Technical Notes and Correspondence

On Continued Fraction Inversion by Routh's Algorithm

R. PARTHASARATHY AND HARPREET SINGH

Abstract—This note points out that the generalized algorithm proposed recently by Chao *et al.* for Cauer second form can be applied for inverting a continued fraction in the Cauer first form by merely writing the transfer function as a ratio of two polynomials arranged in the descending powers of "S."

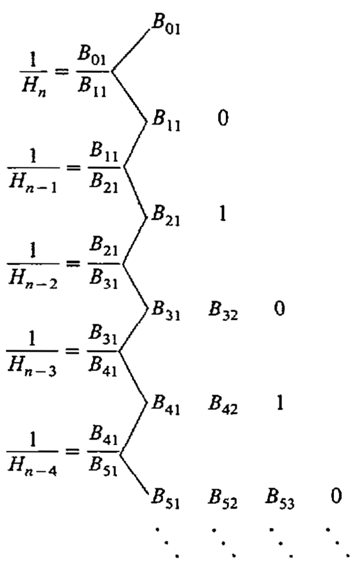
In [1] Rao and Lamba presented an algorithm for inverting a continued fraction in the Cauer second form by developing the Routh array from an even number of partial coefficients. Once the array is so constructed, it is noticed that the first two rows of the Routh table give the corresponding transfer function. The authors have recently extended [3] the procedure of Rao and Lamba for the continued fraction inversion for the case of Cauer first form.

Chao *et al.* have, in a recent note [2], proposed a generalization of the result presented in [1] when the continued fraction expansion is given in the Cauer second form and an odd or even number of partial coefficients are known.

The object of this note is to point out that this generalized algorithm [2], based on a backward expansion of the Routh array, can be applied as well to the inversion of the continued fraction in Cauer first form which is represented by

$$G(s)=\cfrac{1}{H_1S+\cfrac{1}{H_2+\cfrac{1}{H_3S+\cfrac{1}{\ddots}}}}. \tag{1}$$

Given an arbitrary number of partial coefficients $H_n, H_{n-1}, \cdots, H_1$, the inverse table is constructed following the same procedure as in [2]. The table, after correcting the typographical error in the partial coefficients as given in [2], is reproduced below for ready reference.



(2)

It is to be noted that $B_{01}=1$ and that the end elements of all the rows can be written by inspection as 1 or 0, accordingly, as

$$B_{2i,i+1}=1, \qquad i=1,2,3,\cdots$$
$$B_{2i-1,i+1}=0, \qquad i=1,2,3,\cdots. \tag{3}$$

Once the array is completely written, it will be seen that the coefficients in the nth and $(n+1)$th rows give the numerator coefficient a_i and the denominator coefficient b_i, respectively, of the resulting transfer function.

It turns out that these are indeed the coefficients of the corresponding powers of S of the numerator and denominator polynomials, arranged in the descending order as denoted by

$$G(S)=\frac{a_0S^M+a_1S^{M-1}+\cdots+a_M}{b_0S^N+b_1S^{N-1}+\cdots+b_N} \tag{4}$$

where

$$\left.\begin{array}{l} N=n/2 \\ M=(n-1)/2 \end{array}\right\} \quad N>M, \qquad \text{for any } n, \text{ even or odd}$$

whereas a continued fraction in the Cauer second form, on inversion, will give the resulting transfer function representation as [2]

$$G(S)=\frac{a_0+a_1S+\cdots+a_MS^M}{b_0+b_1S+\cdots+b_NS^N} \tag{5}$$

where

$$N=n/2 \quad N>M, \qquad n \text{ even}$$
$$M=(n-1)/2 \quad N>M, \qquad n \text{ odd}$$

and

$$a_i=B_{n-1,i+1}, \qquad i=0,1,2,\cdots,M$$
$$b_i=B_{n,i+1}, \qquad i=0,1,2,\cdots,N. \tag{6}$$

ACKNOWLEDGMENT

The authors wish to thank Prof. A. K. Kamal for his advice and encouragement.

REFERENCES

[1] S. Vittal Rao and S. S. Lamba, "A note on continued fraction inversion by Routh's algorithm," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 283–284, June 1974.
[2] K. S. Chao, K. S. Lu, and I. P. S. Madan, "On the inversion table of continued fraction expansion," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 136–137, Feb. 1975.
[3] R. Parthasarathy and H. Singh, "On continued fraction inversion by Routh's algorithm," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 278–279, Apr. 1975.

