A Deterministic Model of a Service Station (Fluid View)

Primitives

- \( Z(0) \) initial content
- \( \alpha(t) \) input rate
- \( \mu(t) \) potential service rate

\[ \begin{align*}
\text{in} & \quad \longrightarrow \quad \text{Delay} \quad \longrightarrow \quad \text{out}
\end{align*} \]

Model: (Think cumulants)

Inflow: \( A(t) = \int_0^t \alpha(u)du, \quad t \geq 0 \);

Potential Outflow: \( M(t) = \int_0^t \mu(u)du, \quad t \geq 0 \).

- We could start with primitives \( A, M \), in which case they need not be continuous; for example, they could be counting processes.

Netflow: \( X(t) = Z(0) + A(t) - M(t), \quad t \geq 0 \).

Introduce \( Y(t) = \) cumulative potential lost during \([0, t]\).

⇒ Outflow: \( D = M - Y \) (A arrivals; D departures)

⇒ Balance:

\[ \begin{align*}
Z(t) &= Z(0) + A(t) - D(t) \\
&= Z(0) + A(t) - [M(t) - Y(t)] \\
&= X(t) + Y(t), \quad t \geq 0.
\end{align*} \]

Model

\( Z = X + Y \)

Feasible \( Z \geq 0, \ Y \uparrow 0 \) \( Y(0) = 0 \);

Efficient \( Y \) least (hence, \( Y \) unique);

Existence: \( Y = (-X)^+ \) \( Y = -\underline{X} \), when \( Z(0) = 0 \);

\( \underline{X}(t) = \inf_{0 \leq u \leq t} X(u) \), which is called the lower envelope of \( X \).
“Proof”

Least $Y \uparrow 0$
s.t. $Y \geq -X$

When $Z(0) = 0$:

$Z = X - X$
$X = $ lower envelope.

Equivalent characterization via complementarity: (LCP/DCP)

$Y$ least $\iff ZdY = 0$, i.e. $Y$ increases at $t$
only when $Z(t) = 0$.

In words: potential lost due to idleness.

Claim (Skorohod)  Given $X \in \text{RCLL (Right Continuous Left Limit)}$,

there exists a unique $(Y, Z)$ such that

\[
\begin{align*}
Z &= X + Y, \\
Z &\geq 0, \quad Y \uparrow 0, \\
ZdY &= 0.
\end{align*}
\]

Proof  Existence by checking $Y = (\overline{-X})^+ \ (= -X \land 0)$.

Uniqueness by Lyapunov-function argument:

(Note: if minimality is established, then uniqueness is automatic.)

If $(Y_i, Z_i), \ i = 1, 2,$ are two solutions, then consider

\[
\eta = \frac{1}{2}(Y_1 - Y_2)^2.
\]
Assume, for simplicity, continuous \( Y_i \)'s, in which case differentiate:

\[
d\eta = (Y_1 - Y_2)(dY_1 - dY_2) = (Z_1 - Z_2)(dY_1 - dY_2) = -Z_1dY_2 - Z_2dY_1 \leq 0.
\]

Deduce that \( \eta \) decreases, but also

\[
\eta(0) = 0 \Rightarrow \eta \equiv 0 \\
\Rightarrow Y_1 \equiv Y_2.
\]

**Outflow**  
\[
D(t) = M(t) - Y(t) = \int_0^t \delta(u)du, \quad \text{where } \delta(u) = \text{outflow rate},
\]

\[
\Rightarrow Y(t) = \int_0^t [\mu(u) - \delta(u)]du.
\]

In terms of rates: \( dY \geq 0 \) implies \( \delta \leq \mu \).

Now, either  
\[
\delta = \mu \quad \text{or} \quad \delta < \mu \Leftrightarrow dY > 0,
\]

\[
\Rightarrow Z = 0 \quad (\text{since } ZdY = 0),
\]

\[
\Rightarrow d(X + Y) = 0 \quad (\text{consider a neighbourhood and differentiate}),
\]

\[
\Rightarrow (\alpha - \mu) + (\mu - \delta) = \alpha - \delta = 0.
\]

Thus (Hall, pg. 190, Def. 6.6),

\[
\delta(t) = \begin{cases} 
\mu(t) & \text{when } Z(t) > 0, \\
\alpha(t) & \text{when } Z(t) = 0.
\end{cases}
\]

**Note** that the above is not a direct definition of \( \delta \), since it uses \( Z \), which is defined in terms of \( \delta \).
How to calculate Delay?

Define

\[ W(t) = \text{work-load at time } t \]
\[ (= \text{time to process all that is present at time } t) \]
\[ = \text{under FCFS, virtual waiting time.} \]

Defining relation for \( W \):

\[ D(t + W(t)) = Z(0) + A(t) \]

Hence, \( Z(t + W(t)) = Z(0) + A(t + W(t)) - A(t) \).

