CanQueue2010

M. Hlynka

Department of Mathematics and Statistics
University of Windsor
Windsor, ON, Canada.

August 27, 2010

In Memory of Marvin Mandelbaum

- MARVIN MANDELBAUM Aug 19, 1941-Sept. 8, 2009
- Marvin developed split and match (or fork/join) queues for his Master's degree (1968) at the Technion in Israel under the supervision of Benjamin Avi-Itzhak.
- Marvin received a PhD in Industrial Engineering from the University of Toronto in 1978.
  His supervisor was A.A. Cunningham.
  PhD thesis title was Flexibility in Decision Making: An Exploration and Unification.

Marvin Mandelbaum, continued

- He was President of the Toronto Section of the Canadian Operations Research Society from 1993-1999. He was on the national council of CORS from 1995-97. In 1998, Marvin was a winner of a CORS service award.
- Marvin organized CanQueue 2003 at the Fields Institute in Toronto.

THANK YOU, MARVIN!
REST IN PEACE

Transient Queues; Queue and Task Model
by Myron Hlynka and Shan XU
2010

- OUTLINE
- Transient queue formulae for M/M/1
- Graphs (probabilities and expected queue length)
- Task and Queue Model
- Graphical Explanation

The earliest formulas for the $M/M/1$ transient queueing probabilities are given by Clarke (1953) and by Lederman and Reuter (1954). Gross, Shortle, Thompson and Harris (2008) and Kleinrock (1975) present the Lederman and Reuter result in terms of modified Bessel functions. Their expression is:

$$p_{ij}(t) = e^{-(\lambda+\mu)t}[\rho^{(j-i)/2} I_{j-i}(2t\sqrt{\lambda\mu}) + \rho^{(j-i-1)/2} I_{j+i+1}(2t\sqrt{\lambda\mu}) + (1-\rho)\rho^{j} \sum_{k=j+i+2}^{\infty} \rho^{-k/2} I_k(2t\sqrt{\lambda\mu})]$$

where

$$I_k(x) = \sum_{m=0}^{\infty} \frac{(x/2)^{k+2m}}{(k+m)!m!}$$

is the modified Bessel function of the first kind of order k.

For an M/M/1 queue, van de Coevering (1994) presents the result

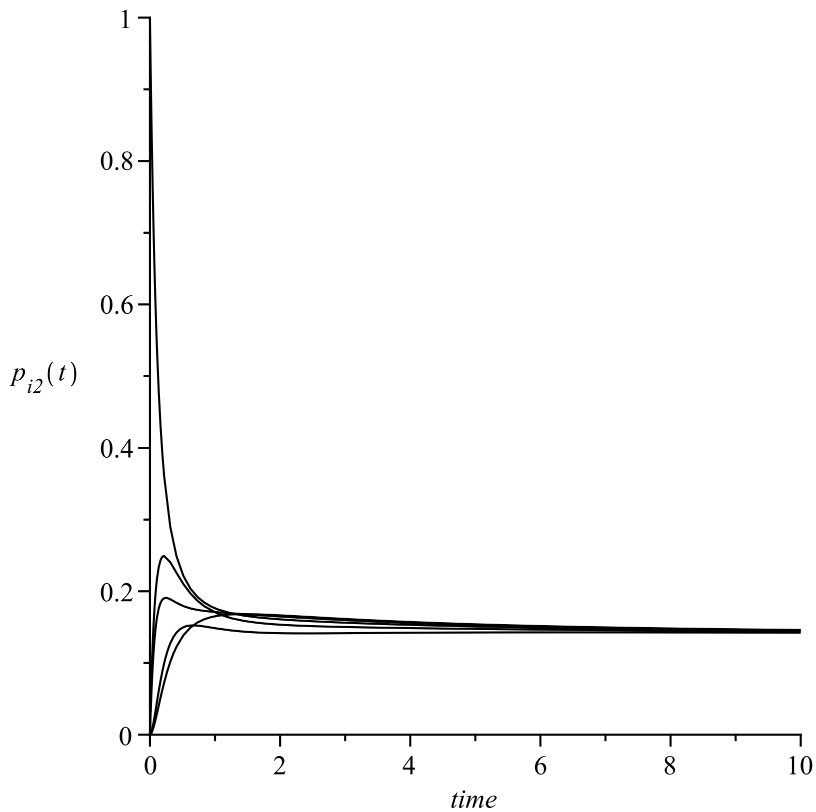
$$p_{ij}(t) = \frac{2}{\pi}\rho^{(j-i)/2}\int_0^{\pi} \frac{e^{-\mu t\gamma(y)}}{\gamma(y)} a_i(y)a_j(y)\,dy + \begin{cases}(1-\rho)\rho^j & \rho < 1\\ 0 & \rho \geq 1,\end{cases} \quad (1)$$

