LONG RUN DYNAMICS OF QUEUES: STABILITY AND CHAOS

by

Eric J. Friedman and A. S. Landsberg

Memorandum No. UCB/ERL/IGCT M93/78

29 October 1993

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

LONG RUN DYNAMICS OF QUEUES: STABILITY AND CHAOS

by

Eric J. Friedman and A. S. Landsberg

Memorandum No. UCB/ERL/IGCT M93/78

29 October 1993

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Long Run Dynamics of Queues: Stability and Chaos

Eric J. Friedman*
Department of Industrial Engineering and Operations Research,
University of California, Berkeley, CA 94720

A. S. Landsberg
Department of Physics
University of California, Berkeley, CA 94720

October 29, 1993

Abstract

We analyze the long run dynamics of queues in which customers undergo self selection. We describe the structure and local stability of equilibria for the various capacity adjustment procedures and solve the problem of global stability for the limiting cases. The intermediate cases can be quite complicated. We show that one such case leads to chaotic dynamics.

Keywords: Queues, Dynamics, Stability, Chaos, Service Facilities.

1 Introduction

In this note we analyze the dynamics of queues with customers undergoing self-selection. In a previous paper [7] we examined the short-run dynamics of such queueing systems. In this paper we focus on the long-run dynamics, in which the operator of the queue varies the processing rate of the queue in order to induce efficient operations. We show that if the operator updates the processing rate either very often or very rarely then the analysis is straightforward, and the system is well behaved; however in the intermediate cases the analysis can be quite complicated, and the dynamics can be chaotic.

In the following section we give a brief review of the problem and its motivation. However, for a more complete survey we refer the reader to our previous work [7] and Sidham's original analysis [12].

*E.J.F. thanks the National Science Foundation Grant IRI-8902813 for support.

2 The Model

Consider the case of a queueing system in which customers decide whether or not to enter based on a entry charge p and a delay cost which depends on the number of other customers also requesting service. We represent these customers by their arrival rate λ. The customers' delay cost $D(\tau)$ is a convex increasing function of their (subjective) expected waiting time τ.

The operator of this system chooses the service rate μ (e.g. the number of workers in a repair shop, or the speed of the CPU or network in a computer system) in order to maximize the total social utility of the queue. This service rate affects customers through the queueing delay $G(\lambda, \mu)$ which increases for higher arrival rate of customers λ and decreases as the capacity (μ) is increased. However, there is a cost to the operator associated with increased capacity, $B(\mu)$. Thus the operator must balance the cost of increased capacity against the benefits received by the customers, while at the same time trying infer the preferences of the customers from their arrival rate.

The question of how customers predict their expected delay can be very complicated; the subject of learning has received much study both in psychology and economics (e.g. [3, 9, 2]). In this note we follow Stidham [12] and our previous work [7] in assuming static expectations. This implies that the expected delay in period $n+1$ denoted τ_{n+1} is simply the observed delay in period n, e.g. $\tau_{n+1} = G(\lambda_n, \mu_n)$. In economics this is known as Best-Reply (or Cournot) dynamics, and has been extensively studied [4, 1, 10]. Recently Rump and Stidham [11] have analyzed the case of adaptive expectations, for the short run model, and show that the dynamics in this model may be much more complex than that of static expectations. In fact they have shown that the system may become chaotic. However, Friedman [6] has shown that if the static expectations model is stable, then any model with 'weakly-rational' expectations is also stable for the short run model.

Formally we define the following functions:

1. The density of job values: $r(v)$ is continuous, positive, and bounded with bounded support.

2. The arrival rate function: $L(c) = \int_c^\infty r(v)dv$.

3. The 'demand function':

$$V(\lambda) = \int_{v_m(\lambda)}^{\infty} v\, r(v)\, dv$$

where $\lambda = L(v_m(\lambda))$ defines $v_m(\lambda)$. Define the marginal value function $v(\lambda) = V'(\lambda)$.

