STRUCTURED PARTITIONING PROBLEMS

S. ANILY

University of British Columbia, Vancouver, Canada, and Tel Aviv University, Tel Aviv, Israel

A. FEDERGRUEN

Columbia University, New York, New York

(Received March 1987; revisions received September 1988, November 1989; accepted December 1989)

In many important combinatorial optimization problems, such as bin packing, allocating customer classes to queueing facilities, vehicle routing, multi-item inventory replenishment and combined routing/inventory control, an optimal partition into groups needs to be determined for a finite collection of objects; each is characterized by a single attribute. The cost is often separable in the groups and the group cost often depends on the cardinality and some aggregate measure of the attributes, such as the sum or the maximum element. An upper bound (*capacity*) may be specified for the cardinality of each group and the number of groups in the partition may either be fixed or variable. The objects are indexed in nondecreasing order of their attribute values and within a given partition the groups are indexed in nondecreasing order of their cardinalities. We identify easily verifiable analytical properties of the group cost function under which it is shown that an optimal partition exists of one of three increasingly special structures, thus allowing for increasingly simple solution methods. We give examples of all the above listed types of planning problems, and apply our results for the identification of efficient solution methods (wherever possible).

In many important combinatorial optimization problems, an optimal partition into groups needs to be determined for a finite collection of objects, each of which is characterized by a single attribute. The cost of a given partition is often separable in the groups and the group cost often depends on the cardinality and some aggregate measure of the attribute values in the group, e.g., the sum or the maximum element. Upper bounds (*capacities*) may be specified for the cardinalities of the groups and the number of groups in the partition may either be fixed or variable.

It is well known that this class of partitioning problem is NP-complete for *general* choices of the group cost function. See, for example, Chakravarty, Orlin and Rothblum (1982) who base this observation on the following example.

Example 1. (This is one of the first problems to be identified as NP-complete; see Karp (1972).) Given N integers $r_1, \ldots, r_N$, verify whether a subset $S \subset \{1, \ldots, N\}$ exists for which $\sum_{i \in S} r_i = R/2$ where $R = \sum_{i=1}^{N} r_j$. An equivalent formulation of this problem is to verify whether a *partition* of the collection $\{1, \ldots, N\}$ into *two* sets X_1, X_2 exists with

$$f\left(\sum_{i \in X_1} r_i\right) + f\left(\sum_{i \in X_2} r_i\right) = 0$$

where the group cost function $f(\cdot)$ is defined by $f(x) = (x - R/2)^2$. (This example shows, in addition, that the above defined class of partitioning problems remains NP-complete when the number of groups is restricted to be equal to *two* and when the group cost is *independent* of the cardinality of the group.)

Examples 10 and 11 (discussed in Section 7) describe bin packing problems (that, e.g., arise in the allocation of records on computer auxiliary storage devices) and a problem of allocating customer classes in general queueing systems with multiple service pools. Both represent special cases of the defined class of partitioning problems for which no efficient exact solution method appears to exist. On the other hand, Chakravarty, Orlin and Rothblum (1982) treat the problem of determining optimal groupings of items in a multicommodity inventory system with joint replenishment costs, and derive an efficient solution method for this special case of the class of partitioning problems (Example 12). Examples 13, 14 and 15 cover partitioning problems (of the above type) that arise in vehicle routing, multi-item, two-stage inventory/production and combined vehicle routing/inventory models, and for these we derive even faster solution methods.

The objective of this paper is to identify a nested set of simple conditions for the group cost function under which an optimal partition may be determined by increasingly simple, efficient algorithms that exploit increasingly stronger structural properties of this partition.

Thus, let $X = \{x_1, \ldots, x_N\}$ be a collection of objects. Each object x_i is characterized by a single attribute r_i and the objects are numbered in ascending

Subject classifications: Queues: structured set partitioning problems; algorithms.

0030-364X/91/3901-0129 $01.25

order of their attribute values, i.e., $r_1 \leqq r_2 \leqq \ldots \leqq r_N$. Let L denote the (fixed or variable) number of groups in the desired partition. Within a given partition we number the groups in nondecreasing order of their cardinalities. Let M_l^* denote the capacity of the lth group ($l = 1, \ldots, L$). (In view of our numbering convention, $M_1^* \leqq M_2^* \leqq \ldots \leqq M_L^*$.) In a given partition, let $l(i)$ denote the index of the group to which the ith object is assigned; we refer to the index function $l(\cdot)$ as the *group index function*.

A partition is called *consecutive* if it consists of consecutive sets; i.e., sets in which the indices of the elements are consecutive integers. For example, $\chi = \{X_1, X_2\} = \{\{4,5\}; \{1,2,3\}\}$ is a consecutive partition of $X = \{1, \ldots, 5\}$. A partition is called *monotone* if the group index function is nondecreasing. Note that a monotone partition is consecutive; the partition χ fails to be monotone, but $\chi^* = \{X_1^*, X_2^*\} = \{\{1,2\}; \{3,4,5\}\}$ is.

In this paper, we specifically identify easily verifiable conditions with respect to the group cost function under which the partitioning problem is:

i. optimized by a *consecutive* partition;
ii. optimized by a *monotone* partition;
iii. *extremal*, i.e., a monotone optimal partition exists and the cost of any monotone partition $\chi = \{X_1, \ldots, X_L\}$ does not increase by shifting the highest indexed object in any of its groups to the next group, i.e., by transferring the highest indexed element of some set X_l to X_{l+1}, $1 \leqq l \leqq L$.

We also show how increasingly simple and efficient algorithms may be employed when the partitioning problem satisfies conditions i–iii, respectively.

Chakravarty, Orlin and Rothblum (1982) consider uncapacitated partitioning problems of the above defined type, in which the group cost depends on the attribute values in the group through their *sum*. It is shown that an optimal consecutive partition exists if the group cost function is *concave* in the attribute value sum, and may thus be determined by computing the shortest path in an acyclic network with N nodes. As pointed out, their paper was motivated by a multifacility (or multi-item) inventory replenishment problem with joint setup costs. Chakravarty, Orlin and Rothblum (1985) consider a generalization where each object is characterized by *two* attributes and where the group cost function is concave in the sum of the values of each of these two attributes, but otherwise is independent of the number of elements in the group. It is shown that in the absence of constraints on the groups' cardinalities an optimal partition may be determined by a similar shortest path calculation. The generalization allows for the treatment of more general, joint setup cost structures in the above mentioned multi-item inventory replenishment problem.

The latter paper also discusses the case where the cost of a partition is a nonseparable function of the sums of the values of the two attributes in each group. Barnes, Hoffman and Rothblum (1989) consider further generalizations where the objects are characterized by an arbitrary number (p) of attributes. Each object is thus characterized by a point in the p-dimensional attribute space. The authors show that an optimal partition exists whose groups have (pairwise) disjoint conic hulls in the attribute space. A weaker property holds if the cardinality of each group is prespecified; namely, there exists an optimal partition whose groups have disjoint (pairwise) convex hulls. These characterizations do not—as of yet—result in general efficient solution methods except for the separable, two attribute case identified by Chakravarty, Orlin and Rothblum (1985).

There are, of course, many partitioning problems in which the objects are characterized by one or a limited number of attributes, but in which the group cost depends on the attribute values in the group according to (aggregate) measures that are different from the attribute sum or maximum.

Examples arise in the areas of clustering (see, e.g., the excellent survey text of Späth (1985); see also Hwang 1981 and Hwang, Sun and Yao 1985), graph partitioning, layout of circuits on computer boards and computer program segmentation (Danath and Hoffman 1973, Barnes and Hoffman 1984, Barnes 1982, 1985).

We conclude this section with an outline of the paper. In Section 1, we introduce some notation. Next, in Sections 2–4, we show how properties i–iii allow for increasingly simpler solution methods. In Sections 5 and 6 we obtain sufficient conditions for properties i–iii to apply to partitioning problems in which the group cost depends, respectively, on the sum and the maximum element of the attribute values in the group. In Section 7 we apply our results to several example models. Section 8 completes the paper with a discussion of related partitioning problems and properties.

1. NOTATION

For any partition $\chi = \{X_1, \ldots, X_L\}$ of X, let $m_l = |X_l|$, $l = 1, \ldots, L$. A partition $\chi = \{X_1, \ldots, X_L\}$ is *feasible* if and only if the number of elements in X_l does not exceed the capacity bound M_l^*, $l = 1, \ldots, L$. If $|X_l| < M_l^*$, the set X_l is said to have *slack*. Let $M^* = M_L^* = \max_l M_l^*$. As pointed out in the Introduction, L may be treated as a given parameter or as a variable.

We assume that the cost of a partition $\chi =$

$\{X_1, \ldots, X_L\}$ is given by a *separable* functional of the form

$$U^1(\chi) = \sum_{l=1}^{L} f\Big(\sum_{j \in X_l} r_j / m_l, m_l\Big)$$

or

$$U^2(\chi) = \sum_{l=1}^{L} f\Big(\max_{j \in X_l} r_j, m_l\Big)$$

where $f: \mathbb{R}_+^2 \to \mathbb{R}$ is a general real-valued function.

Next we define the two partitioning problems.

Problem $\mathbf{P^1}$

$$V^1(X) = \min\{U^1(\chi): \chi = \{X_1, \ldots, X_L\} \text{ partitions } X \text{ and } m_l \leqq M_l^*, l = 1, \ldots, L\}. \quad (1a)$$

Problem $\mathbf{P^2}$

$$V^2(X) = \min\{U^2(\chi): \chi = \{X_1, \ldots, X_L\} \text{ partitions } X \text{ and } m_l \leqq M_l^*, l = 1, \ldots, L\}. \quad (1b)$$

2. WHEN AN OPTIMAL CONSECUTIVE PARTITION EXISTS

Assume that an optimal consecutive partition exists for $\mathbf{P^1}$ or $\mathbf{P^2}$. We need to distinguish between four cases:

Case 1. L is variable and $M_l^* = M^*$, $l = 1, \ldots, L$. In this case an optimal partition may be determined by computing a shortest path in an acyclic network. Let $F^i(j) = V^i(\{x_{j+1}, \ldots, x_N\})$, $i = 1, 2$ and for any $Y \subset X$ define $g_1(Y) = \sum_{j \in Y} r_j / |Y|$ and $g_2(Y) = \max_{j \in Y} r_j$. Note that $F^i(0) = V^i(X)$, $i = 1, 2$.

Clearly, $F^i(0)$ and the corresponding optimal partition χ^i may be determined from the dynamic programming recursion

$$F^i(j) = \min_{1 \leqq j' - j \leqq M^*} \{f(g_i(\{x_{j+1}, \ldots, x_{j'}\}), j' - j) + F^i(j')\} \quad (2)$$

with $F^i(N) = 0$.

In case M^* is a constant independent of N it is easily observed that NM^* operations are required to solve the recursion and we conclude that the complexity of the algorithm is *linear* in N. If $M^* = N$ (the uncapacitated case), the complexity is $0(N^2)$.

Case Ib. L is a given constant and $M_l^* = M^*$, $l = 1, \ldots, L$. Let $F^i(j, l) = \min\{U^i(\chi): \chi = \{X_{l+1}, \ldots, X_L\}$ partitions $\{x_{j+1}, \ldots, x_N\}$ and $|X_k| \leqq M^*$, $k = l+1, \ldots, L\}$. Note that $F^i(0, 0) = V^i(X)$, $(i = 1, 2)$, which together with the optimal partition χ^i, may be determined via the dynamic programming recursion

$$F^i(j, l) = \min_{1 \leqq j' - j \leqq M^*} \{f(g_i(\{x_{j+1}, \ldots, x_{j'}\}), j' - j) + F^i(j', l+1)\} \quad l \leqq j \text{ and } j = 0, \ldots, N \quad (3)$$

with $F^i(N, l) = \infty$, $l = 0, \ldots, L-1$; $F^i(j, L) = \infty$ for $j < N$ and $F^i(N, L) = 0$.

In case M^* is a finite constant independent of N the recursion is solved with Dijkstra's algorithm in $0(NLM^*)$ operations. If $M^* = N$ (the uncapacitated case), the complexity is $0(N^2L)$.

Case Ic. General capacities and L is variable.

Case Id. General capacities and L is constant.

In the case of general, i.e., *nonidentical* capacities, it does not appear that the restriction to consecutive partitions, by itself, allows for efficient solution methods. For example, in a straightforward dynamic programming formulation, one would have to keep track in the *state description* of the capacities that are available at any stage. In general, this results in an exponential number of states. Only if C, the number of *distinct* capacity levels, is small (e.g., $C = 2$ or $C = 3$) does the straightforward dynamic programming approach result in an algorithm whose complexity bound is a polynomial of reasonable degree $(C + 1)$.

