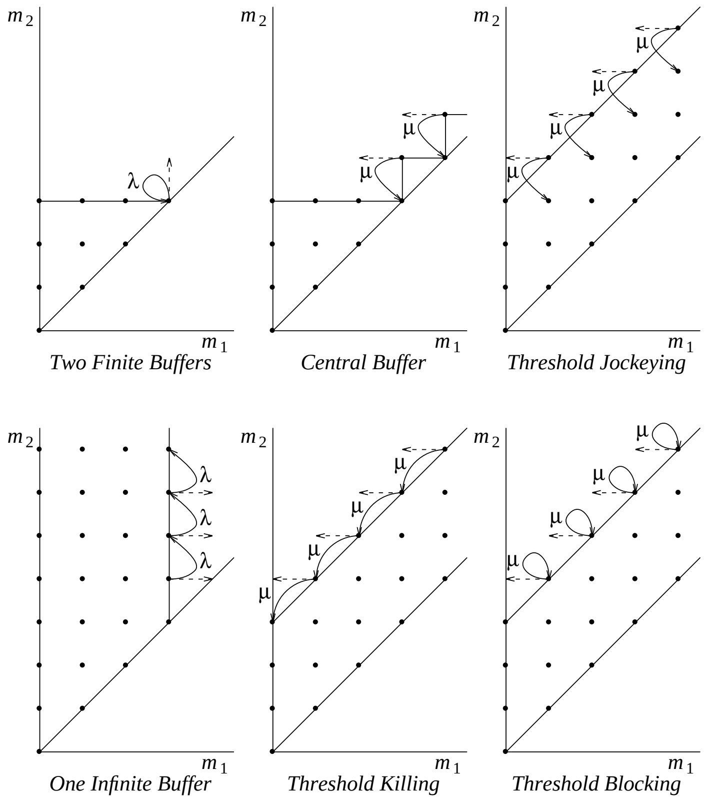
# Shortest queue: Bounds






- Redirect some transitions to more attractive states
- Then $g \geq \tilde{g}$ where $\tilde{g}$ is long-run average cost of new chain

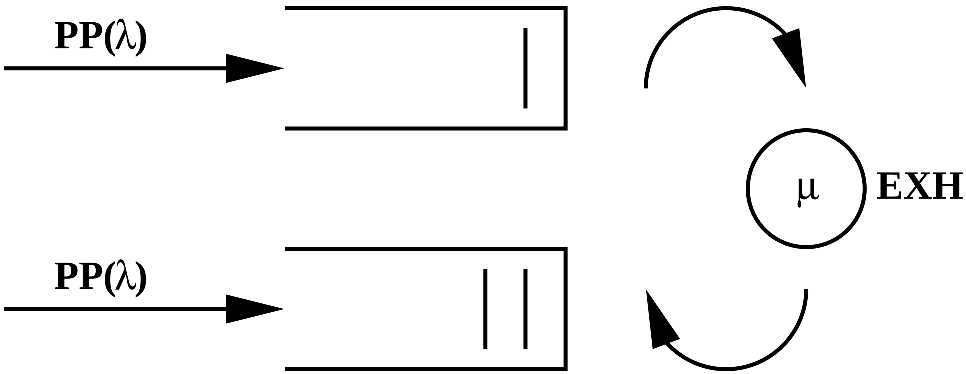
# Shortest queue: Bounds




Bounds work well, even for very large systems (say 50 queues)!

# Exhaustive polling





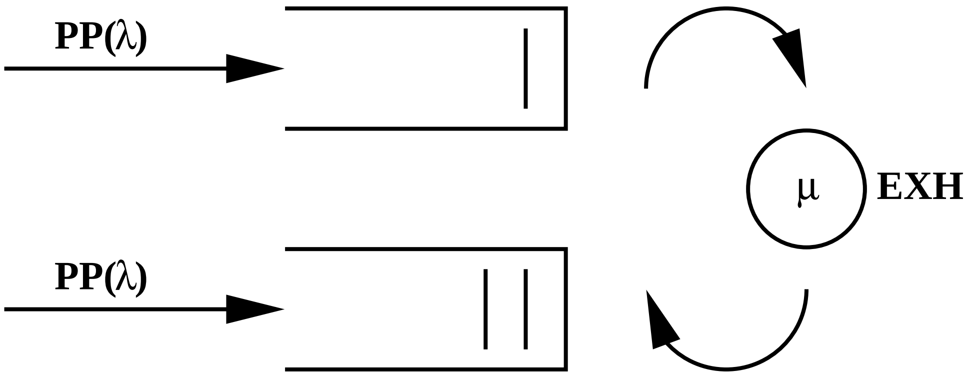

What is steady-state queue length distribution?