for

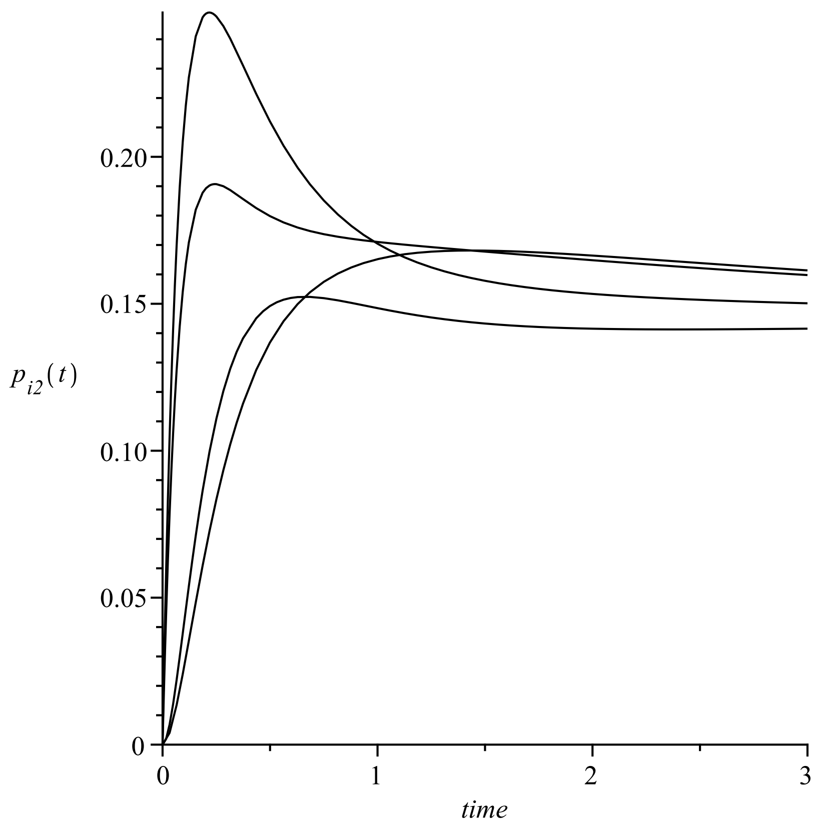
$$\gamma(y) = 1 + \rho - 2\sqrt{\rho}\cos(y) \text{ and } a_k(y) = \sin(ky) - \sqrt{\rho}\sin((k+1)y).$$

He credits Morse (1955) and Takacs (1962, page 23) for this trigonometric version.

