AN ABSTRACT OF THE THESIS OF

MARK THOMAS HEHNEN for the degree of DOCTOR OF PHILOSOPHY

in STATISTICS presented on 18 November 1977

Title: THE GENERIC KNAPSACK FACTORY: DECISION RULES FOR SEQUENTIAL KNAPSACK PROBLEMS WITH AN ADDITIONAL CONSTRAINT

Redacted for Privacy

Abstract approved: Mark R. Lembersky

Consider the situation where one is faced with a sequence of knapsack type problems which involves additional constraints on the permissible number of certain types of knapsack fillers to be used over the sequence. This research is concerned with determining good and perhaps optimal solutions which maximize the total value realized over the sequence and which can be easily computed. Specifically, simple and good, in the sense of nearly optimal, solutions are obtained for the deterministic cases where 1) the knapsack problems in the sequence are identical and 2) the knapsack problems in the sequence are not identical. In both cases, we require that the number of knapsacks in the sequence is known and that the form of each individual knapsack problem is known and hence its solution is available.

The Generic Knapsack Factory: Decision Rules for Sequential Knapsack Problems with an Additional Constraint

by

Mark Thomas Hehnen

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1978

APPROVED:

Redacted for Privacy

Associate Professor of Statistics

Redacted for Privacy

Chairman of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented 18 November 1977

Typed by Clover Redfern for Mark Thomas Hehnen

To and for Vicki, Kathleen, and Steven

TABLE OF CONTENTS

LIST OF FIGURES

THE GENERIC KNAPSACK FACTORY: DECISION RULES FOR SEQUENTIAL KNAPSACK PROBLEMS WITH AN ADDITIONAL CONSTRAINT

I. INTRODUCTION

1. An Illustrative Scenario

The familiar knapsack problem of operations research literature is representative of a large variety of practical problems but is not considered a serious problem in its literal formulation of packing a knapsack for maximum value. In the same way, we will illustrate a recurring industrial problem in the literal context of a sequence of knapsacks packed in a "factory."

The Western Knapsack Company has been in business for several years, providing prepacked knapsacks for outdoorsmen. Their principal product consists of a backpack which has no inherent value but whose price is determined by the value of the individual items with which it is filled. Based on current prices, the firm is able to determine the optimal set of items with which to completely fill their pack. This solution set includes four honey and cereal bars and has a total value of $100.00. At the knapsack factory, the knapsack stuffer duly packs each backpack with four such bars, which heretofore have always been available as needed. In a given shift, the knapsack stuffer fills 30 packs using 120 bars. Unfortunately, current scarcity has reduced the availability of this popular but expensive item so that the company's management has been forced to allocate only 36 bars to each shift. As a consequence, the stuffer has a dilemma; should he fill the first 9 packs with four bars and the rest with none, filling the space with a spare first aid kit, thereby reducing the value of the last 21 knapsacks to \$93.00 apiece or should he fill the first 18 packs with two bars and one can of soup reducing the knapsack's value to \$97.00 and fill the remaining 12 knapsacks with the first aid kit in lieu of any cereal bars. Of course, there are many more alternatives than these and while our stuffer is pondering the problem he receives the discouraging news that the number of cereal bars that he can expect on subsequent shifts may vary. He realizes the answer for a 36 bar constraint may not hold for future shifts. What is needed is a decision rule or set of decision rules which can address the problem parametrically based on information available at the beginning of each shift. This scenario illustrates sufficiently the nature of the problem addressed in this paper and we will proceed to formalize the problem.

2. Generic Knapsack Factory Problems

We shall use the term Generic Knapsack Factory Problem to include situations similar to the one just discussed where the problem faced in an industrial application consists of a sequence of individual "knapsack type" problems for which there are additional constraints on the entire sequence linking the individual problems together. We shall generalize and formulate the individual generic Knapsack Problem in Chapter II. This problem permits an extensive set of variations and permutations. We point out at the outset that the results obtained in this research are for the specific case where there is a single additional constraint tying the sequence together. Nonetheless, we consider it useful to discuss the Knapsack Factory Problem in general at this time in order to provide insight into potential applications and the meaning of certain mathematical expressions and assumptions made in the body of this paper. To do so, let us consider the following situation:

A knapsack factory is under contract to provide a purchaser with knapsacks of several varieties stuffed with fillers of various types. The purchaser accepts all the knapsacks it is offered and pays the knapsack factory an amount equal to the market value of the fillers in each knapsack. The knapsacks themselves have no value or what is the same thing, a constant value which is independent of their eventual contents. The profit generated by the knapsack is this revenue minus the cost of stuffing. We can assume the stuffing cost is either a constant for a given knapsack or is directly proportional to the numbers of each individual filler type, that is, it costs a certain amount to stuff one unit of any particular filler type. We can define the value of a filler type to be the price paid by the purchaser minus this stuffing cost, so that we can consider the profit for a knapsack as a linear function of the value of the fillers. These assumptions are not really necessary in our formulation but serve to make the ideas more concrete.

The knapsacks are obtained from a knapsack supplier and stored in a large warehouse. During each shift in the factory, a set of M knapsacks is delivered to the knapsack stuffer who packs them one by one. It may be that the stuffer can or cannot control the order in which he receives the knapsacks. There are K different types of knapsacks. For example, in the usual linear integer programming knapsack problem, it may be that the "volume" of each knapsack type is different. Within each shift of M knapsacks there is a mix of the K types of knapsacks, $(M_1, M_2, \ldots, M_K)$ and the sum of the M_i is M. The stuffer has available to him N types of fillers, where N is finite, some of which are limited in availability as in our little scenario and others which are essentially unlimited.

The stuffer knows the value of each type of filler and is able to solve a generic knapsack problem for any given knapsack type as it arrives at his station. That is, he can find the optimal combination of fillers which fits and which maximizes the profit for that specific knapsack. This assumption is the one which makes assumptions about the underlying cost structure unnecessary. The difficulty is that he may not be able to fill each knapsack in this optimal manner when it arrives at his station because of various constraints which involve the entire set of M knapsacks. These constraints can include limited availability of fillers as well as constraints on total filler output over a shift or constraints generated by production limitations. Thus, the problem facing the stuffer is to maximize the profit over the entire set of knapsacks subject to meeting certain constraints. In addition, the problem may be viewed as deterministic or stochastic. For example, the number of knapsacks in a shift, M, or the mix of knapsack types may be random variables.

There are any number of variations possible within this framework. Rather than consider specific cases, the remainder of this section will discuss elements of the problem which can enter into specific variations. The case we consider in this paper will be formulated in Chapter II.

We should note for completeness that certain real industrial applications of knapsack type problems are in some sense the reverse of the case where the problem is to put things into a knapsack. That is, the plant receives raw material as discrete entities and the problem is to take things out, in the sense that the factory allocates or divides the total raw material entity into discrete products or subproducts in some optimal way. Examples of this sort would include forest products where the basic entity is a tree stem of given length which can only be cut into specific lengths from a finite set of possible lengths for further processing. That is, simply speaking, one can only cut a 50 foot log into smaller logs with lengths of 8 feet or 10 feet or 12 feet, etc. up to some finite limit, corresponding to the acceptable lumber lengths that will eventually result.

Elements of the Problem

We shall occasionally discuss these elements in the context of the second interpretation of the knapsack type problem, that where an entity is to be decomposed into smaller pieces. The discussion should make it clear which we are talking about. In any event, with the appropriate interpretation, the two approaches should be interchangeable.

A. Knapsacks (Raw Material).

1. A knapsack is a discrete entity with finite dimension.
2. Each entity can be filled (allocated) in a finite number of different ways, each of which realizes a positive return to the factory.
3. Each filling (allocation), i. e., a feasible solution to the underlying knapsack type problem, has a value which can be calculated.
4. An optimal allocation for an individual knapsack exists for a given set of prices and this solution is available to the

operator. That is, the stuffer knows or can compute the solution to the underlying generic knapsack problem. This optimal solution does not take into account constraints linking all the knapsacks in a sequence.

5. There are K different types of knapsacks, where K is not necessarily finite and there are M knapsacks to be processed in a given period of time, called a shift.
6. There exists a slack filler with a volume unit sufficiently small to ensure that the knapsack is completely full when packed (the raw material is completely used up). For example, scrap.

B. Value.

1. The values of the fillers are known and are non-negative including the slack filler type.
2. The cost of the knapsack (raw material) does not depend on the way it is filled (allocated) and thus is not involved in the decision to maximize the return realized.
3. If desired, the fixed costs of the factory can be assigned to each knapsack in a shift as part of the constant cost and so do not affect the decisions.
4. The variable production costs are either assignable to knapsacks independent of the manner in which they are filled or can be assigned directly to a knapsack in such a way that the

return or profitability of an individual filled knapsack can be calculated as a function of the fillers used. For example, as the sum of the values of each filler minus the cost of inserting that filler in the knapsack.

C. Constraints. There are two general classes of constraints that we consider. The first class comprises those constraints which pertain solely to each individual problem and do not link one knapsack to another. In addition to constraints similar to the overall volume constraint of the standard knapsack problem, one could imagine additional constraints such as not permitting filler type four in knapsack type five even though it will "fit" or requiring at least seven units of filler type 14 in a given knapsack. The point is that these constraints do not require information to be carried over from one knapsack to the next and thus only need be considered for each knapsack as it arrives, independent of the rest. We shall not concern ourselves with this type of constraint but assume that the solution to the underlying knapsack problem has taken all such constraints into account.

The second type of constraint includes those that link all M knapsacks together, which we shall refer to hereafter as factory constraints. This type of constraint can include:

1. Filler availability constraints. For example, there may only be 50 units of filler type four available per shift. In the context of allocating raw materials, such a constraint could

occur if a high level planning model assigns production quotas to a plant, thus directly setting a quota on the raw material allocation required by that plant. For example, the constraint may be to deliver no more than 400 units of filler type seven to plant A in a shift.

2. Demand requirements. For example, pack at least 500 units of filler type two per shift in order to satisfy demand. Alternatively, allocate 500 or more such units to satisfy the production goals of a plant which uses filler type two as raw material in its output.

3. Production constraints. Such constraints can take many forms and we assume that this type of constraint can be stated in terms of numbers of fillers of a certain type per shift. For example, consider constraints such as:

   a. Time limitations. It may take such a long time to fill one knapsack with filler type three that even though there are sufficient numbers of filler type three available, you cannot fill all M knapsacks optimally in the available time.

   b. A requirement to keep selected production facilities efficiently utilized. To illustrate, it may be that, as a result of higher level planning, a requirement exists for certain production facilities to be working full shifts, despite the relative lack of profit in their output. Such a constraint would require raw material to be allocated to these facilities in lieu of being allocated to more profitable production possibilities and could be expressed in terms of requirements for a raw material allocation different than that which would optimize each individual raw material entity.

Again, we consider only factory constraints which can be stated in terms of the number of certain filler types to be used during a given shift. This is not an unrealistic limitation since production limitations can be expressed in terms of limits on the amount of raw materials to be processed in a shift and demand or output requirements can generally be expressed as raw material input requirements.

D. Information. One of the elements that can characterize the problem faced by the knapsack stuffer is the amount of information available to him at each stage of the process. For example, we assume that the knapsack stuffer can solve the underlying knapsack problem for each knapsack as it arrives at his station so that he knows the optimal thing to do if he did not have to consider factory constraints. Clearly, he also knows the number of knapsacks he has already processed and thus, if M is known, he knows how many knapsacks remain to be processed. Furthermore, the operator could constantly be aware of the status of the factory constraints. For example, if there were only 100 units of filler type nine available and after 20 knapsacks the stuffer has used 38, then, of course, he is aware that only 62 more are available. In addition, as a consequence of the assumption that the stuffer can calculate the value of any feasible packing of a given knapsack (which may only be a simple addition process), the stuffer can be aware of the total value realized by his decisions up to that point. Therefore, it appears reasonable that, at any given point, the knapsack stuffer could know how many knapsacks he has processed, how much value he has realized, the optimal solution to the current knapsack problem disregarding the factory constraints and the status of those constraints; in particular, whether or not he has violated any of them.

E. Decision. The final element of the problem is the decision facing the knapsack factory, which is how to allocate fillers to the set of knapsacks being processed in a fixed period of time so as to maximize the return to the factory while satisfying factory constraints.

3. Scope of the Research

The research reported in this dissertation focussed on one of the many possible alternatives of the Generic Knapsack Factory Problem. Specifically, we consider here the deterministic case where the parameters of the problem are known. That is, the number of knapsacks, M, is specified as is K, the number of different types and M_k, the number of knapsacks of type k included in the shift of M

knapsacks. Associated costs and values are also assumed known. A few results applicable when the number of knapsacks is a random variable are presented. Furthermore, we only consider the case where there is one factory constraint expressed in terms of limitations on the numbers of a specified class of fillers. Finally, we consider two situations. The first, which involves only one type of knapsack is studied under the heading of the Basic Knapsack Factory Problem. This is perhaps the simplest of all knapsack factory problems and is useful in providing insights to more complicated situations. Chapter IV addresses the second case, allowing more than one type of knapsack, under the heading of the Complex Knapsack Factory Problem.

As will become clear in Chapter II and following sections we can formulate these problems as integer programming problems. As such they are amenable to solution, given specific values of the parameters, by existing techniques of integer programming which are discussed throughout the literature and in texts such as Garfinkel and Nemhauser [1]. Of course, there exists an extensive body of results and solution algorithms for the classical knapsack problem. For a current survey, see Salkin and deKluyver [7]. Garfinkel and Nemhauser devote a chapter in the reference cited above to the Knapsack Problem and Gilmore and Gomory [2] discuss the problem at length in their paper on the theory of Knapsack Functions. However, we point out at the outset that this paper does not concern itself with the theory, techniques or algorithms of integer programming. The focus of this paper is to investigate decision rules which can exploit the sequential nature of the problem, which can be stated in terms of the parameters and which can be practically applied in the sense that they are simple and easily computed. The paper will point out when these decision rules are optimal, when they might not be and the extent of the deviation from an optimal solution when appropriate.

While integer programming algorithms can be applied to our problem when specific values of the parameters are given, such algorithms do not necessarily give insight into the behavior of optimal or even good integer solutions as the parameters change. In this work we are interested in the case alluded to in the opening scenario, where our knapsack stuffer may need to know what to do under changing conditions and may not have the time to compute optimal or nearly optimal solutions using notoriously lengthy integer programming algorithms. Consequently, we wish to be able to state results in terms of the parameters of the problem and investigate the effects changes in the parameters have on the decision rules. The results obtained point out such effects and, of course, provide specific solutions given specific values. Furthermore, the interaction among the various parameters is an essential part of the analysis. In addition, we are not interested in integer programming approaches to the problem because in the applied environment it may be more desirable to implement a nearly optimal or "good" decision quickly arrived at rather than a perhaps costly optimal one which requires extensive computation only to realize a slight increase in return. The original thrust of the research was to determine if such decision rules were possible in the cases studied. Of course, we need a working definition of "simple" and "good". We consider, for the purposes of this paper, a "good" solution to be one where the return realized over the set of M knapsacks is within a specified margin of the optimal solution to the problem. Since we will formulate the problems as linear integer programs, we can say little in general about the value of the optimal solution. Therefore, we use as our standard of comparison the optimal solution to a linear programming (LP) relaxation of the integer program (IP) which is, of course, an upper bound on the optimal IP solution. We show that one can easily find an integer solution which realizes a return which has a difference from the optimal LP solution strictly less than the optimal value of one knapsack. This potential loss is independent of M so that the percentage potential loss goes to zero as the number of knapsacks increases.

As for "simple" we consider any solution which can be quickly computed by hand as simple, since such a solution clearly lends itself to extremely rapid computation on an electronic machine. We show that there are "good" decision rules which can be obtained by hand calculations even in the Complex Factory Case.

Thus, the flavor of this paper is that of heuristic decision rules which can be applied in practice with little difficulty.

4. Organization and Notation

Organization

We begin Chapter II by formulating the generic knapsack problem and function. It is this problem, which occurs sequentially in the factory problem, that we shall refer to as the underlying knapsack problem in future discussion. We remind the reader that we have assumed that the factory can solve this problem as it occurs. Then, the Basic Generic Knapsack Factory Problem, which in the interest of brevity we refer to as the Basic Problem, and the Complex Generic Knapsack Factory Problem, which we shall call the Complex Problem, are formulated in Sections 2 and 3 of Chapter II, respectively. Section 4 of Chapter II presents some useful preliminary results concerning the parameterized generic knapsack function defined in Section 1.

The Basic Problem is studied in Chapter III and the Complex Problem is studied in Chapter IV. The results obtained provide good and simple solutions to those problems in the spirit of Section 2 of Chapter I.

Referencing of results stated in different chapters will be accomplished by giving the chapter number followed by the section number followed by the designation of the result. For example Lemma II.4.1 refers to Lemma 1 of Section 4 of Chapter II. Of course, results will be numbered by type sequentially within a section. For results in the same chapter, the chapter reference will be omitted. Problem statements will be given a three digit number, the first digit of which specifies the chapter and the next two are ordered in sequence. Thus, problem (301) refers to the first problem formulated in Chapter III.

Notation

In those sections which involve the techniques of linear programming, we adopt notation similar to that used in Hadley [3]. Throughout this paper, for a, b integer we shall mean by $a \bmod b$ the unique integer defined by the following relations:

$$0 \leq a \bmod b < b$$

$$a = b[a/b] + a \bmod b,$$

where $[a/b]$ denotes the greatest integer less than or equal to a/b. Thus, for any integers a and b, $a \geq 0$, $b > 0$,

$$a = [a/b]b + a \bmod b$$

and so

$$[a/b] = a/b - a \bmod b/b.$$

When the divisor is clear from the discussion, the notation r_a will sometimes be used for $a \bmod b$.

II. PROBLEM FORMULATION AND PRELIMINARY RESULTS

1. A Generic Knapsack Problem

We shall formulate a rather general mathematical programming problem which we call the underlying knapsack problem. That is to say, it is this problem which occurs repeatedly during a given period of time (a "shift"). When viewed in its entirety, the sequence constitutes the knapsack factory problem when additional constraints are imposed on the sequence of individual problems. The following formulation was motivated originally by the classical knapsack problem:

$$\text{(201)} \qquad \text{maximize} \quad \sum_{i=1}^{N} c_i x_i$$

$$\text{subject to:} \quad \sum_{i=1}^{N} a_i x_i \leq V$$

$$a_i \geq 0, \quad x_i \geq 0, \quad x_i \text{ integer}, \quad i = 1, \ldots, N$$

where V is the "volume" of a knapsack, a_i is the "volume" of item i, c_i is the value of item i and the problem can be considered to represent that of maximizing the total value of a set of discrete items which "fit" into the knapsack. However, the results of this paper pertain to a much more general class of mixed integer programming problems which can underly the factory problem. Hence, we employ the phrase, generic knapsack problem. It will be noted that the classical knapsack problem is a special case of the following:

Let R^{N+} be the set of N dimensional, nonnegative vectors and let $y = (y_1, \ldots, y_N) \in R^{N+}$.

We can consider y_n to be the number of fillers of type n, $n = 1, \ldots, N$. Let $f(y)$ be the value of a knapsack with y_n of fillter type n in it, $n = 1, \ldots, N$ $(f: R^{N+} \rightarrow R)$. It makes sense to require $f(0) = 0$. Let the bounded set $Y \subset R^{N+}$ represent, in the abstract, the "volume" of a knapsack in the sense that $y \in Y$ implies that y fits. Of course, in more usual terminology, Y is some feasible region. Furthermore, $y \in Y$ implies that y_i is finite, $i = 1, \ldots, N$. Then we can write a generic knapsack problem as:

$$\text{(202)} \qquad \text{Find} \qquad Z(Y) = \max_{y \in Y} f(y)$$

Since we are concerned with industrial anologies to the traditional knapsack problem, we shall assume that the maximum exists in order to preclude discussions about pathological cases. Certainly, in the case where $y \in Y$ implies that y_i is an integer for all i, this assumption need not be made since the maximum does exist (Y is bounded). The set function specified by (202) will be called the generic knapsack function.

We have an interest in a subset of the filler types. Accordingly, let I be a special index set, $I \subset \{1, \ldots, N\}$. We require that $i \in I$ implies y_i is integer valued. We do not require y_i to be an integer if $i \notin I$, since we have assumed the maximum exists in (202). We shall make use of the fact that the $\max_{y \in Y} y_i$ exists for all $i \in I$, since Y is bounded.