Manuscript received May 19, 1975; revised December 30, 1975.
R. Parthasarathy is with the Department of Electronics and Communication Engineering, University of Roorkee, Roorkee, India, on leave from the Indian Institute of Technology, Madras, India.
H. Singh is with the Department of Electronics and Communication Engineering, University of Roorkee, Roorkee, India.

On Error Bounds for Successive Approximation Methods

DIMITRI P. BERTSEKAS

Abstract—**This note considers a class of contraction mappings and the successive approximation method for obtaining the associated fixed points. Some error bounds are provided which generalize and strengthen those given by McQueen [1] and Denardo [2] for dynamic programming algorithms.**

Manuscript received July 7, 1975. This work was supported by the Joint Services Electronics Program under Contract DAAB-07-72-C-0259 and by the National Science Foundation under Grant ENG-74-19332.
The author is with the Coordinated Science Laboratory and the Department of Electrical Engineering, University of Illinois, Urbana, IL 61801.

I. Main Result

Let X be a set and $B(X)$ be the set of all bounded real valued functions on X. For any two functions $f,f' \in B(X)$ we write

$$f=f', \quad \text{if } f(x)=f'(x), \quad \forall x \in X$$
$$f \leqslant f', \quad \text{if } f(x) \leqslant f'(x), \quad \forall x \in X.$$

Let also $T: B(X) \to B(X)$ be a mapping on $B(X)$ having the following two properties:

$$f \leqslant f' \Rightarrow T(f) \leqslant T(f'), \qquad \forall f,f' \in B(X) \tag{1}$$

$$\alpha_1 \leqslant \frac{T(f+re)(x)-T(f)(x)}{r} \leqslant \alpha_2, \qquad \forall r \neq 0, f \in B(X), x \in X \tag{2}$$

where e is the unit function on X

$$e(x)=1, \qquad \forall x \in X \tag{3}$$

and α_1, α_2 are two scalars with

$$0<\alpha_1 \leqslant \alpha_2 <1. \tag{4}$$

Notice that (2) may also be written as

$$T(f)+\min[\alpha_1 r, \alpha_2 r]e \leqslant T(f+re) \leqslant T(f)+\max[\alpha_1 r,\alpha_2 r]e, \qquad \forall r \neq 0, f \in B(X), x \in X. \tag{5}$$

We shall make frequent use of the expression above.

It is easy to show that relations (1) and (2) imply that T is a contraction mapping on $B(X)$ viewed as a normed space with the sup norm. Indeed for any two functions $f,f' \in B(X)$ and every $x \in X$ we have

$$f'(x)-\sup_{x\in X}|f(x)-f'(x)| \leqslant f(x) \leqslant f'(x)+\sup_{x\in X}|f(x)-f'(x)|$$

and by applying T above and using (1), (2),

$$T(f')(x)-\alpha_2 \sup_{x\in X}|f(x)-f'(x)| \leqslant T(f)(x) \leqslant T(f')(x)+\alpha_2 \sup_{x\in X}|f(x)-f'(x)|$$

or equivalently,

$$\sup_{x\in X}|T(f)(x)-T(f')(x)| \leqslant \alpha_2 \sup_{x \in X}|f(x)-f'(x)|.$$

Since $B(X)$ with the sup norm is complete it follows that T has a unique fixed point $f^* \in B(X)$

$$f^*=T(f^*).$$

Furthermore, the successive approximation method which generates $T(f),\ldots,T^k(f),\cdots$ starting from an arbitrary function $f \in B(X)$ has the convergence property

$$\lim_{k\to\infty} T^k(f)(x)=f^*(x), \qquad \forall x \in X.$$

The following proposition provides monotonic upper and lower bounds on the difference $T^k(f)(x)-f^*(x)$.