3. WHEN AN OPTIMAL MONOTONE PARTITION EXISTS

The monotonicity property of an optimal partition may be exploited to derive an efficient solution method for the most general case with nonidentical capacities. For the case of identical capacities, the monotonicity property may be exploited to simplify the dynamic programming recursions.

Case Ia. (Identical capacities, L is variable)

For any $j = 0, \ldots, N-1$ consider a monotone optimal partition of $\{x_{j+1}, \ldots, x_N\}$ with *maximal* cardinality for the first set (i.e., the set to which x_{j+1} is assigned) and let $m(j)$ denote this cardinality. We define $m(N) = M^*$. If optimal monotone partitions exist for all sets $\{x_j, \ldots, x_N\}$ $(j = 1, \ldots, N)$ it is possible to implement the dynamic programming recursion as follows.

DP for Monotone Optimal Partitions (DPMOP)

Step 0. $j := N$; $F^i(N) := 0$; $m[N] := M^*$.

Step 1. While $j > 0$ *do*
begin
$j' := j$; $j := j-1$; $F^i(j) := f(g_i(\{x_{j+1}\}), 1) + F^i(j+1)$; $m(j) := 1$,
while $j' \leqq N$ *do*
begin
$j' := j'+1$; *if* $j'-j > m[j']$ *then go to end do*;
$y := f(g_i(\{x_{j+1}, \ldots, x_{j'}\}), j'-j) + F^i(j')$;
if $y \leqq F^i(j)$ *then begin* $F^i(j) := y$; $m(j) := j'-j$ *end*;
end do
end.

It is easy to verify that the DPMOP algorithm generates a monotone optimal partition with $V^i(X) = F^i(0)$.

With the help of the simple test

$$j' - j \leqq m(j') \tag{4}$$

the DPMOP algorithm thus allows for the elimination of a significant number of values of j' when evaluating the minimum in (2). However, since the inequality (4) needs to be checked for all $j' = j+1, \ldots, \min(j+M^*, N)$, no reduction in the above mentioned worst case complexity bounds can be achieved. The DPMOP algorithm ensures, on the other hand, in contrast to standard implementations of the dynamic programming recursion (2), that a monotone optimal partition is generated.

Case Ib. (Identical capacities, L is fixed) Assume that for any fixed value of L and l, $l = 0, \ldots, L-1$ and any $j = 0, \ldots, N-(L-l)$ an optimal partition of $\{x_{j+1}, \ldots, x_N\}$ into exactly $(L-l)$ sets exists which is monotone. Among all such optimal monotone partitions, consider one with maximal cardinality of the first set (i.e, the set to which x_{j+1} is assigned) and let $m(j, l)$ denote this cardinality.

Assume that an optimal (monotone) partition of X into $L = L^*$ sets is required. A straightforward adaptation of the DPMOP algorithm results in an efficient implementation of the dynamic programming recursion (3), which is, in addition, guaranteed to generate monotone (optimal) partitions: Evaluate the vectors $\{F^i(\cdot, l)\}$ and $\{m(\cdot, l)\}$ recursively for $l = 0, 1, \ldots, L^*-1$. For a given value of l, the vectors $\{F^i(\cdot, l)\}$ and $\{m(\cdot, l)\}$ may be computed by an obvious modification of the DPMOP procedure, replacing $F^i(j)$ by $F^i(j, l)$, $m(j)$ by $m(j, l)$, $F^i(j')$ by $F^i(j', l+1)$, $m(j')$ by $m(j', l+1)$ and setting $F^i(j', 0) = \infty$ if $1 \leqq j' \leqq N$, $F^i(N, L) = 0$, $F^i(N, l) = \infty$ if $l < L$ and $m(N, L) = M^*$.

Case Ic. (General nonidentical capacities and L is variable) Assume that for every integer $l = 0, \ldots, N$ and all $j = l, \ldots, N-l$ the partitioning problem

$$F^i(j, l) \overset{\text{def}}{=} \min\{U^i(\chi): \chi = \{X_{l+1}, \ldots, X_L\} \text{ partitions } \{x_{j+1}, \ldots, x_N\} \text{ and } |X_k| \leqq M_k^*,\ k = l+1, \ldots, L\}$$

is optimized by a monotone partition. Among all such monotone optimal partitions, consider one with maximal cardinality for the first set (X_{l+1}) and let $m(j, l)$ denote this cardinality. As before, $F^i(0, 0) = V^i(X)$ $(i = 1, 2)$, which together with the optimal partition χ^i, may be determined via the recursion

$$F^i(j, l) = \min_{1 \leqq j'-j \leqq \min\{M_{l+1}^*,\, m(j', l+1)\}} \{f(g_i(\{x_{j+1}, \ldots, x_{j'}\}, j'-j) + F^i(j', l+1)\}$$
$$l \leqq j \quad \text{and} \quad j = 0, \ldots, N$$

with

$$F^i(N, l) = 0, \quad l = 1, \ldots, N,$$
$$m(N, l) = M_{l+1}^* \quad \text{if } 1 \leqq l \leqq N$$

and $F^i(j, 0) = \infty$ if $1 \leqq j \leqq N$. The vectors $\{F^i(\cdot, l)\}$ and $\{m(\cdot, l)\}$ are determined recursively for $l = N, N-1, \ldots, 1$; for any given value of l, the vectors $\{F^i(\cdot, l)\}$ and $\{m(\cdot, l)\}$ are obtained by a straightforward adaptation of the DPMOP algorithm. In case M^* is a finite constant independent of N, the recursion is solved in $0(N^2M^*)$ operations.

Case Id. (General capacities: L is a given constant, independent of N) Define $F^i(j, l)$ as above for $l = 0, \ldots, L-1$ and $j = 0, \ldots, L-l$. The dynamic programming recursion for Ic may be applied in this case as well, restricting the values of l to the set $\{0, \ldots, L-1\}$ and with boundary conditions $F^i(j, L) = \infty$ for $j < N$, $F^i(N, l) = \infty$ for $l < L$, $m(N, L) = M_L^*$ and $F^i(N, L) = 0$. Assuming once again that M^* is a constant that is independent of N, the recursion is solved in $0(NLM^*)$ operations.

4. WHEN THE PARTITIONING PROBLEM IS EXTREMAL

In this section, we assume that for all $n = 1, \ldots, N$ and any combination of capacities $\{M_l^*: l = 1, \ldots, L\}$, the partitioning problem $\min\{U^i(\chi): \chi = \{X_1, \ldots, X_L\}$ partitions X and $m_l \leqq M_l^*$, $l = 1, \ldots, L\}$ is extremal.

Examples 14 and 15 in Section 7 discuss several logistical planning models in which partitioning problems need to be solved, which can be shown to be

extremal. Sections 5 and 6 exhibit several easily verifiable analytical properties of the group function $f(\cdot)$ under which extremality is guaranteed. In this section, we demonstrate that a partition exists that is optimal under *any* group cost function $f(\cdot)$ under which the partitioning problem is extremal, and this partition may be determined by the Extremal Partitioning Algorithm, which is considerably simpler than the dynamic programming algorithms (2) and (3). We first consider Cases Ia, b and d. The insensitivity of this optimal partition with respect to the specific choice of the group cost function $f(\cdot)$ plays an essential role in the analyses of the models discussed in Examples 14 and 15.

Extremal Partitioning Algorithm (EPA)

Step 0. Initialize $n := N$;

$$l := \begin{cases} L & \text{if } L \text{ a given constant} \\ 1 & \text{if } L \text{ variable} \end{cases}$$

Step 1. If $1 < M_l^* < n - l + 1$ *then*
begin the set $\{n - M_l^* + 1, \ldots, n\}$ is the next set to be added to the partition; $n := n - M_l^*$; if $l > 1$ then $l := l - 1$; repeat Step 1;
end
else
begin if $M_l^* = 1$ add the sets $\{1\}, \ldots, \{n\}$ to the partition; exit; if $M_l^* \geqq 2$ and $l > 1$, add the sets $\{1\}, \ldots, \{l-1\}, \{l, \ldots, n\}$ to the partition; otherwise add the single set $\{1, \ldots, n\}$ to the partition; exit;
end.

The EPA thus generates a partition that consists of a possibly empty collection of singletons followed by at most one set with slack and full sets thereafter. The EPA partitions the elements of X in descending order of their indices. At the beginning of each iteration of Step 1, n denotes the number of objects not yet partitioned; likewise, in case L is a given number, l denotes the number of sets that need to be added to the partition. (If L is variable, $l = 1$ throughout the algorithm. Note also that if the major test, $1 < M_l^* < n - l + 1$, in Step 1 is satisfied, $n > M_l^*$ so that additional objects remain to be partitioned after the current execution of Step 1.)

Note that the partition generated by the EPA depends merely on the value of L and the capacities M_l^* $(l = 1, \ldots, L)$; this partition is, in particular, *independent* of the function f, a robustness property suggested by the definition of extremality.

To assess the complexity of the EPA, note that as long as the test in Step 1 is satisfied, $M_l^* \geqq 2$. Thus, Step 1 is repeated at most $\lceil N/2 \rceil$ times and in each iteration at most 4 operations (comparisons or subtractions) are performed. The complexity is thus bounded by $4 \mid N/2 \mid$ while only the integers n, l and $\{M_1^*, \ldots, M_L^*\}$ need to be kept in memory. Alternatively, for the case where L is fixed the complexity is bounded by $4 \min \{L; \mid N/2 \mid\}$, which is 0(1) as $N \to \infty$! These bounds compare very favorably with the bounds for the dynamic programming recursion.

Theorem 1 proves that the EPA generates an optimal partition. We first need the following lemma.

Lemma 0. *Assume that the partitioning problem* $\mathbf{P}^i$ $(i = 1, 2)$ *is extremal. There exists a monotone optimal partition with the property*

if for some $l = 1, \ldots, L$ X_l *has slack* $(\mid X_l \mid < M_l^*)$

either $l = 1$ *or* X_{l-1} *is a singleton.* (5)

Proof. Assume to the contrary that in each monotone optimal partition (5) is violated for some $l = 1, \ldots, L$. Let χ be a monotone optimal partition lexicographically: i) minimizing the index, and ii) maximizing the cardinality of the highest indexed set violating (5). Let $l \geqq 2$ be the index of the largest indexed set in χ violating (5). Let $\chi' = \{X_1, \ldots, X_{l-2}, X'_{l-1}, X'_l, X_{l+1}, \ldots, X_L\}$ be the partition obtained by transferring the highest indexed element of X_{l-1} to X_l, which is feasible since $1 < \mid X_{l-1} \mid \leqq \mid X_l \mid < M_l^*$. In view of the extremality properties of the problem, $U^i(\chi') \leqq U^i(\chi)$; hence, χ' is optimal. Decompose $X = X^{(1)} \cup X^{(2)}$ and χ' as $\chi' = (\chi'^{(1)}, \chi'^{(2)})$ where $\chi'^{(1)} = \{X_1, \ldots, X'_{l-1}\}$, $\chi'^{(2)} = \{X'_l, \ldots, X_L\}$ and $\chi'^{(1)}$ partitions $X^{(1)}$. Note that $\chi'^{(2)}$ is monotone since $\mid X'_l \mid \leqq M_l^* \leqq M_{l+1}^* = \mid X_{l+1} \mid$. In view of the optimality of χ and χ', $\chi'^{(1)}$ is an optimal partition for the partitioning problem of $X^{(1)}$ using $l - 1$ sets with $\hat{M}_r = \mid X_r \mid$, $r = 1, \ldots, l-2$ and $\hat{M}_{l-1} = \mid X'_{l-1} \mid$. Next, consider the relaxation of this partitioning problem where the capacity of the $l-1$st set is increased to $\mid X_{l-1} \mid = \hat{M}_{l-1} + 1$. A monotone partition $\hat{\chi}^{(1)}$ of $X^{(1)}$ can be found that optimizes this relaxed partitioning problem of $X^{(1)}$, hence, with $U^{(i)}(\hat{\chi}^{(1)}) \leqq U^i(\chi'^{(1)})$. Hence, let $\hat{\chi} = \{\hat{\chi}^{(1)}, \chi'^{(2)}\}$; note that $\hat{\chi}$ is monotone and $U^i(\hat{\chi}) = U^i(\hat{\chi}^{(1)}) + U^i(\chi'^{(2)}) \leqq U^i(\chi'^{(1)}) + U^i(\chi'^{(2)}) = U^i(\chi') = U^i(\chi)$, so $\hat{\chi}$ is a monotone optimal partition for the original partitioning problem. Note that if in this partition some set violates (5), the highest indexed such set has an index $\leqq l$ while the lth set in this partition (X'_l) has a larger cardinality than the lth set of the partition χ, thus contradicting the definition of χ.