What is the waiting time distribution?

What is the mean waiting time?

# Exhaustive polling





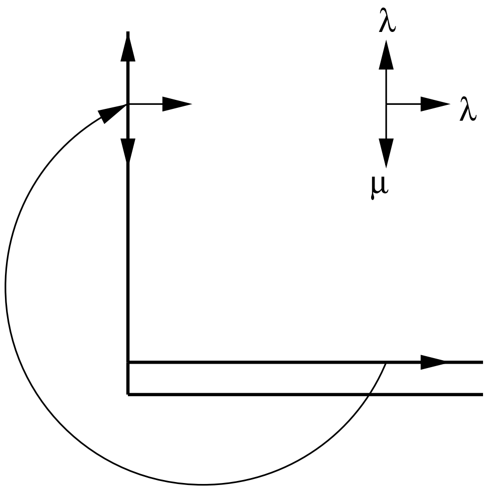

What is steady-state queue length distribution?

What is the waiting time distribution?

What is the mean waiting time? $E(W) = \frac{\rho}{(1-\rho)\mu}$ with $\rho = \frac{2\lambda}{\mu}$

# Exhaustive polling





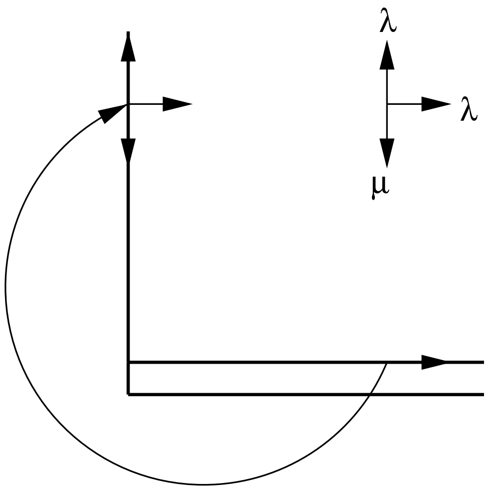

States are grid points $(m, n)$ where

- $m$ is number in idle queue
- $n$ is number in service queue

What is steady-state queue length distribution $p_{m,n}$?

# Exhaustive polling



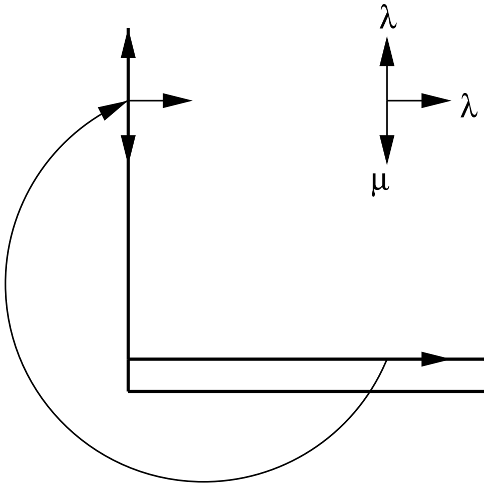



Balance equations ($2\lambda + \mu = 1$)

$$
\begin{aligned}
p_{m,n} &= p_{m,n+1}\mu + p_{m-1,n}\lambda + p_{m,n-1}\lambda \\
p_{m,1} &= p_{m,2}\mu + p_{m-1,1}\lambda \\
p_{0,n} &= p_{0,n+1}\mu + p_{n,1}\mu + p_{0,n-1}\lambda
\end{aligned}
$$

# Exhaustive polling






Basis solutions are formed by products $\alpha^m \beta^{n-1}$ satisfying

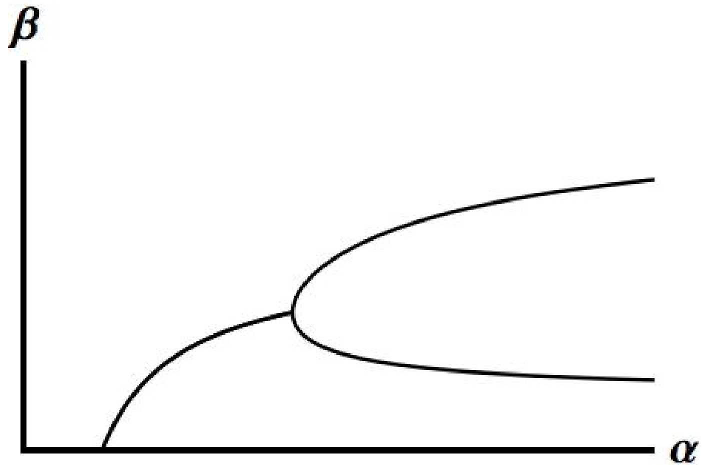
$$p_{m,n} = p_{m,n+1}\mu + p_{m-1,n}\lambda + p_{m,n-1}\lambda$$

