

1 Stability of queueing models via fluid limits

1.1 Background

Let $f_n : \mathbf{R} \rightarrow \mathbf{R}$, $n = 1, 2, \dots$, and $f : \mathbf{R} \rightarrow \mathbf{R}$.

Definition 1 $f_n \xrightarrow{\text{u.o.c.}} f$ as $n \rightarrow \infty$, if, for any compact (bounded and closed) set K ,

$$\sup_{x \in K} |f_n(x) - f(x)| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

For any f , for any $m = 1, 2, \dots$, define a norm $\|f\|_m$:

$$\|f\|_m = \sup_{|x| \leq m} |f(x)|,$$

and a norm $\|f\|$:

$$\|f\| = \sum_{m=1}^{\infty} 2^{-m} \cdot \frac{\|f\|_m}{1 + \|f\|_m}.$$

Lemma 1 Then $f_n \xrightarrow{\text{u.o.c.}} f \iff \|f_n - f\| \rightarrow 0$ as $n \rightarrow \infty$.

Indeed,

($\Rightarrow$). $\forall \varepsilon > 0$, choose m_0 :

$$\sum_{m=m_0+1}^{\infty} 2^{-m} \equiv 2^{-m_0} \leq \frac{\varepsilon}{2}$$

Then, take $k_1 = [-1, 1]$, $k_2 = [-2, 2]$, $\dots$, $k_{m_0} = [-m_0, m_0]$ and choose n_0 :

$$\sup_{x \in k_{m_0}} |f_n(x) - f(x)| \leq \frac{\varepsilon}{2m_0} \quad \forall n \geq n_0.$$

$$\implies \|f_n - f\| \leq \sum_{m=1}^{m_0} \|f_n - f\|_m + 2^{-m_0} \leq m_0 \cdot \frac{\varepsilon}{2m_0} + \frac{\varepsilon}{2} = \varepsilon.$$

($\Leftarrow$). Take $m_0 \gg 1$: $K \subseteq [-m_0, m_0]$. Set $g_n = f_n - f$.

$$\|f_n - f\| \rightarrow 0 \implies 2^{-m_0} \cdot \frac{\|g_n\|_{m_0}}{1 + \|g_n\|_{m_0}} \rightarrow 0 \implies \|g_n\|_{m_0} \rightarrow 0 \implies \|g_n\|_{"K"} \rightarrow 0.$$

□

Remark 1 When $\|f\|$ is finite? Some examples:

- (a) f is continuous;
- (b) f is non-decreasing (non-increasing) and $|f(x)| < \infty \quad \forall x$;
- (c)

Remark 2 If f is defined on some measurable subset $B \subset \mathbf{R}$ only, we put $f(x) = 0 \quad \forall x \notin B$. Or, if B is closed, ...

Remark 3 Any continuous function is determined by its values at all rational points (the same — for right- (left-) continuous).

Lemma 2 Let all f_n be non-decreasing functions and f be continuous. Then the following are equivalent:

- (a) $f_n \xrightarrow{\text{u.o.c.}} f$;
- (b) $f_n(x) \rightarrow f(x) \quad \forall \text{ rational point } x$.

Proof.

($\Leftarrow$ Only!). Take any m and prove that $\|f_n - f\|_m \rightarrow 0$. Assume

$$\limsup_{n \rightarrow \infty} \|f_n - f\|_m = C > 0 \quad (\text{Note: } C < \infty!)$$

$\Rightarrow \exists$ a subsequence $\{n_l\}$: $\|f_{n_l} - f\|_m \rightarrow C$,

$\Rightarrow \exists$ a sequence $\{x_l\}$: $|f_{n_l}(x_l) - f(x_l)| \geq C/2 \quad \forall l, x_l \in [-m, m]$

($\{x_l\}$ is bounded) $\Rightarrow \exists$ a subsequence $x'_l \rightarrow x \in [-m, m]$, take if for $\{x_l\}$ (skip primes).

Choose $\delta > 0$:

- (i) δ is rational;
- (ii) $|f(x + \delta) - f(x)| \leq C/4$.

Then

$$f_{n_l}(x_l) - f(x_l) = f_{n_l}(x_l) - f_{n_l}(x + \delta) + f_{n_l}(x + \delta) - f(x + \delta) + f(x + \delta) - f(x) + f(x) - f(x_l)$$

and

$$\limsup_{l \rightarrow \infty} (f_{n_l}(x_l) - f(x_l)) \leq 0 + 0 + C/4 + 0.$$

Similarly, choose $\tilde{\delta} > 0$:

- (i) $\tilde{\delta}$ is rational;
- (ii) $f(x - \tilde{\delta}) - f(x) \leq -C/4$.

Then

$$\liminf_{l \rightarrow \infty} (f_{n_l}(x_l) - f(x_l)) \geq -C/4 \text{ — a contradiction!}$$

□

Lemma 3 (*The “triangular” scheme*).

Let $D, d > 0$ and $\{f_n\} : \sup_{|x| \leq d} |f_n(x)| \leq D \forall n$.

Then $\exists$ a subsequence $\{n_r\} : \forall$ rational x ,

$$\exists \text{ a limit } \lim_{r \rightarrow \infty} f_{n_r}(x) := f(x).$$

Proof. Number all rational points in $[-d, d]$ in an arbitrary order: $x_1, x_2, \dots$. Start with the procedure:

Step 1. Since $|f_n(x_1)| \leq D \forall n$, $\exists$ a subsequence $\{n_{l,1}\} : \{f_{n_{l,1}}(x_1)\}$ converges (denote the limit by $f(x_1)$).

...

Step $r+1$. Assume we have defined a sequence $\{n_{l,r}\} : f_{n_{l,r}}(x_i) \rightarrow f(x_i) \forall i = 1, \dots, r$. Since $|f_{n_{l,r}}(x_{r+1})| \leq D \forall l$, $\implies \exists$ a subsequence $\{n_{l,r+1}\} : \{f_{n_{l,r+1}}(x_{r+1})\}$ converges (denote the limit by $f(x_{r+1})$).

...

Therefore, $\forall r = 1, 2, \dots$, we defined a sequence $\{n_{l,r}\} : f_{n_{l,r}}(x_i) \rightarrow f(x_i) \forall i = 1, \dots, r$.

Now: set $n_r = n_{r,r}$.