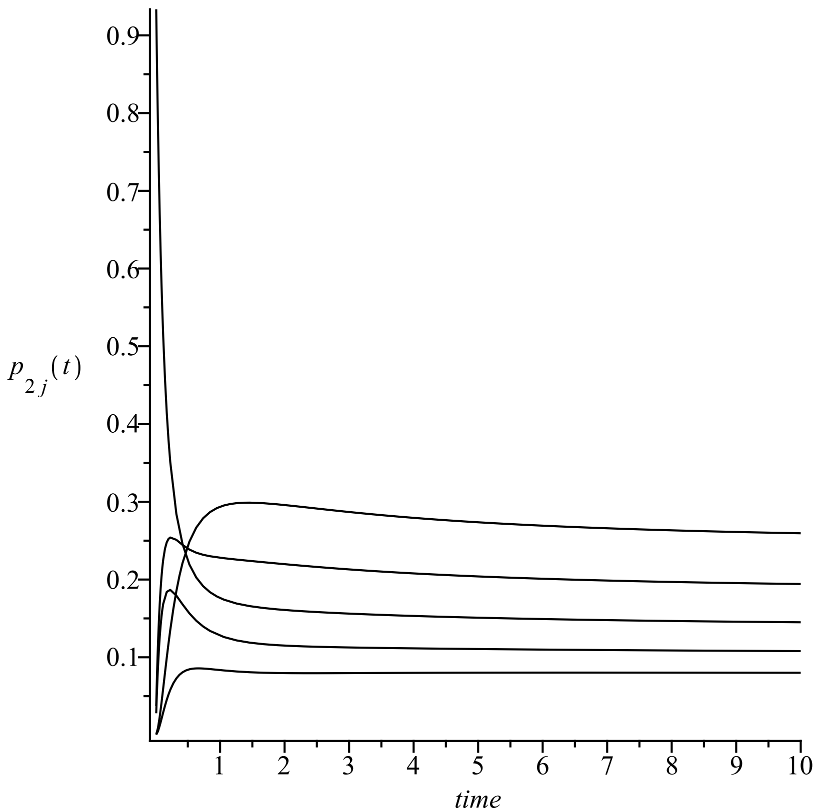
D1: $p_{i2}(t)$, $i = 0, \ldots, 4$, $\lambda = 3, \mu = 4, 0 < t < 10$