so $\alpha$ and $\beta$ are on the curve

$$\alpha\beta = \alpha\beta^2\mu + \beta\lambda + \alpha\lambda$$

# Exhaustive polling






The curve of basis solutions

$$\alpha\beta = \beta\lambda + \alpha\beta^2\mu + \alpha\lambda$$

Iteration? Initial term? Cannot find one!

# Exhaustive polling



## Different iteration:

Let

$$P(z) = \sum_{m=0}^{\infty} p_{m,1} z^m$$

then

$$P(z) = P(h(z)) - p_{0,1}(1 - h(z))$$

where $h(z) = \eta(\lambda(1 - z))$ and $\eta(\cdot)$ is LST of BP of $M(\lambda)/M(\mu)/1$

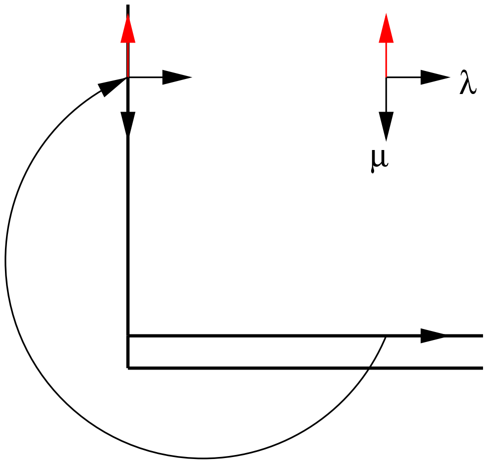
Iterating yields

$$P(z) = \sum_{m=0}^{\infty} p_{m,1} - p_{0,1} \sum_{k=1}^{\infty} (1 - h^{(k)}(z))$$

where $h^{(k)}(z) = h(h^{(k-1)}(z)), \quad h^{(0)}(z) = z$

# Smart gated polling

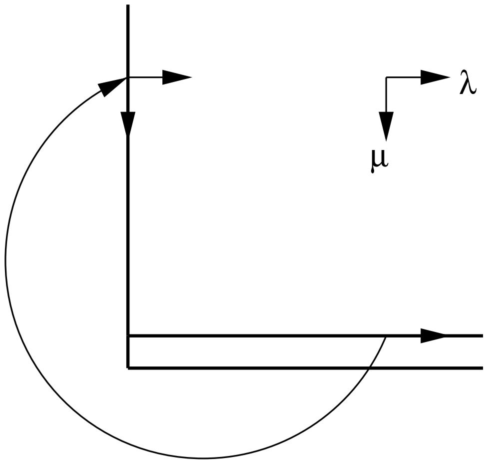





- Gated service
- PP(λ) joins idle queue

# Smart gated polling





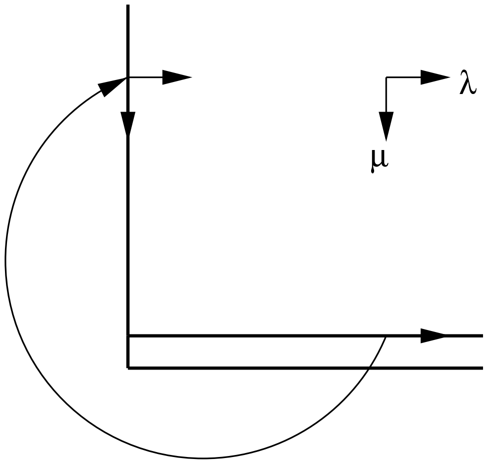

States are grid points $(m, n)$ where

- $m$ is number in idle queue
- $n$ is number in service queue

What is steady-state queue length distribution $p_{m,n}$?