Theorem 1. *Assume that the partitioning problem* $\mathbf{P}^i$, $(i = 1, 2)$, *is extremal.*

a. *If L is variable and $M_l^* = M^*$, $(l = 1, \ldots, L)$, the EPA generates a monotone optimal partition using the lowest possible number $\lceil N/M^* \rceil$ of sets. In this partition, all sets, with the possible exception of the first one, are full. $X_1 = \{x_1, \ldots, x_n\}$, $X_l = \{x_{n+(l-2)M^*+1}, \ldots, x_{n+(l-1)M^*}\}$, $l = 2, \ldots, \lceil N/M^* \rceil$ where $n = M^*$ if N is a multiple of M^* and $n = N - \lceil N/M^* \rceil M^*$ otherwise.*

b. *If L is fixed, the EPA generates the unique monotone partition χ^* which satisfies (5) and this partition is optimal.*

Proof. a. One easily verifies that the EPA generates the specific partition χ, that this partition is monotone and that only the first set may have slack. Assume to the contrary that χ is *not* optimal. Let χ' be a monotone optimal partition which lexicographically: i) minimizes the index, and ii) maximizes the cardinality of the highest indexed set with slack, among all monotone optimal partitions that satisfy (5). Such a partition exists in view of $\mathbf{P}^i$ $(i = 1, 2)$ being extremal. (If an optimal monotone partition exists in which all sets are full, then this partition is equal to χ, thus contradicting the nonoptimality of χ.) Let l be the index of the highest indexed set in χ' with slack. If $l = 1$, $\chi' = \chi$, contradicting the nonoptimality of χ. Thus, $l \geqq 2$ and X'_{l-1} is a singleton, by Lemma 0.

In view of the extremality of the partitioning problem, the partition χ'' obtained by merging X'_{l-1} and X'_l into a single set, has $U^i(\chi'') \leqq U^i(\chi')$, hence, χ'' is also optimal, which contradicts the definition of χ'.

b. By induction with respect to $l = L, L-1, \ldots, 0$. Assume that the last $(L-l)$ sets in any monotone partition that satisfy (5) coincide with the last $(L-l)$ sets of χ^*. The claim is clearly true for $l = L$. Let $n = \max\{i: x_i \in X_l\}$, which in view of the induction assumption, is identical for all monotone partitions that satisfy (5). If $M_l^* \leqq n - l + 1$, then the lth set of any monotone partition satisfying (5) must be full, so that the last $(L-l+1)$ sets are identical to the corresponding sets in χ^*. Similarly, if $M_l^* > n - l + 1$ the lth set of any monotone partition that satisfies (5) must have cardinality $(n-l+1)$ to allow for a feasible choice of $X_1, \ldots, X_{l-1}$. We conclude that in this case as well, the last $(L-l+1)$ sets of the partition coincide with those of χ^*. Thus, there exists a unique monotone partition that satisfies (5) and in view of Lemma 0 this partition is optimal.

The remaining Case Ic, where L is variable and the capacities are nonidentical, is treated by solving the problem repeatedly with a fixed value of L for $L = 1, \ldots, N$. The resulting algorithm is thus completed in at most $3N^2/2$ operations which, again, compares favorably with the complexity bound obtained for the dynamic programming recursion (3).

The following procedure may be used as an alternative to EPA for the case where L is fixed. (Assume without loss of generality that $L \leqq N \leqq \sum_{l=1}^{L} M_l^*$; if N is outside this range, no feasible solution exists.)

Let $C(l) = l - 1 + \sum_{k=l}^{L} M_k^*$ $(l = 1, \ldots, L)$ and $C(L+1) = 0$. Find the unique integer l^* $(1 \leqq l^* \leqq L)$ with $C(l^*+1) < N \leqq C(l^*)$. ($l^*$ exists because $C(L+1) = 0$ and $C(1) = \sum_{k=1}^{L} M_k^*$.) Let χ^* be the (unique) consecutive partition with $m_l = 1$, $l < l^*$; $m_l = M_l^*$, $l > l^*$ and $m_{l^*} = M_{l^*}^* - (C(l^*) - N)$. The value of l^* (and the optimal partition) may be efficiently computed by the following algorithm.

Extremal Partitioning Algorithm II (EPA II)

Step 0. $C := L - 1 + M_L^*$; $l := L$

Step 1. While $C < N$ *do*
begin $l := l - 1$; $C := C + M_l^* - 1$
end

Step 2. $l^* := l$; $m := M_{l^*}^* - (C - N)$; an optimal partition consists of the sets $\{1\}, \ldots, \{l^* - 1\}, \{l^*, \ldots, l^* + m - 1\}$ and $(L - l^*)$ *full-size* sets thereafter.

Step 1 is repeated at most L times, each time requiring three elementary operations. In Step 2 we obtain the boundary indices of the groups in the optimal partition with $L+1$ additions. The overall complexity is thus bounded by $(4L+1)$ elementary operations, which is quite similar to that of the original version. (The latter is slightly more efficient if $L > N/2$.) Alternatively, l^* may be found by bisection on the interval $[1, L]$ after computing all $\{C(l): l = 1, \ldots, L\}$ recursively. This alternative implementation of EPA II requires $3L + 4\log_2 L$ elementary operations.

For the case where all capacities are identical, i.e., $M_1^* = \ldots = M_L^* = M^*$, l^* is the largest integer l to satisfy the inequality $C(l) = l - 1 + (L - l + 1)M^* \geqq N$, that is

$$l^* = \left\lfloor \frac{((L+1)M^* - N - 1)}{(M^* - 1)} \right\rfloor.$$

The results of Sections 2–4 are summarized in Table I.

5. SUFFICIENT CONDITIONS FOR $\mathbf{P}^1$

In this section, we derive sufficient conditions with respect to f under which $\mathbf{P}^1$: i) has a consecutive optimal partition, ii) has a monotone optimal partition,

Table I
The Complexity of Computing the Optimal Partition[a]

Problem Type	No. of Sets	Capacity Constraints		No. of Elementary Operations	No. of Evaluations of the Cost Function
Optimal consecutive partition exists	Given	Yes	Identical	$4NLM^*$	NM^*
			Nonidentical	No efficient algorithm is known	
		No		$N^2L/2$	$N^2/2$
	Variable	Yes	Identical	NM^*	NM^*
			Nonidentical	No efficient algorithm is known	
		No		N^2	$N^2/2$
Optimal monotone partition exists	Given	Yes	Identical	$4NLM^*$	NM^*
			Nonidentical	$4NLM^*$	NM^*
		No		$N^2L/2$	$N^2/2$
	Variable	Yes	Identical	NM^*	NM^*
			Nonidentical	$2N^2M^*$	NM^*
		No		N^2	$N^2/2$
Problem is extremal	Given	Yes	Identical	$4\min\{L, \lceil N/M^* \rceil\}$	0
			Nonidentical	$4\min\{L, \lceil N/2 \rceil\}$	0
		No			0
	Variable	Yes	Identical	$4\lceil N/M^* \rceil$	0
			Nonidentical	$4\lceil N/2 \rceil$	0
		No		1	0

[a]The complexity counts given in Table I assume the points are numbered in ascending order of their attribute values.

and iii) is extremal. Let

$$F_0 = \{\phi\colon \phi(\theta, m) \text{ is nondecreasing in } \theta\}$$

and

$$F_1 = \{\phi \in F_0\colon \phi(\theta, m) \text{ is concave in } \theta\}.$$

The next result is based on a simple modification of the lemma in Chakravarty, Orlin and Rothblum.

Lemma 1. *Let $f \in F_1$. There exists an optimal partition for $\mathbf{P}^1$ which is consecutive.*

Proof. For each subset S of X, let the span $\text{sp}(S) = \max\{i\colon x_i \in S\} - \min\{i\colon x_i \in S\}$. Let $\chi^* = \{X_1^*, \ldots, X_L^*\}$ be an optimal partition that minimizes $\sum_{l=1}^{L} \text{sp}(X_l)$ among all such partitions. Assume that χ^* is not consecutive. Then there exist two sets $X_{l_1}^*$ and $X_{l_2}^*$ with $\min\{i\colon x_i \in X_{l_1}\} < \min\{i\colon x_i \in X_{l_2}\} < \max\{i\colon i \in X_{l_1}\}$. Employing the interchange argument in Chakravarty, Orlin and Rothblum we may repartition $X_{l_1}^* \cup X_{l_2}^*$ into two sets with *cardinalities* $|X_{l_1}^*|$ and $|X_{l_2}^*|$, such that the resulting partition of X has an optimal cost value and a *lower* value for $\sum_l \text{sp}(X_l)$, thus contradicting the definition of χ^*.

It is worth mentioning that Lemma 1 remains valid when the function $f \notin F_0$, i.e., fails to be monotone in θ.

It is often convenient to verify sufficient conditions for the existence of a monotone partition, or for the extremality of the problem, in terms of the function $h\colon R_+^2 \to R$, defined by $h(\Theta, m) = f(\Theta/m, m)$. We also define

$$R^* = \sum_{i=1}^{N} r_i, \quad R_{m,m+l} = \sum_{i=m}^{m+l} r_i$$

$\bar{R}_{m,m+l} = R_{m,m+l}/(l+1)$ for all $m = 1, \ldots, N$ and all integers l, with the convention that $R_{m,m+l} = \bar{R}_{m,m+l} = 0$ for l negative.

Definition 1. A function $\phi\colon R_+^2 \to R$ has *antitone* differences if $\phi(\theta + \Delta, m) - \phi(\theta, m)$ is nonincreasing in m for $\Delta > 0$.

The condition in Definition 1 does not distinguish between the first and second variables because $\phi(\theta + \Delta, m) - \phi(\theta, m) \geq \phi(\theta + \Delta, m + l) - \phi(\theta, m + l)$ if and only if $\phi(\theta, m + l) - \phi(\theta, m) \geq \phi(\theta + \Delta, m + l) - \phi(\theta + \Delta, m)$ for all $\Delta, l > 0$.

Remark 1. Let ϕ: $\mathbb{R}_+^2 \to \mathbb{R}$ be twice differentiable with ϕ_1 and ϕ_2 as its partial derivatives. Then ϕ has antitone differences if and only if (iff) $\partial^2\phi/\partial\theta\,\partial m \leqq 0$ for all $\theta > 0$, $m > 0$. (Observe that $\phi(\theta+\Delta, m) - \phi(\theta, m) - (\phi(\theta+\Delta, m+l) - \phi(\theta, m+l)) \geqq 0$ iff $\int_0^\Delta \phi_1(\theta+u, m)\,du - \int_0^\Delta \phi_1(\theta+u, m+l)\,du \geqq 0$ which can be written as $-\int_0^\Delta \int_0^l \phi_{12}(\theta+u, m+\lambda)\,d\lambda\,du \geqq 0$ and holds for all Δ, $l > 0$ iff $\phi_{12} \leqq 0$.)

The following are examples of functions with antitone differences: i) $\phi(\theta, m) = m^{-a}\theta^b$ with $ab \geqq 0$; ii) let $P(m)$ and $Q(\theta)$ be monotone, nonnegative polynomial functions in m and θ, respectively. Let $\phi(\theta, m) = (P(m))^{-a}(Q(\theta))^b$ and assume that $P'(m)Q'(\theta)ab \geqq 0$; iii) $\phi(\theta, m) = P(\theta)c^{am+b}$ where $P(\theta)$ is a polynomial in θ and $aP'(\theta) \leqq 0$.

The term "antitone differences" was, to our knowledge, introduced by Topkis (1968, 1978); see also Topkis and Veinott (1972). The antitone differences property is equivalent to *submodularity* (see Theorem 3.2 in Topkis 1978): a function $f(\theta, m)$ is submodular if $f((\theta_1, m_1) \wedge (\theta_2, m_2)) + f((\theta_1, m_1) \vee (\theta_2, m_2)) \leqq f(\theta_1, m_1) + f(\theta_2, m_2)$ where $(\theta_1, m_1) \wedge (\theta_2, m_2) = (\min(\theta_1, \theta_2), \min(m_1, m_2))$ and $(\theta_1, m_1) \vee (\theta_2, m_2) = (\max(\theta_1, \theta_2), \max(m_1, m_2))$. The economic interpretation of the antitone differences (or submodularity) property is that the marginal increase in the group cost due to an increase of the aggregate attribute measure (the cardinality) of the group, is smaller for groups of larger size (aggregate attribute measure) than for groups of smaller size (aggregate attribute measure).