Note: $\forall r_0, \{n_r, r \geq r_0\}$ is a subsequence of $\{n_{l,r_0}\}$

$$\implies f_{n_r}(x_{r_0}) \rightarrow f(x_{r_0}).$$

□

Corollary 1 (I). Assume, in addition, that $\exists C < \infty: \forall x, y \in [-d, d]$,

$$\limsup_{n \rightarrow \infty} |f_n(x) - f_n(y)| \leq C|x - y|.$$

If f is any limit from Lemma 3, then

$$|f(x) - f(y)| \leq C|x - y| \quad \forall \text{ rational } x, y \in [-d, d].$$

Therefore,

- (i) $\forall t \in [-d, d]$, one can define $f(t) = \lim_{x_l \rightarrow t} f(x_l)$, where $\{x_l\}$ are rational;
- (ii) this limit does not depend on $\{x_l\}$;
- (iii) $f(t)$ is continuous in $[-d, d]$ and

$$|f(t_1) - f(t_2)| \leq C|t_1 - t_2|.$$

(II). If, in addition, all $\{f_n\}$ are non-decreasing, and conditions of (I) are satisfied $\forall d$, then

$$f_{n_r} \xrightarrow{\text{u.o.c.}} f.$$

Problem No 1. Prove it.

Definition 2 A function f is Lipschitz continuous with parameter $C < \infty$, if $\forall x, y$

$$|f(x) - f(y)| \leq C|x - y|.$$

Theorem 1 Let $f : \mathbf{R} \rightarrow \mathbf{R}$ (or $f : \mathbf{R}_+ \rightarrow \mathbf{R}$) be Lipschitz continuous. Then

- (a) $\exists$ a measurable set $B \equiv B(f) \subseteq \mathbf{R}$ (or $\subseteq \mathbf{R}_+$) : $\lambda(B) = 0$ and

$$\forall x \notin B, \quad \exists f'(x) = \lim_{\Delta \rightarrow 0} \frac{f(x + \Delta) - f(x)}{\Delta}, \quad |f'(x)| \leq C.$$

- (b) Put $f'(x) = 0$ for $x \in B$. Then $\forall x, \quad \forall t > 0$

$$f(x + t) - f(x) = \int_x^{x+t} f'(z) dz.$$

Definition 3 Any point $x \notin B$ is a regular point of f .

Corollary 2 Let $f : \mathbf{R}_+ \rightarrow \mathbf{R}$ be Lipschitz continuous, $f(0) > 0$, $f(t) \geq 0$ $\forall t \geq 0$.

Assume $\exists \varepsilon > 0 : \forall$ regular point $t \geq 0$, if $f(t) > 0$, then $f'(t) \leq -\varepsilon$.

Then $\exists t_0 \leq \frac{f(0)}{\varepsilon} : f(t) = 0 \quad \forall t \geq t_0$.

Proof.

First, if $f(t_0) = 0$, then $f(t) = 0 \quad \forall t \geq t_0$, since $f(t) - f(t_0) = \int_{t_0}^t f'(z) dz \leq 0$.

Second, if $f(t) > 0 \quad \forall t \leq \frac{f(0)}{\varepsilon}$, then

$$f\left(\frac{f(0)}{\varepsilon}\right) \leq \int_0^{f(0)/\varepsilon} (-\varepsilon) dz + f(0) \leq 0.$$

□

1.2 Use of SLLN

1). Consider a renewal process $\{\eta(t), t \geq 0\}$, a sequence $\{\xi_n\}$ — i.i.d., with $\mathbf{E}\xi_1 = a > 0$ (and $< \infty$). Put

$$S_0 = \xi_0, \quad S_n = \xi_1 + \dots + \xi_n + \xi_0.$$

Consider an addition parameter $N > 0$ and a sequence of renewal processes $\{\eta^{(N)}(t), t \geq 0\}$, where $\eta^{(N)}(t)$ is defined by

$$S_0^{(N)} \equiv s(N) = \text{const}, \quad S_n^{(N)} = s(N) + \xi_1 + \dots + \xi_n, \quad n \geq 1$$

and a sequence $\{\bar{\eta}^{(N)}(t), t \geq 0\}$, where

$$\bar{\eta}^{(N)}(t) = \frac{\eta^{(N)}(Nt)}{N}.$$

Assume $\frac{s(N)}{N} \rightarrow \bar{s} \in [0, 1)$ as $N \rightarrow \infty$.

What one can say about the limit behaviour of $\{\bar{\eta}^{(N)}(t), t \geq 0\}$, when $N \rightarrow \infty$?

Let $\{\eta(t), t \geq 0\}$ be a renewal process with $s_0 \equiv 0$.

Then, for $t > \bar{s}$,

$$\bar{\eta}^{(N)}(t) = \frac{\eta_0(Nt - s(N))}{N} \equiv \frac{\eta_0(Nt - s(N))}{Nt - s(N)} \cdot \frac{Nt - s(N)}{N} \rightarrow \frac{1}{a}(t - \bar{s}) \text{ a.s.}$$

Monotonicity $\implies$

$$\bar{\eta}^{(N)}(t) \rightarrow \frac{1}{a}(t - \bar{s})^+ \text{ a.s. } \forall t.$$

2). Now: change notations. $\eta(t) \mapsto A(t)$. Meaning: $\eta(t)$ — arrival process (Arrival). (I'll use “ η ” and “ A ” in parallel).

$$\bar{A}^{(N)}(t) \rightarrow \frac{1}{a}(t - \bar{s})^+ \equiv \bar{A}(t).$$

3). Assume there are K classes of customers, and any arriving customer becomes a “class k customer”, $k = 1, \dots, K$ with probability p_k (independently on everything else), $\sum_{k=1}^K p_k = 1$. One can introduce an i.i.d. sequence $\{\alpha(n)\} \perp \{\xi_n\}$,

$$\mathbf{P}(\alpha(n) = k) = p_k, \quad k = 1, \dots, K.$$

Then

$$A_k(t) = \sum_{n=0}^{\eta(t)-1} I(\alpha(n) = k)$$

represents a number of class k customers that arrive on $[0, t]$, and (for $t > \bar{s}$)

$$\begin{aligned} \bar{A}_k^{(N)}(t) &= \frac{A_k^{(N)}(Nt)}{N} = \frac{A_{k,0}(Nt - s(N))}{Nt - s(N)} \cdot \frac{Nt - s(N)}{N} \\ &= \frac{\sum_{n=0}^{\eta_0(Nt - s(N))} I(\alpha(n) = k)}{\eta_0(Nt - s(N))} \cdot \frac{\eta_0(Nt - s(N))}{Nt - s(N)} \cdot \frac{Nt - s(N)}{N} \rightarrow \frac{p_k}{a}(t - \bar{s}) \text{ a.s.} \end{aligned}$$