Proposition: For any $f \in B(X)$, $x \in X$, and $k=1,2,\cdots$ there holds

$$T^k(f)(x)+b_k \leqslant T^{k+1}(f)(x)+b_{k+1} \leqslant f^*(x) \leqslant T^{k+1}(f)(x)+\bar{b}_{k+1} \leqslant T^k(f)(x)+\bar{b}_k \tag{6}$$

where

$$b_k=\min\left[\frac{\alpha_1}{1-\alpha_1}d_k, \frac{\alpha_2}{1-\alpha_2}d_k\right] \tag{7}$$

$$\bar{b}_k=\max\left[\frac{\alpha_1}{1-\alpha_1}\bar{d}_k, \frac{\alpha_2}{1-\alpha_2}\bar{d}_k\right] \tag{8}$$

$$d_k=\inf_{x\in X}[T^k(f)(x)-T^{k-1}(f)(x)] \tag{9}$$

$$\bar{d}_k=\sup_{x\in X}[T^k(f)(x)-T^{k-1}(f)(x)]. \tag{10}$$

Proof: It is sufficient to prove (6) for $k=1$ since the result for $k>1$ then follows by replacing f by $T^{k-1}(f)$. For notational convenience denote

$$d_1=d, \quad \bar{d}_1=\bar{d}, \quad d_2=d', \quad \bar{d}_2=\bar{d}'.$$

We have

$$f(x)+d \leqslant T(f)(x), \quad \forall x \in X. \tag{11}$$

Applying T on both sides of the above and using (1), (5), (11)

$$f(x)+\min[d+\alpha_1 d, d+\alpha_2 d] \leqslant T(f)(x)+\min[\alpha_1 d,\alpha_2 d] \leqslant T(f+de)(x) \leqslant T^2(f)(x), \qquad \forall x \in X. \tag{12}$$

Applying T and using (1), (5), (11), (12)

$$\begin{aligned} f(x)&+\min[d+\alpha_1 d+\alpha_1^2 d, d+\alpha_2 d+\alpha_2^2 d] \\ &\leqslant T(f)(x)+\min[\alpha_1 d+\alpha_1^2 d, \alpha_2 d+\alpha_2^2 d] \\ &\leqslant T^2(f)(x)+\min[\alpha_1^2 d,\alpha_2^2 d] \\ &\leqslant T(T(f)+\min[\alpha_1 d,\alpha_2 d]e)(x) \leqslant T^3(f)(x). \end{aligned}$$

Proceeding similarly we have for every $k=1,2,\cdots$

$$f(x)+\min\left[\sum_{i=0}^k \alpha_1^i d, \sum_{i=0}^k \alpha_2^i d\right] \leqslant T(f)(x)+\min\left[\sum_{i=1}^k \alpha_1^i d, \sum_{i=1}^k \alpha_2^i d\right] \leqslant \cdots \leqslant T^k(f)(x)+\min[\alpha_1^k d,\alpha_2^k d] \leqslant T^{k+1}(f)(x).$$

Taking the limit as $k \to \infty$

$$f(x)+\min\left[\frac{1}{1-\alpha_1}d, \frac{1}{1-\alpha_2}d\right] \leqslant T(f)(x)+\min\left[\frac{\alpha_1}{1-\alpha_1}d,\frac{\alpha_2}{1-\alpha_2}d\right] \leqslant T^2(f)(x)+\min\left[\frac{\alpha_1^2}{1-\alpha_1}d,\frac{\alpha_2^2}{1-\alpha_2}d\right] \leqslant f^*. \tag{13}$$

Also we have from (12)

$$\min[\alpha_1 d,\alpha_2 d] \leqslant T^2(f)(x)-T(f)(x)$$

and by taking the infimum over $x \in X$ above

$$\min[\alpha_1 d, \alpha_2 d] \leqslant d'.$$

It is easy to see that the above relation implies

$$\min\left[\frac{\alpha_1^2}{1-\alpha_1}d,\frac{\alpha_2^2}{1-\alpha_2}d\right] \leqslant \min\left[\frac{\alpha_1}{1-\alpha_1}d',\frac{\alpha_2}{1-\alpha_2}d'\right]. \tag{14}$$

Combining (13), (14) and using the definitions (7), (9) we obtain

$$T(f)(x)+b_1 \leqslant T^2(f)(x)+b_2.$$

Also from (13) we have $T(f)(x)+b_1 \leqslant f^*(x)$, $\forall x \in X$ and an identical argument shows that $T^2(f)(x)+b_2 \leqslant f^*(x)$, $\forall x \in X$. Hence, the left part of (6) is proved for $k=1$. The right part follows by an entirely similar argument. Q.E.D.