The antitone differences (submodularity) property plays a central role in the theory of lattice programming, which is used to verify whether optimal solutions to certain types of optimization models are monotone in some of the model's parameter. The theory is, e.g., used in dynamic programming problems to establish that optimal actions are monotone in the (certain) state variable(s) (see Topkis 1978), the excellent survey on the topic in Chapter 8 of Heyman and Sobel (1982) and the discussion in Section 8. While the term "antitone differences" is due to Topkis, the underlying concept has played a central role in microeconomics, in particular, in the classical theories of production and consumer choice. Production functions are functions of the levels of employed production factors (e.g., labor, capital) and utility functions have the consumption levels of different products as arguments. When these functions have antitone differences, the production factors (products) are called *substitutes* because more of the one decreases the marginal benefit of the other.

Theorem 2 shows that the existence of a monotone partition is guaranteed if f or h belongs to the class $F_2 = \{\phi \in F_1$: ϕ has antitone differences$\}$. We first need the following lemma.

Lemma 2. *Assume that* $f \in F_2$ *or* $h \in F_2$. *If* $1 \leqq k < n-k$

$$U^1(\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_n\}) \leqq U^1(\{x_1, \ldots, x_{n-k}\}, \{x_{n-k+1}, \ldots, x_n\}). \quad (6)$$

Proof. Assume first that f has antitone differences.

$$\begin{aligned} &U^1(\{x_1, \ldots, x_{n-k}\}, \{x_{n-k+1}, \ldots, x_n\}) \\ &\quad - U^1(\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_n\}) \\ &= f(\bar{R}_{1,n-k}, n-k) + f(\bar{R}_{n-k+1,n}, k) \\ &\quad - f(\bar{R}_{1,k}, k) - f(\bar{R}_{k+1,n}, n-k) \qquad (7) \\ &\geqq f(\bar{R}_{1,k}, n-k) + f(\bar{R}_{k+1,n}, k) \\ &\quad - f(\bar{R}_{1,k}, k) - f(\bar{R}_{k+1,n}, n-k) \\ &\geqq 0. \end{aligned}$$

(The first inequality follows from $k < n-k$, and hence, $\bar{R}_{1,n-k} \geqq \bar{R}_{1,k}$ and $\bar{R}_{n-k+1,n} \geqq \bar{R}_{k+1,n}$, using the fact that $f \in F_0$. The second inequality holds since $f(\bar{R}_{1,k}, n-k) - f(\bar{R}_{1,k}, k) \geqq f(\bar{R}_{k+1,n}, n-k) - f(\bar{R}_{k+1,n}, k)$ in view of f having antitone differences and $\bar{R}_{1,k} \leqq \bar{R}_{k+1,n}$.)

Assume next that $h \in F_2$. Since h has antitone differences and $R_{1,k} \leqq R^* - R_{1,n-k}$ it follows that

$$\begin{aligned} &h(R^* - R_{1,n-k}, k) - h(R_{1,k}, k) \\ &\quad \geqq h(R^* - R_{1,n-k}, n-k) - h(R_{1,k}, n-k). \quad (8) \end{aligned}$$

Since $R^* - R_{1,k} \geqq R^* - R_{1,n-k}$ and h is concave in its first argument, we have

$$\begin{aligned} &h(R^* - R_{1,n-k}, n-k) - h(R_{1,k}, n-k) \\ &\quad \geqq h(R^* - R_{1,k}, n-k) - h(R_{1,n-k}, n-k). \end{aligned}$$

This inequality together with (8) proves the lemma.

Theorem 2. *There exists an optimal monotone partition for* $\mathbf{P^1}$ *provided that* $f \in F_2$ *or* $h \in F_2$.

Proof. Assume to the contrary that every optimal consecutive partition fails to be monotone. For the purpose of this proof only, assume that in any given consecutive partition the groups are indexed to achieve a nondecreasing group index function, i.e., not necessarily in nondecreasing order of their cardinalities. Thus, let χ^* be an optimal consecutive partition that maximizes (among all consecutive optimal partitions) the index of the lowest indexed set whose cardinality is larger than that of its successor in the partition; i.e., $\chi^* = \{X_1^*, X_2^*, \ldots,$

$X_L^*\}$ achieves the maximum in

$$\max_{\chi:\ \chi \text{ consecutive optimal}} \min\{l: |X_l| > |X_{l+1}|, \quad l = 1, \ldots, L-1\}. \quad (9)$$

Let l^* be the value of l that achieves the minimum in (9). We show that the following repartitioning procedure transforms χ^* into a new partition χ which is optimal, consecutive and with $|X_1| \leqq |X_2| \leqq \ldots |X_{l^*}| \leqq |X_{l^*+1}|$, thus contradicting the definition of l^* and χ^*. We conclude that an optimal monotone partition must exist.

Repartitioning Procedure

Step 0. Initialize $\chi = \chi^*$; $k = l^*$.

Step 1. If $|X_k| > |X_{k+1}|$, then repartition $X_k \cup X_{k+1}$ into the pair of consecutive sets X_k' and X_{k+1}' with $|X_k'| = |X_{k+1}| < |X_k| = |X_{k+1}'|$. Otherwise, terminate.

Step 2. $X_k := X_k'$; $X_{k+1} = X_{k+1}'$; $k := k-1$. If $k > 1$, return to Step 1.

The repartitioning procedure clearly generates a consecutive partition in each iteration, which in view of the monotonicity of the bounds M_l^* $(l = 1 \ldots L)$ is feasible, in view of Lemma 2 is at least as good as the previous partition, and hence (in view of the optimality of χ^*) is optimal. One easily verifies using the definition of l^* that at the end of Step 2, $|X_{k+1}| \leqq \ldots \leqq |X_{l^*}| \leqq |X_{l^*+1}|$.

The next example shows that the antitone differences property is indeed required to guarantee the existence of a monotone optimal partition.

Example 2. Let $f(\theta, m) = m^{3/4}\theta^{1/2}$. Note that $f \in F_1$. In fact f is even concave in both of its arguments. However, f fails to have antitone differences and neither does $h(\Theta, m) = f(\Theta/m, m) = m^{1/4}\Theta^{1/2}$.

Consider the set X with $N = 11$, $r_i = 0.01$ $(i = 1, \ldots, 10)$ and $r_{11} = 900$. Let L be variable and $M_l^* = 11$ $(l = 1, \ldots, L)$. Clearly in view of Lemma 1 a consecutive optimal partition exists. The cost associated with the single set partition $\{X\}$ is given by $11^{1/4}(900.1)^{1/2} = 54.63$. Next consider any set $\{x_1, \ldots, x_k\}$ with $k < 11$ and note that the cost associated with any partition $\chi = \{X_1, \ldots, X_L\}$ of this set is given by

$$\sum_{l=1}^{L} |X_l|^{3/4}(0.1) \geqq 0.1\left(\sum_{l=1}^{L} |X_l|\right)^{3/4} = 0.1k^{3/4}$$

with strict equality only for the single set partition $\{\{x_1, \ldots, x_k\}\}$. Also, the cost of the partition $\{\{x_1, \ldots, x_{10}\}, \{x_{11}\}\}$ is given by $0.1(10)^{3/4} + 30 = 30.56$. We conclude that the only consecutive partitions of X that may be optimal are of the form $\{\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_{11}\}\}$ with $1 \leqq k \leqq 10$. The cost of the latter partition is given by $g(k) = 0.1k^{3/4} + (11-k)^{1/4}(900 + (10-k)0.01)^{1/2}$. Note that $g'(k) \leqq 3/4(0.1)k^{-1/4} - 1/4(11-k)^{-3/4}(900 + 0.01(10-k))^{1/2} \leqq 0.075 - 1/4(10^{-3/4})(30) < 0$ for all $1 \leqq k \leqq 10$. (The second inequality follows from maximizing each term and each factor of the second term separately.) We conclude that the partition $\{\{x_1, \ldots, x_{10}\}, \{x_{11}\}\}$ is the unique optimal partition of X.

In Theorems 3–6 we identify sufficient conditions with respect to f (or h) guaranteeing that $\mathbf{P}^1$ is extremal. Let

$$F_3 = \{\phi \in F_2;\ \phi(\theta, m) \text{ is concave in } m\}.$$

Theorem 3. $\mathbf{P}^1$ *is extremal if* $h \in F_3$.

Proof. Since $h \in F_2$ it follows from Theorem 2 that there exists an optimal partition which is monotone. Let $\chi = \{X_1, \ldots, X_L\}$ denote such a partition. It remains to be shown that the total cost value associated with any two subsets X_j and X_{j+1} in χ does not increase by shifting the highest indexed element from X_j to X_{j+1}. Without loss of generality we assume that $X_j = \{1, \ldots, k\}$ and $X_{j+1} = \{k+1, \ldots, n\}$. Also, $n - k \geqq k$ by the monotonicity of χ, and define $R^* = R_{1,n}$. Consider these expressions

$$H_1 = h(R_{1,k-1}, k-1) + h(R_{k,n}, n-k+1)$$

$$H_2 = h(R_{1,k-1}, k) + h(R_{k,n}, n-k)$$

$$H_3 = h(R_{1,k}, k) + h(R_{k+1,n}, n-k).$$

We show that $H_1 \leqq H_2 \leqq H_3$.

$H_1 \leqq H_2$

Note that

$$h(R_{k,n}, n-k+1) - h(R_{k,n}, n-k) \leqq h(R_{k,n}, k) - h(R_{k,n}, k-1) \leqq h(R_{1,k-1}, k) - h(R_{1,k-1}, k-1).$$

(The first inequality follows from the concavity of $h(R_{k,n}, \cdot)$ and the second from the antitone differences and the fact that $R_{1,k-1} \leqq R_{k,n}$.) We conclude that

$$H_1 = h(R_{1,k-1}, k-1) + h(R_{k,n}, n-k+1) \leqq h(R_{1,k-1}, k) + h(R_{k,n}, n-k) = H_2.$$

$H_2 \leqq H_3$

Note that

$$h(R_{k,n}, n-k) - h(R_{k+1,n}, n-k)$$
$$\leqq h(R_{k,n}, k) - h(R_{k+1,n}, k)$$
$$\leqq h(R_{1,k}, k) - h(R_{1,k-1}, k).$$

(The first inequality follows from h having antitone differences and the second from the concavity of $h(\cdot, k)$.) Thus

$$H_2 = h(R_{1,k-1}, k) + h(R_{k,n}, n-k)$$
$$\leqq h(R_{1,k}, k) + h(R_{k+1,n}, n-k) = H_3.$$

Theorem 4. $\mathbf{P}^1$ *is extremal if* $f \in F_3$.

Proof. Since $f \in F_2$, there exists an optimal monotone partition χ; see Theorem 2. It thus remains to be shown that the total cost value associated with any two subsets X_j and X_{j+1} in χ is not increased by shifting the highest indexed element from X_j to X_{j+1}. Without loss of generality let $X_j = \{1, \ldots, k\}$ and $X_{j+1} = \{k+1, \ldots, n\}$ where $k \leqq n-k$. Consider the expressions

$$H_1 = f(\bar{R}_{1,k-1}, k-1) + f(\bar{R}_{k,n}, n-k+1)$$
$$H_2 = f(\bar{R}_{1,k-1}, k) + f(\bar{R}_{k,n}, n-k)$$
$$H_3 = f(\bar{R}_{1,k}, k) + f(\bar{R}_{k+1,n}, n-k).$$

We show that $H_1 \leqq H_2 \leqq H_3$.

$H_1 \leqq H_2$

Note that

$$f(\bar{R}_{1,k-1}, k) - f(\bar{R}_{1,k-1}, k-1)$$
$$\geqq f(\bar{R}_{k,n}, k) - f(\bar{R}_{k,n}, k-1)$$
$$\geqq f(\bar{R}_{k,n}, n-k+1) - f(\bar{R}_{k,n}, n-k).$$

(The first inequality follows from f having antitone differences, using the fact that $\bar{R}_{1,k-1} \leqq \bar{R}_{k,n}$. The second inequality follows from the concavity of $f(\theta, \cdot)$ for any given θ using the fact that $n-k+1 > k$.) Thus

$$H_1 = f(\bar{R}_{1,k-1}, k-1) + f(\bar{R}_{k,n}, n-k+1)$$
$$\leqq f(\bar{R}_{1,k-1}, k) + f(\bar{R}_{k,n}, n-k) = H_2.$$

$H_2 \leqq H_3$

This is immediate from $f \in F_0$ and the inequalities $\bar{R}_{1,k-1} \leqq \bar{R}_{1,k}$ and $\bar{R}_{k,n} \leqq \bar{R}_{k+1,n}$.