# Smart gated polling

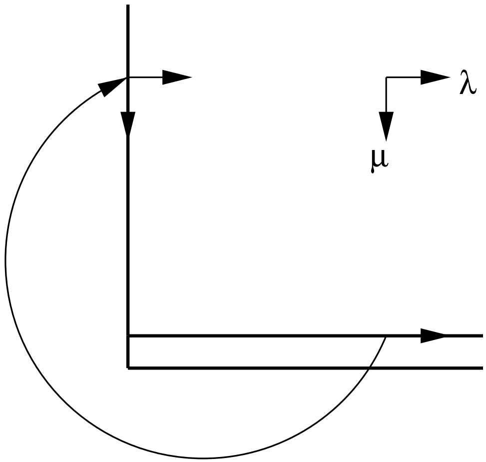





Balance equations ($\lambda + \mu = 1$)

$$p_{m,n} = p_{m-1,n}\lambda + p_{m,n+1}\mu$$
$$p_{0,n} = p_{n,1}\mu + p_{0,n+1}\mu$$

# Smart gated polling




Basis solutions are products $\alpha^m \beta^{n-1}$ satisfying

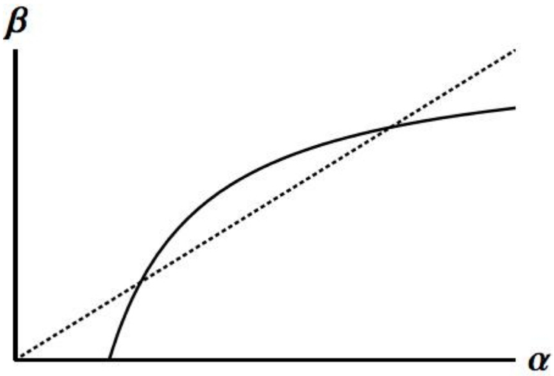
$$p_{m,n} = p_{m-1,n}\lambda + p_{m,n+1}\mu$$

so $\alpha$ and $\beta$ are on the curve

$$\alpha = \lambda + \alpha\beta\mu$$

# Smart gated polling



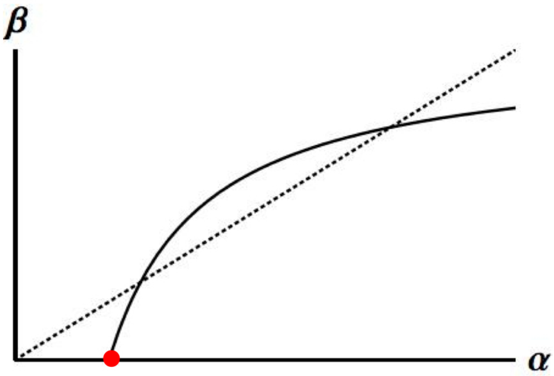



The curve of basis solutions

$$\alpha = \lambda + \alpha\beta\mu$$

# Smart gated polling






Initial product satisfies balance equations for $m > 0$:

$$p_{m,n} \approx c_0 \alpha_0^m \beta_0^{n-1}$$

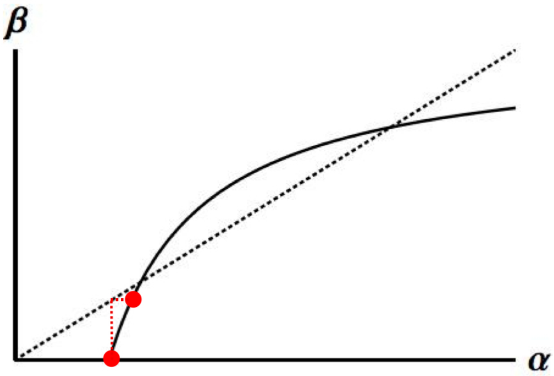
with $\alpha_0 = \lambda$, $\beta_0 = 0$



TU/e Technische Universiteit Eindhoven University of Technology

# Smart gated polling






Initial product $c_0 \alpha_0^m \beta_0^{n-1}$ violates balance equations for $m = 0$