4. The delay function: $G(\lambda, \mu)$ which is positive, decreasing and convex. As in Stidham [12] we assume that $\frac{\partial G(\lambda,\mu)}{\partial \lambda} + \frac{\partial G(\lambda,\mu)}{\partial \mu} = 0$, which is satisfied by many queues including the M/M/1 and M/GI/1. We also assume that $\lim_{\mu\to\infty} G(\lambda, \mu) = 0$ and $\lim_{\mu\to\infty} \partial G(\lambda,\mu)/\partial\mu = 0$.

5. The delay cost function: $D(\tau)$ is convex, increasing, and non-negative.

6. Linear capacity cost: $B(\mu) = b\mu$, with $b > 0$.

First we note that we can simplify the form of $G(\lambda, \mu)$.

Lemma 1 *Define $\hat{G}(\mu - \lambda) = G(0, \mu - \lambda)$. Then $\hat{G}(\mu - \lambda) = G(\lambda, \mu)$.*

Proof: Integration along the line from $(\mu - \lambda, 0)$ yields (μ, λ) we get

$$G(\lambda, \mu) = G(\mu - \lambda, 0) + \int_0^\lambda \left(\frac{\partial G(\mu - \lambda + t, t)}{\partial \mu} + \frac{\partial G(\mu - \lambda + t, t)}{\partial \lambda}\right) dt.$$

However the integrand is zero by assumption and therefore $G(\lambda, \mu) = G(\mu - \lambda, 0)$. □

In order to simplify notation we define

$$H(x) = D(\hat{G}(x))$$

and note that $H(x)$ is positive, decreasing, and convex. We also let $h(x) = H'(x)$.

The short run dynamics is defined by the customers' self-selection problem in which customers whose job value (v) is greater than the total cost ($H(\mu - \lambda) + p$) will enter the queue. Under the assumption of static expectations we see that

$$\lambda_{n+1} = L(H(\mu - \lambda_n) + p) = T_\mu(\lambda_n) \tag{1}$$

The key property of the mapping T_μ is that it is positive and decreasing. In [7] we presented a detailed study of the properties of T_μ. We restate here a result from that paper for later use.

Theorem 1 *There exists an $\hat{\mu}$ such that for $\mu > \hat{\mu}$ the dynamics of (1) always converges to the equilibrium arrival rate $E(\mu)$ defined by*

$$E(\mu) = T_\mu(E(\mu)). \tag{2}$$

Proof: See [7].

2.1 Stidham's Example

In [12], Stidham considered a model with an m/m/1 queue ($G(\lambda, \mu) = (\mu - \lambda)^{-1}$), linear delay cost ($D(\tau) = h\tau$), and a uniform distribution of customer values ($r(v) = \Lambda/a$ for $v \in [0, a]$ and $r(v) = 0$ otherwise). This leads to

$$T_\mu(\lambda) = \Lambda'(1 - \frac{h/a'}{\mu - \lambda}) \tag{3}$$

where $\Lambda' = [a'/a]\Lambda$ and $a' = a - p$. He shows that for $\mu > \Lambda'$ the system has a unique globally attracting equilibrium. In [7] we show that for $\mu < \Lambda'$ all orbits (except for those starting at the equilibrium) will run into the boundary. We note that this second property only holds for the specific mapping (3) and is not typical of mappings of the type (1).

3 Long Run Dynamics

The long run dynamics arise in the problem of maximizing social welfare in both arrival rate (λ) and capacity (μ):

$$\max_{\mu,\lambda} [V(\lambda) - \lambda H(\mu - \lambda) - b\mu]. \tag{4}$$

In this case the system operator chooses service rate (μ) and entry charge (p) and the customers enter the queue based on self selection, i.e. a customer will enter the queue when his value v exceeds his cost $H(\mu - \lambda) + p$. It is shown in Dewan and Mendelson [5] and Stidham [12] that the operator should charge a price equal to the marginal cost of additional capacity ($p = b$) when $\mu = \mu^*$ in order that the optimal arrival rate λ^* is the equilibrium of the short run problem. Thus it makes sense for the operator to charge $p = b$. From Equation (2) we see that for $p = b$ the operator's problem reduces to

$$\max_{\mu} [V(E(\mu)) - E(\mu)H(\mu - E(\mu)) - b\mu] \tag{5}$$

First we examine some properties of the optimal solution.