The following example shows that $\mathbf{P}^1$ may fail to be extremal if neither the function $f \in F_2$ nor the function $h \in F_2$ is concave in their second argument (m).

Example 3. Let $f(\theta, m) = \theta / m$. One easily verifies that $f \in F_2$. Also, $h(\Theta, m) = f(\Theta/m, m) = \Theta/m^2 \in F_2$. Neither f nor h is concave in m. Let $L = 2$ and all $r_i = 1$ $(i = 1, \ldots, N)$. Consider a consecutive partition $\{\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_N\}\}$ with $1 \leqq k \leqq N$; its cost is given by $1/k + 1/(N-k)$ which is a strictly convex function in k achieving its minimum for $k = \lceil N/2 \rceil$ and $k = \lceil N/2 \rceil$ only. The problem thus fails to be extremal.

Theorem 5. *Assume that* $f \in F_0$ *is concave in both arguments and the function* h *has antitone differences. Then* $\mathbf{P}^1$ *is extremal.*

Proof. Note that h is concave in its first argument, since f is as such. Thus, in view of Theorem 2 there exists an optimal partition χ which is monotone. It thus remains to be shown that the total cost value associated with any two subsets X_j and X_{j+1} in χ is not increased by shifting the highest indexed element from X_j to X_{j+1}. Without loss of generality let $X_j = \{1, \ldots, k\}$ and $X_{j+1} = \{k+1, \ldots n\}$ where $k \leqq n-k$. Also let $R^* = R_{1,n}$.

Consider the expressions

$$H_1 = f(\bar{R}_{1,k-1}, k-1) + f(\bar{R}_{k,n}, n-k+1)$$
$$H_2 = f(\bar{R}_{1,k-1}, k) + f((R^* - k\bar{R}_{1,k-1})/(n-k), n-k)$$
$$H_3 = f(\bar{R}_{1,k}, k) + f(\bar{R}_{k+1,n}, n-k).$$

We show that $H_1 \leqq H_2 \leqq H_3$.

$H_1 \leqq H_2$

Note that

$$f(\bar{R}_{1,k-1}, k) - f(\bar{R}_{1,k-1}, k-1)$$
$$\geqq f(\bar{R}_{1,k-1}, n-k+1) - f(\bar{R}_{1,k-1}, n-k)$$
$$= h((n-k+1)\bar{R}_{1,k-1}, n-k+1) - h((n-k)\bar{R}_{1,k-1}, n-k)$$
$$\geqq h(R_{k,n}, n-k+1) - h(R^* - k\bar{R}_{1,k-1}, n-k)$$
$$= f(\bar{R}_{k,n}, n-k+1) - f((R^* - k\bar{R}_{1,k-1})/(n-k), n-k).$$

The first inequality follows from the concavity of $f(\theta, \cdot)$ for any given θ. The second follows from h having antitone differences and the facts

$$(n-k+1)\bar{R}_{1,k-1} \leqq R_{k,n}$$

and

$$\begin{aligned}(n-k+1)\bar{R}_{1,k-1} &- (n-k)\bar{R}_{1,k-1} \\ &= \bar{R}_{1,k-1} \\ &= k\bar{R}_{1,k-1} - R_{1,k-1} \\ &= R_{k,n} - \left(R^* - k\bar{R}_{1,k-1}\right).\end{aligned}$$

This shows that

$$\begin{aligned}H_1 &= f\left(\bar{R}_{1,k-1}, k-1\right) + f\left(\bar{R}_{k,n}, n-k+1\right) \\ &\leqq f\left(\bar{R}_{1,k-1}, k\right) \\ &\quad + f\left(\left(R^* - k\bar{R}_{1,k-1}\right)/(n-k), n-k\right) \\ &= H_2.\end{aligned}$$

$H_2 \leqq H_3$

$$\begin{aligned}f\left(\bar{R}_{1,k}, k\right) &- f\left(\bar{R}_{1,k-1}, k\right) \\ &= h\left(R_{1,k}, k\right) - h\left(k\bar{R}_{1,k-1}, k\right) \\ &\geqq h\left(R_{1,k}, n-k\right) - h\left(k\bar{R}_{1,k-1}, n-k\right) \\ &\geqq h\left(R^* - k\bar{R}_{1,k-1}, n-k\right) \\ &\quad - h\left(R_{k+1,n}, n-k\right) \\ &= f\left(\left(R^* - k\bar{R}_{1,k-1}\right)/(n-k), n-k\right) \\ &\quad - f\left(\bar{R}_{k+1,n}, n-k\right).\end{aligned}$$

The first inequality follows from h having antitone differences and the fact that $k \leqq n-k$. The second inequality follows from the concavity of $h(\cdot, m)$ and the facts

$$k\bar{R}_{1,k-1} \leqq R_{1,k}$$

$$k\bar{R}_{1,k-1} \leqq R_{k+1,n}$$

$$\left(\text{since } R_{k+1,n} \geqq (n-k)r_{k+1} \geqq k\bar{R}_{1,k-1}\right)$$

and

$$\begin{aligned}r_k - \bar{R}_{1,k-1} &= R_{1,k} - k\bar{R}_{1,k-1} \\ &= R^* - k\bar{R}_{1,k-1} - R_{k+1,n}.\end{aligned}$$

We conclude that

$$\begin{aligned}H_2 &= f\left(\bar{R}_{1,k-1}, k\right) \\ &\quad + f\left(\left(R^* - k\bar{R}_{1,k-1}\right)/(n-k), n-k\right) \\ &\leqq f\left(\bar{R}_{1,k}, k\right) + f\left(\bar{R}_{k+1,n}, n-k\right) \\ &= H_3.\end{aligned}$$

Theorem 6. *Let $h \in F_0$ be jointly concave and twice differentiable while f has antitone differences. Then* $\mathbf{P}^1$ *is extremal.*

Proof. Since h is twice differentiable and concave in its first argument, so is f. Let f_i $(i = 1, 2)$ be the partial derivative of f with respect to its ith argument and f_{ij} $(i, j = 1, 2)$ be the partial derivative of f_i with respect to its jth argument. Similar definitions apply to h_i and $h_{i,j}$ $(i, j = 1, 2)$. One easily verifies that $f_2 = \theta h_1 + h_2$ and $f_{22} = \theta^2 h_{11} + 2\theta h_{12} + h_{22} \leqq \max\{x^2 h_{11} + 2xh_{12} + h_{22}: x \in R\}$. Since h is jointly concave we have $h_{11}h_{22} - h_{12}^2 \geqq 0$, $h_{11} \leqq 0$ and $h_{22} \leqq 0$. Thus, if $h_{11} < 0$, the quadratic form in x is nonpositive, while if $h_{11} = 0$ we have $h_{12} = 0$, and since $h_{22} \leqq 0$ the quadratic form is nonpositive as well. We conclude that $f_{22} \leqq 0$, i.e, f is concave in its second argument as well. The theorem follows from Theorem 4.

It remains an open question whether concavity of $h \in F_0$, in both of its arguments separately, in combination with f having antitone differences, is sufficient for $\mathbf{P}^1$ to be extremal. If this result were true, Theorems 3–5 could be summarized by stating that $\mathbf{P}^1$ is extremal provided that one of the two functions f or h is concave in both of its arguments and one of the two functions has antitone differences. The sufficient conditions in Theorems 3 and 4 are stated with respect to a single function (f or h) and, therefore, appear simpler and more natural than their counterparts in Theorems 5 and 6. It is worth noting, however, that extremality of the partitioning problems, which arise in Anily (1986) and Anily and Federgruen (1988a, b, c), is most easily verified via the conditions in Theorem 5. All four theorems and their proofs appear necessary in view of the following observations.

Example 4. Let $f(\theta, m) = \theta\sqrt{m}$. While f is concave in θ and m (separately), the function

$$h(\Theta, m) = f\left(\frac{\Theta}{m}, m\right) = \frac{\Theta}{\sqrt{m}} \text{ is not.}$$

Example 5. Let $h(\Theta, m) = \Theta m$. While h is concave in Θ and m (separately), the function $f(\theta, m) = h(\theta m, m) = \theta m^2$ is not.

Example 6. Let $h(\Theta, m) = \Theta m^{-1/2}$. While h has antitone differences, $f(\theta, m) = \theta m^{1/2}$ does not.

Example 7. Let $f(\theta, m) = \log\log\theta$. While f has antitone differences, $h(\Theta, m) = \log[\log\Theta - \log m]$ does not. To verify the latter, note that

$$\frac{\partial h}{\partial \Theta} = \frac{1}{\Theta(\log\Theta - \log m)}$$

and

$$\frac{\partial^2 h}{\partial m\,\partial \Theta} = \frac{1}{m\Theta(\log \Theta - \log m)^2} > 0.$$

Our results are summarized in Table II.

6. SUFFICIENT CONDITIONS FOR $\mathbf{P}^2$

In this section, we derive sufficient conditions with respect to f (or h) under which $\mathbf{P}^2$: i) has a consecutive optimal partition, ii) has a monotone optimal partition, and iii) is extremal.

Theorem 7. *If $f \in F_0$, there exists an optimal consecutive partition for $\mathbf{P}^2$.*

Proof. Assume to the contrary that every optimal partition fails to be consecutive. Let χ be an optimal partition which maximizes, among all optimal partitions, the index of the lowest indexed nonconsecutive set. (Let p be the index of the lowest indexed nonconsecutive set in χ.) Since X_p is nonconsecutive there exists a set X_t $(t > p)$ with $\min_{x_i \in X_t} r_i < \max_{x_i \in X_p} r_i$. Consider these two cases:

i. $\max_{x_i \in X_p} r_i \leqq \max_{x_i \in X_t} r_i$

Repartition $X_p \cup X_t$ into two consecutive sets $\tilde{X}_p$ and $\tilde{X}_t$ where $|\tilde{X}_p| = |X_p|$, $|\tilde{X}_t| = |X_t|$ and $\tilde{X}_p$ consists of the $|X_p|$ lowest indexed elements of $X_p \cup X_t$. Clearly

$$\max_{x_i \in \tilde{X}_p} r_i \leqq \max_{x_i \in X_p} r_i \quad \text{and} \quad \max_{x_i \in \tilde{X}_t} r_i = \max_{x_i \in X_p} r_i. \tag{10}$$

ii. $\max_{x_i \in X_p} r_i > \max_{x_i \in X_t} r_i$

Repartition $X_p \cup X_t$ into two consecutive sets $\tilde{X}_p$ and $\tilde{X}_t$ where $|\tilde{X}_p| = |X_t|$, $|\tilde{X}_t| = |X_p|$ and $\tilde{X}_p$ consists of the $|X_t|$ lowest indexed elements of $X_p \cup X_t$. Clearly

$$\max_{x_i \in \tilde{X}_p} r_i \leqq \max_{x_i \in X_t} r_i \quad \text{and} \quad \max_{x_i \in \tilde{X}_t} r_i = \max_{x_i \in X_p} r_i. \tag{11}$$

Thus, in both cases $U^2(\{\tilde{X}_p, \tilde{X}_t\}) \leqq U^2(\{X_p, X_t\})$ in view of (10) and (11). The partition χ^1 of X which

Table II
Summary of Results for P^1

Case	$f/h \in F_1$	f Antitone Differences	h Antitone Differences	f Concave in m	h Concave in m	Conclusion
1	+[a]	−	−	−	−	Optimal consecutive partition
2	+	−	−	+	−	(see Lemma 1) *no* monotone
3	+	−	−	−	+	optimal needs to exist
4	+	−	−	+	+	(Example 2)
5	+	+	−	−	−	Monotone optimal partition
6	+	−	+	−	−	exists (see Theorem 2);
7	+	+	+	−	−	Partitioning problem not extremal (see Example 3)
8	+	−	+	−	+	Partitioning problem extremal (Theorem 3)
9	+	+	−	+	−	Partitioning problem extremal (Theorem 4)
10	+	−	+	+	−	Partitioning problem extremal (Theorem 5)
11	+	−	+	+	+	Partitioning problem extremal (Theorems 3, 5)
12	+	+	−	−	+	Optimal monotone partition exists, open question whether partitioning problem extremal (see, however, Theorem 6)
13	+	+	−	+	+	Partitioning problem extremal (Theorem 4)
14	+	+	+	−	+	Partitioning problem extremal (Theorem 3)
15	+	+	+	+	−	Partitioning problem extremal (Theorem 5)
16	+	+	+	+	+	Partitioning problem extremal (Theorems 3, 4, 5)

[a] A "+" ("−") denotes that the property holds (may fail to hold).

results from this partitioning is thus optimal since χ is optimal. Moreover, the first p sets in χ^1 are consecutive, thus contradicting the definition of χ.

