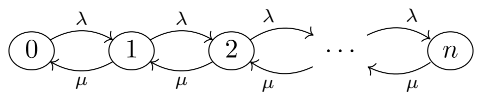
Proof for stationary distribution. The transition diagram is as follows.



Let $\{p_i\}_{i=0}^n$ be the stationary distribution. Then $\sum_{i=0}^n p_i = 1$ and

$$\begin{aligned}
\lambda p_0 &= \mu p_1, \\
\lambda p_0 + \mu p_2 &= \mu p_1 + \lambda p_1 \iff \mu p_2 = \lambda p_1, \\
\lambda p_1 + \mu p_3 &= \mu p_2 + \lambda p_2 \iff \mu p_3 = \lambda p_2, \\
&\vdots \\
\lambda p_{n-1} &= \mu p_n.
\end{aligned}$$

So

$$\lambda p_i = \mu p_{i+1}, \ \forall\ 0 \le i \le n.$$

Thus,

$$p_i \rho = p_{i+1}, \ \forall\ 0 \le i < n$$

Then

$$p_i = p_{i-1}\rho = p_{i-2}\rho^2 = p_0\rho^i.$$

So

$$\sum_{i=0}^n p_0 \rho^i = 1.$$

Then

$$p_0 = \frac{1}{\sum_{i=0}^n \rho^i},$$

and thus

$$p_i = p_0\rho^i = \frac{\rho^i}{\sum_{i=0}^n \rho^i} = \frac{\rho^i(1-\rho)}{1-\rho^{n+1}}.$$

Remark. When the capacity of the service station is infinity, i.e., when $n \to \infty$, and when $0 < \rho < 1$, i.e., $\lambda < \mu$, then

$$p_i \to \rho^i(1-\rho).$$

Proof of expectation q. Assume I is a random variable with distribution $P(I = i) = p_i$, $\forall\ 0 \le i \le n$. Then by definiton, the probability generating function

$$g(z) = E[z^I] = \sum_{i=0}^{\infty} P(I=i) z^i = \sum_{i=0}^n p_i z^i = \frac{1-\rho}{1-\rho^{n+1}} \sum_{i=0}^n (\rho z)^i = \frac{1-\rho}{1-\rho^{n+1}} \frac{1-(\rho z)^{n+1}}{1-\rho z}.$$

So

$$
\begin{aligned}
E[I] = \left.\frac{dg(z)}{dz}\right|_{z=1} &= \left.\frac{1-\rho}{1-\rho^{n+1}}\frac{-(n+1)\rho(\rho z)^n(1-\rho z)-(1-(\rho z)^{n+1})(-\rho)}{(1-\rho z)^2}\right|_{z=1} \\
&= \frac{1-\rho}{1-\rho^{n+1}}\frac{-(n+1)\rho^{n+1}(1-\rho)+\rho(1-\rho^{n+1})}{(1-\rho)^2} \\
&= \frac{\rho[-(n+1)\rho^n(1-\rho)+1-\rho^{n+1}]}{(1-\rho)(1-\rho^{n+1})} \\
&= \frac{\rho[-(n+1)\rho^n+(n+1)\rho^{n+1}+1-\rho^{n+1}]}{(1-\rho)(1-\rho^{n+1})} \\
&= \frac{\rho[1-(n+1)\rho^n+n\rho^{n+1}]}{(1-\rho)(1-\rho^{n+1})}.
\end{aligned}
$$

Remark. When the capacity of the service station is infinity and $\lambda < \mu$, then the utilization b (also the probability that the server is busy or the proportion of time the server is busy) is calculated by

$$
b = \frac{\text{mean time for service}}{\text{mean time between arrivals}} = \frac{\frac{1}{\mu}}{\frac{1}{\lambda}} = \frac{\lambda}{\mu} = \rho.
$$

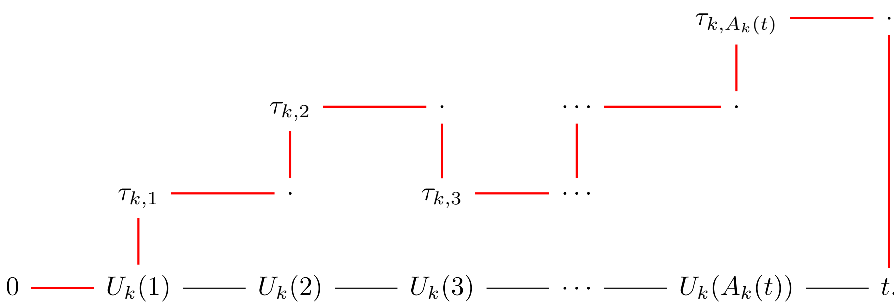
This is consistent with

$$
\lim_{n\to\infty} b = \lim_{n\to\infty} \frac{\rho(1-\rho^n)}{(1-\rho^{n+1})} = \rho.
$$

So when $n \to \infty$, and $\lambda \geq \mu$, then $b = 100\%$.