II. Discussion

Error bounds involving successive differences of iterates of methods of successive approximation utilizing monotone contraction mappings of the type considered here have been given by MacQueen [1] and Denardo

[2]. Reference [1] considers the case where X is a finite set $X=\{1,2,\cdots,n\}$, $\alpha_1=\alpha_2=\alpha$ and obtains the error bounds

$$T^{k-1}(f)(i)+\frac{1}{1-\alpha}d_k \leqslant T^k(f)(i)+\frac{1}{1-\alpha}d_{k+1}$$

$$\leqslant f^*(i) \leqslant T^k(f)(i)+\frac{1}{1-\alpha}\bar{d}_{k+1}$$

$$\leqslant T^{k-1}(f)(i)+\frac{1}{1-\alpha}\bar{d}_k, \qquad i=1,2,\cdots,n \quad (15)$$

for the case of the mapping $T:B(X)\to B(X)$ defined by

$$T(f)(i)=\min_{u\in U(i)}\left\{g(i,u)+\alpha\sum_{j=1}^{n}p_{ij}(u)f(j)\right\} \quad (16)$$

where $U(i)$ is a subset of a finite (control) set for each i, $\alpha\in(0,1)$ is the discount factor, and

$$p_{ij}(u)\geqslant 0, \quad \sum_{j=1}^{n}p_{ij}(u)=1, \qquad \forall i,u\in U(i).$$

This is the mapping associated with the usual discounted finite state Markovian decision problems (see, e.g., [3]–[5]). The error bounds (15) are known to be extremely useful in practice. The mapping (16) falls within our framework with $\alpha_1=\alpha_2=\alpha$. Relation (13) shows that our error bounds are sharper than those of (15) although the improvement is admittedly small.

The error bounds of Denardo [2, p. 171] are given by

$$\sup_{x\in X}|T^k(f)(x)-f^*(x)|\leqslant E_k, \qquad k=1,2,\cdots$$

$$E_k=\min\left[\alpha_2E_{k-1},\frac{\alpha_2}{1-\alpha_2}\sup_{x\in X}|T^k(f)(x)-T^{k-1}(f)(x)|\right]. \quad (17)$$

A comparison of the bounds above with the bounds of the proposition of the past section reveals that if $T^k(f)\leqslant T^{k-1}(f)$ then (17) provides a lower bound to f^* which is equivalent to ours and an upper bound which is less sharp than ours. Similarly if $T^{k-1}(f)\leqslant T^k(f)$ then (17) provides an upper bound to f^* equivalent to ours and a lower bound which is less sharp. When neither $T^k(f)\leqslant T^{k-1}(f)$ nor $T^{k-1}(f)\leqslant T^k(f)$ holds then either our lower or our upper bound is sharper. The improvement can be attributed to the fact that Denardo's bounds take into account only the right part of inequality (2).

Concerning the applicability of the result given it should be pointed out that it may be applied in sequential decision problems where the (effective) discount factor may depend on the current state and the control applied, such as for example in semi-Markov decision problems [3]. It may also be applied in Markovian decision problems with constant discount factors which are solved by successive approximation methods of the Gauss–Seidel type such as those described by Kushner [5] and Bertsekas [6].

References

[1] J. McQueen, "A modified dynamic programming method for Markovian decision problems," *J. Math. Anal. Appl.*, vol. 14, pp. 38–43, 1966.
[2] E. V. Denardo, "Contraction mappings in the theory underlying dynamic programming," *SIAM Rev.*, vol. 9, pp. 165–177, 1967.
[3] S. M. Ross, *Applied Probability Models with Optimization Applications.* San Francisco, CA: Holden-Day, 1970.
[4] C. Derman, *Finite State Markovian Decision Problems.* New York: Academic, 1970.
[5] H. Kushner, *Introduction to Stochastic Control.* New York: Holt, Rinehart and Winston, 1971.
[6] D. P. Bertsekas, *Dynamic Programming and Stochastic Control.* New York: Academic, 1976.

Suboptimality Bounds and Stability in the Control of Nonlinear Dynamic Systems

A. J. LAUB AND F. N. BAILEY

Abstract—Certain results paralleling those of Bailey and Ramapriyan [1] for linear dynamic systems are presented for nonlinear systems. For "weakly perturbed" systems certain suboptimal controls are shown to be stabilizing and upper and lower bounds on the suboptimality are investigated.