Note that $f \in F_0$ iff $h \in F_0$. The following theorem shows that an optimal monotone partition for $\mathbf{P}^2$ exists if $f \in F_0$ and has antitone differences; see Theorem 2 for the corresponding result with respect to $\mathbf{P}^1$.

Theorem 8. *Assume that $f \in F_0$ and f has antitone differences.*

a. If $1 \leqq k \leqq n-k$, $U^2(\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_n\}) \leqq U^2(\{x_1, \ldots, x_{n-k}\}, \{x_{n-k+1}, \ldots, x_n\})$.

b. There exists an optimal monotone partition for $\mathbf{P}^2$.

Proof. a. $U^2(\{x_1, \ldots, x_{n-k}\}, \{x_{n-k+1}, \ldots, x_n\}) = f(r_{n-k}, n-k) + f(r_n, k) \geqslant f(r_k, n-k) + f(r_n, k) \geqq f(r_k, k) + f(r_n, n-k) = U^2(\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_n\})$. The first inequality follows from $f \in F_0$ and the second from f having antitone differences.

b. Given part a the proof is identical to that of Theorem 2.

The following example shows that a monotone optimal partition may fail to exist if f fails to have antitone differences.

Example 8. Let X and f be as defined in Example 2. Following the discussion there, one easily verifies that only the following two partitions may be optimal among all consecutive partitions.

i. $\{X\}$ and ii. $\{\{x_1, \ldots, x_{10}\}, \{x_{11}\}\}$. The former has a cost $11^{3/4}\sqrt{900} = 181.2$ while the latter has a cost $(10^{3/4})\sqrt{0.01} + 30 = 30.56$. Thus, no monotone optimal partition exists.

Theorem 9 shows that the conditions $f \in F_0$, f has antitone differences and f is concave in m, guarantee that $\mathbf{P}^2$ is extremal. Recall from Theorem 4 that to guarantee the extremality of $\mathbf{P}^1$ we require in addition that f be concave in θ, that is, $f \in F_3$.

Theorem 9. *$\mathbf{P}^2$ is extremal if $f(\theta, m) \in F_0$ is concave in m and has antitone differences.*

Proof. Since $f \in F_0$ and has antitone differences, there exists an optimal monotone partition χ; see Theorem 8. It thus remains to be shown that the total cost value of two subsets X_j and X_{j+1} in χ is not increased by shifting the highest indexed element from X_j to X_{j+1}. Without loss of generality let $X_j = \{1, \ldots, k\}$ and $X_{j+1} = \{k+1, \ldots, n\}$ where $k \leqq n-k$. Note that

$$\begin{aligned} &U^2(\{x_1, \ldots x_k\}, \{x_{k+1}, \ldots, x_n\}) \\ &\quad = f(r_k, k) + f(r_n, n-k) \\ &\quad \geqq f(r_{k-1}, k) + f(r_n, n-k). \end{aligned}$$

Also

$$\begin{aligned} &f(r_n, n-k+1) - f(r_n, n-k) \\ &\quad \leqq f(r_n, k) - f(r_n, k-1) \\ &\quad \leqq f(r_{k-1}, k) - f(r_{k-1}, k-1). \end{aligned} \tag{12}$$

The first inequality in (12) follows from f being concave in its second argument and the second inequality from f having antitone differences. Equation 12 implies that

$$\begin{aligned} &U^2(\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_n\}) \\ &\quad \geqq f(r_{k-1}, k) + f(r_n, n-k) \\ &\quad \geqq f(r_{k-1}, k-1) + f(r_n, n-k+1) \\ &\quad = U^2(\{x_1, \ldots, x_{k-1}\}, \{x_k, \ldots, x_n\}). \end{aligned}$$

Example 3 shows that $\mathbf{P}^2$ may fail to be extremal if $f \in F_2$ fails to be concave in its second argument m. Example 9 shows that the conditions $h \in F_2$ and $h \in F_3$ are insufficient to guarantee the existence of a monotone optimal partition or the extremality of $\mathbf{P}^2$, respectively. (This is in contrast to the sufficiency of these conditions with respect to $\mathbf{P}^1$; see Theorems 2 and 3.) This example shows in fact that even the condition $h \in F_3$ is insufficient for the existence of a monotone optimal partition in $\mathbf{P}^2$.

Example 9. Let $h(\Theta, M) = \Theta^{1/4}$. Clearly $h \in F_3$. (In fact, h is even jointly concave in Θ and m.) Note also that $f(\theta, m) = h(\theta m, m) = \theta^{1/4} m^{1/4}$ is jointly concave in both arguments. Consider the set $X = \{x_1, \ldots, x_N\}$ with $N = 10^4 + 1$, $r_1 = \ldots = r_{N-1} = 1$ and $r_N = 10^4$. Let L be variable and all $M_l^* = N$. Clearly $h \in F_0$, so that a consecutive optimal partition exists. Fix $k < N$ and observe first that for any partition $\chi = \{X_1, \ldots, X_L\}$ of $\{x_1, \ldots, x_k\}$, $U^2(\chi) = \sum_{l=1}^{L} |X_l|^{1/4} \geqq (\sum_{l=1}^{L} |X_l|)^{1/4} = k^{1/4}$ with a strict inequality whenever $L \geqq 2$. Thus, the single set partition $\{\{x_1, \ldots, x_k\}\}$ is the *unique* optimal partition of $\{x_1, \ldots, x_k\}$. Moreover, partitioning X into a single set results in a higher total cost value ($\cong 100$) than the partition $\{\{x_1, \ldots, x_{N-1}\}, \{x_N\}\}$, which results in a cost value of 20. Thus, the optimal partition is of the form $\chi(k) = \{\{x_1, \ldots, x_k\}, \{x_{k+1}, \ldots, x_n\}\}$ for some k, $1 \leqq k \leqq N-1$. The cost value of $\chi(k)$ is given by $g(k) = k^{1/4} + ((10^4 + 1 - k)10^4)^{1/4}$ which is a strictly concave function in k; thus it attains its minimum only in one of the extreme points $k = 1$ or $k = 10^4$. A simple comparison shows that $g(10^4) < g(1)$. Thus, the nonmonotone partition $\{\{x_1, \ldots, x_{N-1}\}, \{x_N\}\}$ is the unique optimal partition.

Thus, in terms of conditions with respect to the function h, the only sufficient conditions (for extremality) are given by the next theorem.

Theorem 10. *Assume that $h \in F_0$ is jointly concave and twice differentiable, while f has antitone differences. Then* $\mathbf{P}^2$ *is extremal.*

Proof. As in the proof of Theorem 6 one verifies that f is concave in its second argument. Thus, Theorem 9 can be applied.

We summarize the results of this section in Table III.

7. APPLICATIONS

In this section we apply our results to several optimization models.

Example 10 (Bin Packing Problems). Many combinatorial optimization problems may be formulated as one-dimensional bin packing problems in which a list of objects $\{x_1, \ldots, x_n\}$ is to be packed into a set of bins. Each object x_i has an (integer) size r_i. Applications include storage problems, packing trucks with a given weight limit, assigning commercials to station breaks on television (see Brown 1971) as well as cutting stock and machine scheduling problems. We refer to Coffman, Garey and Johnson (1988) for a survey covering more than one hundred papers on this class of partitioning problems. Here we mention a few versions of the bin packing problem which may directly be formulated as partitioning problems of the type $\mathbf{P}^1$. Chandra and Wang (1975) and Cody and Coffman (1976) study bin packing problems that arise in the allocation of records on computer auxiliary storage devices. The former models paging drums where a given set of pages is to be partitioned among L sectors (bins/groups) of the drum to minimize

Table III
Summary Results for P^2

Case	$f/h \in F_0$	f Antitone Differences	f Concave in m	h Concave in Θ, Antitone Differences	h Concave in Θ, m; Antitone Differences	Possible Combination	Conclusion
1	+[a]	−	−	−	−		Optimal consecutive partition exists (Theorem 7); does not need to be monotone, see Example 8
2	+	−	−	+	−		Optimal consecutive partition exists; does not need to be monotone; see Example 9
3	+	−	−	−	+	Cannot occur	
4	+	−	−	+	+		Optimal consecutive partition exists; does not need to be monotone; see Example 9
5	+	+	−	−	−		Optimal monotone partition exists (Theorem 8)
6	+	+	−	+	−		Partitioning problem does not need to be extremal (Examples 3, 9)
7	+	+	−	−	+	Cannot occur	
8	+	+	−	+	+	Cannot occur in view of Theorem 10	
9	+	−	+	−	−	Cannot occur	Optimal consecutive partition exists; does not need to be monotone; see Examples 8, 9
10	+	−	+	+	−		
11	+	−	+	−	+		
12	+	−	+	+	+		
13	+	+	+	−	−	Cannot occur	Partitioning problem is extremal
14	+	+	+	+	−		
15	+	+	+	−	+		
16	+	+	+	+	+		

[a] A "+" ("−") denotes that the property holds (may fail to hold).

average access time; the latter models arm contention in disk-pack computer storage. In this case, the objective is to minimize the contention that occurs whenever two items from the same bin are requested at the same time. Both models reduce to solving a partitioning problem of the type $\mathbf{P}^1$ with $f(\theta, m) = m^2\theta^2$ or $h(\Theta, m) = \Theta^2$. It is assumed that any number of items may be stored in a bin; Eastman and Wong (1975) consider capacitated bins, i.e., $M_1^* = \ldots = M_L^* < \infty$. Since the functions f and h are *convex* in θ and Θ, respectively, the results of this paper fail to apply. In fact, using a straightforward adaptation of the proof in Example 1, one easily verifies that both partitioning problems are *NP*-complete. The above authors have, however, demonstrated that a relatively simple heuristic is guaranteed to result in solutions whose *cost value* comes within a few percentage points of the optimal value; see Coffman, Garey and Johnson for details.

Example 11. Queueing models are increasingly used to model a wide variety of systems in which users (customers) compete for limited capacity. These include, for example, production systems, telecommunication processes and service facilities. Many such systems consist of a number of parallel servers or server pools and deal with multiple classes of customers, each of which is to be assigned to one of the server pools.

Thus, let $\{1, \ldots, N\}$ be the collection of customer classes. Assume, for example, that all customer classes arrive according to independent Poisson processes, and let λ_i denote the arrival rate of class i $(i = 1, \ldots, N)$. All L server pools consist of c $(\geqq 1)$ identical servers. Customers have independent and identically distributed work loads. The service time distribution of a given server pool may depend on the number of distinct customer classes it is assigned to: The service rate of a server (pool) typically decreases with the number of distinct classes it is responsible for, with fully specialized or dedicated servers (dealing with a single class) and *general purpose, flexible servers* (dealing with a large number of classes) as extremes. Thus, let $G(\cdot, m)$ denote the general service time distribution of a server that deals with m distinct customer classes.

Note that under any given assignment of classes, each server pool operates as an $M/G/c$ system. If the objective is to minimize the expected total number of customers (in the queue or in the system) or the overall average of the expected waiting times experienced by all customers, the assignment problem is easily translated into a partitioning problem of the type $\mathbf{P}^1$. The *cost* of assigning a group $S \subset \{1, \ldots, N\}$ of customer classes to a server pool depends on $\sum_{i \in S} \lambda_i$ and $|S|$.

For example, when minimizing the total queue size, the group cost function is given by

$$\infty \quad \text{if } \sum_{i \in S} \lambda_i \int_0^\infty (1 - G(u, |S|))\, du \geqq 1$$

$$h\Big(\sum_{i \in S} \lambda_i, |S|\Big) = \text{the expected number of customers in an } M/G/c \text{ system with arrival rate } \sum_{i \in S} \lambda_i \text{ and service time distribution } G(\cdot, |S|).$$

For general $c \geqq 1$ and service time distributions G no exact characterizations of the $f(\cdot, \cdot)$ function can be obtained. However, it is well known that the function is convex when the service time distributions are exponential (see Lee and Cohen 1983, Grassmann 1983), when $c = 1$ and G general (see the Pollatzhek-Khintchine formula), and in heavy traffic (see Boxma, Cohen and Nuffels 1979). For the general case, a widely used and extremely accurate approximation formula (see Krampe, Kubat and Runge 1973, Maaløe 1973, Stoyan 1976, Nozaki and Ross 1978, Hokstad 1978 and Tijms, Van Hoorn and Federgruen 1981) for the expected queue size exhibits a convex dependency on the (total) arrival rate as well. As in Example 10, since the group cost function is convex in the total assigned arrival rate, the results of this paper fail to apply. In fact, by using a simple adaptation of the proof of Example 11, one easily verifies that the problem is NP-complete even when the service time distribution is independent of the number of customer classes assigned to a given server and when the number of server pools $L = 2$.