D1.1 $p_{i2}(t)$, $i = 0, 1, 3, 4$, $\lambda = 3, \mu = 4, 0 < t < 3$



D2: $p_{2j}(t)$, $j = 0, \ldots, 4$, $\lambda = 3, \mu = 4, 0 < t < 10$

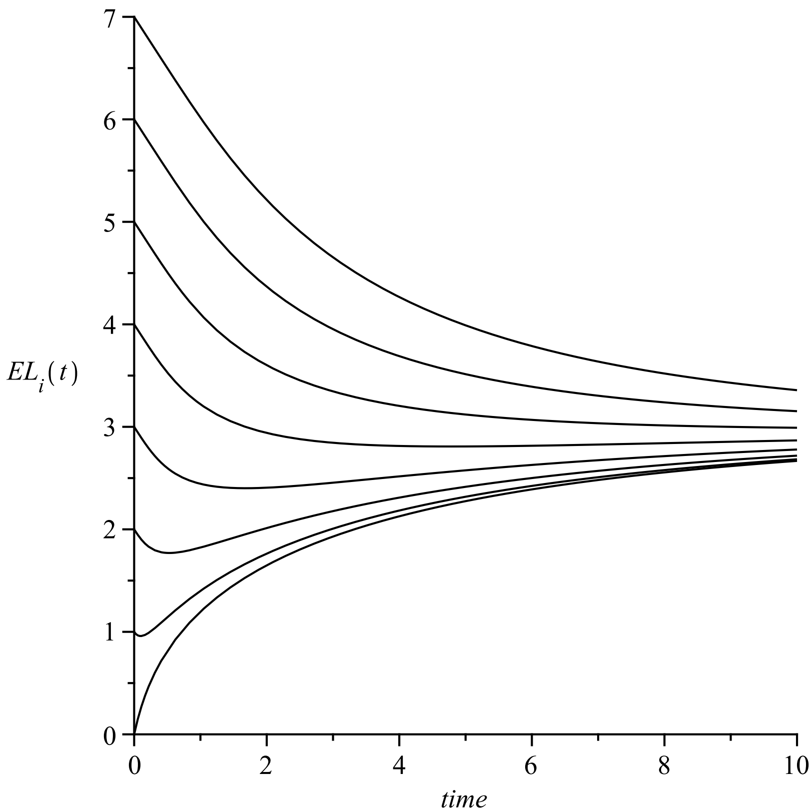


Theorem

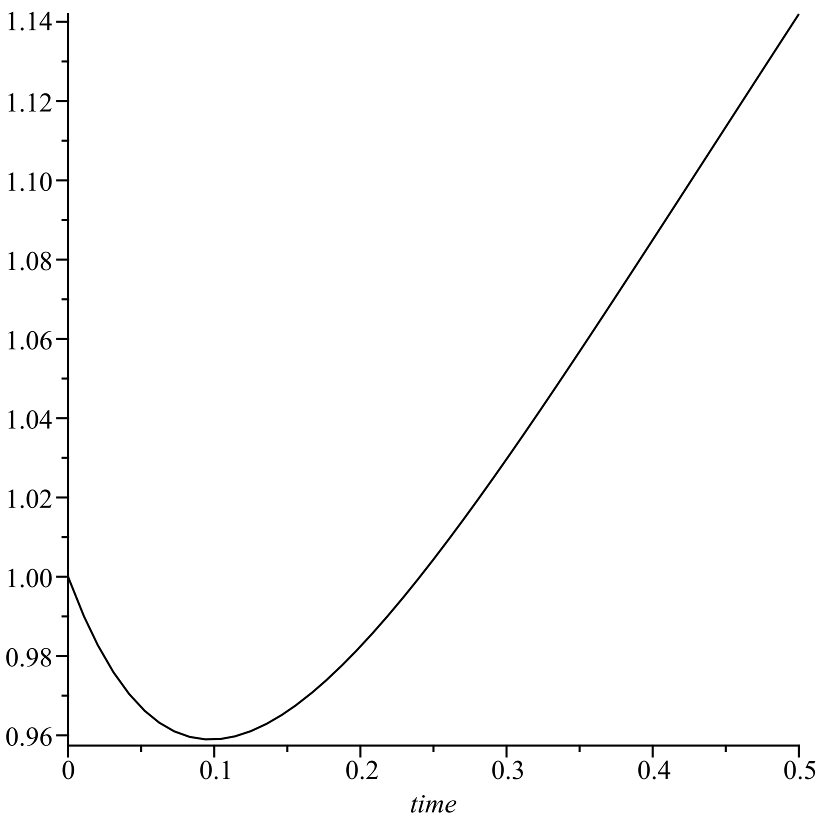
(van de Coevering) For an $M/M/1$ queueing system, let $E_i(t)$ be the expected number of customers at time t if there are i customers at time 0. Then

$$EL_i(t) = \frac{2}{\pi}\rho^{(j-i)/2}\int_0^{\pi} \frac{e^{-\mu t\gamma(y)}}{\gamma(y)^2} a_i(y)\sin(y)\,dy + \begin{cases} \rho/(1-\rho) & \rho < 1 \\ i + (\lambda-\mu)t + \rho^{-1}/(\rho-1) & \rho > 1. \end{cases}$$

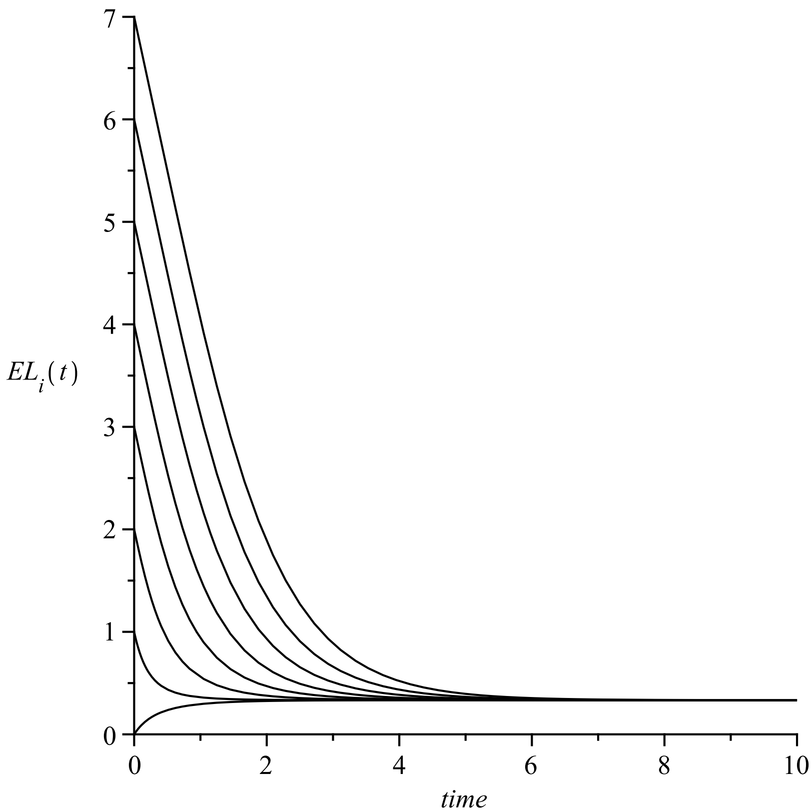
D3: $EL_i(t)$, $i = 0, \ldots, 7$, $\lambda = 3, \mu = 4, 0 < t < 10$