I. Introduction

Bailey and Ramapriyan [1] have presented some results concerning the derivation of upper and lower bounds on suboptimality when using a "naive" decentralized control for the control of an autonomous linear dynamic system with quadratic performance index on the time interval $[0,\infty)$. When the subsystems comprising this overall system are "weakly coupled" the naive decentralized control is shown to be stabilizing and the bounds can be used to derive an estimate of the performance degradation. Moreover, this estimate is calculable without explicit knowledge of the solution to the overall (coupled) regulator problem. Results similar in spirit to these are presented below for the control on $[t_0,\infty)$ of a nonautonomous nonlinear dynamic system with a general integral performance index. However, only a crude lower bound is given because of the unavailability of duality results analogous to those for linear systems as given by McClamroch [4]. Another type of lower bound is also computed in [2] but again this result makes heavy use of linearity. Thus, no convenient estimate of performance degradation can be given so far in the general nonlinear case.

While the upper bound results presented below are similar to those given by Rissanen [6] and Rissanen and Durbeck [7] in the finite time case (and generalized slightly by McClamroch [3]), a new and hopefully more intuitive derivation is offered here. Moreover, this discussion emphasizes the infinite time problem and stability questions but is also applicable to the finite time case considered by the above authors.

II. Problem Formulation

Consider a process modeled by the following dynamical system:

$$\dot{x}=f(t,x,u), \qquad x(t_0)=x_0 \quad (1)$$

where $x(t)\in\boldsymbol{R}^n$, $u(t)\in\boldsymbol{R}^m$ and $f:[t_0,\infty)\times\boldsymbol{R}^n\times\boldsymbol{R}^m\to\boldsymbol{R}^n$. Let $G=[t_0,+\infty)\times S$ where S is some subset of $\boldsymbol{R}^n$ containing the zero vector. It is assumed that there exists a unique (feedback) control law

$$u(t)=k_0(t,x(t)), \qquad t\geqslant t_0, \quad (2)$$

continuous in t, differentiable in x, which minimizes the performance criterion

$$J(t_0,x_0,u)=\int_{t_0}^{\infty}h(t,x,u)\,dt \quad (3)$$

subject to the constraint (1) for all initial values $x_0\in S$ where $h:[t_0,\infty)\times\boldsymbol{R}^n\times\boldsymbol{R}^m\to\boldsymbol{R}$ is a nonnegative definite function on G, continuously differentiable in all arguments. Denote solutions of (1) by $\phi(t;t_0,x_0,u)$. Assume that f is smooth enough (say, f and its first partials continuous) to ensure that for u as given in (2), solutions of (1) exist through any point $(t_1,x_1)\in G$, are unique, depend continuously on the initial data, and $(t,\phi(t;t_1,x_1,k_0))\in G$ for all $t\geqslant t_1$. Assume, furthermore, that $k_0(t,0)=0$ and $f(t,0,0)=0$ for all $t\geqslant t_0$ so $x=0$ is an isolated equilibrium state.

Manuscript received August 11, 1975.
A. J. Laub is with the Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada.
F. N. Bailey is with the Department of Electrical Engineering and Center for Comparative Studies in Technological Development and Social Change, University of Minnesota, Minneapolis, MN 55455.

functional equation for $J^*: S \to R$:

$$J^*(x) = \max_{u \in C} \mathop{E}_{w} \{g(x, u, w) + \alpha J^*[f(x, u, w)]\}$$

where $0 < \alpha < 1$ and g, f, S, C, and w satisfy continuity, compactness, and finiteness assumptions analogous to Assumptions A of Section 5.2. Let $S^1, S^2, \ldots, S^n$ be mutually disjoint sets with $S = \bigcup_{i=1}^n S^i$, select arbitrary points $x^i \in S^i$, $i = 1, \ldots, n$, and consider the discretized functional equation

$$\hat{J}^*(x) = \begin{cases} \max_{u \in C} \mathop{E}_{w} \{g(x, u, w) + \alpha \hat{J}^*[f(x, u, w)]\} \\ \qquad\qquad \text{if} \quad x = x^i, \quad i = 1, \ldots, n, \\ \hat{J}^*(x^i) \qquad \text{if} \quad x \in S^i, \quad i = 1, \ldots, n. \end{cases}$$

(a) Show that both equations have unique solutions J^* and $\hat{J}^*$ within the class of all bounded functions $J: S \to R$ and furthermore

$$\lim_{d_s \to 0} \sup_{x \in S} |J^*(x) - \hat{J}^*(x)| = 0,$$

where

$$d_s = \max_{i=1,\ldots,n} \sup_{x \in S^i} \|x - x^i\|.$$

(b) Provide a discretization procedure and prove a similar result under assumptions analogous to Assumption **B** of Section 5.2.