Lemma 2 *$E(\mu)$ is increasing in μ.*

Proof: Differentiate (2) with respect to μ to get

$$E'(\mu) = L' \cdot h \cdot [1 - E'(\mu)].$$

Rearranging terms we see that

$$E'(\mu) = \frac{L'h}{1 + L'h}$$

and note that both the numerator and denominator are positive. □

Now assume that λ is fixed and the operator wishes to compute the optimal capacity. This requires the operator to solve

$$\max_{\mu} \, [V(\lambda) - \lambda H(\mu - \lambda) - b\mu].$$

for μ. Call this maximum $M(\lambda)$.

Lemma 3 *$M(\lambda) - \lambda$ is increasing in λ.*

Proof: The first order condition for $M(\lambda)$ to be a maximum is

$$\lambda h(M(\lambda) - \lambda) + b = 0$$

Differentiate this with respect to λ to get

$$\frac{d}{d\lambda}[M(\lambda) - \lambda] = \frac{-h(M(\lambda) - \lambda)}{\lambda h'(M(\lambda) - \lambda)}$$

which is positive since both the numerator and denominator are positive. □

Note that for Stidham's example we have $M(\lambda) = \lambda + \sqrt{\lambda h/b}$ and $\Lambda(\mu) = [\mu + \Lambda' - \sqrt{(\mu + \Lambda')^2 - 4\Lambda'(\mu - h/a')}]/2$, which agree with the above lemmas.

Now there are several possibilities for the operator's behavior. (See [12] for a more detailed description.) For example, every period the operator could adjust the capacity to be optimal assuming that the arrival rate will remain constant. This corresponds to setting $\mu = M(\lambda)$.

Alternatively, the operator might prefer (due to technological restrictions) to update the capacity less often, for instance once every k periods. If we track the arrival rate only when the operator changes the capacity, we see that the dynamics is defined by the new mapping

$$P_k(\lambda) = T^k_{M(\lambda)}(\lambda). \tag{6}$$

where T^k_μ represents the composition of T_μ with itself k times.

For example, using the techniques in [7] we can show that for Stidham's example

$$P_k(\lambda) = \lambda + \sqrt{\lambda h/b} - \frac{(\gamma_-^{n+1} - \gamma_+^{n+1})\sqrt{\lambda h/b} - (\gamma_+\gamma_-^{n+1} - \gamma_-\gamma_+^{n+1})}{(\gamma_-^n - \gamma_+^n)\sqrt{\lambda h/b} - (\gamma_+\gamma_-^n - \gamma_-\gamma_+^n)}$$

where

$$\gamma_\pm = \frac{\lambda + \sqrt{\lambda h/b} - \Lambda'}{2} \pm \sqrt{(\frac{\lambda + \sqrt{\lambda h/b} - \Lambda'}{2})^2 + \frac{\Lambda' h}{a'}}.$$

For $k = 1$ the behavior of this map is quite simple. This is because the map $P_1 = \Lambda'(1 - \sqrt{bh/\lambda a'^2})$ is strictly increasing in this example. However, this is also true in general.