Monotonicity $\implies$

$$\bar{A}_k^{(N)}(t) \rightarrow \frac{p_k}{a}(t - \bar{s})^+ \equiv \bar{A}_k(t) \text{ a.s.}$$

4. Note!! $\forall$ fixed t , $\bar{A}_k^{(N)}(t)$ are UI !

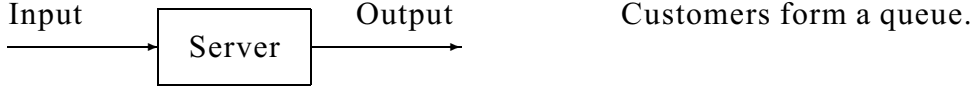
$$\implies \mathbf{E}\bar{A}_k^{(N)}(t) \rightarrow \mathbf{E}\bar{A}_k(t).$$

Definition 4 $\bar{A}_k(t)$ is a fluid limit in continuous time for a class k arrival (or input) process.

In fact, we have a family of fluid limits that depend on initial delays.

We like to introduce a concept of fluid limit for some other processes (throughput, output or “departure”) in a very general class of queueing models.

1.3 Multi-class single-server queue



Service policies

A server serves customers in some order that depends on a current state of queue only (a policy is assumed to be “non-idling” !). Possible variants: FCFS, LCFS, random, priority, ...

Initial conditions

At $t = 0$, there is given an “initial queue”

$$\vec{Q}(0) = (Q_1(0), \dots, Q_K(0)),$$

and $\chi(0) \geq 0$ — a residual service time of some “additional” customer;
 $\xi_0 \geq 0$ — a residual interarrival time.

Service times

With each $k = 1, \dots, K$, there is associated an i.i.d. sequence $\{\sigma_k(n)\}_{n \geq 1}$ of service times with mean $b_k < \infty$. In other words, n -th class k customer has a service time $\sigma_k(n)$.

Set:

$Q_k(t)$ — a queue length of class k customers at $t + 0$;

$D_k(t)$ — a number of service completions of class k customers on $[0, t]$,

$$Q_k(t) = Q_k(0) + A_k(t) - D_k(t)$$

$Y_k(t)$ — the total amount of work which arrived in class k by time t , i.e.

$$Y_k(t) = \sum_{n=1}^{Q_k(0)+A_k(t)} \sigma_k(n)$$

$T_k(t)$ — the total amount of time that server spent on class k customers by time t , i.e.

$$T_k(t) = \sum_{n=1}^{D_k(t)} \sigma_k(n) + \phi_k(t)$$

Here: $\phi_k(t)$ is the time that server spent on the $(D_k(t) + 1)$ -st service of class k customers by time t ;

$W_k(t) = Y_k(t) - T_k(t)$ — the workload vector;

$I(t) = t - \sum_{k=1}^K T_k(t)$ — the cumulative amount of time that the server has been idle by time t .

Construction of fluid limits for $D_k(t)$

Put

$$N = \sum_{k=1}^K Q_k(0) + \chi(0) + \xi_0, \quad Q_k(0) = Q_k^{(N)}(0), \chi(0) = \chi^{(N)}(0), \xi_0 = s(N).$$

Assume:

$$\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow a; \quad \frac{\sigma_k(1) + \dots + \sigma_k(n)}{n} \rightarrow b_k \quad \forall k \quad \forall \omega \in \Omega!$$

Take any $t > 0$, any $\omega \in \Omega$ and any sequence $\{N_l\}$, $N_l \rightarrow \infty$. For any $N = N_l$, $0 \leq u \leq t$, put (write $N \rightarrow \infty$ instead of “ $N_l \rightarrow \infty$ as $l \rightarrow \infty$ ”)

$$\begin{aligned} \bar{A}_k^{(N)}(u) &= \frac{A_k^{(N)}(Nu)}{N}, \quad \bar{D}_k^{(N)}(u) = \frac{D_k^{(N)}(Nu)}{N}, \\ \bar{Q}_k^{(N)}(0) &= \frac{Q_k^{(N)}(0)}{N}, \quad \bar{\chi}^{(N)}(0) = \frac{\chi^{(N)}(0)}{N}, \quad \bar{s}(N) = \frac{s(N)}{N}. \end{aligned}$$

Choose any subsequence:

$$\bar{\chi}^{(N)}(0) \rightarrow \bar{\chi}(0), \bar{s}(N) \rightarrow \bar{s}, 0 \leq \bar{\chi}(0) + \bar{s} \leq 1. \quad (1)$$

Then:

$$\bar{Q}_k^{(N)}(0) \rightarrow \bar{Q}_k(0), 0 \leq \sum_{k=1}^K \bar{Q}_k(0) \leq 1. \quad (2)$$

Lemma 4 For any subsequence satisfying (1)-(2), $\forall k = 1, \dots, K, \forall 0 \leq u \leq t$,

$$\limsup_{N \rightarrow \infty} \left[\bar{D}_k^{(N)}(t) - \bar{D}_k^{(N)}(u) \right] \leq \frac{t - u}{b_k}.$$

Proof.

Let $\chi^{(N)}(u)$ be a residual service time at u ($\chi(u) = 0$, if no service);

$\mu_k^{(N)}(u)$ — a total number of services of class k customers that were started by u ; ($\chi(u) = 0$, if no service);

For $j \geq 1$, $\tilde{\sigma}_k^{(N)}(j) = \sigma_k(\mu_k^{(N)}(Nu) + j)$.

Then

$$D_k^{(N)}(Nt) - D_k^{(N)}(Nu) \leq \tilde{\eta}_k^{(N)}(N(t - u)),$$

where

$$\tilde{\eta}_k^{(N)}(v) = 1 + \max \left\{ n \geq 1 : \sum_{j=1}^n \tilde{\sigma}_k^{(N)}(j) \leq v \right\} = \min \left\{ n \geq 1 : \sum_{j=1}^n \tilde{\sigma}_k^{(N)}(j) > v \right\}$$

— an undelayed renewal process for $\{\tilde{\sigma}_k^{(N)}(j)\}_{j \geq 1}$.