The nondiverted customers $\lambda - \zeta = \lambda(1-p_n) = \mu b$. Since $b \leq 1$, we have the nondiverted customers $\lambda(1-p_n) \leq \mu$. (Even if λ is not necessarily less than μ).

Van's paper



Between $U_k(1)$ and $U_k(2)$, $\tau_k(t)$ is nonincreasing. Define

$$\tau_k(s) = \sum_{i=1}^{A_k(s)-1} \tau_{k,i} \mathbb{1}_{[U_k(i), U_k(i+1))}(s) + \tau_{k,A_k(s)} \mathbb{1}_{[U_k(A_k(s)),s]}(s), \ \forall\ 0 \le s \le t.$$

Then

$$C_k(\tau_k(s)) = \sum_{i=1}^{A_k(s)-1} C_k(\tau_{k,i}) \mathbb{1}_{[U_k(i), U_k(i+1))}(s) + C_k(\tau_{k,A_k(s)}) \mathbb{1}_{[U_k(A_k(s)),s]}(s), \ \forall\ 0 \le s \le t.$$

So

$$\begin{aligned}
J(t) &= \sum_{k=1}^{d} \underbrace{\int_0^t C_k(\tau_k(s)) dA_k(s)}_{\text{Stochastic integral}} \\
&= \sum_{k=1}^{d} \left[\sum_{i=1}^{A_k(t)-1} C_k(\tau_{k,i}) \big(A_k(U_k(i+1)) - A_k(U_k(i))\big) + C_k(\tau_{k,A_k(t)}) \big(A_k(t) - A_k(U_k(A_k(t)))\big) \right] \\
&= \sum_{k=1}^{d} \left[\sum_{i=1}^{A_k(t)-1} C_k(\tau_{k,i}) \big((i+1) - i\big) + C_k(\tau_{k,A_k(t)})(A_k(t)) \right] \\
&= \sum_{k=1}^{d} \sum_{i=1}^{A_k(t)-1} C_k(\tau_{k,i}).
\end{aligned}$$

Theorem 0.1 (FSSLN)**.**

$$A^n(nt) = A^n(t) + \underbrace{A^n(2t) - A^n(t)}_{2^{ed}} + \cdots + \underbrace{A^n(nt) - A^n((n-1)t)}_{n^{th}}.$$

Fix time t, then by SLLN,

$$\overline{A}^n(t) := \frac{A^n(nt)}{n} \xrightarrow{\textit{converges almost surely}} E[A^n(t)] = \underbrace{\lambda t =: \overline{A}^*(t)}_{\textit{fluid limit}} \ \textit{as } n \to \infty.$$

Theorem 0.2 (FCLT)**.** *For large t, we have*

$$A^n(t) \overset{d}{\approx} N\left(\lambda t, \lambda C_a^2 t\right),$$

where

$$C_a = \frac{\textit{standard deviation of interarrival time}}{\textit{mean of interarrivlal time}}.$$

So for large t,

$$A^n(nt) \overset{d}{\approx} N\left(\lambda nt, \lambda C_a^2 nt\right).$$

When $n \to \infty$,

$$\underbrace{\tilde{A}^n(t) := \frac{A^n(nt) - n\lambda t}{n^{1/2}}}_{\textit{(diffusion scaling)}} \xrightarrow{\textit{converges in distribution}} N(0, \lambda C_a^2 t).$$

Then the stochastic process $\{\tilde{A}^n(t), t \geq 0\}$ converges to a Brownian motion with drift 0 and variance term $\lambda C_a^2 t$, when $n \to \infty$, i.e., it converges to $\lambda C_a^2 t B(t)$, where $B(t)$ is a standard brownian motion.

$$
\begin{aligned}
P(x_3, x_2, x_1) &= P(x_3, x_2 \mid x_1)P(x_1) \\
&= P(x_3 \mid x_2, x_1)P(x_2 \mid x_1)P(x_1) \\
&= P(x_3 \mid x_2, x_1)P(x_2, x_1).
\end{aligned}
$$