We now apply the results of this paper to several physical distribution management models.

Example 12 (Joint Replenishment Problems). One of the major complications in managing multi-item inventory systems stems from the fact that various cost components, in particular, setup costs, are often *jointly* incurred between several items. The joint cost structure often reflects *economies of scale* which may be exploited by combining different items in the same production batch or delivery order.

Chakravarty, Orlin and Rothblum (1982) consider a variant of the Joint Replenishment Problem, in which demands are assumed to occur continuously, at item-specific but time-homogeneous rates. The authors consider strategies that employ a fixed partition of items into groups; each time the inventory of a given item is replenished, it is replenished jointly with the other members of the group, and the setup cost of that group is incurred. The joint setup cost is assumed to be given by a function $K(m)$ where m denotes the number of items in the group.

Without loss of generality, the units in which the items are measured may be chosen such that each item's demand rate equals two. Each item i is thus characterized by a single attribute, its holding cost H_i $(i = 1, \ldots, N)$. One easily verifies that the optimal inventory strategy in the above defined class is determined by solving a partitioning problem of the type $\mathbf{P^1}$ with group cost function $h(\Theta, m) = 2(K(m)\Theta)^{1/2}$. An optimal consecutive partition exists in view of Lemma 1; note, however, that neither $h(\cdot, \cdot)$ nor $f(\theta, m) = 2(mK(m)\theta)^{1/2}$ have antitone differences, i.e., one of the Cases 1-4 in Table II applies and therefore no monotone partition needs to be optimal. The dynamic programming algorithms of Section 2 may be used to solve these models.

Example 13. In the classical Vehicle Routing Problem (VRP), a set of deliveries to a given collection of customers is to be assigned to a fixed or variable sized fleet of vehicles, each of which is of limited capacity. The objective is to find a set of *routes*, where each route starts at the depot and returns there after visiting a subset of customers, so that each customer is visited exactly once, the capacity of the vehicle is not exceeded and the total length of all routes is as small as possible. Haimovich and Rinnooy Kan (1985) considered a stylized version of the VRP in which the distances between customers are given by the Euclidean distances between the corresponding points, and deliveries to a given customer may be split between several vehicles. Thus, assuming that all delivery sizes are integers, a customer with delivery size d may be treated as d customers with demand 1, all located at the same location; the capacity of a vehicle may then be stated as an upper bound on the *number* of *customers* that may be included in a single route.

For the above version of the VRP, and considering a fleet of *identical* vehicles, Haimovich and Rinnooy Kan derive easily computable lower and upper bounds for the optimal solution value $R^*(X)$ which are shown to be asymptotically accurate under mild probabilistic assumptions. Following the basic approach of these authors we address the more general case with a fleet of L (possibly) *nonidentical* vehicles with capacities $M_1^* \leqq \cdots \leqq M_L^*$, respectively.

Let $\{1, \ldots, N\}$ be the collection of customers and r_i denote the radial distance of customer i from the depot. For any collection of customers $S \subset \{1, \ldots, N\}$, let $R(S)$ denote the length of the optimal route starting and terminating at the depot and visiting all customers in S exactly once. Clearly

$$R(S) \geqq R^{(2)}(S) \geqq R^{(1)}(S) \tag{12}$$

where

$$R^{(1)}(S) = 2\Big(\sum_{i \in S} r_i\Big) / |S|; \quad R^{(2)}(S) = 2 \max_{i \in S} r_i.$$

($R(S)$ is at least as large as twice the distance to the farthest point in S, thus verifying the first inequality in (12); the second inequality is immediate.) Two lower bounds for $R^*(X)$—the optimal routing cost—may thus be determined by computing $V^1(X)$ and $V^2(X)$, the solutions of $\mathbf{P^1}$ and $\mathbf{P^2}$ with $f(\theta, m) = 2\theta$. Note $V^1(X) \leqq V^2(X) \leqq R^*(X)$.

It follows from Theorems 4 and 9, that both $\mathbf{P^1}$ and $\mathbf{P^2}$ are *extremal*. Thus, the partition χ^* generated by the EPA achieves both $V^1(X)$ and $V^2(X)$ in $\mathbf{P^1}$ and $\mathbf{P^2}$, respectively. Using a minor adaptation to the proofs of Haimovich and Rinnooy Kan one verifies that both lower bounds are asymptotically accurate as the number of customers tends to infinity (under the same probabilistic assumptions as *ibid*). The partition χ^* may also be used as a basis for several regional partitioning schemes that result in (heuristic) sets of routes which can be shown to be asymptotically optimal as well; see Anily and Federgruen (1988b, c) and Example 14.

Example 14. Consider a one-warehouse multiple retailer system in which at each retailer x_i customer demands for a given product occur at a constant deterministic rate μ_i, with $\mu_i = k_i\mu$ for integers $k_i \geqq 1$ and a given base rate $\mu > 0$ $(i = 1, \ldots, N)$. All stock enters the system through the depot from where it is distributed, in efficient routes, to (some of) the retailers via a fleet of vehicles, each with a given load capacity of b units.

Inventories are kept at the retailers but not at the depot. (A different, somewhat more complex, application of our class of routing models arises in systems with central inventories; see Anily 1987, Chapter 5). Inventory carrying costs are incurred at a rate h^+ per unit, per unit of time. Transportation costs include a fixed cost c per route driven and variable costs proportional with the total (Euclidean) distance driven. (The cost per mile is normalized as one.) We wish to determine replenishment strategies that enable all retailers to meet their demands while minimizing long-run average transportation and inventory carrying costs. We refer to Anily and Federgruen (1988c) for a survey of related models on combined inventory control and vehicle routing problems.

Define a demand point as a point in the plane that faces a demand rate of μ. Each retailer x_i $(i = 1, \ldots, N)$ with demand rate $k_i\mu$ may be replaced by k_i independent demand points, all located at the same geographic

point. (This modeling device is similar to the one employed in Example 13.) We restrict ourselves to the class of strategies that partitions the demand points into a collection of L regions such that each time one of the demand points in a given region receives a delivery, this delivery is made by one of the vehicles visiting all other demand points in the region as well. (See Anily and Federgruen 1988c for a discussion of this restriction.) In view of the limited vehicle fleet sizes and other considerations pointed out in Anily and Federgruen (1988c), we specify that a vehicle may be dispatched to a given region at most f^* times per unit of time. This consideration implies, in itself, an upper bound $M^* = bf^*/\mu$ for the number of demand points that may be assigned to a single region (route). See Anily and Federgruen (1988c) for other constraints which may be translated into (possibly nonidentical) upper bounds $M_1^*, \ldots, M_L^*$ for the regions.

Anily and Federgruen (1988c) derive two lower bounds V^1 and V^2 for the optimal solution value V^* which may be determined by solving a partitioning problem of the type $\mathbf{P^1}$ and $\mathbf{P^2}$, respectively, in both cases with

$$f(\theta, m) = \begin{cases} h^+\mu m/(2f^*) + f^*(\theta + c) & \\ \qquad \text{if } \theta + c \leqq \mu m h^+/(2f^*) & \\ \left(2h^+\mu m(\theta + c)\right)^{1/2} & \\ \qquad \text{if } \mu m h^+/(2f^*)^2 \leqq \vartheta + c \leqq b^2 h^+/(2\mu m) & \\ h^+ b/2 + (\mu m/b)(\theta + c) \qquad \text{otherwise.} & \end{cases}$$

Clearly, $f \in F_1$ so that both lower bounds V^1 and V^2 are achieved by *consecutive* partitions. The function $f(\cdot, \cdot)$ also satisfies the assumptions in Theorem 5, so that $\mathbf{P^1}$ is extremal and V^1 is achieved by the partition generated by the EPA. Note that none of the simpler theorems 3 or 4 could be employed.

On the other hand, the fact that f and $h \in F_1$, that f and h are concave in m and h has antitone differences (but f does *not*) is insufficient to demonstrate that an optimal monotone partition exists for $\mathbf{P^2}$ as well, as follows from Examples 8 and 9; see also case 12 in Table III. Thus, the computation of V^2 is more complex than that of V^1, and in the case of nonidentical regional capacities $\{M_1^*, \ldots, M_L^*\}$ no efficient evaluation methods for V^2 appear to be known; see Section 3.

Both lower bounds V^1 and V^2 are shown to be asymptotically accurate, under mild probabilistic assumptions with respect to the distribution of radial distances. It is also shown how the partition that optimizes $\mathbf{P^1}$ or $\mathbf{P^2}$ may be used as the basis for the construction of a collection of regions and associated inventory strategy, the cost of which is asymptotically optimal (under the above referred-to conditions).

Example 15. Consider a continuous-time, two-stage production/inventory system. In the first stage a common intermediate product is produced in batches and possibly stored. In the second phase the intermediate product is fabricated into N distinct finished products; several finished products may be included in a single production batch to exploit economies of scale. In particular, assume that a fixed cost c is incurred for any (second stage) production run. Likewise, a fixed cost K_0 is incurred whenever a new production run for the intermediate product is initiated. Inventories of the intermediate product incur carrying costs at a rate h_0 per unit of time while inventories of end item i are charged at a rate h_i, $i = 1, \ldots, N$. Let $h_i' = h_i - h_0$ $h_i' \geqq 0$ $(i = 1, \ldots, N)$. Since holding cost rates usually increase with the (cumulative) value added, this assumption is almost always satisfied.

Demands for the end items occur at deterministic, constant rates per unit of time, all expressed in a common unit (pounds, gallons, etc.). These demands must be filled from available inventories, i.e., backlogging is not allowed. While different items may be combined in a single production batch, the total production volume per batch cannot exceed a *capacity limit* of b units.

We are interested in determining a production/inventory strategy that minimizes long-run average costs. We assume that the variable production costs (in both stages) are linear in the production volumes; hence, these cost components may be ignored because their long-run average value is identical for all relevant replenishment strategies, with long-run average production rates equal to the demand rates.

Assuming (as above) that all demand rates are integer multiples of some base rate μ, each finished product may be viewed as representing an integer number of independent *demand-items*, each with a demand rate μ.

Optimal policies may be very complex even without joint setup costs (see, e.g., Roundy 1985), and their complexity makes them difficult to implement even if they could be computed efficiently. As a consequence, we restrict ourselves to replenishment strategies which partition the demand items into a collection of families $\chi = \{X_1, \ldots, X_L\}$ such that in any (second-stage) production batch only one of the families in χ is produced. Note that a finished product may be included in several families and may thus be produced by itself as well as in conjunction with different combinations of other end products. See Anily and Federgruen (1988a) for a discussion of this restriction. As in the previous example, various constraints imply an upper bound M^* for the

number of demand items which may be assigned to a given family.

In Anily and Federgruen (1988a), we derive a lower bound $\underline{V}$ for V^*, the minimum long-run average cost among all strategies in the above defined class. This lower bound is of the form

$$\underline{V} = \min_{T>0} \underline{V}(T)$$

where

$$\underline{V}(T) = \min\left\{ \sum_{l=1} h_T\left(\sum_{i\in X_l} h'_i, \mid X_l \mid \right) : \chi = \{X_1, \ldots, X_L\} \text{ is a feasible partition} \right\}. \tag{13}$$

See Anily and Federgruen for a specification of the function $h_T(\cdot,\cdot)$.

For any $T>0$, the corresponding function $f_T(\cdot,\cdot)$ satisfies the conditions of Theorem 5, so that the corresponding partitioning problem (of type $\mathbf{P}^1$) is *extremal*, and therefore optimized by the (unique) partition $\chi^* = \{X_1^*, \ldots, X_L^*\}$ generated by the EPA. We conclude that the same partition χ^* achieves the minimum in (13) for all $T>0$! Hence

$$\underline{V}(T) = \min_{T>0} \sum_{l=1}^{L} h'_T\left(\sum_{i\in X_l^*} h'_i, \mid X_l^* \mid \right). \tag{14}$$

The function to the right of (13) is clearly convex in T and its minimum may be computed in closed form! Note that if we had failed to recognize that the partitioning problem in (13) is extremal, we would suspect that the optimal partition in (13) depends on T and that the function $\underline{V}(T)$ fails to be convex, leaving us with a complex minimization problem over T.