Theorem 2 *$P_1(\lambda)$ is strictly increasing in λ.*

Proof: Note that $P_1(\lambda) = L(H(M(\lambda) - \lambda) + b)$. Differentiate with respect to λ we get

$$P_1' = L' \cdot h \cdot \frac{d}{d\lambda}(M(\lambda) - \lambda)$$

which is positive by Lemma 3. □

The properties of increasing maps are quite simple as has been shown in [6]. Everything is completely determined by the maps fixed points and local stability. In Stidham's model the mapping P_1 is concave and is always less than Λ'. In Figure 1 we show the three possibilities for maps of this type. Case (a) is trivial as all trajectories go to zero. In case (b) it is clear that the larger fixed point λ'' is stable, while the smaller λ' one must be unstable (since its slope is greater than 1). Globally, this implies (see [6]) that for all $\lambda_0 > \lambda'$ the arrival rate will converge to λ'' while for $\lambda_0 < \lambda'$ the arrival rate will converge to zero. Finally, the degenerate case (c) has a single fixed point λ' and for $\lambda_0 \geq \lambda'$ the arrival rate converges to λ' and zero otherwise. Thus it is easy to understand the mapping when $k = 1$.

In the case where $T_{M(\lambda)}$ has a globally stable equilibrium (which is always true for λ sufficiently large, as show in Lemma 3) we can define the map for $k = \infty$ as $P_\infty(\lambda) = \Lambda(M(\lambda))$. Thus

$$\lim_{k\to\infty} P_k(\lambda) = P_\infty(\lambda)$$

only when $T_{M(\lambda)}$ has a unique globally attracting fixed point.

However, on the region for which it is defined P_∞ is an increasing map, and its behavior is identical to that of P_1.

Theorem 3 *$P_\infty(\lambda)$ is increasing in λ.*

Proof: Differentiate with respect to λ and use lemmas 2 and 3. □

Thus the analysis of P_∞ is straightforward. (In fact, using the techniques we develop in the next section, the proof becomes trivial.)

When $2 \leq k < \infty$, P_k is not monotonic and the behavior can be quite complicated as we discuss in the next section.

3.1 Chaotic Dynamics

Consider the mapping P_2 for Stidham's example with $\Lambda' = 1$, $a' = 1$, $h = .1$, and $b = 1.2$. This mapping is plotted in Figure 2a. We plot the third iterate of this map in Figure 2b. From this figure we see that the map P_2 must have four 3-cycles. It is well known that any mapping of the interval $[0,1]$ into itself with a 3-cycle is chaotic (in the topological sense) [8]. However, this chaos is not strictly relevant since it occurs on a set of measure zero. It remains an open question whether measurable chaos occurs in these mappings. (For example in [11] it is shown that a simple model of queueing system dynamics leads to measurable chaos.)

However, in all the cases the properties of the fixed points is important; particularly for the cases when $k = 1$ or $k = \infty$ in which the local properties of the fixed points determine the global behavior of the map. We study this structure in the next section.

4 Fixed Points and Stability

In this section we show some properties of the fixed points of the various mappings P_k for different k's.

Define the set of fixed points for P_k by $S_k = \{\lambda \mid P_k(\lambda) = \lambda\}$. Note that this set can contain a large number of points. Note that by definition $\lambda \in S_k$ if and only if λ is a k-cycle of $T_{M(\lambda)}$; however, if k is odd then all k-cycles of T_μ are fixed points since T_μ is a decreasing map. If k is even then the k-cycles of $T_{M(\lambda)}$ are either fixed points or 2-cycles. Thus we have the following theorem.

Theorem 4

1) If k is odd then $S_k = S_1$.

2) If k is even then $S_k = S_2$.

3) $S_\infty = S_1$.

4) $S_1 \subseteq S_2$.

Note that there is a significant difference depending on whether k is odd or even. Now we analyze the stability of these fixed points.

For Stidham's specific example we see that $S_1 = S_2 \cup \lambda'$ where $M(\lambda') = \Lambda'$ since this is the only case where $T_M(\lambda)$ has a 2-cycle.

Let $\hat{\lambda}$ be a fixed point of P_1 ($\hat{\lambda} \in S_1$). Then we know that $\hat{\lambda}$ must be a fixed point for all P_k ($\hat{\lambda} \in S_k$) by Theorem 4. We can compute the stability of this fixed point $\hat{\lambda}$ for all of the mappings P_k.