We define a parameterized set (knapsack "volume") as follows:

$$Y(j) = \{y : y \in Y \text{ and } \sum_{i \in I} y_i = j\}, \quad j = 0, 1, \ldots, \quad .$$

As an immediate consequence of the finiteness of the y_i, $i \in I$ there exists a finite j_{max} such that for all j' greater than j_{max}, $\sum_{i \in I} y_i = j'$ implies that $y \notin Y$. In other words, the class $\{Y(j)\}$ is finite, also. Let us define the parameterized generic knapsack function

$$Z(j, Y) = \max_{y \in Y(j)} f(y) \tag{203}$$

or alternately,

$$Z(j, Y) = \max f(y)$$

subject to

$$y \in Y$$

and

$$\sum_{i \in I} y_i = j \; ,$$

which we assume to exist for each feasible j, $0 \leq j \leq j_{max}$. By the existence of j_{max} and the integer nature of y_i for $i \in I$, there exists a j, call it j^*, such that $Z(j^*, Y) = Z(Y)$. That is, the solution to (202) must occur at some j between 0 and j_{max}. Hereafter, we shall use $Z(j^*)$ to indicate this global optimum. Since we have assumed that the solution to (203) can be obtained for a specific knapsack, i.e. one whose Y is known, we shall suppress the dependence of the solution on the particular volume, Y, and write $Z(j, Y)$ as $Z(j)$. We see that j_{max} represents the maximum aggregate number of fillers in our special class, I, which can "fit" in our knapsack of finite dimension. The interpretation placed on any $Z(j)$ is this: given that we require the sum of the units of the filler types in the special class be equal to j, we have otherwise "filled" the knapsack in an optimal manner. We shall hereafter use the phrase: "fill the knapsack with j" to mean fill the knapsack with some set of fillers which corresponds to the solution to (203). Any such solution therefore contains exactly j units of the filler types which are in the special set.

The reader can anticipate that our real interest is in the special class of fillers, stemming from the fact that these will represent the types which are involved in the factory constraint imposed on the sequence of underlying knapsack problems. We further note that we can consider a knapsack type to correspond to a "Volume," Y. That is, our generic knapsacks are characterized by their feasible set, Y, and so, if we have K different types of knapsacks, we mean that there are K distinct sets Y_k, $k = 1, \ldots, K$, with which any given knapsack can be associated. In the case where we are interested in more than one type of knapsack, we shall only refer to this dependence on the associated "Volume" by a subscript, $Z_k(j)$. That is $Z_k(j) = Z(j, Y_k)$. Furthermore, there is nothing which prevents the use of different value functions, say $f_k(\cdot)$, for different knapsack types but we shall not indicate this distinction in our notation.

EXAMPLE: The following example may serve to illustrate the formulation and to demonstrate that the usual knapsack function is a special case. Let $K = 1$ and let

$$Y = \{y : \sum_{i=1}^{N} a_i y_i = b,\ y_i \geq 0,\ y_i \text{ integer } i = 1, \ldots, N\}$$

where $a_i \geq 0$, $i = 1, \ldots, N$. Furthermore, let

$$f(y) = \sum_{i=1}^{N} c_i y_i$$

where $c_i \geq 0$, $i = 1, \ldots, N$. Then

$$Z = \max \sum_{i=1}^{N} c_i y_i$$

$$\text{s.t.} \qquad \sum_{i=1}^{N} a_i y_i = b$$

$$y_i \geq 0, \; y_i \text{ integer}, \; i = 1, \ldots, N$$

is our generic knapsack function. Gilmore and Gomory [2] make explicit the dependence on b by denoting $Z(b)$ as the knapsack function in their paper. Then, let $I = \{1\}$. The implication is we are interested in filler type 1 and so we parameterize according to the value of y_1 as follows:

$$Z(j) = \max \sum_{i=2}^{N} c_i y_i + j c_1$$

$$\text{s.t.} \qquad \sum_{i=2}^{N} a_i y_i = b - j a_1$$

$$y_i \geq 0, \; y_i \text{ integer}, \; i = 2, \ldots, N$$

we have that $j_{max} = [b/a_1]$. Specifically, look at

$$\begin{array}{ll} \text{max.} & 25y_1 + 30y_2 + y_3 \\ \text{s.t.} & 2y_1 + 3y_2 + y_3 = 10 \\ & y_i \geq 0,\ y_i \text{ integer},\ i = 1, 2, 3. \end{array}$$

Now $Z(j^*) = 125$ and the optimal solution corresponding to this value is $y^* = (5, 0, 0)$. We set $I = \{1\}$ so we have, for the parameterized generic knapsack function, $Z(j)$:

j	$Z(j)$	y^*
0	91	(0, 3, 1)
1	87	(1, 2, 2)
2	110	(2, 2, 0)
3	106	(3, 1, 1)
4	102	(4, 0, 2)
5	125	(5, 0, 0)

In this example, $j_{max} = 5$ and $j^* = 5$. Certainly, this may not always be the case. In fact, if we let $I = \{2\}$ we see that $j_{max} = 3$ but $j^* = 0$. The example serves to illustrate that in general very little can be specified about the form of the parameterized knapsack function even in the case of the usual linear knapsack problem. Note that here, even though $j^* = 5 > 0$, that $Z(0) > Z(1)$ (and $Z(2) > Z(3) > Z(4)$). We shall study further the form of the parameterized generic knapsack function in Section 4.

2. The Basic Knapsack Factory Problem

Now that we have defined the underlying knapsack problem, we are in a position to formulate the simplest case of a factory problem which we will consider. The mathematical model describes a situation where the knapsack stuffer receives M identical knapsacks in a shift. For each knapsack, he is able to calculate the parameterized generic knapsack function, $Z(j)$, $j = 0, \ldots, j_{max}$. A constraint has been imposed on the aggregate number of units of fillers of a certain type which are to be used over the sequence of M knapsacks.

Given a constraint on the allowable numbers of filler types in a restricted class over a shift, the stuffer wishes to maximize the value of the M knapsacks while satisfying the constraint. We will distinguish two types of factory constraint, a lower bound and an upper bound on the total number of restricted filler types. We denote the value of this bound on fillers in the class I by b. We will discuss the formulation in the context of an upper bound and merely state the alternate formulation at the conclusion of this section. We will show that under realistic assumptions, an equality constraint, as in a production goal, falls into one or the other of these two categories depending on the nature of the parameterized generic knapsack function and the value of M.

Consider a direct formulation. Let y^m, $y^m \in R^{N+}$, be the solution used in knapsack m, $m = 1, \ldots, M$. Then we wish to maximize

$$(204) \qquad \sum_{m=1}^{M} f(y^m)$$

subject to $\qquad y^m \in Y, \quad m = 1, \ldots, M$

and $\qquad \sum_{m=1}^{M} \sum_{i \in I} y_i^m \leq b .$

Recall that $y^m \in Y$ implies $y^m \geq 0$ and $i \in I$ implies y_i^m integer, $m = 1, \ldots, M$. Fix m and note that if we wish to maximize $f(y^m)$ we will certainly choose a solution from among the set of y^ms corresponding to $Z(j)$, $j = 0, \ldots, j_{max}$, since the class $\{Y_j = \{y : y \in Y$ and $\sum_{i \in I} y_i = j\} \}$ for $j = 0, \ldots, j_{max}$ forms a partition of Y. To see this, assume we have an optimal solution to the factory problem, (204), say $(y^{1*}, y^{2*}, \ldots, y^{M*})$. Certainly, for each $m = 1, \ldots, M$ there exists some integer, j^{m*}, $0 \leq j^{m*} \leq j_{max}$ such that $\sum_{i \in I} y_i^{m*} = j^{m*}$. This follows from the integer requirement on y_i^m for all $i \in I$ and the finite dimension of Y. Assume that y^{m*} does not correspond to any solution resulting in $Z(j^{m*})$. Certainly, $f(y^{m*}) < Z(j^{m*})$ by definition. So without affecting the feasibility of the solution set, $(y^{1*}, \ldots, y^{M*})$, we can substitute the solution which does correspond to $Z(j^{m*})$ for y^{m*} and realize an improvement in the objective function. Therefore, y^{m*} could not have been a component of the optimal solution to (204), contrary to our assumption.

Now, let $t_j^m = 1$ if the m^{th} knapsack is filled with j, i.e. $f(y^m) = Z(j)$, $y^m \in Y$ and $\sum_{i \in I} y_i^m = j$. Let $t_j^m = 0$, otherwise.

Then (204) can be written as:

(204a) maximize $\displaystyle\sum_{m=1}^{M} \sum_{j=0}^{j_{max}} t_j^m Z(j)$

subject to $\displaystyle\sum_{j=0}^{j_{max}} t_j^m = 1, \quad m = 1, \ldots, M$

(which implies $\displaystyle\sum_{m=1}^{M} \sum_{j=0}^{j_{max}} t_j^m = M$) and

$$\sum_{m=1}^{M} \sum_{j=0}^{j_{max}} j t_j^m \leq b$$

$$t_j^m = 0, 1.$$

Considering the optimal solution to this problem, we observe that if

$t_k^1 = 1$ and $t_\ell^2 = 1$, we can realize another optimal solution by setting $t_\ell^1 = 1$, $t_k^1 = 0$ and $t_k^2 = 1$, $t_\ell^2 = 0$. The point is that the order of the solution is unimportant in the sense that we can arbitrarily interchange the superscript indices. In fact, there are always numerous alternative optima. Since M and j_{max} are both finite, we can interchange the order of summation and, for fixed j, we can set $\sum_{m=1}^{M} t_j^m = X_j$ where X_j can be interpreted as the number of knapsacks "filled with j." Since the constraint $\sum_{j=0}^{j_{max}} t_j^m = 1$ assures that there is only one solution for each knapsack, we can achieve the same end by setting $\sum_{j=0}^{j_{max}} X_j = M$ and arbitrarily assigning knapsacks filled with j to any X_j of the M knapsacks.

The point of the preceding discussion is to motivate the following equivalent formulation of the Basic Problem, which is the form we will use because of its direct interpretation and tractability. Let X_j be the number of knapsacks "filled with j." The precise meaning of "filled with j" is given in Section 1. Then the Basic Problem is

$$\text{(205)} \qquad \text{maximize} \quad \sum_{j=0}^{j_{max}} X_j Z(j)$$

$$\text{subject to} \quad \sum_{j=0}^{j_{max}} X_j = M$$

$$\sum_{j=1}^{j_{max}} jX_j \leq b$$

$$X_j \geq 0, \ X_j \text{ integer}$$

which we see is a pure integer programming problem with two constraints.

Further, as a consequence of the fact that the stuffer knows the value of M and of $Z(j)$ for all j, this formulation shows us that the order in which the stuffer implements his solution is not relevant. We shall assume here that he will implement the solution in decreasing order of $Z(j)$. That is, he will fill the first knapsack with the largest value of $Z(j)$ in the solution and proceed towards the smallest. This procedure would make sense if M in fact were random.

Now we observe that the number of knapsacks filled with 0, X_0, does not affect the feasibility with respect to the factory constraint of any particular solution, so that if we set

$$X_0 = M - \sum_{j=1}^{j_{max}} X_j \geq 0$$ and write (205) as

(205a) maximize $$\sum_{j=1}^{j_{max}} X_j(Z(j)-Z(0)) + MZ(0)$$

subject to $$\sum_{j=1}^{j_{max}} X_j \leq M$$

and $$\sum_{j=1}^{j_{max}} jX_j \leq b$$

$$X_j \geq 0,\ X_j \text{ integer},\ j = 1,\ldots,j_{max} .$$

This is an equivalent formulation which reflects the fact that $X_0 = M$, $X_j = 0 \;\; j \geq 1$, is a feasible solution to (205). We can look for better solutions by searching over X_j, $j \geq 1$ for feasible improvements according to the incremental increase $Z(j) - Z(0)$, and replacing a knapsack filled with 0 by one filled with j as long as $Z(j) - Z(0) > 0$ and the second constraint is not violated. Finally, we note that if $M \geq b$, certainly $\sum_{j=1}^{j_{max}} X_j \leq \sum_{j=1}^{j_{max}} jX_j \leq b \leq M$ so that, in this case, the first constraint is implied by the second and we can write (205a) as a classical knapsack problem:

$$\text{maximize} \quad \sum_{j=1}^{j_{max}} X_j(Z(j)-Z(0))$$

$$\text{subject to} \quad \sum_{j=1}^{j_{max}} jX_j \leq b$$

$$X_j \geq 0,\ X_j \text{ integer},\ j = 1,\ldots,j_{max}\ .$$

In the lower bound case, we write (205) as

$$\text{(206)} \qquad \text{maximize} \quad \sum_{j=0}^{j_{max}} X_j Z(j)$$

$$\text{subject to} \quad \sum_{j=0}^{j_{max}} X_j = M$$

$$\text{and} \quad \sum_{j=1}^{j_{max}} jX_j \geq b$$

$$X_j \geq 0,\ X_j \text{ integer for all } j.$$

3. The Complex Knapsack Factory Problem

In this case, we have K distinct types of knapsacks and the factory processes M knapsacks in a shift as before, but here, M_k

is the number of knapsacks of type k processed in a shift, $k = 1, \ldots, K$ and of course $\sum_{k=1}^{K} M_k = M$. We must distinguish between types of knapsacks when we consider the parameterized knapsack function and so we should write $Z_k(j_k)$ where $0 \leq j_k \leq j_{max}(k)$. That is, we allow the possibility for $j_{max}(k)$ to be different for each k. However, since the subscript on j is redundant, we shall merely write $Z_k(j)$. Thus, the expression $Z_k(j^*)$ shall be understood to mean $Z_k(j_k^*)$, etc. We will still write j_k when the context does not make the meaning clear. Let X_{kj} be the number of knapsacks of type k filled with j. Then we can write the Complex Problem as:

(207) maximize
$$\sum_{k=1}^{K} \sum_{j=0}^{j_{max}(k)} X_{kj} Z_k(j)$$

subject to
$$\sum_{j=0}^{j_{max}(k)} X_{kj} = M_k, \qquad k = 1, \ldots, K$$

and
$$\sum_{k=1}^{K} \sum_{j=1}^{j_{max}(k)} j X_{kj} \leq b \quad (\geq b)$$

$$X_{kj} \geq 0, \; X_{kj} \text{ integer for all } k \text{ and } j, \; b > 0.$$

4. Preliminary Results

We shall see in later sections that the shape of the parameterized knapsack function has a major role in the search for good or optimal solutions. This function maps a finite set of integers into the real numbers. Figure II. 4. 1 provides a representation in a graph displaying the ordered pairs $(j, Z(j))$ for $j = 0, 1, \ldots, j_{max}$. We shall shortly show that in the upper bound case we need only consider solutions corresponding to $j \leq j^*$. Since we have assumed that $f(y) \geq 0$ for all y in R^{N+} then certainly, if $Y(0) \neq \phi$, the null set, $Z(0) \geq 0$. We can define $Z(0) = 0$ in the case where $Y(0) = \phi$. In fact, we shall arbitrarily define $Z(j) = 0$ whenever $Y(j) = \phi$ without affecting the analysis. Also, it is clear that $Z(j^*) \geq Z(0)$.

We need to establish the following results with respect to an optimal solution to the Basic Problem in order to motivate further study of the parameterized knapsack function. In the first case, the upper bound case, the result simply establishes the intuitive idea that we would never fill a knapsack with more of the restricted units than the number which appear in the optimal solution to the underlying problem. Similarly, in the lower bound case, we would never fill a knapsack with less than the number of restricted units called for in the optimal solution to the underlying knapsack problem.

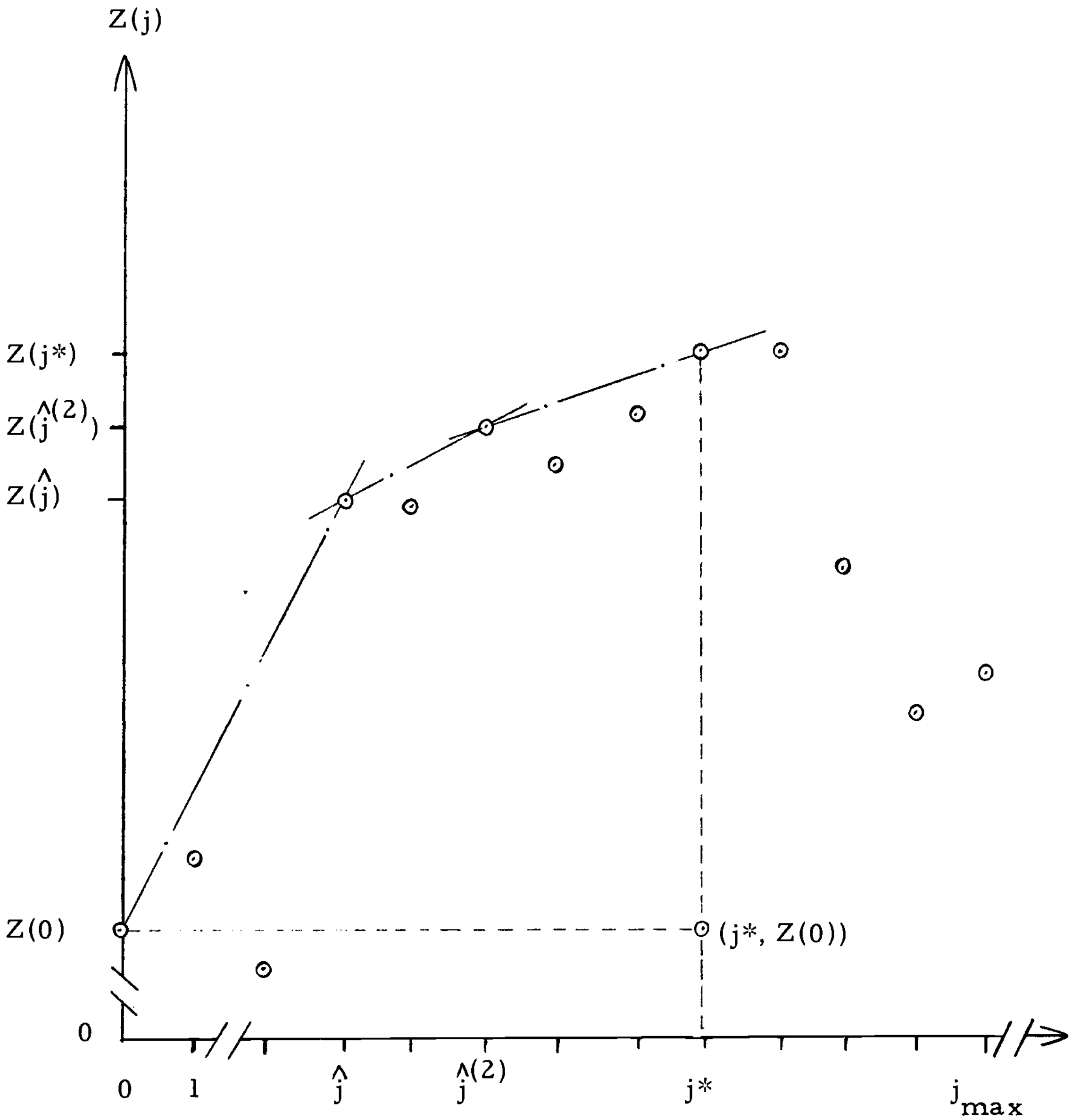


Figure II. 4. 1. Parameterized knapsack function.

We establish that an optimal solution exists to (205) in the upper bound case. Clearly, $X_0 = M$, $X_j = 0$, $j \geq 0$ is a feasible solution. Since the number of possible combinations which satisfy the constraints is finite, in theory we can find the optimal solution by exhaustive (and perhaps exhausting) search. We are forced to add an additional assumption to our formulation of the lower bound case, Basic Problem (207), namely, that the product Mj_{max} exceeds b. Otherwise, no feasible solution exists and the problem has no content in the sense that nothing the stuffer can do will satisfy the lower bound requirement. Given the assumption, then a feasible solution to (206) exists, $X_{j_{max}} = M$, $X_j = 0$, $j < j_{max}$ and thus an optimal solution exists.

First, we must choose a precise value of j^* in the event that there are multiple values of j which correspond to the optimal solution, Z. In the upper bound case, we will always choose the smallest such j. That is, we choose j^* such that $j < j^*$ implies $Z(j) < Z(j^*)$. Alternately, in the lower bound case, we choose the largest such j so that for $j > j^*$, $Z(j) < Z(j^*)$. Consider the upper bound case first.

PROPOSITION 4. 1. There exists an optimal solution to (205) in which $X_j = 0$ for all $j > j^*$.

PROOF. Obvious from the definition of j^*.

COROLLARY 1. If j^* is unique, then there exists no optimal solution in which $X_j > 0$ for $j > j^*$.

PROOF. Obvious.

Consequently, our search for an optimal solution to (205) need only involve filling knapsacks with $j \leq j^*$ and we can ignore solutions which involve filling knapsacks with $j > j^*$. By extension, we will restrict our search for good solutions to the same candidates.

Because of the symmetry, we can establish the following results for the lower bound case. Recall that here j^* is the largest j such that $Z(j) = Z$.

PROPOSITION 4. 2. There exists an optimal solution to (206) in which $X_j = 0$ for all $j < j^*$ (if we assume that $Mj_{max} \geq b$. Otherwise there exists no solution to (206)).