Add new product to compensate for this error:

$$p_{m,n} \approx c_0 \alpha_0^m \beta_0^{n-1} + c_1 \alpha_1^m \beta_1^{n-1}$$

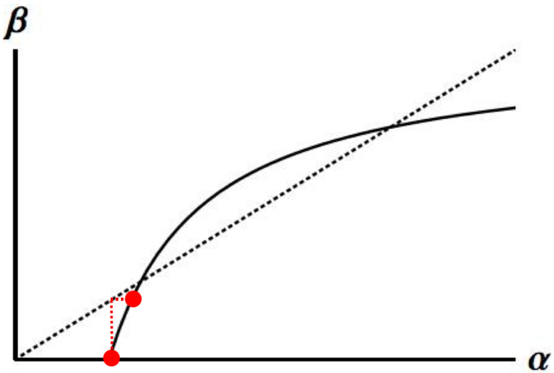
where $\beta_1 = \alpha_0$



Technische Universiteit Eindhoven University of Technology

# Smart gated polling






Steady-state queue length distribution is given by:

$$p_{m,n} = c_0\alpha_0^m\beta_0^{n-1} + c_1\alpha_1^m\beta_1^{n-1} + c_2\alpha_2^m\beta_2^{n-1} + \cdots$$

where $\alpha_i \to \lambda/\mu$ and $\beta_i \to \lambda/\mu$ as $i \to \infty$

and this random walk in the quarter plane has transitions to the East

# Smart gated polling



## Different iteration:

Let

$$P(z) = \sum_{m=0}^{\infty} p_{m,1} z^n$$

then

$$P(z) = P(h(z)) - p_{0,1}(1 - h(z))$$

where

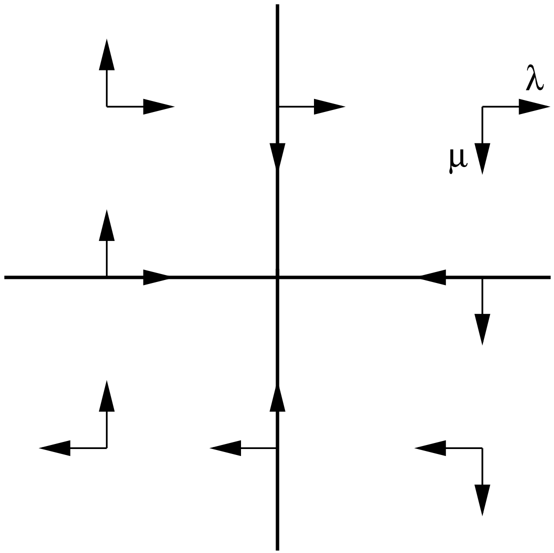
$$h(z) = \frac{\mu}{\mu + \lambda(1 - z)}$$

Iterating yields

$$P(z) = \sum_{m=0}^{\infty} p_{m,1} - p_{0,1} \sum_{k=1}^{\infty} (1 - h^{(k)}(z))$$

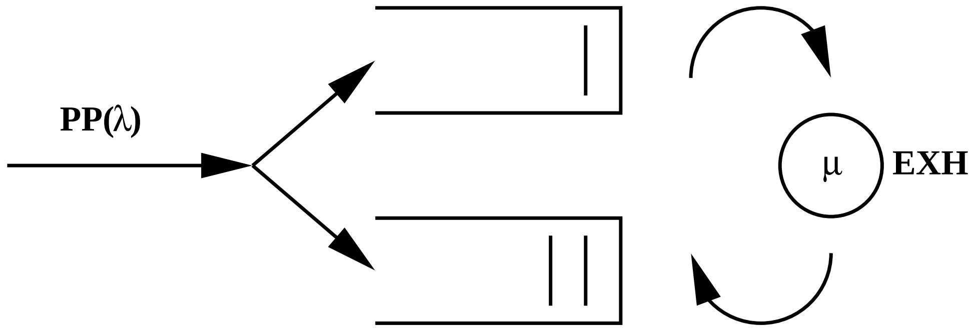
# Smart gated polling: Another look






which is a particle randomly circulating in the plane

# Shortest queue and polling



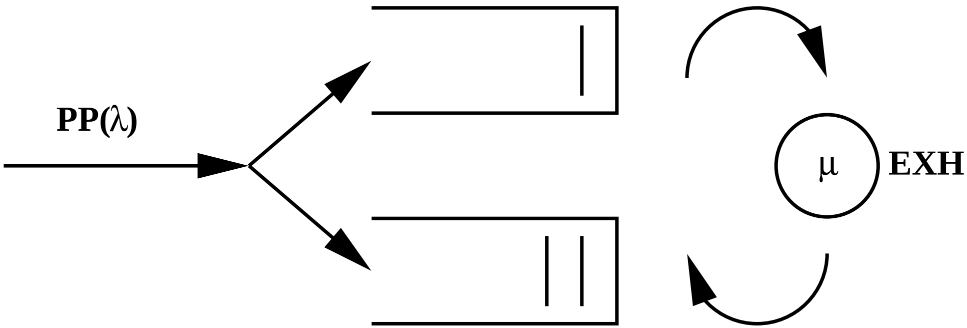

What is steady-state queue length distribution?