Lemma 4 *Let $\alpha = T'_{M(\hat{\lambda})}(\hat{\lambda})$ and $\beta = -\alpha M'(\hat{\lambda})$, then*

$$P'_k(\hat{\lambda}) = \alpha^k + \beta\frac{1-\alpha^k}{1-\alpha}$$

and note that $P'_\infty = \beta/(1-\alpha)$ when it is defined.

Proof: Note the following identity among the P_k's

$$P_k(\lambda) = T^k_{M(\lambda)}(P_{k-1}(\lambda)).$$

It follows by differentiating that

$$P_k'(\hat{\lambda}) = \alpha P_{k-1}'(\hat{\lambda}) + \beta. \tag{7}$$

This is a decreasing linear mapping for $P_k'(\hat{\lambda})$ which we can solve analytically with the initial condition $P_1'(\hat{\lambda}) = \alpha + \beta$ to prove the above results. □

Note that (4) is a linear decreasing mapping. Thus we can easily understand its dynamics to prove further relations between the stability fixed points of the various mappings P_k. First recall that the condition that a fixed point $\hat{\lambda}$ be stable is $|P_k'(\hat{\lambda})| < 1$. Thus we can show the following:

Theorem 5

1)If $\hat{\lambda}$ is stable for $k = 1$ and $k = \infty$ then $\hat{\lambda}$ is stable for all odd k, but not necessarily if k is even.

2) There exists a $\hat{k}$ such that for $k \geq \hat{k}$ the stability of $\hat{\lambda}$ under P_k is the same as for P_∞.

3) If $\hat{\lambda}$ is unstable for P_∞ and stable for P_1 then it is unstable under P_k for all even k.

4) If $\hat{\lambda}$ is stable for P_∞ and unstable for P_1 then $P_k'(\hat{\lambda})$ is positive for all odd k, but may be negative for certain even k.

Proof: First note that both $s_1 \equiv P_1'(\hat{\lambda}) \geq 0$ and $s_\infty \equiv P_\infty(\hat{\lambda}) \geq 0$ by Theorems 2 and 3.

1,4) All even iterates of (7) must lie between s_1 and s_∞.

2) This is true because the iterates of (7) converge to its equilibrium s_∞.

3) All even iterates of (7) must be larger than s_∞.

Once again the differences between even and odd k are significant. For example part (4) of the above theorem has an interesting corollary.

Corollary 1 *Assume that k is odd. Then for all $\hat{\lambda} \in S_k = S_1$, $P_k'(\hat{\lambda}) > 0$.*

The above statement may be false when k is even. Thus, while the limiting cases ($k = 1$ and $k = \infty$) are well behaved, the intermediate cases ($2 \leq k < \infty$) can be quite complicated, even chaotic, and the dynamics depends precisely on the specific model chosen.

References

[1] B. D. Bernheim. Rationalizable strategic behavior. *Econometrica*, 52:1007–1028, 1984.

[2] G.H. Bower and E.R. Hilgard. *Theories of Learning*. Appleton-Century-Crofts, New York, 1966.

[3] Carter and Maddock. *Rational Expectations*. MacMilan, London, 1984.

[4] A. Cournot. *Recherches sur les Principes Mathematics de la Theorie de la Richesse*. Hachette, Paris, 1838.

[5] S. Dewan and H. Mendelson. User delay costs and internal pricing for a service facility. *Management Science*, 36(12):1502–17, December 1989.

[6] E. J. Friedman. Dynamics and rationality in ordered externality games. mimeo, 1993.

[7] E. J. Friedman and A. S. Landsberg. Short run dyanmics of multi-class queues. *Operations Research Letters*, 1993. Forthcoming.

[8] T.Y. Li and J.A. Yorke. Period three implies chaos. *Amer. Math. Monthly*, 82:985–992, 1975.

[9] P. Milgrom and J. Roberts. Adaptive and sophisticated learning in repeated normal form games. *Games and Economic Behavior*, 3:82–100, 1991.