Therefore (**Problem No 2**. To prove more accurately)

$$\limsup_{N \rightarrow \infty} [\dots] \leq \limsup_{N \rightarrow \infty} \frac{\tilde{\eta}_k^{(N)}(N(t - u))}{N} \rightarrow \frac{t - u}{b_k}.$$

□

Now: $\forall$ fixed $t > 0$, $\forall$ rational point $0 \leq u \leq t$ consider a subsequence:

$$\bar{D}_k^{(N)}(u) \rightarrow \bar{D}_k(u).$$

Note: $\bar{D}_k(u)$ is Lipschitz continuous, $\bar{D}_k^{(N)}(u)$ are non-decreasing $\implies \bar{D}_k^{(N)}(u) \xrightarrow{\text{u.o.c.}} \bar{D}_k(u)$ (“uniformly on compact sets”).

Definition 5 Any function $\bar{D}_k$ is called a fluid limit for departure process. The family of fluid limits is called a fluid model.

Some useful equations

$\forall t > 0, \forall k, \forall$ fluid limit

$$\begin{aligned}
\bar{Q}_k(t) &= \bar{Q}_k(0) + \bar{A}_k(t) - \bar{D}_k(t), \\
\bar{Y}_k(t) &= b_k \cdot (\bar{Q}_k(0) + \bar{A}_k(t)), \\
\bar{T}_k(t) &= b_k \cdot \bar{D}_k(t), \\
\bar{W}_k(t) &= b_k \cdot \bar{Q}_k(t), \\
\bar{I}(t) &= t - \sum_{k=1}^K b_k \cdot \bar{D}_k(t), \\
\int_0^\infty \sum_{k=1}^K \bar{Q}_k(t) d\bar{I}(t) &= 0,
\end{aligned} \tag{*}$$

And $\forall$ regular point $t, \forall k$

$$\begin{aligned}
\dot{\bar{Q}}_k(t) &= \dot{\bar{A}}_k(t) - \dot{\bar{D}}_k(t); \dot{\bar{Y}}_k(t) = b_k \cdot \dot{\bar{A}}_k(t); \\
\dot{\bar{T}}_k(t) &= b_k \cdot \dot{\bar{D}}_k(t); \dot{\bar{W}}_k(t) = b_k \cdot \dot{\bar{Q}}_k(t); \\
1 \geq \dot{\bar{I}}(t) &= 1 - \sum_{k=1}^K b_k \cdot \dot{\bar{D}}_k(t) \geq 0; \\
\sum_{k=1}^K \bar{Q}_k(t) > 0 &\implies \dot{\bar{I}}(t) = 0
\end{aligned}$$

Lemma 5 If $\bar{Q}_k(t) = 0$ and t — regular point, then $\dot{\bar{Q}}_k(t) = 0$.

Indeed,

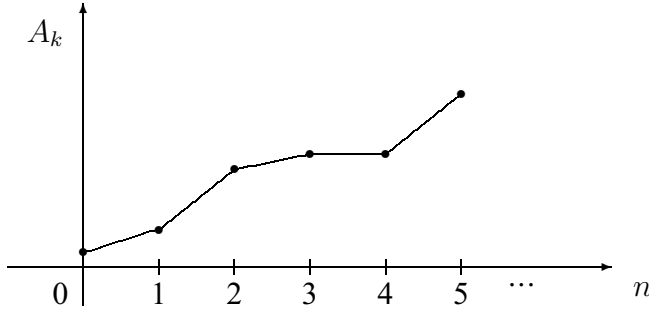
(a) $\dot{\bar{Q}}_k(t)$ can not be negative: $\dot{\bar{Q}}_k(t) = \lim_{\Delta \rightarrow 0, \Delta > 0} \frac{\bar{Q}_k(t + \Delta) - \bar{Q}_k(t)}{\Delta} \geq 0;$

(b) $\dot{\bar{Q}}_k(t)$ can not be positive: ...

□

1.4 Multi-class single-server queue: fluid limits in discrete time

Put, for $n \geq 0$, $A_{k,n} = A_k(S_n)$, $D_{k,n} = D_k(S_n)$, etc.



Note: $A_{k,n} = \sum_{j=0}^n I(\alpha(j) = k)$

Connect points
by straight lines...

Consider fluid limits. $\check{\bar{A}}_k(t) = tp_k = \bar{A}_k(ta)$ $\xi_0 \equiv 0 \implies \bar{s} = 0$ $\check{\bar{D}}_k(t) = \bar{D}_k(ta)$, etc. (Only this limits).

1.5 Multi-class queueing network

Introduce the sequence of r.v.'s $\{\alpha_k(n)\}_{n=1}^\infty$, $k = 1, \dots, K$: they are independent (and independent of everything else); $\mathbf{P}(\alpha_k(n) = l) = p_{k,l}$.

Then we have the following dynamics for the queue-length process:

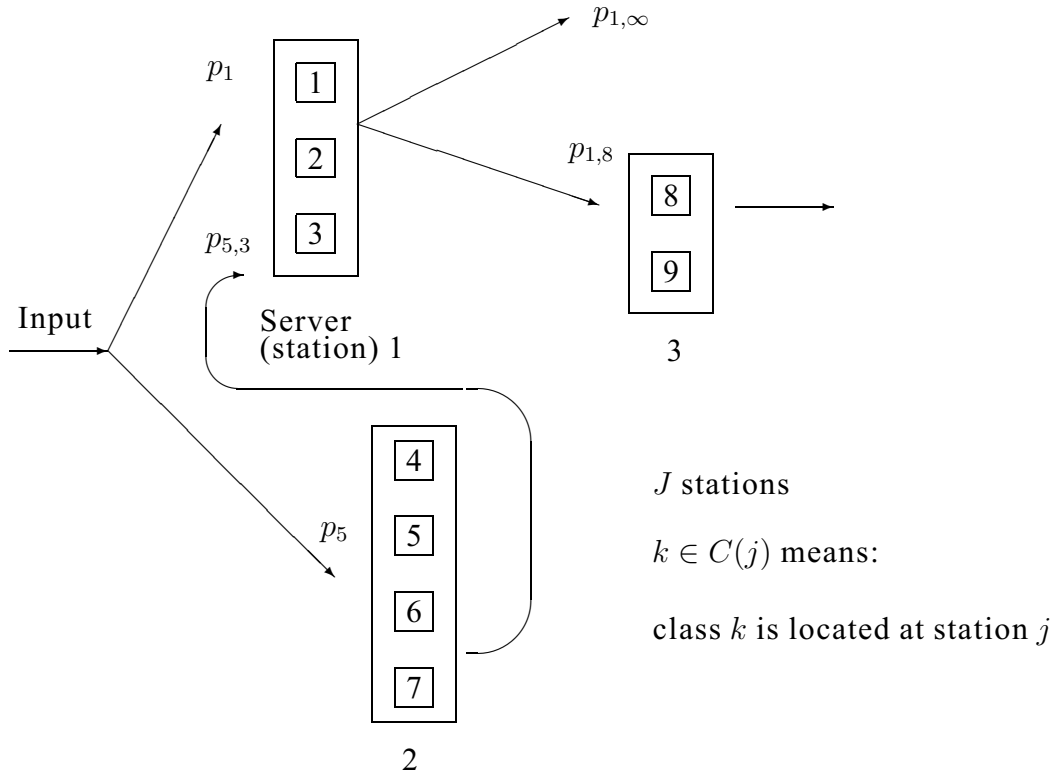
$$Q_k(t) = Q_k(0) + A_k(t) + \sum_{l=1}^K \sum_{i=1}^{D_l(t)} I(\alpha_l(i) = k) - D_k(t)$$