What is the waiting time distribution?

What is the mean waiting time?

# Shortest queue and polling





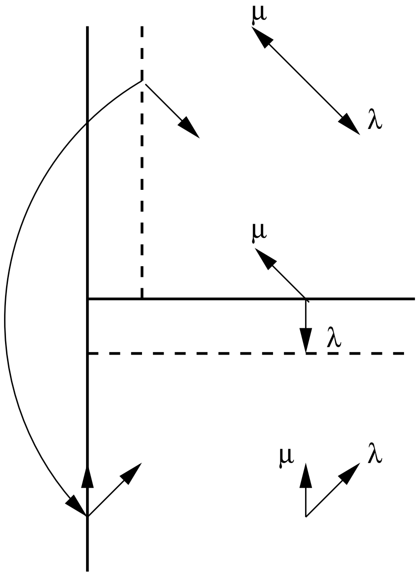

What is steady-state queue length distribution?

What is the waiting time distribution?

What is the mean waiting time? $E(W) = \frac{\rho}{(1-\rho)\mu}$ with $\rho = \frac{\lambda}{\mu}$

# Shortest queue and polling



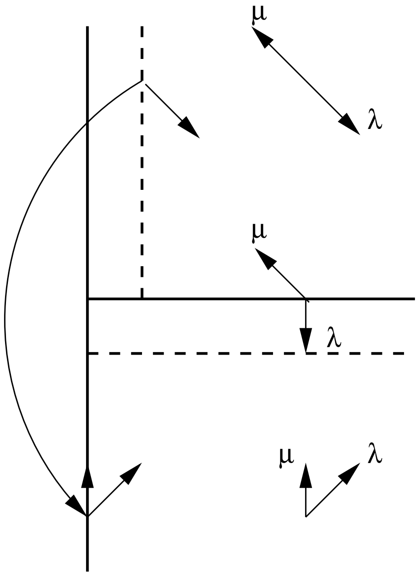



States are grid points $(m, n)$ in the right half plane where

- $m$ is length of shortest queue
- $n$ is difference between idle and service queue

What is steady-state queue length distribution $p_{m,n}$?

# Shortest queue and polling

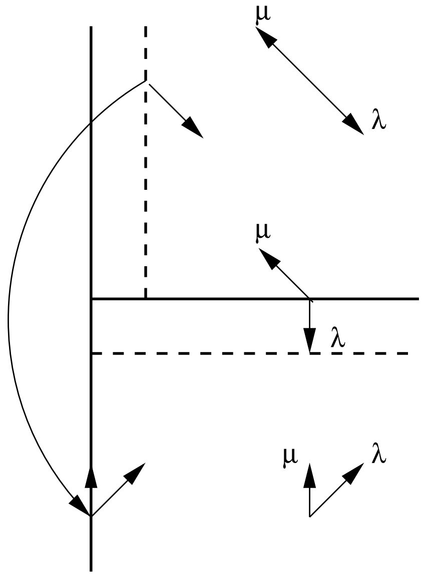



Balance equations ($\lambda + \mu = 1$)

$$
\begin{aligned}
p_{m,n} &= p_{m-1,n+1}\lambda + p_{m+1,n-1}\mu \\
p_{m,-n} &= p_{m-1,-n-1}\lambda + p_{m,-n-1}\mu
\end{aligned}
$$

# Shortest queue and polling






Basis solutions for $n \geq 0$ are products $\alpha^m \beta^n$ satisfying

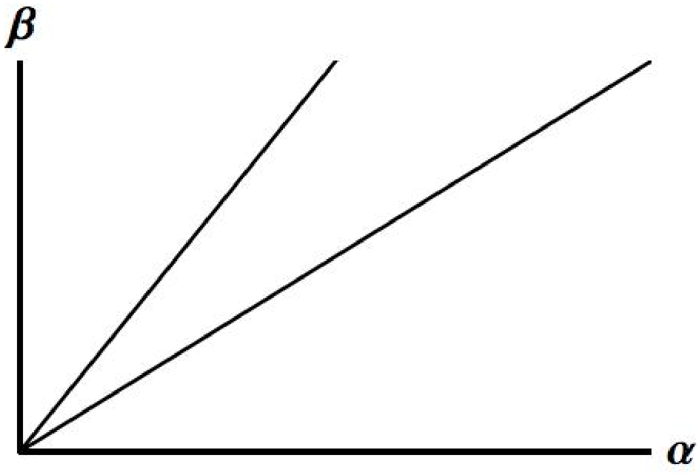
$$p_{m,n} = p_{m-1,n+1}\lambda + p_{m+1,n-1}\mu$$

so $\alpha$ and $\beta$ are on the curve

$$\alpha\beta = \beta^2\lambda + \alpha^2\mu \quad \text{or} \quad (\alpha - \beta)(\alpha\mu - \beta\lambda) = 0$$