D4: $EL_1(t)$, $\lambda = 3, \mu = 4, 0 < t < .5$



D4.5: $EL_i(t)$, $i = 0, \ldots, 7$, $\lambda = 1, \mu = 4, 0 < t < 10$



Theorem

If $\lambda < \mu$*, then*

$$\left.\frac{d(E(L_i(t)))}{dt}\right|_{t=0} = \lambda - \mu < 0,$$

for $i = 1, 2 \ldots$

- PROOF
- $E(L_i(0)) = i$.

$$\begin{aligned} E(L_i(\Delta t)) &= \sum_{\text{all j}} jp_{ij}(\Delta t) = (i+1)(\lambda\Delta t + o(\Delta t)) + (i-1)(\mu\Delta t + \\ &\quad + i(1 - \lambda\Delta t - \mu\Delta t + o(\Delta t)) + o(\Delta t) \\ &= i + (\lambda - \mu)\Delta t + o(\Delta t) \\ \frac{d(E(L_i(t)))}{dt}\Big|_{t=0} &= \lim_{\Delta t \to 0} \frac{E(L_i(\Delta t)) - E(L_i(0))}{\Delta t} \\ &= \lambda - \mu < 0. \end{aligned}$$

- MODEL: Task vs Queue
- We have an $M/M/1$ queueing system. A customer arrives and sees i people ($i = 0, 1, 2, \ldots$) in the system. This customer must receive service from the server and has an additional task of fixed length D. The customer has two choices:
- (a) Let T_{QD} be the total system time if the customer joins the queue first, then does the task.
- (b) Let T_{DQ} be the total system time if the customer does the task first, then joins the queue.

- Assume i is the number of customers initially observed.
  Let $p_{ij}(t) = \text{Prob}(j \text{ customers at time } t | i \text{ customers at time } 0)$.
- Let T_{QD} be the total system time of the arriving customer if the queue is done first.
- Let T_{DQ} be the total system time of the arriving customer if the task is done first.
- When is $ET_{QD} < ET_{DQ}$?

- Our intuition suggests that if the initial number i is large (relative to $E(L)$, the average system length), then the arriving customer should perform the task first (and hope that the length decreases) and that if the initial number i is small (relative to $E(L)$), then the arriving customer should join the queue first for fear that it might increase by the time the task is completed.

Theorem

Given the setting above,

(a) $$E(T_{QD}) = \frac{i+1}{\mu} + D$$

(b) $$E(T_{DQ}) = D + \frac{E(L_i(D)) + 1}{\mu}$$

Proof.

Let X_k $(k = 1, 2, \ldots)$ be the independent identically distributed service times of the customer in the system, including the customer already in service (if any) and the newly arriving customer. Since we have an $M/M/1$ system, the memoryless property of the exponential distribution allows us to includ the customer in service in this way. Then

$$T_{QD} = \sum_{j=1}^{i+1} X_j + D.$$

Note that the upper limit is $i + 1$ since we need to service the newly arriving customer in addition to those in the system. So $E(T_{QD}) = E(\sum_{j=1}^{i+1} X_j + D) = \sum_{j=1}^{i+1} E(X_j) + E(D) = \sum_{j=1}^{i+1} \frac{1}{\mu} + D = \frac{i+1}{\mu} + D.$ □

Proof.