$$\implies \bar{Q}_k(t) = \bar{Q}_k(0) + \bar{A}_k(t) + \sum_{l=1}^K \bar{D}_l(t) p_{l,k} - \bar{D}_k(t).$$

Indeed,

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{D_l^{(N)}(t)} I(\alpha_l(i) = k) &= \frac{D_l^{(N)}(t)}{N} \cdot \frac{1}{D_l^{(N)}(t)} \sum_{i=1}^{D_l^{(N)}(t)} I(\alpha_l(i) = k) \\ &\quad \downarrow \quad \quad \downarrow \\ &\quad \bar{D}_l(t) \quad p_{l,k} \quad (\text{if } \bar{D}_l(t) \neq 0) \end{aligned}$$

All other equations are the same.



1.6 A stability study of queueing systems via fluid models

Definition 6 A fluid model is stable, if $\exists t_0 > 0 : \forall t \geq t_0, \forall \text{ fluid limit},$

$$\overline{Q}(t) \equiv \sum_{k=1}^K \overline{Q}_k(t) = 0 \quad (\text{or } \check{\overline{Q}}(ta) = 0).$$

Introduce (for a queueing network) a Markov Chain

$$\{X_n\} = \{({}^1X_n, \dots, {}^JX_n)\}$$

at embedded arrival epochs:

$${}^jX_n = \{{}^jQ_n; ({}^jx_{n,1}, \dots, {}^jx_{n,{}^jQ_n}); {}^j\chi_n\},$$

where ${}^jQ_n = \sum_{k \in C(j)} Q_{k,n}$; ${}^jx_{n,l}$ — a class number of l -th customer in queue;

${}^j\chi_n$ — a residual service time of the first customer (=0 if ${}^jQ_n = 0$).

Assume that, at all stations, service policies depend on a current state of a queue only. Then $\{X_n\}$ is a time-homogeneous M.C. Put

$$|X_n| \equiv \sum_j |X_j| \equiv \sum_j ({}^jQ_n + {}^j\chi_n) \equiv \sum_{k=1}^K Q_{n,k} + \sum_{j=1}^J {}^j\chi_n.$$

Theorem 2.

(Dai) *If a fluid model is stable, then $\exists N < \infty$ such that the set $B_N \equiv \{x : |x| \leq N\}$ is positive recurrent for $\{X_n\}$.*

Proof.

(a) Preliminary “step”.

Let $a_1, \dots, a_n$ be a sequence of numbers:

$$\frac{a_1 + \dots + a_n}{n} \rightarrow c.$$

Then

$$\frac{a_n}{n} = \frac{a_1 + \dots + a_n}{n} - \frac{n-1}{n} \cdot \frac{a_1 + \dots + a_{n-1}}{n-1} \rightarrow 0.$$

Furthermore, if ν_n is any sequence of positive integers such that $\limsup_{n \rightarrow \infty} \frac{\nu_n}{n} < \infty$,

then $b_n \equiv \frac{a_{\nu_n}}{n} \rightarrow 0$.

Indeed, what means “ $b_n \rightarrow 0$ ”? There exists the following

Criterion (*) $b_n \rightarrow 0$ if and only if $\forall$ subsequence $\{n_m\} \ni$ a subsequence $\{n_{m_l}\} : b_{n_{m_l}} \rightarrow 0$ as $l \rightarrow \infty$.

Make use of it:

Take any subsequence $\{n_m\}$ and consider $b_{n_m} \equiv \frac{a_{\nu_{n_m}}}{n_m}$.

There are two possibilities: either

(i) $\sup_m \nu_{n_m} \equiv L < \infty$ or

(ii) $\exists \{n_{m_l}\}_{l \geq 1} : \nu_{n_{m_l}} \rightarrow \infty$ as $l \rightarrow \infty$.

In the first case,

$$|b_{n_m}| \leq \frac{\max_{1 \leq i \leq L} |a_i|}{n_m} \rightarrow 0, \text{ since } n \rightarrow \infty.$$

In the second case,

$$|b_{n_m}| \equiv \frac{|a_{\nu_{n_{m_l}}}|}{n_{m_l}} \equiv \frac{\nu_{n_{m_l}}}{n_{m_l}} \cdot \frac{|a_{\nu_{n_{m_l}}}|}{\nu_{n_{m_l}}} \leq 1 \cdot \left| \frac{a_{\nu_{n_{m_l}}}}{\nu_{n_{m_l}}} \right| \rightarrow 0,$$

since $\{\nu_{n_{m_l}}\}$ is a subsequence of $\{n\}$.