# Shortest queue and polling





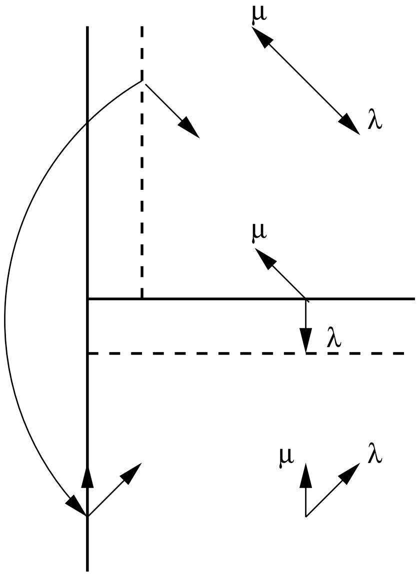

The curves of basis solutions for $n \geq 0$

$$\alpha = \beta, \qquad \mu\alpha = \lambda\beta$$

Curve of basis solutions for $n < -1$

$$\alpha = \frac{\lambda\beta}{1 - \mu\beta}$$

# Shortest queue and polling



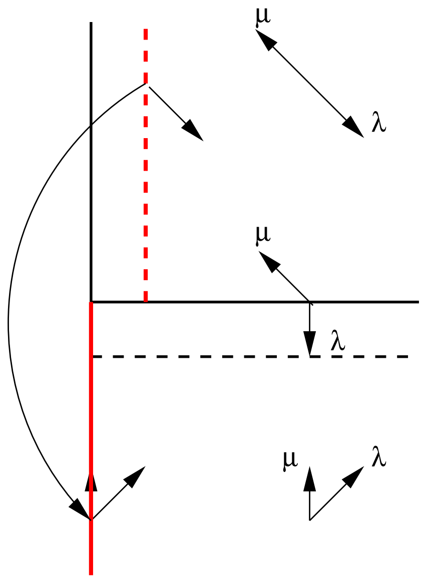

Initial product:

$$p_{m,n} = \begin{cases} c_0 \alpha_0^m \beta_0^n, & m > 0, n \geq 0, \\ d_0 \alpha_0^m, & m > 0, n = -1, \\ 0, & m \geq 0, n < -1, \end{cases}$$

where $\alpha_0 = \beta_0 = \rho^2$

# Shortest queue and polling





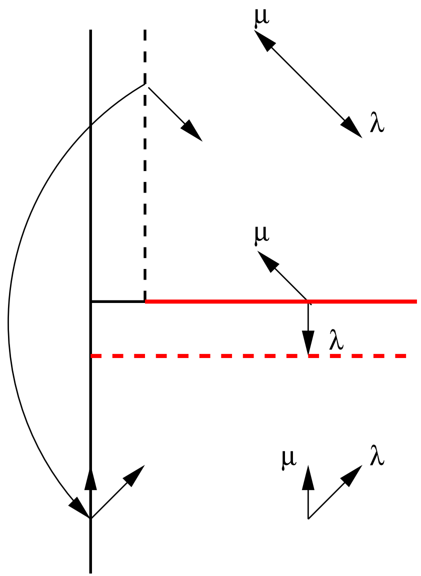

Compensation on vertical boundary of $c_0\alpha_0^m\beta_0^n$, $m > 0$, $n \geq 0$:

$$p_{m,n} = \begin{cases} c_0\alpha_0^m\beta_0^n + c_1\alpha_1^m\beta_1^n, & m > 0, n \geq 0, \\ c_2\alpha_2^m\beta_2^{-n}, & m \geq 0, n < -1, \end{cases}$$

where $\beta_2 = \beta_1 = \beta_0$ (so two new terms!)