(Continued)
Next consider T_{DQ}. Let X_k ($k = 1, 2, \ldots$) be the iid service times of the customer in the system, including the customer already in service (if any) and the newly arriving customer, and any customers arriving during time interval $(0, D)$. While the task of length D is being performed, the length of the queue will change from i to $L_i(D)$. Thus

$$T_{DQ} = D + \sum_{j=1}^{L_i(D)+1} X_j$$

where $L_i(D)$ is a random variable. So

$$\begin{aligned} E(T_{DQ}) &= E(D + \sum_{j=1}^{L_i(D)+1} X_j) = E(D) + E(\sum_{j=1}^{L_i(D)+1} X_j) \\ &= D + (E(L_i(D)) + 1)E(X_j) = D + \frac{E(L_i(D)) + 1}{\mu}. \end{aligned}$$

Theorem

Given the setting of this model, $E(T_{QD}) > E(T_{DQ})$ *iff* $i > EL_i(D)$.

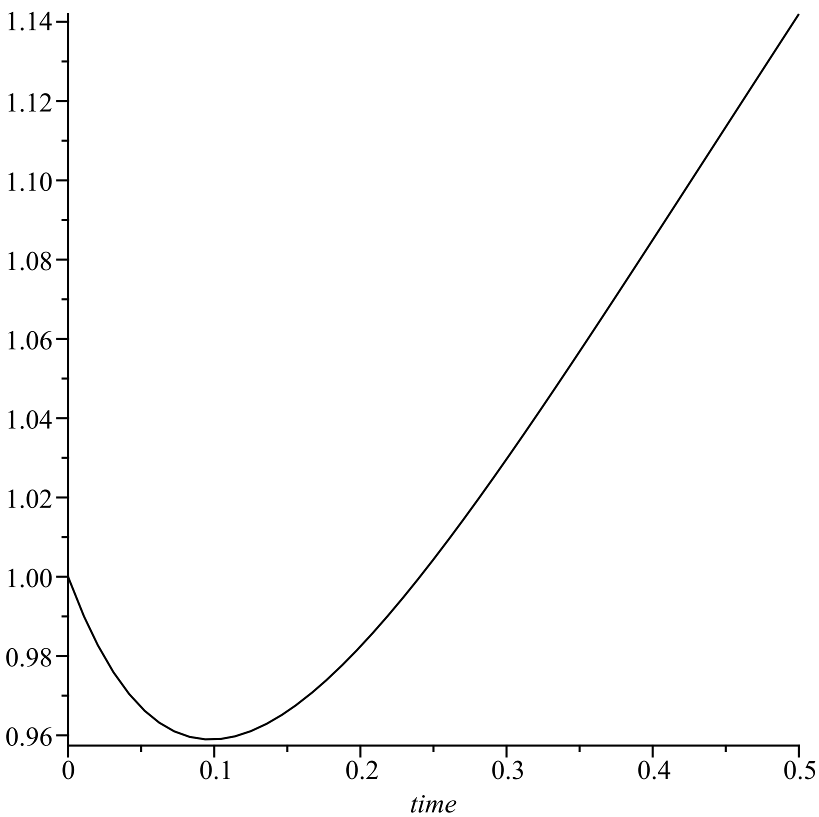
Corollary

Let the task time be D and let the initial number in the queueing system be i. If $EL_i(D) < i$*, then the arrival should perform the task first.*