(b) Prove that, for $\{X_n\}$, for any initial state X_0 ,

$$\frac{\sum_{j=1}^J {}^j\chi_n}{n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Indeed, $\forall j$

$$0 \leq {}^j\chi_n \leq \max_{k \in C(j)} \sigma_k(D_{k,n} + 1) \leq \sum_{k \in C(j)} \sigma_k(D_{k,n} + 1),$$

and $\forall k$

$$\begin{aligned} \frac{\sigma_k(D_{k,n} + 1)}{n} &= \frac{\sigma_k(D_{k,n} + 1)}{D_{k,n} + 1} \cdot \frac{D_{k,n} + 1}{n} \quad \text{and} \\ \frac{\sigma_k(n)}{n} &\rightarrow 0, \quad \limsup_n \frac{D_{k,n} + 1}{n} < \infty. \end{aligned}$$

(c) Take any $0 < \varepsilon < 1$. Prove that

$$\limsup_{N \rightarrow \infty} \frac{\sum_{k=1}^K Q_{k,g(N)}^{(N)}}{N} \leq 1 - \varepsilon,$$

where

$$N \equiv \sum_{k=1}^K Q_{k,0} + \sum_{j=1}^J {}^j\chi_0 \quad \text{and} \quad g(N) = [N \cdot t_0 a].$$

Indeed, if it is not true, then $\exists N_l \rightarrow \infty :$

$$\lim_{l \rightarrow \infty} \frac{\sum_{k=1}^K Q_{k,g(N_l)}^{(N_l)}}{N_l} \rightarrow C > 1 - \varepsilon.$$

Take a subsequence $\models \exists$ a fluid limit: $\check{\bar{Q}}(t_0 a) > 1 - \varepsilon$.

(d) Prove that, $\forall N_0 = \text{const}, \forall k, \forall j \in J$,

$$\text{a family } \left\{ \frac{Q_{k,g(N)}^{(N)}}{N} \right\}_{N \geq N_0} \quad \text{and a family } \left\{ \frac{j \chi_{g(N)}^{(N)}}{N} \right\}_{N \geq N_0}$$

are uniformly integrable (UI).

Indeed,

$$0 \leq \frac{Q_{k,g(N)}^{(N)}}{N} \leq \frac{Q_{k,0}^{(N)} + A_{k,g(N)}^{(N)}}{N} \leq 1 + \frac{A(S_{g(N)})}{N} = 1 + \frac{g(N) + 1}{N},$$

and

$$0 \leq \frac{j \chi_{g(N)}^{(N)}}{N} \leq \frac{j \chi_0^{(N)}}{N} + \max_{k \in C(j)} \frac{\sum_{n=1}^{\eta_k(S_{g(N)})} \sigma_k(n)}{N},$$

where $\eta_k(t)$ — an undelayed renewal process, associated with $\{\sigma_k(n)\}$.

Replace $\max_{k \in C(j)}$ by $\sum_{k \in C(j)} \models$ it is sufficient to prove UI $\forall k$.

Then,

$$\sum_{n=1}^{\eta_k(S_{g(N)})} \sigma_k(n) \leq \sum_{m=1}^{[N]+1} R_k(m) \quad \text{where, for } G(m) = ([t_0 a] + 1)m,$$

$$R_k(1) = \sum_{n=1}^{\eta_k(S_{G(1)})} \sigma_k(n); R_k(2) = \sum_{n=1}^{\tilde{\eta}_k(S_{G(2)-G(1)})} \tilde{\sigma}_k(n); \text{ etc.}$$

and $\tilde{\sigma}_k(n) = \sigma_k(n + \eta_k(S_{G(1)}))$, etc.

Since $\{R_k(m)\}$ are i.i.d. with a finite mean,

$$\frac{\sum_{m=1}^{[N]+1} R_k(m)}{N} \text{ are UI } \models \frac{\sum_{n=1}^{\eta_k(S_{g(N)})} \sigma_k(n)}{N} \text{ are UI } \models \frac{j \chi_{g(N)}^{(N)}}{N} \text{ are UI.}$$

(e): (b), (c) and (d) imply that

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{\sum_{k=1}^K \mathbf{E}\{Q_{k,g(N)}^{(N)}\}}{N} &= \limsup_{N \rightarrow \infty} \frac{\sum_{k=1}^K \mathbf{E}\{Q_{k,g(N)}^{(N)}\} + \sum_{j=1}^J \mathbf{E}\{j \chi_{g(N)}^{(N)}\}}{N} = \\ &= \limsup_{N \rightarrow \infty} \mathbf{E} \left\{ \frac{|X_{g(N)}|}{N} \middle| X_0 = x, |x| \leq N \right\} \leq 1 - \varepsilon. \end{aligned}$$

Therefore, $\exists N_0, \forall N \geq N_0, \forall x, |x| = N$

$$\mathbf{E} \left\{ |X_{g(N)}| \mid X_0 = x \right\} \leq |x| - \frac{\varepsilon}{2} |x|. \quad (3)$$

(3) means that the conditions of Th.1 are satisfied with

$$g(x) = \lfloor |x| \cdot t_0 a \rfloor \text{ and } h(x) = \frac{\varepsilon}{2} |x|, \text{ for } |x| \geq N_0.$$

And, for $|x| = N \leq N_0$,

$$\mathbf{E} \{ {}^j \chi_{g(N)}^{(N)} \} \leq {}^j \chi_0^{(N)} + \sum_k \mathbf{E} \left\{ \sum_{n=1}^{\eta_k(S_{g(N)})} \sigma_k(n) \right\} \leq N_0 + \dots \sum^{\eta_k(S_{g(N_0)})} \dots \equiv c_1(N_0)$$

and

$$\mathbf{E} \{ Q_{k,g(N)}^{(N)} \} \leq Q_{k,0}^{(N)} + A_{k,g(N)}^{(N)} \leq N_0 + g(N_0) + 1 \equiv c_2(N_0).$$

Put $C = J \cdot c_1(N_0) + K \cdot c_2(N_0)$. Then

$$\mathbf{E} \left\{ |X_{g(N)}| \mid X_0 = x, |x| = N \right\} \leq C \equiv |x| - (|x| - C) = |x| - h(x) \text{ for } |x| \leq N.$$

□

Remark 4 Let $c_1, \dots, c_K > 0$ be arbitrary numbers.

Note: $\overline{Q}(t) = \sum_{k=1}^K \overline{Q}_k(t) = 0 \iff \text{all } \overline{Q}_k(t) = 0 \iff \sum_{k=1}^K c_k \overline{Q}_k(t) = 0$.