COROLLARY 1. If j^* is unique, then there exists no optimal solution to (206) in which $X_j > 0$ for $j < j^*$.

Analogously, we have established that we can restrict our search for optimal or good solutions to (206) to those involving filling knapsacks with $j \geq j^*$.

Our study in future sections will show that a quantity of particular importance in the analysis is the incremental return per unit of restricted fillers realized by increasing from j_1 to j_2, which is simply the quantity $(Z(j_2)-Z(j_1))/(j_2-j_1)$. For the upper bound problem, we are particularly interested in the case where $j_1 = 0$, that is we are interested in $(Z(j)-Z(0))/j$ for $j > 0$. To avoid future incongruities, we will arbitrarily define $(Z(j)-Z(0))/j$ to be 0 for $j = 0$. In the upper bound case, we will show that we are really only interested in certain values of j which we will define recursively as follows:

<u>DEFINITION 4.1</u>. Let $\hat{j}^{(1)}$ be the largest j such that

1) $0 \leq \hat{j}^{(1)} \leq j^*$ and

2) $$\frac{Z(\hat{j}^{(1)})-Z(0)}{\hat{j}^{(1)}} \geq \frac{Z(j)-Z(0)}{j} \quad \text{for all } j > 0.$$

From here on we shall drop the particular superscript $n = 1$. We find $\hat{j}$ by searching for the $\max\limits_{0 \leq j \leq j^*} \frac{Z(j)-Z(0)}{j}$.

<u>NOTE</u>: If $j^* = 0$, then $\hat{j} = j^* = 0$. Then, given $\hat{j}$, we define $\hat{j}^{(2)}$ to be the largest j such that

1) $\hat{j} < \hat{j}^{(2)} \leq j^*$ and

2) $$\frac{Z(\hat{j}^{(2)})-Z(\hat{j})}{\hat{j}^{(2)}-\hat{j}} \geq \frac{Z(j)-Z(\hat{j})}{j-\hat{j}} \qquad j > \hat{j}.$$

Similarly, we define $\hat{j}^{(3)}$, $\hat{j}^{(4)}$, etc. The sequence of $\hat{j}^{(n)}$'s is finite, because at some point it must be that $j^* = \hat{j}^{(\bar{n})}$ for some finite $\bar{n} \geq 1$ and we stop.

PROPOSITION 4.3. For any parameterized function $Z(\cdot)$, we are guaranteed the existence of at least $\hat{j}$ (i.e. $\bar{n} \geq 1$).

PROOF. Consider two cases, $j^* = 0$ or $j^* > 0$.

If $j^* = 0$ then, of course $Z(j) \leq Z(0)$ so that $\frac{Z(j)-Z(0)}{j} \leq 0$ for all j. Since we have defined $\frac{Z(0)-Z(0)}{0}$ to be 0 then we can set $\hat{j} = 0$ without violating the definition. Note that our convention is to choose as $\hat{j}$ the largest $j \leq j^*$ such that $\frac{Z(j)-Z(0)}{j} \geq \frac{Z(i)-Z(0)}{i}$ for all $i \leq j^*$ which is consistent here.

If $j^* > 0$ then, of course, we have at least one candidate for $\hat{j}$, namely j^*, since $\frac{Z(j^*)-Z(0)}{j^*} > 0$.

See Figure 4.1 for a graphic description of the $\hat{j}^{(n)}$'s in that example. The figure shows that we are selecting those extreme points of the convex hull generated by a finite set of bounded points in R^2, $\{(0, Z(0)), (1, Z(1)), \ldots, (j^*, Z(j^*)), (j^*, Z(0))\}$ which lie above the line $j = Z(0)$. Since this particular statement is not relevant to future discussions, we will not establish it formally.

We shall provide the analogous definitions for the lower bound case in Section 3 of Chapter III.

PROPOSITION 4.4. Define $\hat{j}^{(0)} = 0$. For $j^* > 0$,

$$\frac{Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} > 0 \qquad \text{for all} \quad n \leq \bar{n}.$$

PROOF. Let $n = 1$. Then $\frac{Z(\hat{j})-Z(0)}{\hat{j}} \geq \frac{Z(j^*)-Z(0)}{j^*}$ and since $Z(j^*) > Z(0)$ by definition the result holds. Suppose $\hat{j} = j^*$, then we are done. So assume $\hat{j} < j^*$. Then for $n = 2$,

$$\frac{Z(\hat{j}^{(2)})-Z(\hat{j})}{\hat{j}^{(2)}-\hat{j}} \geq \frac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}}$$

and again, by the definition of j^* and the fact that $\hat{j} < j^*$, $\frac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}} > 0$. This argument holds for all $\hat{j}^{(n)} < j^*$. Finally, we will reach $\bar{n}$ where $\hat{j}^{(\bar{n})} = j^* > \hat{j}^{(\bar{n}-1)}$ but

$$\frac{Z(\hat{j}^{(\bar{n})})-Z(\hat{j}^{(\bar{n}-1)})}{\hat{j}^{(\bar{n})}-\hat{j}^{(\bar{n}-1)}} = \frac{Z(j^*)-Z(\hat{j}^{(\bar{n}-1)})}{j^*-\hat{j}^{(\bar{n}-1)}} > 0$$

and the result holds.

One of the implications of Proposition 4.4 is that whenever $j^* > 0$, we can always find a better solution than $X_0 = M$, $X_j = 0$, $j > 0$. Another is that we will approach the constraint as close as the integer nature of the variables will permit. That is, loosely speaking, the optimal solution will try to realize equality in the factory constraint on the total number of special fillers. We also note here that $j^* > 0$ is no limitation on the problem because the case where

$j^* = 0$ in problem (205) means the factory constraint is not binding. Thus, we can ignore it by filling each knapsack optimally, i.e. with $j^* = 0$. We shall discuss these points in detail in Chapter III.

We are now ready to establish the most useful tool in our subsequent analysis.

LEMMA 4.1. By definition

$$\frac{Z(\hat{j})-Z(0)}{\hat{j}} \geq \frac{Z(j)-Z(0)}{j} \quad \forall j, \quad 0 \leq j \leq j^*$$

Then

1) $\dfrac{Z(\hat{j})-Z(0)}{\hat{j}} \geq \dfrac{Z(j)-Z(0)}{j} \qquad \forall j \leq j_{max}$

2) If $0 < j < \hat{j}$, then $Z(\hat{j}) - Z(j) > 0$ and

$\dfrac{Z(j)-Z(0)}{j} \leq \dfrac{Z(\hat{j})-Z(0)}{\hat{j}} \leq \dfrac{Z(\hat{j})-Z(j)}{\hat{j}-j}$

3) If $j > \hat{j}$, then $\dfrac{Z(\hat{j})-Z(j)}{\hat{j}-j} < \dfrac{Z(j)-Z(0)}{j} < \dfrac{Z(\hat{j})-Z(0)}{\hat{j}}$

PROOF.

1) $\dfrac{Z(\hat{j})-Z(0)}{\hat{j}} \geq \dfrac{Z(j^*)-Z(0)}{j^*}$ by definition but $Z(j^*) \geq Z(j)$ for all $j \leq j_{max}$ implies that $Z(j^*) - Z(0) \geq Z(j) - Z(0)$ for all such j. Therefore, for any $j > j^*$, $\dfrac{Z(j^*)-Z(0)}{j^*} \geq \dfrac{Z(j)-Z(0)}{j^*} > \dfrac{Z(j)-Z(0)}{j}$, establishing the result. The result is meaningless for $j > j_{max}$.

2) Since $0 < j < \hat{j}$, then $\dfrac{Z(j)-Z(0)}{j} > \dfrac{Z(j)-Z(0)}{\hat{j}}$. Now $\dfrac{Z(\hat{j})-Z(0)}{\hat{j}} \geq \dfrac{Z(j)-Z(0)}{j}$ implies that $\dfrac{Z(\hat{j})-Z(0)}{\hat{j}} > \dfrac{Z(j)-Z(0)}{\hat{j}}$ giving us

that $Z(\hat{j}) - Z(0) > Z(j) - Z(0), \quad \forall j < \hat{j}$ or $Z(\hat{j}) - Z(j) > 0$ which establishes the first part of 2).

Of course, we have by definition that $\frac{Z(\hat{j})-Z(0)}{\hat{j}} \geq \frac{Z(j)-Z(0)}{j}$. Multiplying both sides by the positive quantity $j/(\hat{j}-j)$ gives us that

$$\frac{Z(\hat{j})-Z(0)}{\hat{j}-j}\left(\frac{j}{\hat{j}}\right) \geq \frac{Z(j)-Z(0)}{\hat{j}-j}$$

or, letting $j = \hat{j} + (j-\hat{j})$

$$\frac{Z(\hat{j})-Z(0)}{\hat{j}-j}\left(1 - \frac{(\hat{j}-j)}{\hat{j}}\right) \geq \frac{Z(j)-Z(0)}{\hat{j}-j}$$

subtracting $\frac{Z(\hat{j})-Z(0)}{\hat{j}-j}$ from both sides, we get

$$-\left(\frac{Z(\hat{j})-Z(0)}{\hat{j}}\right) \geq \frac{Z(j)-Z(\hat{j})}{\hat{j}-j}$$

or

$$\frac{Z(\hat{j})-Z(0)}{\hat{j}} \leq \frac{Z(\hat{j})-Z(j)}{\hat{j}-j} ,$$

the desired result.

3) For $j > \hat{j}$, we have $\hat{j} - j < 0$, and by part 1, if $j > j^*$, $\frac{Z(\hat{j})-Z(0)}{\hat{j}} > \frac{Z(j)-Z(0)}{j}$. Multiplying both sides by $\frac{j}{\hat{j}-j} < 0$, we get that

$$\frac{Z(\hat{j})-Z(0)}{\hat{j}}\left(\frac{j}{\hat{j}-j}\right) < \frac{Z(j)-Z(0)}{\hat{j}-j}$$

or

$$\frac{Z(\hat{j})-Z(0)}{\hat{j}-j}\left(\frac{j}{\hat{j}}\right) < \frac{Z(j)-Z(0)}{\hat{j}-j}$$

If we multiply through by $\hat{j}/j > 0$, we get

$$\frac{Z(\hat{j})-Z(0)}{\hat{j}-j} < \frac{\hat{j}}{j}\left(\frac{Z(j)-Z(0)}{\hat{j}-j}\right) = \left(1+\frac{(\hat{j}-j)}{j}\right)\frac{Z(j)-Z(0)}{\hat{j}-j}$$

and subtracting $\frac{Z(j)-Z(0)}{\hat{j}-j}$ from both sides, we have that $\frac{Z(\hat{j})-Z(j)}{\hat{j}-j} < \frac{Z(j)-Z(0)}{j}$. Observing that $\frac{Z(j)-Z(0)}{j} < \frac{Z(\hat{j})-Z(0)}{\hat{j}}$ by our choice of $\hat{j}$, we have the desired result.

We will use the following corollaries in later analysis.

<u>COROLLARY 1.</u> For $j < \hat{j}$, $\frac{Z(\hat{j})-Z(j)}{\hat{j}-j} > 0$.

<u>PROOF.</u> This follows immediately from Proposition 4.4 (i.e. $\frac{Z(\hat{j})-Z(0)}{\hat{j}} > 0$) and 2) above.

<u>COROLLARY 2.</u> If $\frac{Z(\hat{j})-Z(0)}{\hat{j}} \geq A$, then for all $j < \hat{j}$, $\frac{Z(\hat{j})-Z(j)}{\hat{j}-j} \geq A$.

<u>PROOF.</u> The result is an immediate consequence of part 2) above.

This particular corollary will prove useful when we consider the Complex Problem.

COROLLARY 3. $Z(j) \geq Z(\hat{j})$ implies that $j > \hat{j}$.

PROOF. By the contrapositive to the first part of 2) in the lemma.

COROLLARY 4. If $\frac{Z(\hat{j})-Z(0)}{\hat{j}} \leq A$ then there does not exist a $j > \hat{j}$ such that $\frac{Z(j)-Z(\hat{j})}{j-\hat{j}} > A$.

PROOF. Follows from part 3) of the lemma.

Finally, we establish the result for a more general case as follows:

LEMMA 4.2. By definition, for $n \leq \bar{n}$

$$\frac{Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} \geq \frac{Z(j)-Z(\hat{j}^{(n-1)})}{j-\hat{j}^{(n-1)}}$$

for all j such that $\hat{j}^{(n-1)} < j \leq j^*$. As a consequence,

1) $\frac{Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} \geq \frac{Z(j)-Z(\hat{j}^{(n-1)})}{j-\hat{j}^{(n-1)}}$ for all j such that $\hat{j}^{(n-1)} < j \leq j_{max}$.

2) If $\hat{j}^{(n-1)} < j < \hat{j}^{(n)}$, then (a) $Z(\hat{j}^{(n)}) > Z(j)$ and

(b) $\frac{Z(j)-Z(\hat{j}^{(n-1)})}{j-\hat{j}^{(n-1)}} \leq \frac{Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} \leq \frac{Z(\hat{j}^{(n)})-Z(j)}{\hat{j}^{(n)}-j}$

3) If $\hat{j}^{(n)} < j$, then

$$\frac{Z(\hat{j}^{(n)})-Z(j)}{\hat{j}^{(n)}-j} < \frac{Z(j)-Z(\hat{j}^{(n-1)})}{j-\hat{j}^{(n-1)}} < \frac{Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}-\hat{j}^{(n-1)}}$$

PROOF. By substituting $\hat{j}^{(n-1)}$ for 0, $\hat{j}^{(n)}$ for $\hat{j}$ in the argument for Lemma 4.1, the proof, which is merely algebraic manipulation, will go through in precisely the same way.

COROLLARY 1. $\dfrac{Z(j)-Z(\hat{j}^{(n-1)})}{j-\hat{j}^{(n-1)}} > \dfrac{Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}-\hat{j}^{(n-1)}}$ for all $j > \hat{j}^{(n-1)}$.

PROOF. By part 2) letting $\hat{j}^{(n-1)}$ play the role of $\hat{j}^{(n)}$, we have $\dfrac{Z(j)-Z(\hat{j}^{(n-1)})}{j-\hat{j}^{(n-1)}} \geq \dfrac{Z(\hat{j}^{(n-1)})-Z(\hat{j}^{(n-2)})}{\hat{j}^{(n-1)}-\hat{j}^{(n-2)}}$ and by part 3), since $\hat{j}^{(n)} > \hat{j}^{(n-1)}$, we see that

$$\frac{Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} < \frac{Z(\hat{j}^{(n-1)})-Z(\hat{j}^{(n-2)})}{\hat{j}^{(n-1)}-\hat{j}^{(n-2)}}$$

which implies the desired result.

These two lemmas will provide the elementary results we need to analyze the Basic and Complex Problems in later chapters. In general, without more restrictive assumptions, we can say little more about the shape of parameterized generic knapsack functions. Nonetheless, we shall find it more efficient in future discussion to define particular characteristics of such discrete functions in the same spirit as the usual definition of convexity. Consider the following definitions of discrete convexity.

<u>DEFINITION 4.2</u>. The discrete valued function $Z(j)$ defined on integers is locally discretely convex (LDC) at the point $\bar{j}$, if

1) $\bar{j}-1$, $\bar{j}$ and $\bar{j}+1$ are in the domain of $Z(\cdot)$, and

2) $\frac{Z(\bar{j}-1)+Z(\bar{j}+1)}{2} \geq Z(\bar{j})$.

The definition applies to an end point of an interval, e.g. j^*, if it is true at j^*-1 or j^*+1, depending on which is in the domain.

<u>PROPOSITION 4.5.</u> $Z(\cdot)$ is LDC at $\bar{j}$ if and only if $\frac{Z(\bar{j}+1)-Z(\bar{j}-1)}{2} \geq Z(\bar{j})-Z(\bar{j}-1)$.

<u>PROOF.</u> Obvious.

<u>DEFINITION 4.3</u>. $Z(\cdot)$ is discretely convex over the set $\{j, \ldots, j+N\}$ if it is locally discretely convex at every point within the set.

A more useful working statement is:

<u>DEFINITION 4.4</u>. $Z(\cdot)$ is discretely convex at $\bar{j}$ with respect to the set $J = \{\bar{j}, \bar{j}+1, \ldots, \bar{j}+N\}$ if, for any $k \in J$, $k > \bar{j}$,

$$\left(\frac{i-\bar{j}}{k-\bar{j}}\right)Z(k) + \left(\frac{k-i}{k-\bar{j}}\right)Z(\bar{j}) \geq Z(i)$$

for all i such that $\bar{j} \leq i \leq k$.

<u>PROPOSITION 4.6.</u> $Z(\cdot)$ is discretely convex at $\bar{j}$ with respect to the set $J = \{\bar{j}, \ldots, \bar{j}+N\}$ if and only if, for any $k \in J$, $k > \bar{j}$, and all i, $\bar{j} < i < k$

$$\frac{Z(k)-Z(\bar{j})}{k-\bar{j}} > \frac{Z(i)-Z(\bar{j})}{i-\bar{j}} .$$

<u>PROOF.</u> (=>) We have, from the definition, that

$$\frac{(i-\bar{j})Z(k)+(k-i)Z(\bar{j})}{k-\bar{j}} \geq Z(i) \quad \text{for all} \quad i, \bar{j} \leq i \leq k$$

By multiplying through by $(k-\bar{j}) > 0$ and subtracting $(k-\bar{j})Z(\bar{j})$ from both sides, we obtain $(i-\bar{j})(Z(k)-Z(\bar{j})) \geq (k-\bar{j})(Z(i)-Z(\bar{j}))$ which is equivalent to the desired result since $(i-\bar{j}) > 0$.

(<=) The algebraic manipulation proceeds through in reverse. That is, given, for any $k \in J$, $k > \bar{j}$ and any $i, \bar{j} \leq i \leq k$

$$\frac{Z(k)-Z(\bar{j})}{k-\bar{j}} \geq \frac{Z(i)-Z(\bar{j})}{(i-\bar{j})}$$

we cross multiply and add $(k-\bar{j})Z(\bar{j})$ to both sides to obtain

$$\left(\frac{i-\bar{j}}{k-\bar{j}}\right)Z(k) + \left(\frac{k-i}{k-\bar{j}}\right)Z(\bar{j}) \geq Z(i)$$

Since the hypothesis holds for any k and all i we are done.

COROLLARY 1. In particular, if $Z(\cdot)$ is convex at 0 with respect to the set $\{0, \ldots, j^*\}$ then $j^* = \hat{j}$.

PROOF. By the proposition,

$$\frac{Z(j^*)-Z(0)}{j^*} \geq \frac{Z(i)-Z(0)}{i} . \qquad \forall i, \ 0 < i \leq j^*$$

but this is precisely the definition of $\hat{j}$.

However, $j^* = \hat{j}$ does not imply that $Z(\cdot)$ is convex at 0 w.r.t. $\{0, \ldots, j^*\}$ and so we shall establish another less restrictive definition, much in the spirit of quasiconvexity. For a definition of quasiconvexity in the case of numerical functions, see Mangasarian [6]. Before we do so we remark that, as in the more usual case, strict discrete convexity can be defined by substituting the strict inequality symbol in the definitions. Further, discrete concavity can be defined by applying the above definitions to $-Z(\cdot)$. These definitions will satisfy the usual convexity requirements for the continuous function formed by connecting the graph of the values $Z(j)$ with straight line segments to form a continuous piecewise linear function.

We shall see in later chapters that we are not interested so much in convexity at a point with respect to a set as we are in whether or not there is no point in the interval $\{j, \ldots, j+N\}$ such that

$\frac{Z(i)-Z(j)}{i-j} > \frac{Z(j+N)-Z(j)}{N}$. Such a condition we shall define below as DISCRETE QUASICONVEXITY. Our definition is somewhat different from and more restrictive than the one we would obtain from that usually formulated for numerical functions.

DEFINITION 4.5. $Z(\cdot)$ is said to be DISCRETE QUASICONVEX over the set $J = \{j, j+1, \ldots, j+N\}$ if

$$\frac{Z(j+N)-Z(j)}{N} \geq \frac{Z(i)-Z(j)}{i-j} \quad \text{for all} \quad i \in J, \quad i \neq j.$$

The requirement is that, for a function to be quasiconvex over an interval, the graph of the value of the function never lies above the straight line joining the values at the end points. This is a weaker condition than discrete convexity at an end point over the interval as in Definition 4.4 since Definition 4.4 implies Def. 4.5 directly, but one can easily generate a counterexample to show that Definition 4.5 does not imply Definition 4.4. For example, the parameterized knapsack function described by:

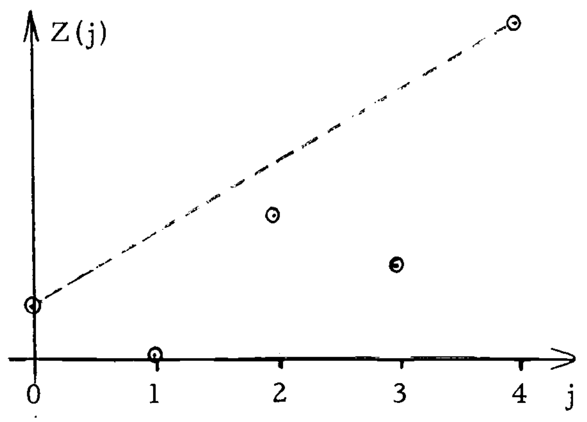
$Z(0) = 91$

$Z(1) = 82$

$Z(2) = 109$

$Z(3) = 100$

$Z(4) = 145$



is discretely quasiconvex over $\{0, 1, 2, 3, 4\}$ but not discretely convex.