The lower bound $\underline{V}$ and the partition χ^* may be used as the basis for the construction of a replenishment strategy whose cost value comes within 6% of $\underline{V}$ (and hence of V^*); see Anily and Federgruen (1988a) for details and even better optimality gaps that apply in special cases.

8. RELATED PARTITIONING PROBLEMS

Grötschel, Lovász and Schrijver (1982) consider the general partitioning problem.

Problem P

Minimize $\{U(\chi)$: χ partitions X and $m_l \leqq M_l^*$, $l = 1, \ldots, L\}$ where $U(\chi) = \sum_{l=1}^{L} g(X_l)$ and $g(\cdot)$ is a general (normalized) monotone *submodular* set function, i.e., $g: 2^X \to R$ with

i. $g(\emptyset) = 0$ (normalization),

ii. $g(S) \leqq g(T)$ if $S \subset T$ (monotonicity),

iii. $g(S\cup T) + g(S\cap T) \leqq g(S) + g(T)$ for all $S, T \subset X$ (*submodularity*). Grötschel, Lovász and Schrijver prove that **P** can be solved by the ellipsoid method. The running time of this algorithm is polynomial, albeit of an unattractively high degree.

Consider the case where $g(S) = h(\sum_{i:x_i\in S} r_i, \mid S \mid)$. If $h\in F_3$, it follows from Lemma 2 in Federgruen and Zheng (1988) that g is a normalized monotone and submodular set function. On the other hand, if $h\notin F_3$, then g will generally fail to be submodular. But if $h\in F_3$, we know that **P** is extremal and the EPA solves the problem in no more than $3N/2$ operations and generates a partition of a strikingly simple structure; see Section 4 and cases 8, 11, 14, and 16 in Table II.

Topkis (1978) deals with partitioning problems in which a consecutive partition is known to be optimal or in which only consecutive partitions are considered. As pointed out in Section 2, with L variable and $M_l^* = M^*$, $l = 1, \ldots, L$ such problems may be solved by computing a shortest path in an acyclic network with N nodes. The latter may, of course, be described by dynamic programming recursions of the form

$$F(j) = \min_{1\leqq l\leqq j-1} \{c(j,l) + F(l)\} \quad 1 \leqq j \leqq N, \qquad F(0) = 0. \tag{18}$$

Let $s(j)$ denote the smallest optimal successor of j, i.e., $s(j)$ is the smallest value of l achieving the minimum in (18). Topkis (1978) was primarily interested in conditions under which $s(j)$ is nondecreasing in j. This property has important implications for the existence of planning and forecasting horizons; see his paper as well as Heyman and Sobel (Chapter 8). He shows that the successor function $s(\cdot)$ is nondecreasing if c has antitone differences, i.e., it is submodular:

$$c(j_2, i_1) + c(j_1, i_2) \leqq c(j_1, i_1) + c(j_2, i_2) \quad \text{if } i_1, i_2 \leqq \min(j_1, j_2).$$

This result comes as a corollary to a significantly more general treatment. Here we present a simple and self-contained proof (which to our knowledge has not appeared in the literature).

Lemma 3. *If $c(i,j)$ is a submodular function (i.e., if $c(\cdot,\cdot)$ has antitone differences), then $s(j)$ is a nondecreasing function of j.*

Proof. The proof is a contradiction. Assume that there exist $j_1 < j_2$ such that $s(j_1) > s(j_2)$.

$$j_2 \leqq j_1 < s(j_2) < s(j_1).$$

Then

$$\begin{aligned} F(j_1)+F(j_2) &= c(j_1, s(j_1)) + F(s(j_1)-1) \\ &\quad + c(j_2, s(j_2)) + F(s(j_2)-1) \\ &\geqq c(j_2, s(j_1)) + F(s(j_1)-1) \\ &\quad + c(j_1, s(j_2)) + F(s(j_2)-1) \\ &\geqq c(j_2, s(j_2)) + F(s(j_2)-1) \\ &\quad + c(j_1, s(j_2)) + F(s(j_2)-1). \end{aligned}$$

The first inequality follows from the submodularity of c and the second from the definition of $s(\cdot)$. We conclude that

$$\begin{aligned} &c(j_1, s(j_2)) + F(s(j_2)-1) \\ &\quad \leqq c(j_1, s(j_1)) + F(s(j_1)-1). \end{aligned}$$

Hence $s(j_1)$ does not achieve the minimum in (18) for $j = j_1$, thus contradicting the definition of $s(j_1)$.

When the partitioning problem is extremal, it follows from Theorem 1 that the EPA in Section 4, when applied to a set $\{x_1, \ldots, x_j\}$ $(1 \leqq j \leqq N)$, generates a partition in which $\hat{s}(j)$, the index of the lowest indexed element of the last set, equals $s(j)$. One also easily verifies that $\hat{s}(j)$ $(= s(j))$ is nondecreasing in j. Thus, the conditions in Theorems 3–6, 9 and 10 provide alternative conditions under which the successor function $s(\cdot)$ is nondecreasing. It is noteworthy that these conditions may hold while the corresponding $c(\cdot,\cdot)$ function *fails* to be submodular when viewed as a function of the lowest and highest index in the group; see Example 16.

Example 16. Consider $\mathbf{P}^1$ and let $f(\theta, m) = \theta$. The problem is clearly extremal. Let $N = 4$, $r_1 = 1$, $r_2 = r_3 = 5$ and $r_4 = 10$. Note that

$$c(3,1) = \sum_{i=1}^{3} r_i/3 = 3\frac{2}{3}$$

$$c(4,2) = \sum_{i=2}^{4} r_i/3 = 6\frac{2}{3}$$

$$c(4,1) = \sum_{i=1}^{4} r_i/4 = 5.25$$

$$c(3,2) = \sum_{i=2}^{3} r_i/2 = 5.$$

Thus $c(3,1) + c(4,2) > c(4,1) + c(3,2)$, which violates submodularity.

ACKNOWLEDGMENT

The research of the first author was partially supported by NSERC grant A4802. The research of the second author was partially supported by NSF grant ECS-8604409 as well as a grant by the Faculty Research Fund of the Graduate School of Business, Columbia University.

REFERENCES

ANILY, S. 1987. Integrating Inventory Control and Transportation Planning. Ph.D. Dissertation, Columbia University, New York.

ANILY, S., AND A. FEDERGRUEN. 1988a. Replenishment Strategies for Capacitated Two-Stage Production/Inventory Systems. Graduate School of Business Working Paper, Columbia University, New York. (To appear in *Opns. Res.*).

ANILY, S., AND A. FEDERGRUEN. 1988b. A Class of Euclidean Routing Problems With General Cost Functions. Working Paper, Faculty of Commerce and Business Administration, University of British Columbia, Vancouver. (To appear in *Math. Opns. Res.*).

ANILY, S., AND A. FEDERGRUEN. 1988c. One Warehouse Multiple Retailer Systems With Vehicle Routing Costs. Working Paper, Faculty of Commerce and Business Administration, University of British Columbia, Vancouver. (To appear in *Mgmt. Sci.*).

BARNES, E. R. 1982. An Algorithm for Partitioning the Nodes of a Graph. *SIAM J. Algebraic and Discrete Math.* **3**, 541–550.

BARNES, E. R. 1985. Partitioning the Nodes of a Graph. Manuscript.

BARNES, E. R., AND A. J. HOFFMAN. 1984. Partitioning Spectra and Linear Programming. In *Progress in Combinatorial Optimization*, W. R. Pulleyblank (ed.).

BARNES, E. R., A. J. HOFFMAN AND U. G. ROTHBLUM. 1989. Optimal Partitions Having Disjoint Convex and Conic Hulls. Working Paper, Faculty of Industrial Engineering and Management, Technion, Haifa, Israel.

BOXMA, O. J., J. W. COHEN AND N. NUFFELS. 1979. Approximations of the Mean Waiting Time in $M/G/s$ Queueing Systems. *Opns. Res.* **27**, 1115–1127.

BROWN, A. 1971. *Optimum Packing and Depletion*. American Elsevier, New York.

CHAKRAVARTY, A. K., J. B. ORLIN AND V. G. ROTHBLUM. 1982. A Partitioning Problem With Additive Objective With an Application to Optimal Inventory Groupings for Joint Replenishment. *Opns. Res.* **30**, 1018–1020.

CHAKRAVARTY, A. K., J. B. ORLIN AND V. G. ROTHBLUM. 1985. Consecutive Optimizers for a Partitioning Problem With Applications to Optimal Inventory Groupings for Joint Replenishment. *Opns. Res.* **33**, 820–834.

CHANDRA, A., AND C. WANG. 1975. Worst Case Analysis

of a Placement Algorithm Related to Storage Allocation. *SIAM J. Comput.* **4**, 249–263.

CODY, R. AND E. COFFMAN. (1976). Record Allocation for Minimizing Expected Retrieval Costs on Drum-Like Storage Devices. *J. Assoc. Comput. Mach.* **23**, 103–115.

COFFMAN, E., M. GAREY AND D. JOHNSON. 1988. Approximation Algorithms For Bin-Packing—An Updated Survey. Unpublished manuscript.

DANATH, W. E., AND A. J. HOFFMAN. 1973. Lower Bounds for Partitioning of Graphs. *IBM J. Res. Dev.* **17**, 420–425.

EASTMAN, M., AND C. WONG. 1975. The Effect of a Capacity Constraint on the Minimal Cost of a Partition. *J. Assoc. Comput. Mach.* **22**, 441–449.

FEDERGRUEN, A., AND Y. ZHENG. 1988. Minimizing Submodular Set Functions: Efficient Algorithms for Special Structures With Applications to Joint Replenishment Problems. Working Paper. Graduate School of Business, Columbia University, New York.

GRASSMANN, W. 1983. The Convexity of the Mean Queue Size of the $M/M/c$ Queue With Respect to the Traffic Intensity. *J. Appl. Prob.* **20**, 920–923.

GRÖTSCHEL, M., L. LOVÁSZ AND A. SCHRIJVER. 1982. The Ellipsoid Method and Its Consequences in Combinatorial Optimization. *Combinatorica* **1**, 169–197.

HAIMOVICH, M., AND A. RINNOOY KAN. 1985. Bounds and Heuristics for Capacitated Routing Problems. *Math. Opns. Res.* **10**, 527–542.

HEYMAN, D. P., AND M. J. SOBEL. 1982. *Stochastic Models in Operations Research*, Vol. II. McGraw-Hill, New York.

HOKSTAD, P. 1978. Approximations for the $M/G/m$ Queue. *Opns. Res.* **26**, 511–523.

HWANG, F. K., 1981. Optimal Partitions. *J. Optim. Theory Appl.* **34**, 1–10.

HWANG, F. K., J. SUM AND E. Y. YAO. 1985. Optimal Set Partitioning. *SIAM J. Algebraic Discrete Math.* **6**, 163–170.

KARP, R. 1972. Reducibility Among Combinatorial Problems. In *Complexity of Computer Computations*, R. Miller and G. Thatcher (eds.). Plenum Press, New York.

KRAMPE, H., J. KUBAT AND W. RUNGE. 1973. *Bedienungsmodelle*. Oldenburg, München.

LEE, H., AND M. COHEN. 1983. A Note on the Convexity of Performance Measures of $M/M/c$ Queueing Systems. *J. Appl. Prob.* **20**, 920–923.

MAALØE, E. 1973. Approximation Formula for Estimation of Waiting Time in Multiple-Channel Queueing Systems. *Mgmt. Sci.* **19**, 703–710.

NOZAKI, S. A., AND M. ROSS. 1978. Approximations in Finite Capacity Multi-Server Queues With Poisson Arrivals. *J. Applied Prob.* **15**, 826–834.

ROUNDY, R. 1985. 98%-Effective Inter-Ratio Lot-Sizing for One-Warehouse Multi-Retailer Systems. *Mgmt. Sci.* **31**, 1416–1430.

SPÄTH, H. 1985. *Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples*. John Wiley, Chichester, England.

STOYAN, D. 1976. Approximations for $M/G/s$ Queues. *Math. Operations-forsch. Statist.* **7**, 587–594.

TIJMS, H., M. VAN HOORN AND A. FEDERGRUEN. 1981. Approximations for the Steady State Probabilities in the $M/G/c$ Queue. *Adv. Applied Prob.*, **13**, 186–206.

TOPKIS, D. 1968. Ordered Optimal Solutions. Ph.D. Dissertation, Stanford University, Stanford, Calif.

TOPKIS, D. M. 1978. Minimizing a Submodular Function on a Lattice. *Opns. Res.* **26**, 305–321.

TOPKIS, D., AND A. VEINOTT. 1972. Isotone Solutions of Extremal Problems on a Lattice. Unpublished manuscript.