Therefore, if we replace

$$\begin{aligned} |X_n| &= \sum_{k=1}^K Q_{n,k} + \sum_{j=1}^J {}^j \chi_n \quad \text{by} \\ |X_n| &= \sum_{k=1}^K c_k Q_{n,k} + \sum_{j=1}^J d_j {}^j \chi_n \quad \text{for some } d_1, \dots, d_J > 0, \end{aligned}$$

then the statement of Theorem 2 still holds!

Examples

(1) GI/GI/1 queue. $\mathbf{E}\tau_n = a, \mathbf{E}\sigma_n = b, b < a$.

Take any fluid limit ($\xi_0 = 0$): $\overline{Q}(0) + \overline{\chi}(0) = 1$

(i) $\forall$ regular $t < \bar{\chi}(0)$, $\dot{\bar{D}}(t) = 0$

$$\begin{aligned} \implies \bar{Q}(\bar{\chi}(0)) = \bar{Q}(0) + \frac{\bar{\chi}(0) \cdot 1}{\underset{\parallel}{a}} &\leq \max(1, \frac{1}{a}). \\ &\bar{A}(\bar{\chi}(0)) \end{aligned}$$

(ii) $\forall$ regular $t > \bar{\chi}(0)$, if $\bar{Q}(t) > 0$ then

$$\dot{\bar{D}}(t) = \frac{1}{b} \text{ and } \dot{\bar{Q}}(t) = \frac{1}{a} - \frac{1}{b} = \frac{b-a}{ab} = -\varepsilon.$$

Take $t_0 = 1 + \max(1, \frac{1}{a})/\varepsilon$.

(2) GI/GI/1 multiclass queue. $\mathbf{E}\tau_n = a$, $\sum_k p_k b_k = b < a$.

$$\sum_{k=1}^K \bar{Q}_k(0) + \bar{\chi}(0) = 1$$

(i) $\forall$ regular $t < \bar{\chi}(0)$, $\forall k$ $\dot{\bar{D}}_k(t) = 0$

$$\implies \bar{Q}_k(\bar{\chi}(0)) = \bar{Q}_k(0) + \frac{\bar{\chi}(0) \cdot p_k}{a}.$$

(ii) $\forall$ regular $t > \bar{\chi}(0)$, if $\bar{Q}(t) > 0$ then

$$\sum_{k=1}^K \dot{\bar{T}}_k(t) = \sum_{k=1}^K b_k \dot{\bar{D}}_k(t) = 1.$$

Take a “norm” $|Q(t)| = \sum_{k=1}^K b_k Q_k(t)$.

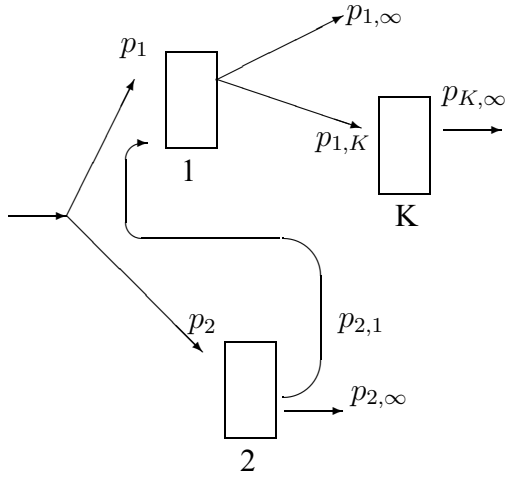
$$\implies |\bar{Q}_k(\bar{\chi}(0))| = |\bar{Q}_k(0)| + \bar{\chi}(0) \cdot \frac{b}{a} \leq \max_k b_k + \frac{b}{a} \equiv C$$

and, $\forall t > \bar{\chi}(0)$, t — regular, if $\bar{Q}(t) > 0$ then

$$|\dot{\bar{Q}}(t)| = \sum_{k=1}^K b_k \dot{\bar{A}}_k(t) - 1 = \sum_{k=1}^K b_k \frac{p_k}{a} - 1 = \frac{b}{a} - 1 = -\varepsilon.$$

Take $t_0 = 1 + C/\varepsilon$.

(3) “Jackson-type” (“generalized Jackson”) networks.



How many times any customer visits, say, station k ?

$$\nu_k : \pi_k = \mathbf{E}\nu_k \implies b_k \cdot \pi_k < a \quad \forall k, \quad \rho = \frac{\max_k b_k \pi_k}{a}.$$

$\rho \leq 1$ is necessary for stability. Show that $\rho < 1$ is sufficient for stability of fluid model.

? What is π_k ? Denote by $\{Y_n\}$ a MC with the state space $\{0, 1, \dots, K, \infty\}$, with initial state $Y_0 = 0$ and with transition probabilities

$$p_{0,k} = p_k; \quad p_{k,l}; \quad p_{\infty,\infty} = 1.$$

Then

$$\pi_k = \mathbf{E} \left\{ \sum_{n=1}^{\infty} I(Y_n = k) \right\} = \sum_{n=1}^{\infty} \mathbf{P}(Y_n = k) = \sum_{l=1}^K \mathbf{P}(Y_1 = l) \sum_{n=1}^{\infty} \mathbf{P}(Y_n = k | Y_1 = l).$$

For $l = k$,

$$\mathbf{P}(Y_1 = k | Y_1 = l) = 1 \text{ and } \mathbf{P}(Y_{n+1} = k | Y_1 = k) = p_{k,k}^{(n)}.$$

For $l \neq k$,

$$\mathbf{P}(Y_1 = k | Y_1 = l) = 0 \text{ and } \mathbf{P}(Y_{n+1} = k | Y_1 = l) = p_{l,k}^{(n)}.$$

Therefore, (for $m = n - 1$)

$$\pi_k = \sum_{l=1}^K p_l \cdot \left(I(l = k) + \sum_{m=1}^{\infty} p_{l,k}^{(m)} \right).$$

Note also:

$$\sum_{k=1}^K \pi_k \cdot p_{k,\infty} = \sum_{k=1}^K \sum_{n=1}^{\infty} \mathbf{P}(Y_n = k, Y_{n+1} = \infty) = \mathbf{P}(\exists n : Y_n = \infty) = 1.$$

Denote $P = \|p_{l,j}\|_{1 \leq l,j \leq K}$. Then $p_{l,j}^{(m)}$ are elements of P^m .

$$\implies \vec{\pi} \equiv (\pi_1, \dots, \pi_K) = (p_1, \dots, p_K) \cdot (I + P + P^2 + \dots) \equiv \vec{p} \cdot (I - P)^{-1}.$$