[10] H. Moulin. Dominance sovability and Cournot stability. *Mathematical Social Sciences*, pages 83–102, 1984.

[11] C. Rump and S. Stidham. Chaotic dynamics in a service facility. mimeo, 1993.

[12] S. Stidham. Pricing and capacity decisions for a service facility: Stability and multiple local optima. *Management Science*, 38(8):1121–1139, 1992.

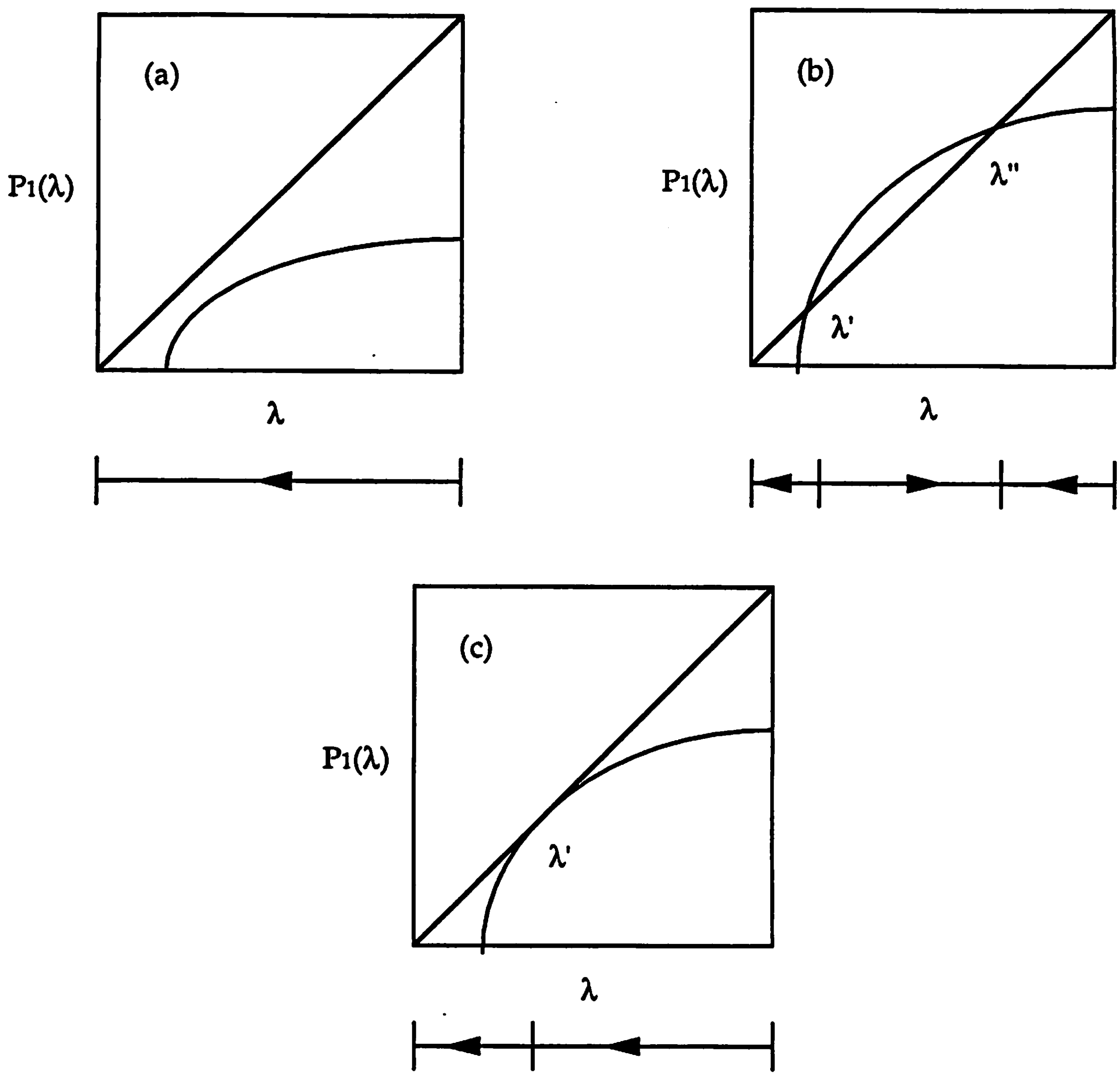

Figure 1: The three possibilities for P1(λ) and their stability diagrams.