MOP’s over a finite horizon \( T \):

**Averages**

- **Inflow**: \( \bar{\alpha} = \frac{1}{T} \int_0^T \alpha(t) dt; \)
- **Outflow**: \( \bar{\delta} = \frac{1}{T} \int_0^T \delta(t) dt; \)
- **Throughput**: \( \lambda, \text{ defined when } \bar{\alpha} = \bar{\delta} \text{ as their common value.} \)

Queue length (Inventory): \( \bar{Z} = \frac{1}{T} \int_0^T Z(t) dt = \frac{1}{T} \times \text{Area.} \)

Delay: \( \bar{W} = \frac{1}{\lambda(T)} \int_0^T W(t) dA(t) \quad \left( = \frac{\int_0^T W(t) \alpha(t) dt}{\int_0^T \alpha(t) dt} \right). \)

\[ \uparrow \text{ Rieman-Stiltjes} \]
Intuition:

- Discrete arrivals ⇒ \( \bar{W} = \frac{1}{A(T)} \sum_{n=1}^{A(T)} W_n \) (as in Hall, Chap. 2);
- Absolutely continuous: \( \alpha(t)\,dt \) arrivals during \((t, t + dt)\), each suffering a delay of \( W(t) \).

Little’s Conservation Law: \( \bar{Z} = \lambda \cdot \bar{W} \).

Cumulative lost potential \( Y(T) \).

Efficiency \( \varepsilon(T) = 1 - \frac{Y(T)}{M(T)} = \)

\[
= \frac{D(T)}{M(T)} \left( \frac{\int_{0}^{T} \delta(t)\,dt}{\int_{0}^{T} \mu(t)\,dt} , \text{when applicable} \right) .
\]

Example \( constant \ rates \quad \alpha(t) \equiv \alpha , \quad \mu(t) \equiv \mu . \)

(linear model)

\[
\begin{array}{c}
\alpha > \mu \\
\alpha = \mu \\
\alpha < \mu \\
\end{array}
\]

overloaded (supercritical) \( \rho > 1 \)
critically loaded (critical) \( \rho = 1 \)
underloaded (subcritical) \( \rho < 1 \)

Definition: \( \rho = \alpha/\mu \) traffic (flow) intensity.

Natural extension: piecewise constant rates, as in National Cranberry (HBS case).

Example \( periodic \ rates \ e.g . \)

(If \( \alpha \) has a period \( T_{\alpha} = 8 \), \( \mu \) has a period \( T_{\mu} = 3 \), take period \( T = T_{\alpha} \cdot T_{\mu} = 24 \).)
Long-run: \[ \bar{\alpha} = \frac{1}{T} \int_{0}^{T} \alpha(t)dt; \quad \bar{\mu} = \frac{1}{T} \int_{0}^{T} \mu(t)dt; \]
\[ \rho = \bar{\alpha}/\bar{\mu} \text{ (Heyman-Whitt)}. \]

Short-run: Phase-transitions (different from Hall, pg. 189–190, that has stagnant \rightarrow growth \rightarrow decline \rightarrow stagnant).

**Short-Run Phase Transitions**

Overloaded at \( t \): \( Z(t) > 0; \)
Underloaded \( \delta(t) < \mu(t) \) (excess capacity, \( dY(t) > 0); \)
Critically loaded \( \delta(t) = \mu(t) \) (balanced capacity, \( dY(t) = 0) \).

The analogue of \( \rho \), traffic intensity, is here (assume \( Z(0) = 0) \):

\[ \rho(t) = \sup_{0 \leq s \leq t} \frac{\int_{s}^{t} \alpha(u)du}{\int_{s}^{t} \mu(u)du} \begin{cases} > 1 & \text{overloaded} \\ = 1 & \text{critically loaded} \\ < 1 & \text{underloaded} \end{cases} \]
For finer approximations, we must acknowledge more phases, as depicted in the following figure.

Phase transition diagram for the asymptotic regions.
(Massey & Mandelbaum.)

References:


Mathematical Framework

Reflection Mapping \( \rightarrow X - X \wedge 0 \)  
(Regulator)  
\((X \rightarrow X - X, \text{ when } X(0) = 0)\).

Fundamental:

- Flow analysis (Fluid Models);
- Economics;
- Stochastic Processes;
  - Skorohod (needed cumulant \( Y' \));
  - Queueing Models (later);
- Approximations.

Idea of Approximations: \[ Z = f(X), \text{ } f \text{ continuous (Lipshitz)}. \]

Hence, \( X \approx X \) implies \( Z \approx Z = f(X) \)

\( X \approx X \) fluid \( \Rightarrow Z = f(X) \) fluid approximations.

\( X \approx X + \dot{X} \) diffusion \( \Rightarrow \dot{Z} = f(X + \dot{X}) \) diffusion refinements.

Reference: Harrison, Chapter 2 (which covers also finite buffers, and two-node networks).