By induction,

$$
\begin{aligned}
P(x_n, \ldots, x_1) &= P(x_1) \cdot \prod_{t=2}^{n} P(x_t \mid x_{t-1}, \ldots, x_1) \\
&= P(x_1) \prod_{t=2}^{n} P(x_t \mid \tilde{x}_{t-1}).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
P(x_n, \ldots, x_1 \mid \theta) &= P(x_1 \mid \theta) \cdot \prod_{t=2}^{n} P(x_t \mid x_{t-1}, \ldots, x_1, \theta) \\
&= P(x_1 \mid \theta) \prod_{t=2}^{n} P(x_t \mid \tilde{x}_{t-1}, \theta).
\end{aligned}
$$

We have

$$
\begin{aligned}
P\left\{ \begin{array}{c} \alpha_{T_i}, y_{T_i}, x_{T_i} \\ \vdots \\ \alpha_{i2}, y_{i2}, x_{i2} \\ \alpha_{i1}, y_{i1}, x_{i1} \end{array} \middle| \, \theta \right\} &= P\left\{ \begin{array}{c} \alpha_{T_i}, y_{T_i}, x_{T_i} \\ \vdots \\ \alpha_{i3}, y_{i3}, x_{i3} \\ \alpha_{i2}, y_{i2}, x_{i2} \end{array} \middle| \, \alpha_{i1}, y_{i1}, x_{i1}, \theta \right\} P\{\alpha_{i1}, y_{i1}, x_{i1} \mid \theta\} \\
&= P\{\alpha_{i1}, y_{i1}, x_{i1} \mid \theta\} \prod_{t=2}^{T_i} P\{\alpha_{it}, y_{it}, x_{it} \mid \tilde{\alpha}_{i,t-1}, \tilde{y}_{i,t-1}, \tilde{x}_{i,t-1}, \theta\} \\
&= P(y_{i1} \mid \alpha_{i1}, x_{i1}, \theta) P(\alpha_{i1} \mid x_{i1}, \theta) P(x_{i1} \mid \theta) \\
&\quad \prod_{t=2}^{T_i} f_Y(y_{it} \mid a_{it}, x_{it}, \theta) \underbrace{P(a_{it} \mid x_{it}, \theta)}_{\text{sufficient statistic}} f_X(x_{it} \mid \underbrace{a_{i,t-1}, x_{i,t-1}, \theta)}_{\text{Markovian}} \\
&= P(x_{i1} \mid \theta) \prod_{t=1}^{T_i} f_Y(y_{it} \mid a_{it}, x_{it}, \theta) P(a_{it} \mid x_{it}, \theta) \prod_{i=2}^{T_i} f_X(x_{it} \mid a_{i,t-1}, x_{i,t-1}, \theta) \\
&= P(x_{i1} \mid \theta) \prod_{t=1}^{T_i} f_Y(y_{it} \mid a_{it}, x_{it}, \theta_Y) P(a_{it} \mid x_{it}, \theta) \prod_{i=2}^{T_i} f_X(x_{it} \mid a_{i,t-1}, x_{i,t-1}, \theta_f)
\end{aligned}
$$

Then

$$
\begin{aligned}
l_i(\theta) &= \log P \left\{ \begin{array}{c|c} \alpha_{T_i}, y_{T_i}, x_{T_i} & \\ \vdots & \\ \alpha_{i2}, y_{i2}, x_{i2} & \theta \\ \alpha_{i1}, y_{i1}, x_{i1} & \end{array} \right\} \\
&= \log \left(P(x_{i1} \mid \theta) \prod_{t=1}^{T_i} f_Y(y_{it} \mid a_{it}, x_{it}, \theta_Y) P(a_{it} \mid x_{it}, \theta) \prod_{i=2}^{T_i} f_X(x_{it} \mid a_{i,t-1}, x_{i,t-1}, \theta_f) \right) \\
&= \log P(x_{i1} \mid \theta) + \sum_{t=1}^{T_i} \log f_Y(y_{it} \mid a_{it}, x_{it}, \theta_Y) + \sum_{t=1}^{T_i} P(a_{it} \mid x_{it}, \theta) + \sum_{i=2}^{T_i} f_X(x_{it} \mid a_{i,t-1}, x_{i,t-1}, \theta_f) \\
&= \log P(x_{i1} \mid \theta) + \sum_{t=1}^{T_i} \log f_Y(y_{it} \mid a_{it}, x_{it}, \theta_Y) + \sum_{t=1}^{T_i} P(a_{it} \mid x_{it}, \theta) + \sum_{t=1}^{T_i-1} f_X(x_{i,t+1} \mid a_{it}, x_{it}, \theta_f).
\end{aligned}
$$