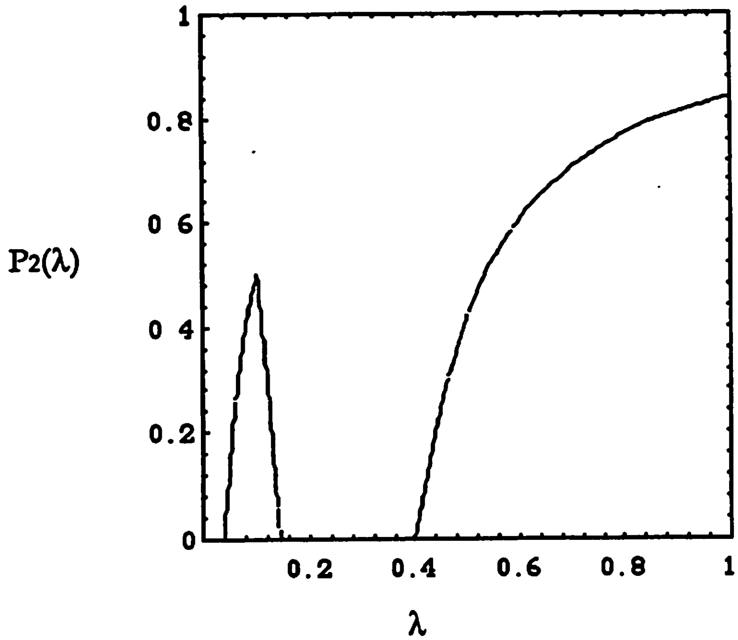

Figure 2a: Plot of P2 for $\Lambda'=1$, $a'=1$, $b=1.2$, $h=.1$.

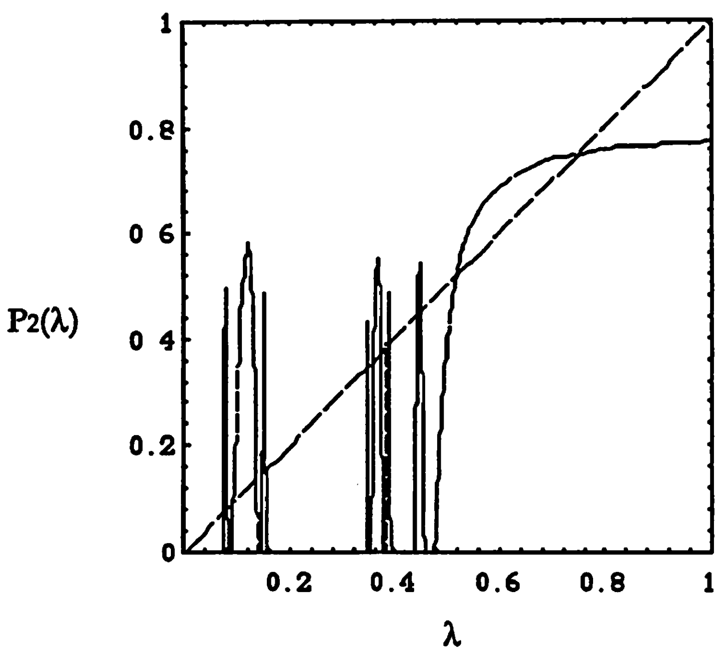

Figure: 2b: Plot of the third iterate of P2 for $\Lambda'=1$, $a'=1$, $b=1.2$, $h=.1$.