Necessary and sufficient condition for $\exists(I - P)^{-1}$:

$$|\max. \text{ eigenvalue of } P| < 1 \iff \mathbf{P}(\exists n : Y_n = \infty) = 1.$$

We are needed in the following auxiliary models:

(A) take any $k = 1, \dots, K$ and replace all service times $\{\sigma_k(n)\}$ at station k by 0's. Then a new model is equivalent to a model with $K - 1$ stations and with transition probabilities:

$$\forall l \neq k,$$

$$\tilde{p}_l = p_l + p_k \cdot p_{k,l}(1 + p_{k,k} + p_{k,k}^2 + \dots) = p_l + p_k \cdot p_{k,l} \frac{1}{1 - p_{k,k}}.$$

$$\forall l, j \neq k,$$

$$\tilde{p}_{l,j} = p_{l,j} + p_{l,k} \cdot \frac{1}{1 - p_{k,k}} p_{k,j}.$$

And if, for original model, $\rho < 1$, then $\tilde{\rho} < 1$, too. Here: $\tilde{\rho} = \max_{l \neq k} \frac{b_l \pi_l}{a}$.

(B) take any subset $J \subseteq \{1, 2, \dots, K\}$ and replace all service times $\{\sigma_k(n)\}$ at all stations $k \in J$ by 0's. Then a new model is equivalent to a model with $K - |J|$ stations and with...

And $\rho < 1 \implies \tilde{\rho} < 1$. Here: $\tilde{\rho} = \max_{l \notin J} \frac{b_l \pi_l}{a}$.

Theorem 3 $\rho < 1 \implies$ the fluid model of Jackson-type queueing network is stable.

Proof. Consider any fluid limit $\overline{Q}_k(0), \overline{\chi}_k(0), k = 1, \dots, K$. Put (we consider $c_1 = \dots = c_K = 1$)

$$\overline{Q}(t) = \sum_{k=1}^K \overline{Q}_k(t), \overline{\chi} = \max_k \overline{\chi}_k(0).$$

Show that, $\forall t > \overline{\chi}, t$ -regular, $\overline{Q}(t) > 0$

$$\dot{\overline{Q}}(t) \leq -\varepsilon \quad \text{for some fixed } \varepsilon > 0. \quad (4)$$

Since $\overline{\chi} \leq 1$,

$$\overline{Q}(\overline{\chi}) \leq \overline{Q}(0) + \overline{A}(1) \equiv \overline{Q}(0) + \frac{1}{a} \leq 1 + \frac{1}{a} \quad \implies t_0 = 1 + (1 + \frac{1}{a})/\varepsilon.$$

Prove (4).

(a) Assume that $\overline{Q}_k(t) > 0 \quad \forall k = 1, \dots, K$. Then

$$\dot{\overline{D}}_k(t) = \frac{1}{b_k} \text{ and}$$

$$\dot{\overline{Q}}_k(t) = \dot{\overline{A}}_k(t) + \sum_{l=1}^K \frac{1}{b_l} \cdot p_{l,k} - \frac{1}{b_k} = \frac{p_k}{a} + \sum_{l=1}^K \frac{1}{b_l} \cdot p_{l,k} - \frac{1}{b_k}.$$

Therefore,

$$\begin{aligned} \dot{\overline{Q}}(t) &= \sum_{k=1}^K \frac{p_k}{a} + \sum_{k=1}^K \sum_{l=1}^K \frac{1}{b_l} \cdot p_{l,k} - \sum_{k=1}^K \frac{1}{b_k} \equiv \\ &\equiv \frac{1}{a} + \sum_{k=1}^K \sum_{l=1}^K \frac{1}{b_l} \cdot p_{l,k} - \sum_{l=1}^K \frac{1}{b_l} = \frac{1}{a} + \sum_{l=1}^K \frac{1}{b_l} \cdot \left(\sum_{k=1}^K p_{l,k} - 1 \right) = \\ &= \frac{1}{a} - \sum_{l=1}^K \frac{1}{b_l} \cdot p_{l,\infty}. \end{aligned}$$

Note: $\pi_l b_l < a \quad \forall l$, or $\frac{1}{b_l} > \frac{\pi_l}{a}; \quad -\frac{1}{b_l} < -\frac{\pi_l}{a}.$

Therefore,

$$-\varepsilon_{\{1,2,\dots,K\}} \equiv \frac{1}{a} - \sum_{l=1}^K \frac{1}{b_l} \cdot p_{l,\infty} < \frac{1}{a} - \sum_{l=1}^K \frac{\pi_l}{a} \cdot p_{l,\infty} = \frac{1}{a} - \frac{1}{a} = 0.$$

(b) Assume that $\bar{Q}_k(t) > 0$ for some, but not all k . Renumber stations in such a way, that

$$\bar{Q}_1(t) > 0, \dots, \bar{Q}_M(t) > 0, \bar{Q}_{M+1}(t) = \dots = \bar{Q}_K(t) = 0.$$

Note:

$$\dot{\bar{Q}}_j(t) = 0; j = M + 1, \dots, K \quad \Longrightarrow \quad \dot{\bar{A}}_j(t) = \dot{\bar{D}}_j(t); j = M + 1, \dots, K.$$

It means (locally) that the whole “fluid”, that comes to station j , leaves this station immediately. It is equivalent to the following: this station behaves (locally) like a station with zero-valued service times!

Therefore,

$$\dot{\bar{Q}}(t) = \sum_{k=1}^M \dot{\bar{Q}}_k(t) = \frac{1}{a} - \sum_{l=1}^M \frac{1}{b_l} \cdot \tilde{p}_{l,\infty} \equiv -\varepsilon_{\{1,\dots,M\}} < 0.$$

Finally, take

$$\varepsilon = \min_{J \subseteq \{1,2,\dots,K\}, J \neq \emptyset} \varepsilon_J > 0 \quad \Longrightarrow \quad \forall t > \bar{\chi}, \quad \dot{\bar{Q}}(t) \leq -\varepsilon.$$

□

Concluding remarks

1) The fluid approximation approach (FAA) simplified (significantly!) proofs of many known stability results.

2) The use of FAA gave a opportunity and tools for stability analysis of many other queueing (and communication, manufacturing, etc.) models, for which stability conditions were unknown before.

3) FAA is not universal. There are many examples where the linear scaling is too “rough” and irrelevant to an essential information...