If we were to simply consider a discrete version of the definition found in Mangasarian [6], we would only require for a point i in the set J that $Z(i) \leq Z(j+N)$ which does not give us the desired characteristic. Discrete quasiconcavity is defined by reversing the inequality in the definition.

One immediate consequence of the definition is the following:

LEMMA 4.3. $Z(\cdot)$ is discretely quasiconvex over the set $\{\hat{j}^{(n-1)}, \ldots, \hat{j}^{(n)}\}$.

PROOF. By Definition 4.1

$$\frac{Z(\hat{j}^{(n)}) - Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} \geq \frac{Z(i) - Z(\hat{j}^{(n-1)})}{i - \hat{j}^{(n-1)}} \quad \forall i$$

such that $\hat{j}^{(n-1)} < i \leq \hat{j}^{(n)}$, the definition of discrete quasiconvexity over the set $\{\hat{j}^{(n-1)}, \ldots, \hat{j}^{(n)}\}$.

LEMMA 4.4. $j^* = \hat{j}^{(n)}$ if and only if $Z(\cdot)$ is discretely quasiconvex over the set $\{\hat{j}^{(n-1)}, \ldots, j^*\}$.

PROOF. (=>) The forward implication follows from the preceding lemma.

(<=) Certainly, j^* is a candidate for $\hat{j}^{(n)}$ since Definition

4.5 implies that

$$\text{(A)} \qquad \frac{Z(j^*)-Z(\hat{j}^{(n-1)})}{j^*-\hat{j}^{(n-1)}} \geq \frac{Z(j)-Z(\hat{j}^{(n-1)})}{j-\hat{j}^{(n-1)}}$$

for all j in the set. The only requirement is for j^* to be the largest $j \leq j_{max}$ such that the above relation holds. But, for any $j' > j^*$ such that $Z(j') \geq Z(j^*)$, the only possibility is that $Z(j') = Z(j^*)$ and therefore

$$\frac{Z(j^*)-Z(\hat{j}^{(n-1)})}{j^*-\hat{j}^{(n-1)}} > \frac{Z(j^*)-Z(\hat{j}^{(n-1)})}{j'-\hat{j}^{(n-1)}} = \frac{Z(j')-Z(\hat{j}^{(n-1)})}{j'-\hat{j}^{(n-1)}}$$

This inequality certainly holds for any $j > j^*$ such that $Z(j) < Z(j^*)$ and so j^* is the largest j such that (A) holds and therefore $j^* = \hat{j}^{(n)}$.

We are now ready to study the Basic Problem.

III. THE BASIC KNAPSACK FACTORY PROBLEM

1. Introduction and Review

Recall that we formulated the Basic Problem as a mixed integer program in problem (204), a form that was not at all instructive. Then, under the assumption that the solution to the underlying knapsack problem is known, an equivalent integer program was formulated as problems (205) and (206). This assumption is not considered unwarranted, because if the factory cannot solve the underlying knapsack problem, then they have little basis for searching for the solution to a sequence of such problems. Then, as a consequence of Proposition II.4.1. (II.4.2) we realize we need only consider solutions involving filling knapsacks with $j \leq j^*$ $(j \geq j^*)$ in the upper bound case (lower bound case). The balance of this chapter will address good solutions to these problems stated in terms of the parameters of the problem. We specifically ignore the fact that for any given set of parameters one may be able to use the techniques of integer programming to generate a good or perhaps optimal numerical solution. Such a course may not be practical in an industrial environment and is not helpful in our search for general, parametric decision rules. We are particularly interested in solutions which lend themselves to dynamic application, that is, which can be applied knapsack by knapsack, taking advantage of the sequential nature of the problem. Also, we hope to find solutions which can be quickly calculated based on the parameters. That is, we would like to establish solutions which can be computed in a small fraction of the time it takes to fill one knapsack, so that the decision rules have some hope of being useful in a real time sense in applications where knapsacks may be processed at a high rate, perhaps hundreds per hour. Furthermore, we are interested in the effect on the solution of M, the number of knapsacks to be processed, and b, the bound on the number of constrained filler types. Investigating these effects by solving a large integer program for each possible pair of values can at best provide a parametric study of specific cases, which does not provide a convincing argument for generality.

The approach taken here is to study the relaxation of an integer program into a linear program to see if the special structure of the problem will lend itself to our search for simple, good decision rules and will provide us with answers to the questions of "How good?" and "How simple?" We shall see that it does.

We will discuss the upper bound case first, problem (205), in Section 2. This analysis will be in greater detail than that of Section 3, where we study the lower bound problem, problem (206), since the results there are analogous and where appropriate, can merely be sketched. We shall see that in the deterministic case we are indifferent to the order of applying the decisions. Nonetheless, the solutions can still be stated in terms of a sequence of instructions to our knapsack stuffer. This set of instructions offers some insight into the problem when the number of knapsacks is random, and Section 4 of this chapter addresses that issue.

2. The Lower Bound Case

A Related Linear Program

The problem has been stated as

$$(205) \qquad \text{maximize} \quad \sum_{j=0}^{j_{max}} X_j Z(j)$$

$$\text{subject to:} \quad \sum_{j=0}^{j_{max}} X_j = M$$

$$\sum_{j=1}^{j_{max}} jX_j \leq b, \quad b > 0$$

$$X_j \geq 0, \; X_j \text{ integer}, \; j = 0, \ldots, j_{max}$$

where X_j represents the number of knapsacks filled with j. As a consequence of Proposition II. 4. 1, we can rewrite (205) as:

$$(205) \qquad \text{maximize} \quad \sum_{j=0}^{j^*} X_j Z(j)$$

$$\text{subject to} \quad \sum_{j=0}^{j^*} X_j = M$$

$$\sum_{j=1}^{j^*} jX_j \leq b, \quad b > 0$$

$$X_j \geq 0, \; X_j \text{ integer}, \; j = 0, \ldots, j^*$$

where $0 \leq j^* \leq j_{max}$.

We can disregard the case where $j^* = 0$ because in this case the second constraint, the factory constraint, is not binding and the optimal solution is clear. That is, the optimal solution is to set $X_0 = M$, $X_j = 0 \quad j > 0$, and the value is $MZ(0) = MZ(j^*)$ which is certainly the maximum value of the objective function. Thus, we assume hereafter that $j^* > 0$. Consider the relaxation of (205):

$$(301) \qquad \text{maximize} \quad \sum_{j=0}^{j^*} X_j Z(j)$$

$$\text{subject to} \quad \sum_{j=0}^{j^*} X_j = M$$

$$\sum_{j=0}^{j*} jX_j \leq b, \quad b > 0$$

$$X_j \geq 0, \; j = 0, 1, \ldots, j*$$

which is a linear program with two constraints. We shall study this problem in three cases as follows:

1) $M \leq \frac{b}{j*}$

2) $M \geq \frac{b}{\hat{j}}$ ($\hat{j}$ as Definition II.4.1)

3) $\frac{b}{j*} < M < \frac{b}{\hat{j}}$

Of course, in the event that $j* = \hat{j}$, the three cases become two.

<u>CASE 1</u> ($M \leq \frac{b}{j*}$). In this situation, the factory constraint is again without content, since it is not binding and does not prevent the factory from filling each knapsack with an optimal solution. That is, the solution to (301) is $X_{j*} = M$, $X_j = 0$, $j \neq j*$ and the objective function takes on the value $MZ(j*)$. This solution is non-negative, $\sum_{j=0}^{j*} X_j = X_{j*} = M$ and $\sum_{j=1}^{j*} jX_j = j*M \leq b$ by assumption, and thus is feasible. Certainly, since $Z(j) < Z(j*)$ for all $j < j*$ as a consequence of our choice of $j*$, one can only decrease the value of

the objective function by making some other X_j positive. Thus, we have

<u>PROPOSITION 2.1.</u> If $Mj^* \leq b$, the optimal solution to (301) is $X_j^* = M$, $X_j = 0$, otherwise.

Note that this solution is integer valued and therefore is the optimal solution to (205).

<u>CASE 2</u> $(M \geq \frac{b}{j^*})$. Let us proceed to solve (301) using the simplex procedure. Consider the problem stated in matrix notation, where transposes are not indicated as they should be apparent.

$$\begin{array}{ll} \max & ZX \\ \text{s.t.} & AX = d \\ & X \geq 0 \end{array}$$

where

$$Z = (Z(0), Z(1), \ldots, Z(j^*), 0)$$

$$X = (X_0, X_1, \ldots, X_{j^*}, X_{s\ell})$$

$$d = (M, b)$$

and

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & \ldots & 1 & 0 \\ 0 & 1 & 2 & 3 & \ldots & \ldots & j^* & 1 \end{bmatrix}.$$

$X_{s\ell}$ is a slack variable. Certainly, any basic feasible solution will consist of two variables. We first establish that under our assumption that $Z(j^*) > Z(0) \geq 0$, there will be zero slack in any optimal solution.

PROPOSITION 2.1. $X_{s\ell} = 0$ in an optimal solution to (301), when $M \geq \frac{b}{j^*}$.

PROOF. Assume it is not. Then our optimal basis B will be $B = \begin{bmatrix} 1 & 0 \\ j' & 1 \end{bmatrix}$ where $0 \leq j' \leq j^*$. Then $B^{-1} = \begin{bmatrix} 1 & 0 \\ -j' & 1 \end{bmatrix}$, $C_B = (Z(j'), 0)$ and

$$\begin{aligned} \pi_j &= C_B B^{-1} a_j - Z(j) \\ &= (Z(j'), 0) \begin{bmatrix} 1 & 0 \\ -j' & 1 \end{bmatrix} \begin{bmatrix} 1 \\ j \end{bmatrix} - Z(j) \\ &= Z(j') - Z(j) \end{aligned}$$

At optimality, we require $\pi_j \geq 0 \quad \forall j$, or $Z(j') \geq Z(j)$ for all j. Therefore $j' = j^*$. Now the optimal basic feasible solution is

$$X_B = B^{-1} d = \begin{bmatrix} 1 & 0 \\ -j^* & 1 \end{bmatrix} \begin{bmatrix} M \\ b \end{bmatrix} = (M, b - Mj^*).$$

However, we have assumed $M \geq \frac{b}{j^*}$ so that $X_{s\ell} = b - Mj^* \leq 0$. Thus, the only feasible possibility is that $X_{s\ell} = 0$, a contradiction.

As a consequence, we shall find optimal solutions to (301) which result in $\sum_{j=1}^{j^*} jX_j = b$. We see then that in the LP relaxation of the Basic Problem, our assumption that $j^* > 0$ and $M \geq \frac{b}{j^*}$ (which implies $Z(j^*) > 0$) causes the optimal solution to use up all of the restricted filler types. It is in this sense that we observe that the integer solution to the Basic Problem will also try to reach equality in the factory constraint, allowing for the integer nature of the arithmetic. That is, if we round off solutions to (301) to the nearest integer, we may not use exactly b units of the restricted fillers. Another consequence is that we shall iterate in the simplex method using bases of the form $B = \begin{bmatrix} 1 & 1 \\ j_1 & j_2 \end{bmatrix}$ where, w.l.o.g., $j_1 < j_2$. In the interest of efficiency, we shall proceed directly to the optimal solution as follows. First, we note that the solution is bounded as a consequence of the fact that the feasible region is bounded and therefore, since a feasible solution exists, namely $X_0 = M$, $X_j = 0$, $j > 0$, an optimal solution exists.

If we were to start with a feasible solution, for example $X_0 = M$, $X_1 = 0$, $X_j = 0$ for $j > 1$ and proceed through simplex iterations we would quickly discover the following:

<u>THEOREM 2.1.</u> The optimal solution to (301), when $M \geq \frac{b}{\hat{j}}$, is $X_{\hat{j}} = \frac{b}{\hat{j}}$, $X_0 = M - \frac{b}{\hat{j}}$, $X_j = 0$ otherwise.

PROOF. We have that the candidate for the optimal basic feasible solution is $X_B = (X_0, X_{\hat{j}}) = (M - \frac{b}{\hat{j}}, \frac{b}{\hat{j}})$ and $X_j = 0$ otherwise. Under the assumption that $M \geq \frac{b}{\hat{j}}$, $X_B \geq 0$ and $\sum_{j=0}^{j^*} X_j = M$, $\sum_{j=1}^{j^*} jX_j = \hat{j}\frac{b}{\hat{j}} = b$ so that our candidate solution is feasible. Now the associated basis is $B = \begin{bmatrix} 1 & 1 \\ 0 & \hat{j} \end{bmatrix}$ so that $B^{-1} = \frac{1}{\hat{j}}\begin{bmatrix} \hat{j} & -1 \\ 0 & 0 \end{bmatrix}$. The cost vector associated with X_B is $C_B = (Z(0), Z(\hat{j}))$ and so

$$\begin{aligned}
\pi_k &= C_B B^{-1} a_k - Z(k) \\
&= \frac{(\hat{j}-k)}{\hat{j}} Z(0) + \frac{k}{\hat{j}} Z(\hat{j}) - Z(k) \\
&= k\left(\frac{Z(\hat{j}) - Z(0)}{\hat{j}}\right) - (Z(k) - Z(0)).
\end{aligned}$$

In order for our candidate to be optimal, we require that $\pi_k \geq 0$ for all k, which implies that $\frac{Z(\hat{j})-Z(0)}{\hat{j}} \geq \frac{Z(k)-Z(0)}{k}$ $\forall k > 0$ but this is precisely the definition of $\hat{j}$, i. e. we find $\hat{j}$ by searching for the largest j which maximizes $\frac{Z(j)-Z(0)}{j}$. We know $\hat{j} > 0$ as a consequence of our assumption that $j^* > 0$.

We note that in the event there is more than one such $\hat{j}$, there are alternative optima for (301). We have chosen the largest such j as $\hat{j}$ in Definition II. 4. 1. in order to reduce the size of the interval $\frac{b}{j^*} < M < \frac{b}{\hat{j}}$ as much as possible, for as we shall see, that case has a somewhat more complicated solution. Of course, this result establishes our interest in the quantity $\frac{Z(j)-Z(0)}{j}$, and is the motivation for Definition II.4.1. Finally, the optimal value of the objective function in this case is $b(\frac{Z(\hat{j})-Z(0)}{\hat{j}}) + MZ(0)$.

COROLLARY 1. If $j^* = \hat{j}$, then the optimal solution to (301) is, for any M,

$$X_{j^*} = \min(M, \frac{b}{j^*})$$

$$X_0 = M - X_{j^*}$$

$$X_j = 0 \qquad \text{otherwise}$$

and the optimal value is $X_{j^*}(Z(j^*)-Z(0)) + MZ(0)$.

PROOF. Since case 3, $\frac{b}{j^*} < M < \frac{b}{\hat{j}}$ does not apply here, the result holds from Proposition 2.1 and Theorem 2.1.

Case 3 ($\frac{b}{j^*} < M < \frac{b}{\hat{j}}$). The first point we shall note is that in this case, we have that $\hat{j} < j^*$, and so $\bar{n} \geq 2$ as in Definition II.4.1. That is, we have the existence of at least $\hat{j}^{(2)}$. The precise value of $\bar{n}$ for the underlying knapsack problem of the basic problem has an impact on the complexity of any solution algorithm, but not on the basic idea, as we shall see. Let us proceed to solve this case by starting with a feasible solution. Consider $X_0 = M, X_{s\ell} = b$, $X_j = 0$ o.w. We have $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = B^{-1}$, $X_B = (M, b)$. So

$\pi_k = Z(0) - Z(k)$, and we know that $Z(0) - Z(\hat{j}) < 0$. We shall bring $X_{\hat{j}}$ into the basis. In the usual way we search for the $\min\limits_{y_{\hat{j}i} > 0} \{\frac{X_{Bi}}{y_{\hat{j}i}}\}$ to determine which variable leaves the basis. Since $y_{\hat{j}} = B^{-1}a_{\hat{j}} = [\begin{smallmatrix}1\\ \hat{j}\end{smallmatrix}] > 0$, we want the $\min(\frac{M}{1}, \frac{b}{\hat{j}})$. We have in this case $M < \frac{b}{\hat{j}}$, so $X_{\hat{j}}$ will replace X_0. The new basis is $B = [\begin{smallmatrix}1 & 0\\ \hat{j} & 1\end{smallmatrix}]$ so $B^{-1} = [\begin{smallmatrix}1 & 0\\ -\hat{j} & 1\end{smallmatrix}]$. Also, the new basic feasible solution is $X_B = (X_{\hat{j}}, X_{s\ell}) = B^{-1}d = (M, b-M\hat{j})$ with $C_B = (Z(\hat{j}), 0)$. Now, $\pi_k = Z(\hat{j}) - Z(k)$. We know there is at least one k such that $Z(k) > Z(\hat{j})$, namely $k = j^*$, so $\pi_{j^*} < 0$, and we will bring X_{j^*} into the basic feasible solution. In this event we are searching for the $\min(M, \frac{b-M\hat{j}}{j^*-\hat{j}})$ since $y_{j^*} = [\begin{smallmatrix}1\\ j^*-\hat{j}\end{smallmatrix}] > 0$. Now $\frac{b}{j^*} < M$ implies $b < Mj^*$, so certainly $b-M\hat{j} < Mj^* - M\hat{j}$ or $\frac{b-M\hat{j}}{j^*-\hat{j}} < M$ so we see that we should replace $X_{s\ell}$ in the solution with X_{j^*}. The new basis is $B = [\begin{smallmatrix}1 & 1\\ \hat{j} & j^*\end{smallmatrix}]$, $B^{-1} = \frac{1}{(j^*-\hat{j})}[\begin{smallmatrix}j^* & -1\\ -\hat{j} & 1\end{smallmatrix}]$, $C_B = (Z(j), Z(j^*))$ and $X_B = (X_{\hat{j}}, X_{j^*}) = B^{-1}d$, so that

$$X_B = \left(\frac{j^*M-b}{j^*-\hat{j}} , \frac{b-Mj}{j^*-\hat{j}} \right) .$$

We can wonder at this point whether we have found the optimal solution to (301) under case 3. As we shall see, the answer depends on the value of $\hat{j}^{(2)}$. Now, $\pi_k = C_B B^{-1} a_k - Z(k)$ so that

$$\pi_k = \frac{(j^*-k)Z(\hat{j})+(k-\hat{j})Z(j^*)}{j^*-\hat{j}} - Z(k). \qquad \text{(EQ. 1)}$$

We want to know if there is a k such that $\pi_k < 0$. If not, then the current solution is optimal.

<u>PROPOSITION 2.2.</u> There is no $k < \hat{j}$ such that $\pi_k < 0$.

<u>PROOF.</u> Suppose there is. Then we have, if we let $j^* - k = j^* - \hat{j} - (k-\hat{j})$,

$$\frac{(j^*-\hat{j})Z(\hat{j})+(k-\hat{j})(Z(j^*)-Z(\hat{j}))}{j^*-\hat{j}} - Z(k) < 0$$

rearranging terms and multiplying through by $(j^*-\hat{j})$, the result is

$$(j^*-\hat{j})(Z(\hat{j})-Z(k)) + (k-\hat{j})(Z(j^*)-Z(\hat{j})) < 0$$

or

$$(j^*-\hat{j})(Z(\hat{j})-Z(k)) < (\hat{j}-k)(Z(j^*)-Z(\hat{j}))$$

and since $(j^*-\hat{j}) > 0$, $(\hat{j}-k) > 0$ we end up with

$$\frac{Z(\hat{j})-Z(k)}{\hat{j}-k} < \frac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}},$$

but Lemma II.4.1 says that

$$\frac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}} < \frac{Z(\hat{j})-Z(0)}{\hat{j}} \leq \frac{Z(\hat{j})-Z(k)}{\hat{j}-k},$$

a contradiction.

The result is that we can restrict our search for an optimal solution to those j's such that $\hat{j} \leq j \leq j^*$. We note immediately therefore, that we never consider j, $0 < j < \hat{j}$ no matter what value M has. The upshot is that we are reducing the list of possible candidates for an optimal solution.

<u>THEOREM 2.2.</u> The basic feasible solution containing $X_{\hat{j}}$ and X_{j^*} (i.e. $X_{\hat{j}} = \frac{j^*M-b}{j^*-\hat{j}}$, $X_{j^*} = \frac{b-M\hat{j}}{j^*-\hat{j}}$, $X_j = 0$ otherwise) is optimal in (301) if and only if $\frac{b}{j^*} \leq M \leq \frac{b}{\hat{j}}$ and $\bar{n} = 2$, that is $j^* = \hat{j}^{(2)}$.