TU/e Technische Universiteit Eindhoven University of Technology




Compensation on horizontal boundary of $c_1\alpha_1^m\beta_1^n$, $m > 0$, $n \geq 0$:

$$p_{m,n} = \begin{cases} c_1\alpha_1^m\beta_1^n + c_3\alpha_3^m\beta_3^n, & m > 0, n \geq 0, \\ d_3\alpha_3^m, & m > 0, n = -1, \\ 0, & m \geq 0, n < -1, \end{cases}$$

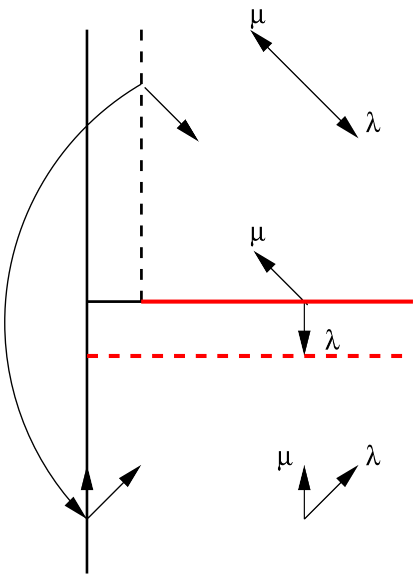
where $\alpha_3 = \alpha_1$




Technische Universiteit Eindhoven University of Technology

# Shortest queue and polling






Compensation on horizontal boundary of $c_2\alpha_2^m\beta_2^{-n}$, $m \geq 0$, $n < -1$:

$$p_{m,n} = \begin{cases} c_4\alpha_4^m\beta_4^n, & m > 0, n \geq 0, \\ d_4\alpha_4^m, & m > 0, n = -1, \\ c_2\alpha_2^m\beta_2^{-n}, & m \geq 0, n < -1, \end{cases}$$

where $\alpha_4 = \alpha_2$

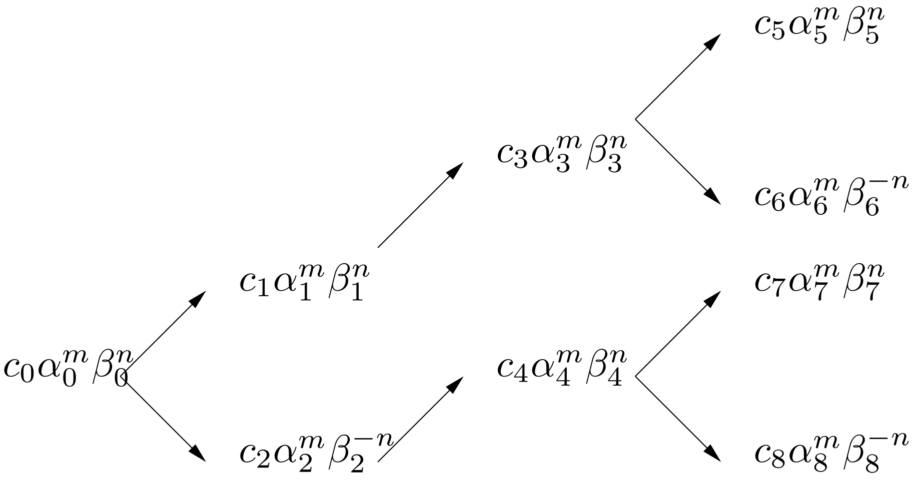


TU/e Technische Universiteit Eindhoven University of Technology

# Shortest queue and polling

Results in tree of terms:

- $c_i \alpha_i^m \beta_i^n$ live in Upper quadrant
- $c_i \alpha_i^m \beta_i^{-n}$ live in Lower quadrant



# Shortest queue and polling



$$p_{m,n} = \sum_{i \in U} c_i \alpha_i^m \beta_i^n, \qquad m \geq 1, n \geq 0$$

$$p_{m,-1} = \sum_{i \in U \cup L} d_i \alpha_i^m, \quad m \geq 1, n = -1$$

$$p_{m,n} = \sum_{i \in L} c_i \alpha_i^m \beta_i^{-n}, \quad m \geq 0, n < -1$$

Observations:

- $\alpha_i, \beta_i \downarrow 0$ as $i \to \infty$
- series converge absolutely, faster further away from origin
- distribution of shortest queue and waiting time (in service queue)

$$p_m = \sum_{i \in U \cup L} D_i (1 - \alpha_i) \alpha_i^m, \quad P(W > t) = \sum_{i \in U \cup L} D_i \alpha_i e^{-\mu(1-\alpha_i)t}$$



Technische Universiteit Eindhoven University of Technology