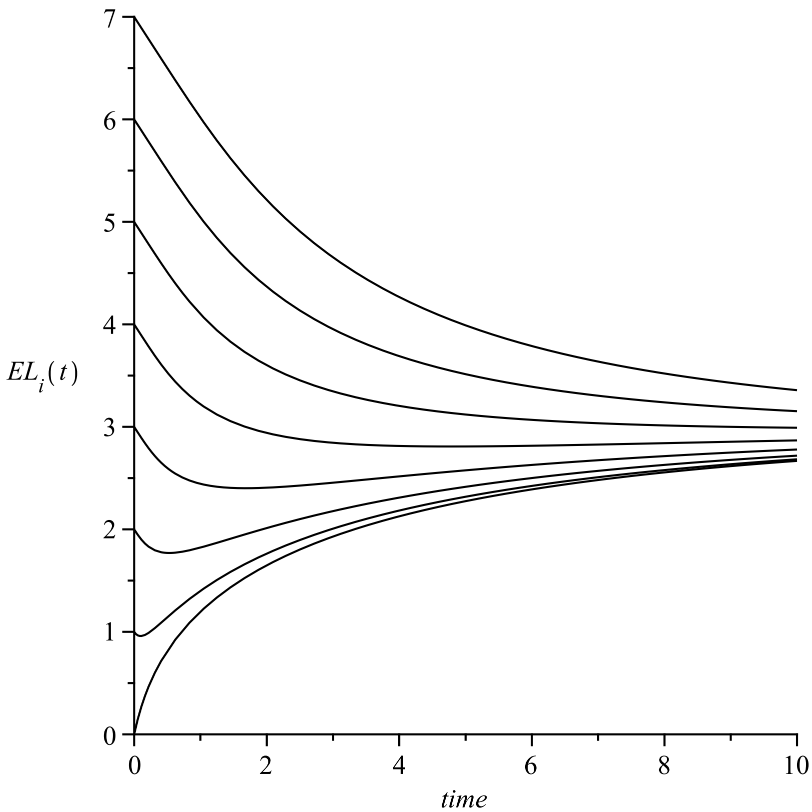
As a result we can get the following nice description of the condition for preferring to do the task first.

- PROCEDURE:
- 1. For given initial count i, draw graph of $E(L_i(t))$.
- 2. Draw horizontal line from $(0, i)$ to until it intersects the graph at some point (D^*, i).
- 3. For any task time D with $D < D^*$, the arriving customer should perform the task first. Otherwise not.

D4: $EL_1(t)$, $i = 0, \ldots, 4$, $\lambda = 3, \mu = 4, 0 < t < .5$



D3: $EL_i(t)$, $i = 0, \ldots, 7$, $\lambda = 3, \mu = 4, 0 < t < 10$



References

- Abate, J. and Whitt, W. (1988) Transient behavior of the M/M/1 queue via Laplace transforms, *Advances in Applied Probability*, 20, 145-178.
- ALFA, Attahiru Sule. (2010) Queueing Theory for Telecommunications: Discrete Time Modelling of a Single Node System. 247 p. 100 illus., 50 in color., ISBN 978-1-4419-7313-9
- Jain, JL. Mohanty, S., and Bohm, W. (2007) A Course on Queueing Models. Chapman and Hall.
- BRILL, Percy. (2008) Level Crossing Methods in Stochastic Models. Springer.
- Champernowne, D. G. (1956) An elementary method of solution of the queueing problem with a single server and constant parameters, *Journal of the Royal Statistical Society*, B18, 125-128.

- Conolly, B. W. and Langaris Christos (1993) On a New Formula for the Transient State Probabilities for M/M/1 Queue and Computational Implications. *Journal of Applied Probability*, Vol. 30, No. 1. Mar., 237-246.
- Gross, D. Shortle, J., Thompson, J. and Harris, C.M., (2008) Fundamentals of Queueing Theory, 4th ed. New York: *Wiley*.
- Hlynka, Myron and Molinaro, Samantha. (2010) Comparing Expected Wait Times of a M/M/1 Queue. University of Windsor, Windsor, ON., N9B 3P4, Canada
- Kleinrock, L. (1975) Queueing Systems, Vol.I. New York: *Wiley*.
- Ledermann, W. and Reuter, G.E.H., (1954) Spectral theory for the differential equations of simple birth and death processes, *Philosophical Transactions of the Royal Society*, London, Ser. A, 246, 321-369.

- Morse, P.M. (1955). Stochastic properties of Waiting Lines. Operations Research, 3, 225-261.
- Morse, P.M. (1958). Queues, Inventories, and Maintenance. Wiley Publications.
- Sharma, O. P. (1990) Markovian Queues. *Ellis Horwood*, Chichester.
- Takacs, L. (1962) Introduction to the theory of queues, *Oxford University Press, London*
- Van de Coevering, M.C.T. (1995) Computing transient performance measures for the M/M/1 queue. OR Spectrum Volume 17, Number 1, 19-22.

- Thank you!