<u>PROOF.</u> (=>) From Theorem 2.1, and Proposition 2.1. $X_{j^*} > 0$ in an optimal solution only if $M < \frac{b}{\hat{j}}$. If $M = \frac{b}{\hat{j}}$, then based on the results in Theorem 2.1, we can find an alternate optimal solution where $X_{j^*} = 0$ and is in the basis. By the same results, we see that $X_{\hat{j}} > 0$ only if $M > \frac{b}{j^*}$ and similarly, an alternate optimal solution with $X_{\hat{j}} = 0$ can be found if $M = \frac{b}{j^*}$. The intersection of these two intervals is the desired region. Now, optimality implies that $\pi_k \geq 0$ for all k. We have seen in Proposition 2.2 that $\pi_k \geq 0$ for all $k \leq \hat{j}$. So we only need consider $k > \hat{j}$. $\pi_k \geq 0$ implies that $\frac{(j^*-k)Z(\hat{j})+(k-\hat{j})Z(j^*)}{(j^*-\hat{j})} - Z(k) \geq 0$. So, since $(j^*-\hat{j}) > 0$, $(j^*-k)Z(\hat{j}) + (k-\hat{j})Z(j^*) \geq (j^*-\hat{j})Z(k)$. If we let $j^* - k = (j^*-\hat{j}) - (k-\hat{j})$, we have, by subtracting $(j^*-\hat{j})Z(\hat{j})$ from both sides, $(k-\hat{j})(Z(j^*)-Z(\hat{j})) \geq (j^*-\hat{j})(Z(k)-Z(\hat{j}))$. Dividing through by

$(k-\hat{j})(j^*-\hat{j}) > 0$, results in $\dfrac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}} \geq \dfrac{Z(k)-Z(\hat{j})}{k-\hat{j}}$ for any $k > \hat{j}$ but this is exactly the requirement for j^* to be $\hat{j}^{(2)}$.

($\Leftarrow$) If $\dfrac{b}{j^*} \leq M \leq \dfrac{b}{\hat{j}}$, the proposed solution is non-negative,

$$\sum_{j=0}^{j^*} X_j = X_{\hat{j}} + X_{j^*} = \frac{j^*M - b + b - \hat{j}M}{j^*-\hat{j}} = M ,$$

$$\sum_{j=0}^{j^*} jX_j = \hat{j}X_{\hat{j}} + j^*X_{j^*} = \frac{\hat{j}j^*M - \hat{j}b + j^*b - \hat{j}j^*M}{j^*-\hat{j}} = b$$

and so it is feasible. Now, since $j^* = \hat{j}^{(2)}$ this means that $\dfrac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}} \geq \dfrac{Z(k)-Z(\hat{j})}{k-\hat{j}}$ for all $k > \hat{j}$. Since the algebraic argument in the first part of the proof works in reverse, we have that $\pi_k \geq 0$ for all $k > \hat{j}$. Proposition 2.2 tells us that $\pi_k \geq 0$ for all $k < \hat{j}$ and so the proposed solution is optimal.

COROLLARY 1. If $Z(\cdot)$ is discretely quasiconvex over the set $\{\hat{j}, \ldots, j^*\}$, and $\dfrac{b}{j^*} \leq M \leq \dfrac{b}{\hat{j}}$, then the solution specified in the theorem is optimal in (301).

PROOF. By Lemma II.4.4, $j^* = \hat{j}^{(2)}$.

Since nothing we have said or assumed guarantees that $j^* = \hat{j}^{(2)}$, it is possible that there exists a k such that $\pi_k < 0$ and that therefore we will bring X_k in to the basic solution. One can anticipate that the simplex procedure will lead us to $\hat{j}^{(2)}$ in the event $j^* \neq \hat{j}^{(2)}$. The variable that leaves the basis depends on the value of M as we shall see. First, we provide a necessary condition for $\pi_k < 0$ where π_k is as defined in Equation 1 of this section.

THEOREM 2.3. In case 3 where $\frac{b}{j^*} < M < \frac{b}{\hat{j}}$, if $\pi_k < 0$ then

1) $k > \hat{j}$ and

2) $\frac{Z(k)-Z(0)}{k} > \frac{Z(j^*)-Z(0)}{j^*}$.

PROOF. 1) is a restatement of Proposition 2.2.

2) From the proof of the previous theorem, we see that $\pi_k < 0$ implies that $\frac{Z(k)-Z(\hat{j})}{k-\hat{j}} \geq \frac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}}$ and since $k > \hat{j}$, we have $j^*Z(k) - j^*Z(\hat{j}) - \hat{j}Z(k) + \hat{j}Z(\hat{j}) \geq kZ(j^*) - kZ(\hat{j}) - \hat{j}Z(j^*) + \hat{j}Z(\hat{j})$ collecting terms and subtracting $j^*Z(0)$ from both sides results in $j^*(Z(k)-Z(0)) \geq k(Z(j^*)-Z(0)) + ((j^*-k)(Z(\hat{j})-Z(0))-\hat{j}(Z(j^*)-Z(k)))$. If we can show that the term in brackets is positive, then we are done.

Now $\frac{Z(k)-Z(\hat{j})}{k-\hat{j}} \geq \frac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}}$ and if we let k play the role of $\hat{j}^{(2)}$ in part 3 of Lemma II.4.2, we see that, since $j^* > k$,

$$\frac{Z(j^*)-Z(k)}{j^*-k} < \frac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}}$$

and so

$$\frac{Z(j^*)-Z(k)}{j^*-k} < \frac{Z(k)-Z(\hat{j})}{k-\hat{j}}$$

but, by part 2 of Lemma II. 4. 1,

$$\frac{Z(k)-Z(\hat{j})}{k-\hat{j}} < \frac{Z(\hat{j})-Z(0)}{\hat{j}}$$

so that

$$\frac{Z(j^*)-Z(k)}{j^*-k} < \frac{Z(\hat{j})-Z(0)}{\hat{j}}$$

so certainly $(j^*-k)(Z(\hat{j})-Z(0)) - \hat{j}(Z(j^*)-Z(k)) > 0$ as was necessary to be shown.

Let us suppose such a k exists. If it does, then of course, there exists a $\hat{j}^{(2)}$, $\hat{j}^{(2)} < j^*$ and $\pi_{\hat{j}^{(2)}} < 0$. We shall bring $\hat{j}^{(2)}$ into the basis. Which variable will leave is determined in the usual way by the $\min\limits_{y_{ji} > 0} \left(\frac{X_{Bi}}{y_{ji}}\right)$, where in this case,

$$y_{ji} = \begin{bmatrix} \frac{j^*-\hat{j}^{(2)}}{j^*-\hat{j}} \\ \\ \frac{\hat{j}^{(2)}-\hat{j}}{j^*-\hat{j}} \end{bmatrix} > 0 ,$$

so we want the

$$\min\left(\frac{j^*M-b}{j^*-\hat{j}^{(2)}}, \frac{b-\hat{j}M}{\hat{j}^{(2)}-\hat{j}}\right) .$$

Let us assume that $M \geq \frac{b}{\hat{j}^{(2)}}$ and, of course, $\frac{b}{\hat{j}^{(2)}} > \frac{b}{j^*}$. Then $\hat{j}^{(2)} M(j^*-\hat{j}) \geq (j^*-\hat{j})b$, so that $\hat{j}^{(2)} j^*M + \hat{j}b \geq j^*b + \hat{j}^{(2)}\hat{j}M$, and subtracting the quantity $\hat{j}j^*M + \hat{j}^{(2)}b$ from both sides we have $(\hat{j}^{(2)}-\hat{j})(j^*M-b) \geq (\hat{j}^{(2)}-j^*)(\hat{j}M-b)$ and since $(\hat{j}^{(2)}-\hat{j}) > 0$ but $(\hat{j}^{(2)}-j^*) < 0$, this results in

$$\frac{(j^*M-b)}{\hat{j}^{(2)}-j^*} \leq \frac{\hat{j}M-b}{\hat{j}^{(2)}-\hat{j}}$$

or

$$\frac{j^*M-b}{j^*-\hat{j}^{(2)}} \geq \frac{b-\hat{j}M}{\hat{j}^{(2)}-\hat{j}},$$

which means that we can replace X_{j^*} with $X_{\hat{j}^{(2)}}$. The new basis is

$$B = \begin{bmatrix} 1 & 1 \\ j & j^{(2)} \end{bmatrix}, \quad X_B = (X_{\hat{j}}, X_{\hat{j}^{(2)}}), \quad C_B = (Z(\hat{j}), Z(\hat{j}^{(2)}))$$

and

$$B^{-1} = \frac{1}{\hat{j}^{(2)}-\hat{j}} \begin{bmatrix} \hat{j}^{(2)} & -1 \\ -\hat{j} & 1 \end{bmatrix}.$$

$$X_B = B^{-1}d = \frac{1}{\hat{j}^{(2)}-\hat{j}} \begin{bmatrix} \hat{j}^{(2)} & -1 \\ -\hat{j} & 1 \end{bmatrix} \begin{bmatrix} M \\ b \end{bmatrix}$$

$$= \left(\frac{M\hat{j}^{(2)}-b}{\hat{j}^{(2)}-\hat{j}}, \frac{b-M\hat{j}}{\hat{j}^{(2)}-\hat{j}} \right) \geq 0.$$

We calculate $\pi_k = C_B B^{-1} a_k - Z(k)$ to check for optimality.

$$\pi_k = \left(\frac{\hat{j}^{(2)}-k}{\hat{j}^{(2)}-\hat{j}}\right)Z(\hat{j}) + \left(\frac{k-\hat{j}}{\hat{j}^{(2)}-\hat{j}}\right)Z(\hat{j}^{(2)}) - Z(k)$$

which we can rewrite as

$$\frac{1}{\hat{j}^{(2)}-\hat{j}}((Z(\hat{j})(\hat{j}^{(2)}-\hat{j}-(k-\hat{j}))+(k-\hat{j})Z(\hat{j}^{(2)}) - Z(k)$$

or

$$Z(\hat{j}) - Z(k) + \frac{(k-\hat{j})(Z(\hat{j}^{(2)})-Z(\hat{j}))}{\hat{j}^{(2)}-\hat{j}}.$$

Suppose $\pi_k < 0$, then, for $k > \hat{j}$

$$\frac{Z(\hat{j}^{(2)})-Z(\hat{j})}{\hat{j}^{(2)}-\hat{j}} < \frac{Z(k)-Z(\hat{j})}{k-\hat{j}}$$

but this contradicts the definition of $\hat{j}^{(2)}$ and we know from Proposition 2.2 that $\pi_k \geq 0$ for $k < \hat{j}$. Thus, $\pi_k \geq 0$ for all k and our solution is optimal. We have just proved the following theorem.

<u>THEOREM 2.4.</u> If $\frac{b}{\hat{j}^{(2)}} \leq M < \frac{b}{\hat{j}}$, then the optimal solution to (301) is

$$X_{\hat{j}} = \frac{M\hat{j}^{(2)}-b}{\hat{j}^{(2)}-\hat{j}}, \quad X_{\hat{j}^{(2)}} = \frac{b-M\hat{j}}{\hat{j}^{(2)}-\hat{j}}, \quad X_j = 0 \text{ otherwise.}$$

The value of the objective function for this solution is

$$(b-M\hat{j})\left(\frac{Z(\hat{j}^{(2)})-Z(\hat{j})}{\hat{j}^{(2)}-\hat{j}}\right) + MZ(\hat{j}).$$

PROOF. After the preceding discussion, the only thing to establish is the value of the objective function. The value is $X_{\hat{j}}Z(\hat{j}) + X_{\hat{j}^{(2)}}Z(\hat{j}^{(2)})$ and when we observe that

$$X_{\hat{j}} = \frac{M\hat{j}^{(2)}-b}{\hat{j}^{(2)}-\hat{j}} = M - \frac{b-M\hat{j}}{\hat{j}^{(2)}-\hat{j}}$$

(more simply, $X_{\hat{j}} = M - X_{\hat{j}^{(2)}}$) the result is clear.

On the other hand, if we assume that $M > b/\hat{j}^{(2)}$, then we replace $X_{\hat{j}}$ by $X_{\hat{j}^{(2)}}$ in the basis and we are back in the same position as before, in that now we are concerned with whether or not $\hat{j}^{(3)} = j^*$. We know that this process is finite, and we have provided sufficient motivation to proceed directly to the following theorem which generalizes the solution to (301) for any M and b.

Let us first define $\hat{j}^{(0)} = 0$ and $\hat{j}^{(\bar{n}+1)} = +\infty$, and $X_{\hat{j}^{(\bar{n}+1)}} = X_0$. Of course, $\frac{b}{\infty} = 0$ for b finite. Then, we can write

THEOREM 2.5. When $\frac{b}{\hat{j}^{(n)}} \leq M < \frac{b}{\hat{j}^{(n-1)}}$, for

$n = 1, \ldots, \bar{n}, \bar{n}+1$, the optimal solution to (301) is given by:

$$X_{\hat{j}(n)} = \frac{b - M\hat{j}(n-1)}{\hat{j}(n) - \hat{j}(n-1)}$$

$$X_{\hat{j}(n-1)} = M - X_{\hat{j}(n)}$$

$$X_j = 0 \qquad \text{otherwise.}$$

The value of the objective function is

$$\frac{b - M\hat{j}(n-1)}{\hat{j}(n) - \hat{j}(n-1)} \left(Z(\hat{j}(n)) - Z(\hat{j}(n-1)) \right) + MZ(\hat{j}(n-1)).$$

<u>PROOF.</u> For $n = 1$, so that $\frac{b}{\hat{j}} \leq M$ and $X_{\hat{j}} = \frac{b}{\hat{j}}$, $X_0 = M - \frac{b}{\hat{j}}$, Theorem 2.1 establishes the result.

For $n = \bar{n}+1$, so that $M < \frac{b}{j^*}$, and $X_0 = 0$, $X_{j^*} = M$, Proposition 2.1 proves the required result.

So for n such that $2 \leq n \leq \bar{n}$, we have that $X_{\hat{j}(n)}$ and $X_{\hat{j}(n-1)}$ are in the basis. Now, under the conditions we know the optimal solution is non-negative and clearly sums to M. Further

$$\sum_{j=1}^{j^*} jX_j = \hat{j}(n-1) X_{\hat{j}(n-1)} + \hat{j}(n) X_{\hat{j}(n)}$$

$$= \hat{j}(n) \left(\frac{b - M\hat{j}(n-1)}{\hat{j}(n) - \hat{j}(n-1)} \right) + \hat{j}(n-1) \left(M - \frac{b - M\hat{j}(n-1)}{\hat{j}(n) - \hat{j}(n-1)} \right) = b \ ,$$

so that the proposed solution is feasible. Now the basis associated with this solution is

$$B = \begin{bmatrix} 1 & 1 \\ \hat{j}^{(n-1)} & \hat{j}^{(n)} \end{bmatrix}$$

so that

$$B^{-1} = \frac{1}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} \begin{bmatrix} \hat{j}^{(n)} & -1 \\ -\hat{j}^{(n-1)} & 1 \end{bmatrix}$$

and

$$\pi_k = \frac{1}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} ((\hat{j}^{(n)} - k)Z(\hat{j}^{(n-1)}) + (k - \hat{j}^{(n-1)})Z(\hat{j}^{(n)})) - Z(k)$$

Suppose $\pi_k < 0$, then in the way we have manipulated this equation before, we can see that for $k > \hat{j}^{(n-1)}$,

$$\frac{Z(k) - Z(\hat{j}^{(n-1)})}{k - \hat{j}^{(n-1)}} > \frac{Z(\hat{j}^{(n)}) - Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)} - \hat{j}^{(n-1)}},$$

which violates the definition of $\hat{j}^{(n)}$. For $k < \hat{j}^{(n-1)}$, and $\pi_k < 0$, if we let $\hat{j}^{(n)} - k = \hat{j}^{(n)} - \hat{j}^{(n-1)} - (k - \hat{j}^{(n-1)})$, we have

$$(k - \hat{j}^{(n-1)}) \frac{(Z(\hat{j}^{(n)}) - Z(\hat{j}^{(n-1)}))}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} < Z(k) - Z(\hat{j}^{(n-1)}) .$$

Since $k - \hat{j}^{(n-1)} < 0$, we see that $\pi_k < 0$ implies

$$\frac{Z(\hat{j}^{(n)}) - Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} > \frac{Z(k) - Z(\hat{j}^{(n-1)})}{k - \hat{j}^{(n-1)}} .$$

If we apply Corollary 1 to Lemma II. 4. 2 to this case we generate a contradiction. Therefore, $\pi_k \geq 0$ for all k and we have the desired result.

The preceding discussion also establishes

<u>COROLLARY 1.</u> The only candidates for positive values in the optimal solution to (301) are X_j for $j = 0, \hat{j}, \hat{j}^{(2)}, \ldots, \hat{j}^{(\bar{n})} = j^*$ which is a subset of the indices $0, 1, \ldots, j^*$.

We note without proof that the set $\{j: j=0, \hat{j}, \hat{j}^{(2)}, \ldots, \hat{j}^{(\bar{n})} = j^*\}$ is a proper subset of the set of all indices unless $Z(\cdot)$ is strictly discretely concave on the set of all indices.

We have provided a solution to (301) in general terms which can be calculated very quickly by hand. If we can establish that we can quickly generate a solution to (205) from this answer which is good in the sense that it is within the value of one knapsack filled optimally of the possible optimal solution to (205) for any M (thus the percentage of potential loss decreases in M) we will have achieved our goal.

Heuristic Solution to the Basic Problem

We now have solutions to the linear program (301) and we shall use these solutions to obtain heuristic solutions to the Basic Problem. We shall do so in the traditional way of obtaining heuristic solutions to an integer program by rounding the answers off to integer values in such a way that feasibility is preserved. We will then establish that these solutions are "close" in a precise sense to the optimal value of (205), when they are not in fact optimal.

Let us write our procedure in terms of a solution algorithm, in which we can summarize the results of the preceding subsection: we propose the following solution procedure for the Basic Problem:

ALGORITHM 2.1. The parameters of the problem are M, b and $Z(j)$ for $j = 0, 1, \ldots, j_{max}$.

STEP 1. Calculate $\hat{j}, \hat{j}^{(2)}, \ldots, \hat{j}^{(\bar{n})} = j^*$ in accordance with Definition II.4.1 and disregard all $j > j^*$. Continue

STEP 2. If $j^* > 0$, go to Step 3. Otherwise, if $j^* = 0$ set $X_0 = M$, $X_j = 0$ for all $j > 0$. The value of the objective function is $MZ(0)$. Stop.

STEP 3. If $M > \frac{b}{j^*}$ go to Step 4. Otherwise, if $M \leq \frac{b}{j^*}$, set $X_{j^*} = M$, $X_j = 0$ for all $j \neq j^*$. The value of the objective function is $MZ(j^*)$. Stop.

STEP 4. Calculate n such that $\frac{b}{\hat{j}^{(n)}} < M \leq \frac{b}{\hat{j}^{(n-1)}}$

$1 \leq n \leq \bar{n}$. Then, set $X_{\hat{j}^{(n)}} = \left[\frac{b-M\hat{j}^{(n-1)}}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} \right]$, (where $[\cdot]$ is the greatest integer notation), $X_{\hat{j}^{(n-1)}} = M - X_{\hat{j}^{(n)}}$ and let $X_j = 0$ otherwise. The value of the objective function is

$$\left[\frac{b-M\hat{j}^{(n-1)}}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} \right] (Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)}))+MZ(\hat{j}^{(n-1)}) .$$

Stop.

This procedure covers all possibilities and we have rounded off in such a way to guarantee feasibility. That is, the number of restricted fillers used is $Mj^* \leq b$ in the case covered by step 2. In step 3, the number of fillers used is $\hat{j}^{n} X_{\hat{j}^{n}} + \hat{j}^{(n-1)} X_{\hat{j}^{(n-1)}}$ or

$$(\hat{j}^{(n)}-\hat{j}^{(n-1)})\left[\frac{b-M\hat{j}^{(n-1)}}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} \right] + \hat{j}^{(n-1)}M .$$

Since

$$\left[\frac{b-M\hat{j}^{(n-1)}}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} \right] \leq \frac{b-M\hat{j}^{(n-1)}}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} ,$$

the number of fillers used is $\leq b$. According to the convention of Section 3 of Chapter I, we will denote $(b-M\hat{j}^{(n-1)})\operatorname{mod}(\hat{j}^{(n)}-\hat{j}^{(n-1)})$ as r_n. Note that

$$r_n = (b - M\hat{j}^{(n-1)}) - (\hat{j}^{(n)} - \hat{j}^{(n-1)})\left[\frac{b - M\hat{j}^{(n-1)}}{\hat{j}^{(n)} - \hat{j}^{(n-1)}}\right]$$

or $0 \leq r_n < \hat{j}^{(n)} - \hat{j}^{(n-1)}$. We see that the number of fillers used by our solution is $b - r_n$ and so the number left over is r_n.