Hint: Use the results already proved in Section 5.2.

3. Let S be a set and $B(S)$ be the set of all bounded real-valued functions on S. Let $T: B(S) \to B(S)$ be a mapping with the following two properties:

(1) $T(J) \leqslant T(J')$ for all $J, J' \in B(S)$ with $J \leqslant J'$.
(2) For every scalar $r \neq 0$ and all $x \in S$

$$\alpha_1 \leqslant [T(J + re)(x) - T(J)(x)]/r \leqslant \alpha_2,$$

where α_1, α_2 are two scalars with $0 \leqslant \alpha_1 \leqslant \alpha_2 < 1$.

(a) Show that T is a contraction mapping on $B(S)$ and hence for every $J \in B(S)$ we have

$$\lim_{k \to \infty} T^k(J)(x) = J^*(x) \qquad \forall x \in S,$$

where J^* is the unique fixed point of T in $B(S)$.

(b) Show that for all $J \in B(S)$, $x \in S$, and $k = 1, 2, \ldots,$

$$T^k(J)(x) + c_k \leqslant T^{k+1}(J)(x) + c_{k+1} \leqslant J^*(x)$$
$$\leqslant T^{k+1}(J)(x) + \bar{c}_{k+1} \leqslant T^k(J)(x) + \bar{c}_k,$$

where for all k

$$c_k = \min\left\{\frac{\alpha_1}{1-\alpha_1}\inf_{x\in S}[T^k(J)(x) - T^{k-1}(J)(x)],\right.$$

$$\left.\frac{\alpha_2}{1-\alpha_2}\inf_{x\in S}[T^k(J)(x) - T^{k-1}(J)(x)]\right\},$$

$$\bar{c}_k = \max\left\{\frac{\alpha_1}{1-\alpha_1}\sup_{x\in S}[T^k(J)(x) - T^{k-1}(J)(x)],\right.$$

$$\left.\frac{\alpha_2}{1-\alpha_2}\sup_{x\in S}[T^k(J)(x) - T^{k-1}(J)(x)]\right\}.$$

A geometric interpretation of these relations for the case where S consists of a single element is provided in Fig. 6.3.

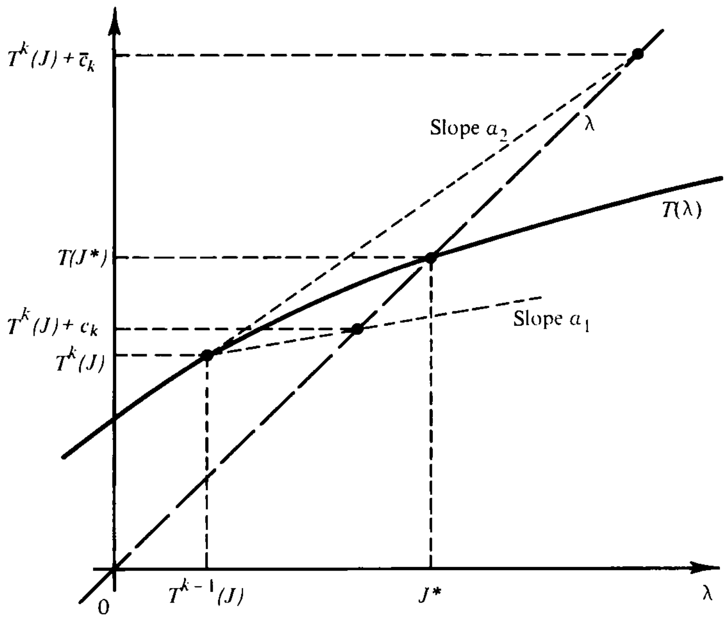


FIGURE 6.3

(c) Show that the mapping F defined by (38)–(41) satisfies

$$\alpha^n \leqslant [F(J+re)(x) - F(J)(x)]/r \leqslant \alpha,$$

where n is the number of elements in S.

Hint: Use a line of argument similar to the one of Section 6.2.