We see that the cases covered by steps two and three provide integer solutions which are in fact optimal. The case where $j^* = 0$ is uninteresting and we shall no longer consider it. The case where $j^* > 0$ and $M \leq b/j^*$ is also uninteresting, but is germane to the discussion in Section 4 when we consider M random and so we shall refer to it again.

We point out here that our solution procedure can be translated in terms of instructions to our knapsack stuffer as follows: for steps 2 and three, "fill all M knapsacks with j^*." In step four, "fill the first $\left[\frac{b - M\hat{j}^{(n-1)}}{\hat{j}^{(n)} - \hat{j}^{(n-1)}}\right]$ knapsacks with $\hat{j}^{(n)}$ and the remaining knapsacks with $\hat{j}^{(n-1)}$." Also, we can argue that the value realized by this procedure can be increased in the event that some restricted fillers are left over. That is, if the number of restricted fillers left over, which we have seen is equal to r_n, is positive, and there exists a j such that $\hat{j}^{(n-1)} < j < \hat{j}^{(n-1)} + r_n$ and $Z(j) > Z(\hat{j}^{(n-1)})$, then a rational stuffer would add $j - \hat{j}^{(n-1)}$ units of the remainder to one or more of the knapsacks filled with $\hat{j}^{(n-1)}$ until he could no longer find such a j. We shall refer to this process as the remainder problem and point out that other than mentioning it, little can be said in general without assuming more than we desire.

Before we address the question of how good this solution is, we shall point out a sufficient condition for its optimality. To do so, we argue that since the linear program is a relaxation of our integer program, the optimal value of the Basic Problem has, as an upper bound, the optimal value of the linear program under the same parameters. Since we can say nothing about the value of the optimal solution to the integer program we shall use the optimal value of the linear program as our standard of comparison. Since the difference between our solution and the optimal solution to (205) is less than the difference between our solution and the LP solution, this latter difference provides us with an upper bound on our potential loss. Thus, it is a liberal estimate of such a potential loss in the sense that the real loss is at worst equal to our estimate. As a consequence, if our loss estimate is zero, then our solution is optimal. This is equivalent to saying either that the solution to the LP results in integer values of the variables or what is the same thing, the value of our solution to the Basic Problem realizes its upper bound, the value of the LP solution. We formalize this argument in the following:

THEOREM 2.6. If $r_n = 0$, where $r_n = (b - M\hat{j}^{(n-1)}) \bmod (\hat{j}^{(n)} - \hat{j}^{(n-1)})$ then the solution to the Basic Problem specified in our procedure, i.e.

$$X_{\hat{j}(n)} = \left[\frac{b - M\hat{j}^{(n-1)}}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} \right]$$

$$X_{\hat{j}(n-1)} = M - X_{\hat{j}(n)}$$

$$X_j = 0 \quad \text{otherwise},$$

is optimal in the Basic Problem.

PROOF. $r_n = 0$ implies that $\left[\frac{b - M\hat{j}^{(n-1)}}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} \right]$ equals $\frac{b - M\hat{j}^{(n-1)}}{\hat{j}^{(n)} - \hat{j}^{(n-1)}}$ so that the value of this solution to the Basic Problem is

$$b - M\hat{j}^{(n-1)} \left(\frac{Z(\hat{j}^{(n)}) - Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} \right) + MZ(\hat{j}^{(n-1)})$$

but this is the value of optimal solution to the linear program (301). Thus, by standard arguments (for example, see Garfinkel and Nemhauser [1]), this solution realizes an upper bound on the value of any solution and so it must be that the solution presented is an optimal solution to the Basic Problem.

Note: Recall that the value of n in this theorem is uniquely specified by the requirement that $\frac{b}{\hat{j}^{(n)}} < M \leq \frac{b}{\hat{j}^{(n-1)}}$.

<u>COROLLARY 1.</u> If $\hat{j}^{(n)}-\hat{j}^{(n-1)} = 1$, then the solution is optimal.

<u>PROOF.</u> Since r_n is an integer less than $\hat{j}^{(n)}-\hat{j}^{(n-1)}$, it must be that $r_n = 0$.

<u>COROLLARY 2.</u> If $\hat{j} = 1$, then for $M \geq b$, the solution is optimal.

<u>PROOF.</u> We have $\hat{j}^{(n)}-\hat{j}^{(n-1)} = 1$ in this case.

<u>Loss Functions</u>

In this subsection we are interested in determining the value of our decision rules in terms of their closeness to the optimal solution to the Basic Problem.

We will use the optimal solution to the associated linear program as our standard. Accordingly, let $LP(M,b)$ represent the value of the optimal solution to the linear program for a given M and b. Let $IP(M,b)$ be the value of the optimal solution to the Basic Problem and let $\overline{IP}(M,b)$ represent the value of our heuristic solution. Since M and b are understood, we shall not continue to indicate the arguments of these functions. We are interested in $IP - \overline{IP}$, but must settle for $LP - \overline{IP}$, in the knowledge that since $IP \leq LP$ we know $0 \leq IP-\overline{IP} \leq LP-\overline{IP}$. We now show that $LP-\overline{IP} < \delta$ where

δ will be specified and is independent of M and b.

THEOREM 2.7. The loss function, $LP - \overline{IP}$, is strictly less than $Z(j^*) - Z(0)$, for any M and b.

PROOF. We will consider three cases:

1) $M \leq \frac{b}{j^*}$. In this case, $LP - \overline{IP} = 0$ which is certainly less than $Z(j^*) - Z(0)$.

2) $M \geq \frac{b}{\hat{j}}$. In this case

$$LP - \overline{IP} = b\left(\frac{Z(\hat{j})-Z(0)}{\hat{j}}\right) + MZ(0) - \left[\frac{b}{\hat{j}}\right]\left(Z(\hat{j}) - Z(0)\right) - MZ(0)$$

$$= \left(\frac{b}{\hat{j}} - \left[\frac{b}{\hat{j}}\right]\right)(Z(\hat{j})-Z(0))$$

and we know $\frac{b}{\hat{j}} - \left[\frac{b}{\hat{j}}\right] = \frac{b \bmod \hat{j}}{\hat{j}} = \frac{r_1}{\hat{j}}$

so that $LP - \overline{IP} = \frac{r_1}{\hat{j}}(Z(\hat{j})-Z(0))$ but $0 \leq r_1 < \hat{j}$ implies that $LP-\overline{IP} < Z(\hat{j})-Z(0) < Z(j^*)-Z(0)$.

3) We have $\frac{b}{\hat{j}^{(n)}} < M \leq \frac{b}{\hat{j}^{(n-1)}}$ for some n, $2 \leq n \leq \bar{n}$. In this case,

$$LP - \overline{IP} = \frac{b-M\hat{j}^{(n-1)}}{\hat{j}^{(n)}-\hat{j}^{(n-1)}} - \left[\frac{b-M\hat{j}^{(n-1)}}{\hat{j}^{(n)}-\hat{j}^{(n-1)}}\right] \quad (Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)}))$$

$$= r_n\left(\frac{Z(\hat{j}^{(n)}-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}-\hat{j}^{(n-1)}}\right)$$

So $LP-\overline{IP} < Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})$. For $2 \leq n \leq \bar{n}$,

$Z(j^*) = Z(j^{(\bar{n})}) \geq Z(\hat{j}^{(n)})$ and $Z(0) < Z(\hat{j}^{(n-1)})$ by Lemma II. 4. 2 and so $LP-\overline{IP} < Z(j^*)-Z(0)$. Since we have covered all possible combinations of M and b, the result holds.

COROLLARY 1. For a given M and b, which determine n, the potential loss, $LP - \overline{IP}$, is between 0 and

$$\left(\frac{\hat{j}^{(n)}\hat{-j}^{(n-1)}-1}{\hat{j}^{(n)}\hat{-j}^{(n-1)}}\right)(Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})) .$$

PROOF. From the proof of the theorem, we see that

$$LP - \overline{IP} = r_n\left(\frac{Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)})}{\hat{j}^{(n)}\hat{-j}^{(n-1)}}\right)$$

for $n \geq 1$ and $0 \leq r_n < \hat{j}^{(n)}\hat{-j}^{(n-1)}$. Since r_n is an integer, we have the desired range.

We define the percentage potential loss to be $\frac{LP-\overline{IP}}{LP}$ which we now show goes to zero as M increases, for fixed b.

PROPOSITION 2.3. Given b, $\lim_{M\to\infty} \frac{LP-\overline{IP}}{LP} = 0$.

PROOF. We have that $LP-\overline{IP} < Z(j^*)-Z(0)$ for any M. Now for $M \geq \frac{b}{\hat{j}}$, the value of the LP is $\frac{b}{\hat{j}}(Z(\hat{j})-Z(0)) + MZ(0)$.

Therefore,

$$\lim_{M \to \infty} \frac{LP - \overline{IP}}{LP} \leq \lim_{M \to \infty} \frac{Z(j^*) - Z(0)}{\frac{b}{\hat{j}}(Z(\hat{j}) - Z(0)) + MZ(0)} = 0$$

so the $\lim\limits_{M \to \infty} \frac{LP-\overline{IP}}{LP} \leq 0$ but $\frac{LP-\overline{IP}}{LP} \geq 0$ for all M, since $LP-\overline{IP} \geq 0$ and the value of the optimal solution to the LP is greater than the value of the feasible solution, $MZ(0) \geq 0$ as a consequence of the fact that $Z(0) \geq 0$. This, it must be that $\lim\limits_{M \to \infty} \frac{LP-\overline{IP}}{LP} = 0$.

In order to illustrate these ideas, consider some examples.

<u>EXAMPLE 2.1.</u> Suppose we have a knapsack characterized by the following parametric generic knapsack function:

$Z(j)$	$\frac{Z(j)-Z(0)}{j}$	$\frac{Z(j)-Z(\hat{j})}{j-\hat{j}}$
$Z(0) = 915$	0	--
$Z(1) = 950$	35	--
$Z(2) = 1025$	55	0
$Z(3) = 1029$	38	4
$Z(4) = 1095$	45	35
$Z(5) = 1115$	40	30
$Z(6) = 1041$	21	4

We see that $\hat{j} = 2$, $\hat{j}^{(2)} = 4$, $\hat{j}^{(3)} = j^* = 5$ and $j_{max} = 6$. Suppose we fix an arbitrary value of b and consider the solution as M varies. Let $b = 120$ and consider the following:

M	Solution	$\overline{IP}$ Value	LP Value	$\delta = LP - \overline{IP}$
10	$X_5 = 10$	11150	11150	0
18	$X_5 = 18$	20070	20070	0
23	$X_5 = 23$	25645	25645	0
24	$X_5 = 24$	26760	26760	0
25	$X_5 = 20 \quad X_4 = 5$	27775	27775	0
26	$X_5 = 17 \quad X_4 = 10$	28790	28790	0
27	$X_5 = 12 \quad X_4 = 15$	29805	29805	0
28	$X_5 = 8 \quad X_4 = 20$	30820	20820	0
...	...	...	...	...
30	$X_4 = 30$	32850	32850	0
31	$X_4 = 29 \quad X_2 = 2$	33805	33805	0
32	$X_4 = 28 \quad X_2 = 4$	34760	34760	0
33	$X_4 = 27 \quad X_2 = 6$	35715	35715	0
34	$X_4 = 26 \quad X_2 = 8$	36670	36670	0
...	...	...	...	...
60	$X_2 = 60$	61500	61500	0
61	$X_2 = 60 \quad X_0 = 1$	62415	62415	0
62	$X_2 = 60 \quad X_0 = 2$	63330	63330	0
63	$X_2 = 60 \quad X_0 = 3$	64245	64245	0
64	$X_2 = 60 \quad X_0 = 4$	65160	65160	0

where we have $\frac{b}{\lambda_j} = 60$, $\frac{b}{\lambda_j(2)} = 30$, $\frac{b}{\lambda_j(3)} = 24$ and all variables not listed equal zero.

This example, of course had parameters which guaranteed that $r_n = 0$ for all n, thus ensuring that the heuristic solution was optimal. This serves to illustrate the value of flexibility in establishing the constraint on the number of fillers. That is, for fixed M, if the factory has some small leeway in establishing b, they can choose b such that $\frac{b-M\hat{j}^{(n-1)}}{\hat{j}^{(n)}-\hat{j}^{(n-1)}}$ is an integer. This involves a change of, at most, the minimum of r_n or $((\hat{j}^{(n)}-\hat{j}^{(n-1)})-r_n)$ units either of which are strictly less than $\hat{j}^{(n)}-\hat{j}^{(n-1)}$. Of course, if $r_n = 0$, then no change is necessary.

EXAMPLE 2.2. In order to demonstrate the effect of a remainder, let us consider the same situation, but require $b = 119$. We note that $Z(\hat{j})-Z(0) = 110$ and $Z(j^*)-Z(0) = 200$. We now have the following (as shown on page 84): where

$$\left[\frac{b}{\hat{j}}\right] = 59, \quad \left[\frac{b}{\hat{j}^{(2)}}\right] = 29,$$

and

$$\left[\frac{b}{\hat{j}^{(3)}}\right] = 23.$$

We note that the loss is always strictly less than $Z(j^*)-Z(0)$. The fact that the loss is constant over any interval $\frac{b}{\hat{j}^{(n)}} < M \leq \frac{b}{\hat{j}^{(n-1)}}$ in this case is a direct consequence of the fact that $\hat{j}^{(n)}-\hat{j}^{(n-1)} = 2$ and so r_n is either 0 or 1. Since b is odd and $\hat{j}^{(n)}$ is even, the remainder is always 1. This does not hold, of course, in general, and the loss in any interval can fluctuate between 0 and $\frac{\hat{j}^{(n)}-\hat{j}^{(n-1)}-1}{\hat{j}^{(n)}-\hat{j}^{(n-1)}}(Z(\hat{j}^{(n)})-Z(\hat{j}^{(n-1)}))$.

M	Solution	$\overline{IP}$ Value	LP Value	$\delta = LP - \overline{IP}$
10	$X_5 = 10$	11150	11150	0
18	$X_5 = 18$	20070	20070	0
23	$X_5 = 23$	25645	25645	0
24	$X_5 = 23 \quad X_4 = 1$	26740	26740	0
25	$X_5 = 19 \quad X_4 = 6$	27755	27755	0
26	$X_5 = 15 \quad X_4 = 11$	28770	28770	0
27	$X_5 = 11 \quad X_4 = 16$	29785	29785	0
28	$X_5 = 7 \quad X_4 = 21$	30800	30800	0
29	$X_5 = 3 \quad X_4 = 26$	31815	31815	0
30	$X_4 = 29 \quad X_2 = 1$	32780	32815	35
31	$X_4 = 28 \quad X_2 = 3$	33735	33770	35
32	$X_4 = 27 \quad X_2 = 5$	34690	34725	35
33	$X_4 = 26 \quad X_2 = 7$	35645	35680	35
34	$X_4 = 25 \quad X_2 = 9$	36600	36635	35
...	...	...	...	...
59	$X_4 = 0 \quad X_2 = 59$	60475	60510	35
60	$X_2 = 59 \quad X_0 = 1$	61390	61445	55
61	$X_2 = 59 \quad X_0 = 2$	62305	62360	55
62	$X_2 = 59 \quad X_0 = 3$	63220	63275	55
63	$X_2 = 59 \quad X_0 = 4$	64135	64190	55
64	$X_2 = 59 \quad X_0 = 5$	65050	65105	55
...	...	...	...	...

It should be clear that while, if $\Delta = LP - \overline{IP}$, $\Delta = 0$ implies that our solution is optimal, $\Delta > 0$ does not necessarily imply that $IP \neq \overline{IP}$. We simply do not know.

On the Greedy Solution

Up to this point we have formulated a simple solution and shown that the potential difference between our solution and the optimal solution to the Basic Problem is relatively small and is bounded by a quantity which does not depend on the number of knapsacks in a shift. We are interested in determining when our heuristic solution is optimal and have determined that it is whenever that solution corresponds to an integer solution to the associated linear program. In general, we can say little else about the optimality of any solution to the Basic Problem in the absence of specific values for the parameters. There is one case in which we can make additional statements about the optimality of the solution, and where we can improve upon our solution with little extra computation. That is the case when $b \leq M$. So for the balance of this section, we assume that $M \geq b$. In this instance, recall from Section 2 of Chapter II that an alternate formulation was possible which transformed the two constraint Basic Problem into a one constraint standard knapsack problem which we repeat here:

$$\text{(205b)} \qquad \text{maximize} \quad \sum_{j=1}^{j^*} X_j(Z(j)-Z(0)) + MZ(0)$$

$$\text{subject to} \quad \sum_{j=1}^{j^*} jX_j \leq b$$

$$X_j \geq 0,\ X_j \text{ integer},\ j = 1, \ldots, j^*.$$

(We know that we need only consider $j \leq j^*$.) Since $M \geq b \geq \hat{j}$, we also know that the optimal solution to the linear program (301) in this case is $X_{\hat{j}} = \frac{b}{\hat{j}}$, $X_0 = M - \frac{b}{\hat{j}}$, $X_j = 0$ otherwise, where $\hat{j}$ is chosen as usual. Clearly, this solution less the X_0 component is the optimal solution to the following relaxation of problem (205b):

$$\text{(302)} \qquad \text{maximize} \quad \sum_{j=1}^{j^*} X_j(Z(j)-Z(0)) + MZ(0)$$

$$\text{subject to} \quad \sum_{j=1}^{j^*} jX_j \leq b$$

$$X_j \geq 0,\ j = 1, \ldots, j^*.$$

In reverse, we can obtain the solution to the equivalent LP (301) by solving (302) and setting $X_0 = M - \sum_{j=1}^{j^*} X_j$. So, for the purposes of this section, we can concentrate on finding optimal solutions to the Basic Problem by searching for optimal solutions to (205b) and then setting $X_0 = M - \sum_{j=1}^{j^*} X_j$, since X_j integer, $j \geq 1$, implies that X_0 is also integer.

A wealth of algorithms are available for solving (205b) given specific parameters. For a well known set of these, see Gilmore and Gomory [2]. We note in passing that Section 5 of their paper also contains an algorithm for solving the Basic Problem since, in the terminology of Gilmore and Gomory, the Basic Problem can be considered as a specialized knapsack function with a "cutting knife" limitation. Also, as justification for our approach, we quote from the authors with regard to their algorithm, "It has the disadvantage, however, of sometimes requiring long calculations."

We wish to be able to make more general statements about optimal solutions to the Basic Problem. To do so, we shall review some recent theoretical work about a heuristic solution to (205b), called the "Greedy Solution" in the literature.

The results stated in following paragraphs are taken from a paper by Magazine, Nemhauser and Trotter [5], subsequently amplified by Hu and Lenard [4]. They consider the following problem: (actually, they discuss a minimization problem but we have made the necessary changes).

(303) $$f_k(y) = \max \sum_{j=1}^{k} C_j X_j$$

$$\text{subject to} \quad \sum_{j=1}^{k} a_j X_j = y$$

$$X_j \geq 0, \text{ integer}, \ 1 \leq j \leq k$$

where y ranges over positive integers and $k = 1, \ldots, n$. They also assume:

1) a_j and C_j are integers, $1 \leq j \leq n$.

2) $\frac{C_1}{a_1} \leq \frac{C_2}{a_2} \leq \ldots \leq \frac{C_n}{a_n}$.

3) $a_j > 0$, $1 \leq j \leq n$.

4) for some $1 \leq j \leq n$, $a_j = 1$.

Assumptions 1) and 2) are made without loss of generality. 4) guarantees a feasible solution for all positive integers and 3) assures that $f_k(y)$ is bounded. Also, assumptions 3) and 4) imply

5) $1 = a_1 \leq a_j$, $2 \leq j \leq n$.

If C_1 is set equal to zero, one obtains an inequality constrained knapsack problem such as (205b). The Magazine paper defines the "greedy solution" recursively by

$$X_j = [\frac{y}{a_k}], \qquad j = k$$

$$X_j = \left[y - \sum_{i=j+1}^{k} a_i X_i \right], \qquad j = k-1, k-2, \ldots, 1$$

where $[\cdot]$ is the greatest integer notation. They denote the value of the greedy solution by $g_k(y)$ and point out that this solution is "good" in our sense of "simple" in comparison to the algorithms in Gilmore and Gomory [2].

They say the greedy solution is optimal when $f_k(y) = g_k(y)$ for all positive integers y (or $f_k = g_k$) and the paper then provides results to establish conditions when $f_k = g_k$. The necessary and sufficient conditions obtained permit one to check if $f_n = g_n$. As we shall note, our heuristic solution is analogous to their greedy solution. Thus, this result will tell us when the greedy solution to (205b) is optimal for any b. The result most useful to us is the following theorem, proved in Magazine et al. [5] and the proof simplified in Hu and Lenard [6]. Obviously $f_1 = g_1$, since $a_1 = 1$.

<u>THEOREM 2.8.</u> Fix $k \geq 1$, let $f_k = g_k$ and assume $a_{k+1} > a_k$. Define m_k and r_k uniquely by the relations $0 \leq r_k < a_k$ and $a_{k+1} = m_k a_k - r_k$ (i.e. m_k is the least integer greater than or equal to a_{k+1}/a_k). Then, the following are equivalent:

a. $f_{k+1} = g_{k+1}$

b. $f_{k+1}(m_k a_k) = g_{k+1}(m_k a_k)$

c. $C_{k+1} + g_k(r_k) \geq m_k C_k$.

Let us establish the correspondence between problems (205b) and (303). We are interested in whether $f_n = g_n$. We have that $C_k = Z(j) - Z(0)$ for some j, $1 \leq j \leq j^*$ and a_k equals the same j. When we rearrange the problem as in assumption 2) of problem (303), it must be that $\frac{C_n}{a_n} = \frac{Z(\hat{j}) - Z(0)}{\hat{j}}$. Our solution says to set $X_n = X_{\hat{j}} = [\frac{b}{\hat{j}}]$ which is equivalent to the first step of the greedy solution procedure. Also, for $k = n-1, n-2,$ if $a_k = j > a_n = \hat{j}$ then $[\frac{b - \hat{j}X_{\hat{j}}}{\hat{j}}] = 0$ since we know $b - \hat{j}X_{\hat{j}} < \hat{j}$. In this event, the treedy solution says to set $X_j = 0$ just as in our solution to (205b). The difference lies in the fact that our solution says to set $X_{\hat{j}} = [\frac{b}{\hat{j}}]$, $X_0 = M - X_{\hat{j}}$ and $X_j = 0$ otherwise, whereas the greedy solution says that if $b - \hat{j}X_{\hat{j}} > 0$, as you recursively proceed down the a_k's, $k = n-1, n-2, \ldots$ the first time you find an $a_k = j'$ which "fits," i.e. $j' \leq b - \hat{j}X_{\hat{j}} < \hat{j}$, you use up as much as possible of the remainder by setting $X_{j'} = [\frac{b - \hat{j}X_{\hat{j}}}{j'}] \geq 1$. One then continues the search for the largest $j \leq b - \hat{j}X_{\hat{j}} - j'X_{j'}$, etc. This greedy procedure is exactly that alluded to when we discussed the remainder problem; that is, a rational knapsack stuffer, after having stuffed $[b/\hat{j}]$ knapsacks with $\hat{j}$ and having some restricted fillers left over, would fill as many of the remaining knapsacks as he could with the next most valuable $j \geq 0$ which uses no more than the amount left over, $b - \hat{j}[\frac{b}{\hat{j}}]$.

He would continue this until he either used up all the remainder or could no longer find a feasible j such that $Z(j) \geq Z(0)$. Only then would he fill the remaining knapsacks with 0. If we were to modify our algorithm to allow for this procedure when $\frac{b}{\hat{j}} \leq M$ (or whenever the integer arithmetic forced us to have a remainder of restricted fillers after we had implemented the solution) then the optimality of the greedy solution is of interest to us, because it shows how to realize an optimal solution to (205b), based on a simple procedure. Unfortunately, the useful result applies to our case only under certain conditions. First of all, the theorem's requirement that $a_k < a_{k+1}$ is equivalent to requiring that $\frac{Z(1)-Z(0)}{1} \leq \frac{Z(2)-Z(0)}{2} \leq \ldots \leq \frac{Z(j^*)-Z(0)}{j^*}$ in our problem. As we shall see, it is still not enough to assume $\frac{Z(j)-Z(0)}{j} \geq \frac{Z(j-1)-Z(0)}{j-1}$ for all j, $1 \leq j \leq j^*$. Note that this assumption implies that $j^* = \hat{j}$. We must assume a stronger condition, that $Z(\cdot)$ is discretely convex on the set $\{0, 1, \ldots, j^*\}$ as defined in Section 4 of Chapter II. Given this assumption, we show first that shifting the origin does not affect the definition. Then we show that the second assumption implies the first and give a counter example to disprove the converse.

LEMMA 2.1. If $Z(\cdot)$ is discretely convex on the set $\{0, 1, \ldots, j^*\}$, then $Z(\cdot) - Z(0)$ is discretely convex on the set $\{0, 1, \ldots, j^*\}$.

PROOF. We have, for $1 \leq j \leq j^*-1$, that $\frac{Z(j-1)+Z(j+1)}{2} \geq Z(j)$. If we subtract $Z(0)$ from both sides, we get $\frac{(Z(j-1)-Z(0))+(Z(j+1)-Z(0))}{2} \geq Z(j) - Z(0)$. To check the end points, 0 and j^*, we investigate the behavior at $j = 1$ and $j = j^*-1$, respectively. For $j = 1$, we must have $\frac{Z(2)-Z(0)}{2} \geq Z(1) - Z(0)$. But we know $\frac{Z(2)}{2} \geq Z(1)$ so the requirement is met. For $j = j^*-1$, the hypothesis tells us that $\frac{Z(j^*-2)+Z(j^*)}{2} \geq Z(j^*-1)$ so subtracting $Z(0)$ from both sides establishes the necessary condition.

LEMMA 2.2. If $Z(\cdot)$ is discretely convex on the set $\{0, 1, \ldots, j^*\}$, then $\frac{Z(j+1)-Z(0)}{j+1} \geq \frac{Z(j)-Z(0)}{j}$ for $j = 1, \ldots, j^*$.

PROOF. We shall proceed by induction. Now, Lemma 2.1 above showed that $Z(j) - Z(0)$ is locally discretely convex at any point j in the interval. Therefore, $\frac{Z(j+1)-Z(0)+Z(j-1)-Z(0)}{2} \geq Z(j) - Z(0)$ for any $j \geq 1$. Let $j = 1$, then the definition says that $\frac{Z(2)-Z(0)}{2} \geq Z(1) - Z(0)$ which establishes the result for $j = 1$. So, assume $\frac{Z(j)-Z(0)}{j} \geq \frac{Z(j-1)-Z(0)}{j-1}$, $j \geq 2$. Then, $Z(j) - Z(0)$ locally discretely convex at j implies

$$\frac{Z(j+1)-Z(0)}{j+1} + \frac{Z(j-1)-Z(0)}{j+1} \geq 2\,\frac{(Z(j)-Z(0))}{j+1}$$

but $Z(j-1) - Z(0) \leq \frac{(j-1)(Z(j)-Z(0))}{j}$ so

$$\frac{Z(j-1)-Z(0)}{j+1} \leq \frac{(j-1)(Z(j)-Z(0))}{j(j+1)}$$

which in turn implies that

$$\frac{Z(j+1)-Z(0)}{j+1} + \frac{(j-1)}{j(j+1)}(Z(j)-Z(0)) \geq 2\,\frac{Z(j)-Z(0)}{j+1}$$

or

$$\frac{Z(j+1)-Z(0)}{j+1} \geq \left(\frac{2}{j+1} - \frac{(j-1)}{j(j+1)}\right)(Z(j) - Z(0)).$$

The term in parens is equal to $\frac{2j-j+1}{j(j+1)} = \frac{j+1}{j(j+1)} = \frac{1}{j}$ and we have the desired result for $2 \leq j \leq j^*$.

We provide the following counter example to show that $\frac{Z(j)-Z(0)}{j} \leq \frac{Z(j+1)-Z(0)}{j+1}$ for $1 \leq j \leq j^*-1$ does not imply that $Z(j) - Z(0)$ is discretely convex over the set $\{0, 1, \ldots, j^*\}$. Let $Z(5) - Z(0) = 50$, $Z(4) - Z(0) = 36$, $Z(3) - Z(0) = 21$. Then

$$\frac{Z(5)-Z(0)}{5} = 10 > \frac{Z(4)-Z(0)}{4} = 9 > \frac{Z(3)-Z(0)}{3} = 7$$

but

$$\frac{Z(5)-Z(0)+Z(3)-Z(0)}{2} = \frac{71}{2} \not\geq Z(4) - Z(0) = 36.$$

We are now in a position to apply Theorem 2.8 to our problem.

THEOREM 2.9. If $Z(\cdot)$ is discretely convex on the set $\{0, 1, \ldots, j^*\}$, then the greedy solution to (205b) is optimal for any b (hence, optimal in (205) where $M \geq b$ when we set $X_0 = M - \sum_{j=1}^{j^*} X_j$).

PROOF. If we can show that $C_{k+1} + g_k(r_k) \geq m_k C_k$ for all k, $1 \leq k \leq j^*-1$, and $a_{k+1} > a_k$ as in Theorem 2.8 we can apply that result to problem (302) to obtain our result. First, we see that Lemma 2.2 shows that under the hypothesis, $\frac{Z(j)-Z(0)}{j} \leq \frac{Z(j+1)-Z(0)}{j+1}$, $1 \leq j \leq j^*-1$ so that we have $C_k = Z(k) - Z(0)$ and $a_k = k$. Clearly $a_{k+1} > a_k$. Also, $k+1 = 2k - (k-1)$, so that $m_k = 2$ and $r_k = (k-1)$ as defined in Theorem 2.8. Consequently, $g_k(r_k) = g_k(k-1)$.

The solution corresponding to $g_k(k-1)$ is $X_k = 0$, $X_{k-1} = 1$, $X_j = 0$ otherwise, giving us that $g_k(r_k) = C_{k-1}$. Therefore, in order for the greedy solution to be optimal for fixed $k+1$, i.e. $g_{k+1} = f_{k+1}$, we require $c_{k+1} + g_k(r_k) \geq m_k C_k$ or $Z(k+1) - Z(0) + Z(k-1) - Z(0) \geq 2(Z(k)-Z(0))$ but this is precisely the definition of local discrete convexity, which is valid for all k, $1 \leq k \leq j^*-1$. Therefore, the greedy solution to (205b) is optimal for any b.

The paper by Magazine et al. [5] also provides a generalized version of Theorem 2.8 which does not require $a_k < a_{k+1}$. This theorem can be applied to our case when specific parameters are cited, but does not provide us with insights in the generality we desire in this paper and so will not be restated here.

Improving the Solution

Let us restrict ourselves to the situation where $M \geq \frac{b}{\hat{j}}$. In this case we chose to set $X_{\hat{j}} = [\frac{b}{\hat{j}}]$, $X_0 = M - [\frac{b}{\hat{j}}]$, $X_j = 0$ otherwise, which results in a value of $[\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) + MZ(0)$. We did so because we know $\frac{b}{\hat{j}}(Z(\hat{j})-Z(0)) \geq \frac{b}{j}(Z(j)-Z(0))$ for all other j by definition, and so, if our problem were a linear program, we would realize the optimal value by setting $X_{\hat{j}} = \frac{b}{\hat{j}}$, $X_0 = M - \frac{b}{\hat{j}}$, $X_j = 0$ otherwise. Unfortunately, the fact that

$$\frac{b}{\hat{j}}(Z(\hat{j})-Z(0)) \geq \frac{b}{j}(Z(j)-Z(0)), \quad 0 \leq j \leq j^* \tag{A}$$

does not guarantee that

$$[\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) \geq [\frac{b}{j}](Z(j)-Z(0)), \quad 0 \leq j \leq j^*, \tag{B}$$

an anomaly of modular arithmetic. In this section we will provide a necessary condition to determine when the above situation does not hold. We will establish the fact that for b large enough A does imply B. Then, we provide a slight modification to our algorithm to take advantage of this situation which can only improve our solution with very little additional complexity.

First we wish to find conditions for the existence of a j such that although $\frac{b}{\hat{j}} Z(\hat{j})-Z(0) \geq \frac{b}{j}(Z(j)-Z(0))$ we have a situation where $[\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) < [\frac{b}{j}](Z(j)-Z(0))$. In order to make subsequent analysis easier we shall make some notational definitions.

<u>DEFINITION 2.1</u>. Let

$$\rho_i = \frac{\frac{Z(i)-Z(0)}{i}}{\frac{Z(\hat{j})-Z(0)}{\hat{j}}} \quad .$$

Note that if $\frac{Z(i)-Z(0)}{i} \geq 0$, which would be the only i we might be interested in, then $0 \leq \rho_i \leq 1$ by the definition of $\hat{j}$.

<u>DEFINITION 2.2</u>. Let $\Delta_i = [\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) - [\frac{b}{i}](Z(i)-Z(0))$.

Δ_i is the difference between the value of our solution and the value of one where we set $X_i = [\frac{b}{i}]$, $X_0 = M - [\frac{b}{i}]$ and $X_j = 0$ otherwise.

<u>DEFINITION 2.3</u>. Let $r_i = b \bmod i$.

<u>PROPOSITION 2.4.</u> If $\Delta_i < 0$ then $r_i < \frac{r_j}{\rho}$.

PROOF. We know $\frac{b}{\hat{j}}(Z(\hat{j}) - Z(0) \geq \frac{b}{i}(Z(i)-Z(0))$. Substituting $[\frac{b}{j}] + \frac{r_j}{j} = \frac{b}{j}$ we have $\Delta_i \geq r_i \frac{Z(i)-Z(0)}{i} - r_{\hat{j}} \frac{Z(\hat{j})-Z(0)}{\hat{j}}$. Thus, if $\Delta_i < 0$, then $r_i \frac{Z(i)-Z(0)}{i} - r_{\hat{j}} \frac{Z(\hat{j})-Z(0)}{\hat{j}} < 0$. Furthermore $\Delta_i < 0$ implies that $[\frac{b}{i}]Z(i) - Z(0) > 0$ since $[\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) \geq 0$ which in turn implies $\frac{Z(i)-Z(0)}{i} > 0$ so that $\rho > 0$ and we have $r_i < r_j(\frac{1}{\rho})$ as was to be shown.

COROLLARY 1. If $\rho = 1$ and $\Delta_i < 0$ then $r_i < r_j$.

PROOF. Trivial.

COROLLARY 2. If $r_{\hat{j}} = 0$, then $\Delta_i \geq 0$.

PROOF. $r_i \geq 0$ implies $r_i \geq \frac{r_{\hat{j}}}{\rho}$ and so the corollary is the contrapositive of the theorem.

This is, of course, a restatement of the fact that $r_{\hat{j}} = 0$ implies our solution is optimal in the Basic Problem.

COROLLARY 3. If $\Delta_i < 0$, then $Z(i) - Z(0) > 0$.

PROOF. From the proof of the theorem.

We can show that in the case where $\rho_i = 1$, i.e., that $\frac{Z(i)-Z(0)}{i} = \frac{Z(\hat{j})-Z(0)}{\hat{j}}$ that the respective remainders determine which is a better solution.

THEOREM 2.10. When $\rho_i = 1$, then $\Delta_i < 0$ if and only if $r_i < r_{\hat{j}}$.

PROOF. (=>) follows from Corollary 1 above. (<=) Since $\rho_i = 1$, $\frac{b}{i}(Z(i)-Z(0)) = \frac{b}{\hat{j}}(Z(\hat{j})-Z(0))$ and we saw in the previous result that this implies $\Delta_i = (r_i - r_{\hat{j}})\left(\frac{Z(\hat{j})-Z(0)}{\hat{j}}\right)$. Since $\frac{Z(\hat{j})-Z(0)}{\hat{j}} \geq 0$, $r_i < r_{\hat{j}}$ establishes that $\Delta_i < 0$.

This result says that if we have an i, necessarily less than $\hat{j}$ because of the definition of $\hat{j}$, such that $\rho_i = 1$, then we choose the solution to use according to the minimum of $b \bmod i$ and $b \bmod j$.

This is easily made a part of the solution algorithm since it involves a trivial calculation.

In the case when $\rho_i < 1$ for all i, we can show that for b sufficiently large, our original choice of $\hat{j}$ is correct, i.e. that $\Delta_i \geq 0$ for any i.

THEOREM 2.11. When $\rho_i < 1$ for all $i < j^*$ such that $\frac{Z(i)-Z(0)}{i} > 0$, let $\rho = \max_i \rho_i$. Certainly $\rho > 0$. If $b \geq \frac{\hat{j}-1}{1-\rho}$ then $\Delta_i \geq 0$ for all $i < j^*$.

PROOF. $b \geq \frac{\hat{j}-1}{1-\rho}$ implies that $b(1-\rho_i) \geq \hat{j}-1$ for all i such that $Z(i)-Z(0) > 0$. Now

$$b(1-\rho) = b\left(\frac{Z(\hat{j})-Z(0)}{\hat{j}} - \frac{Z(i)-Z(0)}{i} \right) \frac{\hat{j}}{Z(\hat{j})-Z(0)} ,$$

so we have

$$b\left(\frac{Z(\hat{j})-Z(0)}{\hat{j}} - \frac{Z(i)-Z(0)}{i} \right) \geq (\hat{j}-1) \frac{Z(\hat{j})-Z(0)}{\hat{j}}$$

for any i such that $Z(i)-Z(0) > 0$. But,

$$\frac{(\hat{j}-1)}{\hat{j}}(Z(\hat{j})-Z(0)) \geq r_{\hat{j}} \frac{Z(\hat{j})-Z(0)}{\hat{j}}$$

since $r_{\hat{j}} \leq \hat{j}-1$. Also, for i such that $Z(i)-Z(0) > 0$,

$$(\hat{j}-1) \frac{Z(\hat{j})-Z(0)}{\hat{j}} \geq r_{\hat{j}} \frac{Z(\hat{j})-Z(0)}{\hat{j}} - \frac{r_i(Z(i)-Z(0))}{i} .$$

Therefore, combining the two inequalities, we have

$$\frac{b}{\hat{j}}(Z(\hat{j})-Z(0)) - \frac{b}{i}(Z(i)-Z(0)) \geq r_{\hat{j}} \frac{Z(\hat{j})-Z(0)}{\hat{j}} - r_i \frac{Z(i)-Z(0)}{i}$$

which is equivalent to $\Delta_i \geq 0$ for i such that $Z(i)-Z(0) > 0$. We can show by the contrapositive to Corollary 3 to Proposition 2.4 that $Z(i)-Z(0) \leq 0$ implies $\Delta_i \geq 0$ for any b, thus completing the proof.

COROLLARY 1. If $\Delta_i < 0$ then $b < \frac{\hat{j}-1}{1-\rho}$.

PROOF. By the contrapositive to the theorem.

We can generate tighter bounds than those presented in the preceding theorem if we note that $r_{\hat{j}}\left(\frac{Z(\hat{j})-Z(0)}{\hat{j}}\right) - r_i\left(\frac{Z(i)-Z(0)}{i}\right)$ is equal to $(r_{\hat{j}}-\rho_i r_i)\frac{Z(\hat{j})-Z(0)}{\hat{j}}$ and that

$$b\left(\frac{Z(\hat{j})-Z(0)}{\hat{j}} - \frac{(Z(i)-Z(0))}{i}\right) = b(1-\rho_i)\frac{Z(\hat{j})-Z(0)}{\hat{j}}$$

Therefore,

$$b \geq \frac{r_{\hat{j}}-\rho_i r_i}{1-\rho_i}$$

for all i such that $Z(i)-Z(0) > 0$ implies that $\Delta_i \geq 0$. The proof of this statement follows the same line as the above theorem and so we state the result as a corollary without proof.

<u>COROLLARY 2.</u> Under the conditions of Theorem 2.11, $\Delta_i \geq 0$ for all $i \leq j^*$ when $b \geq \frac{r_{\hat{j}}-\rho_i r_i}{1-\rho_i}$ for all i.

<u>PROOF.</u> As above with the additional note that if $r_{\hat{j}}-\rho_i r_i \leq 0$ we know $\Delta_i \geq 0$. Since $b > 0$ the result holds trivially.

The net result of the discussion in this subsection is that we can add the following modification to step 3 of our solution procedure for the Basic Problem:

3a. If $M \geq \frac{b}{\hat{j}}$ find $\max_j\{(\min(M, [\frac{b}{j}]))(Z(j)-Z(0))\}$ and, for the i which corresponds to the maximum value, set

$X_i = \min\{M, [\frac{b}{i}]\}$, $X_0 = M - X_i$, $X_j = 0$ otherwise. The value of this solution will be greater than or equal to $[\frac{b}{\hat{j}}](Z(\hat{j}) - Z(0)) + MZ(0)$.

STOP.

3. The Lower Bound Case

The Linear Program Relaxation

The Basic Problem in the case where the constraint requires at least a certain number, b, of the restricted filler types to be used, was formulated in Section 2 of Chapter II as Problem (206):

$$\text{maximize} \quad \sum_{j=0}^{j_{max}} X_j Z(j)$$

$$\text{subject to} \quad \sum_{j=0}^{j_{max}} X_j = M$$

$$\sum_{j=1}^{j_{max}} jX_j \geq b$$

$$X_j \geq 0,\ X_j \text{ integer for all } j = 0, 1, \ldots, j_{max}.$$

As we noted in Theorem II. 4. 2, one can show that you should not consider any X_j such that $j < j^*$ as a candidate for the optimal solution to (206). Recall that in this "lower bound" case we choose the largest alternative optimal value of j to designate as j^* in contrast to choosing the smallest such value in the upper bound case. We do have the additional restriction that an optimal solution to (206) exists if and only if $Mj_{max} \geq b$, since otherwise no feasible solution exists. We note here that if $Mj^* \geq b$, the factory constraint is vacuous, since the optimal solution to (206) is clear in this situation; $X_{j^*} = M$, $X_j = 0$ otherwise. That is, the factory constraint poses no restraint on the knapsack stuffer at all in the sense that he can fill each knapsack with the optimal solution to the underlying knapsack problem and disregard the factory constraint. Therefore, we restrict our attention to the interesting case where $Mj^* < b$. Furthermore, the case where $j^* = j_{max}$ is also uninteresting because there are only two possibilities, either there is no feasible solution or $X_{j_{max}} = M$, $X = 0$ otherwise, is optimal. So, we assume $j^* < j_{max}$. As in the upper bound case, we shall write (206) as

$$\text{maximize} \quad \sum_{j=j^*}^{j_{max}} X_j Z(j)$$

$$\text{subject to} \quad \sum_{j=j^*}^{j_{max}} X_j = M$$

$$\sum_{j=j^*}^{j_{max}} jX_j \geq b, \quad b > 0$$

$$X_j \geq 0, \; X_j \text{ integer}, \; j = j^*, \ldots, j_{max} .$$

We could go to the trouble of converting the indices to some $j' = j - j^*$ for $j = j^*, \ldots, j_{max}$ but will not, in order to keep the underlying knapsack function's role clear.

As before, we shall work with a linear programming relaxation of (206);

$$\text{(304)} \qquad \text{maximize} \quad \sum_{j=j^*}^{j_{max}} X_j Z(j)$$

$$\text{subject to} \quad \sum_{j=j^*}^{j_{max}} X_j = M$$

$$\sum_{j=j^*}^{j_{max}} jX_j \geq b$$

$$X_j \geq 0, \; j = j^*, \ldots, j_{max} .$$

Since we have already elaborated the simplex procedure for this particular problem in Section 2, we shall dispose of the surplus variable rather than the slack variable, which provides the only difference in the A matrix and then proceed to the principal result, a counterpart to Theorem III. 2. 5. Let X_{su} stand for the surplus variable required to put (304) in standard form.

PROPOSITION 3. 1. $X_{su} = 0$ in an optimal solution to (304) when $Mj^* < b$ (and, of course $Mj_{max} \geq b$).

PROOF. Assume $X_{su} > 0$. Then the optimal Basis will be $B = \begin{bmatrix} 1 & 0 \\ j' & -1 \end{bmatrix}$ where $j^* \leq j' \leq j_{max}$. So $B^{-1} = \begin{bmatrix} 1 & 0 \\ j' & -1 \end{bmatrix}$, $C_B = (Z(j'), 0)$ and $\pi_j = Z(j') - Z(j)$, which implies that $j' = j^*$ since optimality requires $\pi_j \geq 0$ for all j. However the corresponding optimal basic feasible solution, $X_B = (X_{j^*}, X_{su})$ is, of course, $B^{-1}d = (M, Mj^*-b)$. We have assumed that $Mj^* < b$, so therefore $X_{su} < 0$, a contradiction. It must be that $X_{su} = 0$.

This result tells us that, just as in the upper bound case, the optimal solution to the linear program relaxation of our problem results in equality in the second constraint. So, in a loose sense, we will try to get as close as we can to equality in the second constraint when we find the optimal solution to (206).

We shall proceed directly to the solution to (304). First, we require the following definition and lemma, which have their counterparts for the upper bound case in Section 4 of Chapter II.

<u>DEFINITION 3.1.</u> Let $\bar{j}^{(1)}$ be the largest j such that

1) $0 \leq j^* < \bar{j}^{(1)} \leq j_{max}$ and

2) $\dfrac{Z(\bar{j}^{(1)})-Z(j^*)}{\bar{j}^{(1)}-j^*} \geq \dfrac{Z(j)-Z(j^*)}{j-j^*}$ for all j, $j^* < j \leq j_{max}$.

Then, given $\bar{j}^{(1)}$, we define sequentially, for $n \geq 2$, $\bar{j}^{(n)}$ to be the largest j such that

1) $\bar{j}^{(n-1)} < \bar{j}^{(n)} \leq j_{max}$ and

2) $\dfrac{Z(\bar{j}^{(n)})-Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)}-\bar{j}^{(n-1)}} \geq \dfrac{Z(j)-Z(\bar{j}^{(n-1)})}{j-\bar{j}^{(n-1)}}$ for all j,

$\bar{j}^{(n-1)} < j \leq j_{max}$. The sequence of $\bar{j}^{(n)}$'s is finite as a consequence of the fact that j_{max} if finite and consequently there exists an $\bar{n} \geq 1$ such that $\bar{j}^{(\bar{n})} = j_{max}$. We shall define $\bar{j}^{(0)} = j^*$ to simplify future discussion.

In order to find $\bar{j}^{(1)}$, which we shall denote hereafter as $\bar{j}$, we search for the largest j which results in the maximum of $\dfrac{Z(j)-Z(j^*)}{j-j^*}$ over $j > j^*$. Note that we are searching over a set of negative quantities. Upon finding $\bar{j}$, we will have found the smallest incremental loss per unit of restricted filler given we have to increase the number of fillers beyond j^*. Further, we are assured of the existence of at least $\bar{j}$, as long as we define $\bar{j} = j^*$ in the case where $j^* = j_{max}$, a case in which we have little real interest.

We defined $\bar{j}^{(0)} = j^*$, and so we can say that for $1 \leq n \leq \bar{n}$,

$$\frac{Z(\bar{j}^{(n)}) - Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} \geq \frac{Z(j) - Z(\bar{j}^{(n-1)})}{j - \bar{j}^{(n-1)}}$$

for all j such that $\bar{j}^{(n-1)} < j \leq j_{max}$. We can establish that:

LEMMA 3.1. For $1 \leq n \leq \bar{n}$,

1) If $\bar{j}^{(n-1)} < j < \bar{j}^{(n)}$, then

$$\frac{Z(j) - Z(\bar{j}^{(n-1)})}{j - \bar{j}^{(n-1)}} \leq \frac{Z(\bar{j}^{(n)}) - Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} \leq \frac{Z(\bar{j}^{(n)}) - Z(j)}{\bar{j}^{(n)} - j}$$

2) If $\bar{j}^{(n)} < j < j_{max}$, then

$$\frac{Z(\bar{j}^{(n)}) - Z(j)}{\bar{j}^{(n)} - j} < \frac{Z(j) - Z(\bar{j}^{(n-1)})}{j - \bar{j}^{(n-1)}} < \frac{Z(\bar{j}^{(n)}) - Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)} - \bar{j}^{(n-1)}}.$$

PROOF. This is exactly the same statement as parts 2b and 3 of Lemma II.4.2, except that we have replaced the symbol $\hat{j}^{(n)}$ with $\bar{j}^{(n)}$ to distinguish the lower bound case.

The reason we cannot just apply Lemma II.4.2 directly is that part 2a no longer holds (and part 1 is redundant). Instead, we can write this result.

LEMMA 3.2. If $j > \bar{j}^{(n)}$ for $1 \leq n \leq \bar{n}-1$, then $Z(j) < Z(\bar{j}^{(n)})$.

PROOF. We have that $Z(j)-Z(j^*) < 0$ and $j-j^* > \bar{j}-j^*$. Thus, $\frac{Z(j)-Z(j^*)}{j-j^*} > \frac{Z(j)-Z(j^*)}{\bar{j}-j^*}$ and it follows from the definition of $\bar{j}$ that $\frac{Z(\bar{j})-Z(j^*)}{\bar{j}-j^*} > \frac{Z(j)-Z(j^*)}{j-j^*}$. Then it must be that $Z(\bar{j})-Z(j^*) > Z(j)-Z(j^*)$ which establishes the result for $n = 1$. So assume it for $n-1$. Then $Z(\bar{j}^{(n)}) < Z(\bar{j}^{(n-1)})$ and $j-\bar{j}^{(n-1)} > \bar{j}^{(n)}-\bar{j}^{(n-1)}$ together imply that

$$\frac{Z(j)-Z(\bar{j}^{(n-1)})}{j-\bar{j}^{(n-1)}} > \frac{Z(j)-Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)}-\bar{j}^{(n-1)}}.$$

The definition of $\bar{j}^{(n)}$ leads to the fact that $Z(\bar{j}^{(n)})-Z(\bar{j}^{(n-1)}) > Z(j)-Z(\bar{j}^{(n-1)})$. This is equivalent to the desired result, which therefore holds for any n.

With these basic elements, we can now show:

THEOREM 3.1. When $\frac{b}{\bar{j}^{(n)}} \leq M < \frac{b}{\bar{j}^{(n-1)}}$, for $1 \leq n \leq \bar{n}$, the optimal solution to (303) is given by:

$$X_{\bar{j}^{(n)}} = \frac{b-M\bar{j}^{(n-1)}}{\bar{j}^{(n)}-\bar{j}^{(n-1)}}$$

$$X_{\bar{j}^{(n-1)}} = M - X_{\bar{j}^{(n)}}$$

$$X_j = 0 \qquad \text{otherwise.}$$

The optimal value of the objective function is

$$(b - M\bar{j}^{(n-1)})\left(\frac{Z(\bar{j}^{(n)}) - Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)} - \bar{j}^{(n-1)}}\right) + MZ(\bar{j}^{(n-1)})$$

which is equal to

$$M\bar{j}^{(n)} - b\left(\frac{Z(\bar{j}^{(n-1)}) - Z(\bar{j}^{(n)})}{\bar{j}^{(n-1)} - \bar{j}^{(n)}}\right) + MZ(\bar{j}^{(n)})$$

<u>PROOF.</u> Consider the basis for (304) containing $X_{\bar{j}^{(n)}}$ and $X_{\bar{j}^{(n-1)}}$,

$$B = \begin{bmatrix} 1 & 1 \\ \bar{j}^{(n-1)} & \bar{j}^{(n)} \end{bmatrix},$$

so that

$$B^{-1} = \frac{1}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} \begin{bmatrix} \bar{j}^{(n)} & -1 \\ -\bar{j}^{(n-1)} & 1 \end{bmatrix},$$

and the basis cost vector is $(Z(\bar{j}^{(n-1)}), Z(\bar{j}^{(n)}))$. Then

$$X_B = B^{-1}d = \frac{M\bar{j}^{(n)} - b}{\bar{j}^{(n)} - \bar{j}^{(n-1)}}, \frac{b - M\bar{j}^{(n-1)}}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} = (X_{\bar{j}^{(n-1)}} X_{\bar{j}^{(n)}}).$$

$X_B \geq 0$ since $\frac{b}{\bar{j}^{(n)}} \leq M < \frac{b}{\bar{j}^{(n-1)}}$ and

$$\sum_{j=j*}^{j_{max}} X_j = X_{\bar{j}^{(n-1)}} + X_{\bar{j}^{(n)}} = M$$

and further

$$\sum_{j=j*}^{j_{max}} jX_j = \bar{j}^{(n-1)} X_{\bar{j}^{(n-1)}} + \bar{j}^{(n)} X_{\bar{j}^{(n)}} = b.$$

So our proposed optimal solution is feasible. In order for our proposed solution to be optimal it must be that $\pi_k = C_B B^{-1} a_k - Z(k) \geq 0$ for all k. Now,

$$\pi_k = \frac{(\bar{j}^{(n)} - j)}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} Z(\bar{j}^{(n-1)}) + \frac{j - \bar{j}^{(n-1)}}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} Z(\bar{j}^{(n)}) - Z(k)$$

Suppose, for $k > \bar{j}^{(n-1)}$ that $\pi_k < 0$. Then by suitable rearrangement (see Theorem III. 2. 4) of the expression for π_k, we have that

$$\frac{Z(k) - Z(\bar{j}^{(n-1)})}{k - \bar{j}^{(n-1)}} > \frac{Z(\bar{j}^{(n)}) - Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)} - \bar{j}^{(n-1)}}$$

which contradicts the way we have chosen $\bar{j}^{(n)} > \bar{j}^{(n-1)}$. Suppose $k < \bar{j}^{(n-1)}$ and $\pi_k < 0$. If we let $\bar{j}^{(n)} - k = \bar{j}^{(n)} - \bar{j}^{(n-1)} - (k - \bar{j}^{(n-1)})$, we can show that $\pi_k < 0$ implies

$$(k-\bar{j}^{(n-1)}) \frac{(Z(\bar{j}^{(n)})-Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)}-\bar{j}^{(n-1)}} < Z(k) - Z(\bar{j}^{(n-1)})$$

and since $k < \bar{j}^{(n-1)}$, we have that

$$\frac{Z(\bar{j}^{(n)})-Z(\bar{j}^{(n-1)})}{\bar{j}^{(n)}-\bar{j}^{(n-1)}} > \frac{Z(k)-Z(\bar{j}^{(n-1)})}{k-\bar{j}^{(n-1)}} .$$

We can apply Corollary 1 to Lemma II. 4. 2 to this case if we let $\bar{j}^{(n)} = \hat{j}^{(n)}$ and see that this is a contradiction. Therefore, $\pi_k \geq 0$ for all k and we have established that our candidate solution is indeed an optimal solution.

Heuristic Solution to the Basic Problem

As before, we propose a heuristic solution to (206) based on Theorem 3. 1 and the discussion leading up to that result. We will state the solution in terms of an algorithm.

ALGORITHM 3. 1.

STEP 1. If $Mj_{max} \geq b$, go to Step 2. Otherwise, stop. There is no feasible solution.

STEP 2. If $Mj^* < b$, go to Step 3. If not, set $X_{j*} = M$ and $X_j = 0$ otherwise. This is the optimal solution to (206) and the value of the objective function is $MZ(j^*)$. STOP.

STEP 3. Determine $\bar{j}, \bar{j}^{(2)}, \ldots, \bar{j}^{(\bar{n})} = j_{max}$. Find n such that $\frac{b}{\bar{j}^{(n)}} \leq M < \frac{b}{\bar{j}^{(n-1)}}$. Set

$$X_{\bar{j}^{(n-1)}} = \left[\frac{M\bar{j}^{(n)} - b}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} \right]$$

$$X_{\bar{j}^{(n)}} = M - X_{\bar{j}^{(n-1)}}$$

$$X_j = 0 \qquad \text{otherwise.}$$

The value of the objective function is

$$\left[\frac{M\bar{j}^{(n)} - b}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} \right](Z(\bar{j}^{(n-1)}) - Z(\bar{j}^{(n)})) + MZ(\bar{j}^{(n)})$$

STOP.

We point out that this rounding off is the opposite of that in the upper bound case in that we have rounded off the direction which will increase the number of restricted fillers used, in order to guarantee that $\sum_{j=j^*}^{j_{max}} jX_j \geq b$. Of course, if there is no remainder when we take the greatest integer less than or equal to $\frac{M\bar{j}^{(n)} - b}{\bar{j}^{(n)} - \bar{j}^{(n-1)}}$ then there is no rounding off. In fact, in this event our solution is an optimal solution to (206).

Just as in the upper bound case, there is an opportunity to improve our solution if there is a remainder, which we will again denote by r_n. In the event that $r_n > 0$ and there is a j such that $j = \bar{j}^{(n)} - r_n$ or $\bar{j}^{(n-1)} - r_n$ where $Z(j) > Z(\bar{j}^{(n)})$ or $Z(j) > Z(\bar{j}^{(n-1)})$, we can instruct our knapsack stuffer to fill one knapsack with j rather than $\bar{j}^{(n)}$ or $\bar{j}^{(n-1)}$ and improve the solution somewhat. Since this possibility is entirely dependent on the specific parameters of a particular situation, we will pursue it no further.

Finally, we will show that this solution procedure has a value which is within a specified amount of the optimal value of the solution to the LP (304) no matter what value M has. We define the loss function to be the difference between the value of the linear program for a given set of parameters (LP) and the value of our solution $(\overline{IP})$ and denote it by $LP - \overline{IP}$. We note that n is determined by the algorithm.

<u>THEOREM 3.2.</u> $LP\text{-}\overline{IP} < Z(j^*)-Z(j_{max})$ for any M and b such that $Mj_{max} \geq b$.

<u>PROOF.</u> We must consider two cases:

1) $Mj^* \geq b$. In this event $LP - \overline{IP} = 0$ so the result holds.

2) $Mj^* < b$. In this case, we have, by simple subtraction, that

$$LP - IP = r_n\left(\frac{Z(\bar{j}^{(n-1)}) - Z(\bar{j}^{(n)})}{\bar{j}^{(n)} - \bar{j}^{(n-1)}} \right)$$

where

$$r_n = (M\bar{j}^{(n)} - b) \bmod (\bar{j}^{(n)} - \bar{j}^{(n-1)}) .$$

Now $r_n < \bar{j}^{(n)} - \bar{j}^{(n-1)}$ so that $LP - \overline{IP} < Z(\bar{j}^{(n-1)}) - Z(\bar{j}^{(n)})$. As a consequence of Lemma 3.2, $Z(\bar{j}^{(n)}) < Z(\bar{j}^{(n-1)})$ and so $Z(\bar{j}^{(\bar{n})}) < Z(\bar{j}^{(n)})$ for all $n < \bar{n}$. Of course $Z(j^*) > Z(j^{(\bar{n}-1)})$ and so $LP - \overline{IP} < Z(j^*) - Z(\bar{j}^{(\bar{n})})$. Since $\bar{j}^{(\bar{n})} = j_{max}$ by definition, we have the desired result.

It is clear that the necessary parallels between the upper bound case and the lower bound case are easily drawn and so we will just content ourselves with the basic results of this section.

Equality in the Second Constraint

The discussion in this section pertains to the case where the second constraint, (the factory constraint) requires the knapsack stuffer to use exactly b units of the restricted filler, no more and no less. We have seen that this poses no problem in the linear program relaxations of the Basic Problem, since the optimal solution to either the upper bound or lower bound case results in exactly b units of the filler being used. As a result, the stuffer would only need to determine if $Mj^* > b$. If so, and equality is required, then he would proceed as if he were in the upper bound case. Conversely, if $Mj^* < b$, he would use the solution to the lower bound case. Of course, if $Mj^* = b$ he has no problem at all. We note that the existence of alternative optima to the underlying knapsack problem create a choice so that if any value of j such that $Z(j) \geq Z(i)$ for all feasible i results in $Mj = b$, the answer is clear. Furthermore, if there are two alternative optimal solutions to the underlying knapsack problem, say j_1^* and j_2^*, such that $Mj_1^* < b < Mj_2^*$, then a linear, integer combination of $X_{j_1^*}$ and $X_{j_2^*}$ may exist where $j_1^* X_{j_1^*} + j_2^* X_{j_2^*} = b$. The resulting solution is certainly optimal. Of course, some linear combination of the X_{j^*}'s exists such that $j_1^* X_{j_1^*} + j_2^* X_{j_2^*} = b$ but it may not be integer. If it is not, we can choose to look at both the upper and lower bound solutions in the manner described below and choose the best.

Of course, we need an integer solution to the problem and we are more concerned with the case where $Mj_i^* < b$ for all j_i^* or where $Mj^* > b$ for all j_i^*. In these cases, the stuffer will choose the appropriate upper or lower bound case and investigate integer solutions to the problem using the algorithms we have derived. If there is no remainder, he is done, since he knows that exactly b units of filler will be used. This follows from the fact that the solution to the linear program is an all integer solution. If a remainder exists, he must fill a knapsack with that remainder to meet the equality constraint. (In the event a remainder exists in the lower bound case, he must "take out" the remainder from some knapsack in reality, but we will ignore the distinction.) This is a problem similar to the remainder problem we have already discussed except that here the requirement to use all b of the fillers may result in a total value less than that indicated by our algorithm. Since the remainder must be less than j^* in the upper bound case, this loss will be less thatn $Z(j^*) - Z(r_n)$ where r_n is the remainder. In the absence of any specific information for a particular problem we see that the total potential loss from our algorithm if we insist on filling the last knapsack in such a way that we use up exactly b fillers is strictly less than $2Z(j^*) - Z(0) - Z(r_n)$ where $r_n > 0$, in the upper bound case. Similarly, the total potential loss in the lower bound case is strictly less than $2\,Z(j^*)-Z(r_n)-Z(j_{max})$. There is no point to further discussion in the absence of specific values of the parameters.

4. Simple Decision Rules for M Random

The purpose of this section will be to investigate the effect on the factory when the number of knapsacks processed in a shift, M, is a random variable. Since an extension to the lower bound case is straightforward, we shall do so solely in the context of the upper bound problem. Further, to avoid the algebraic complications of the discussion in Section 2 , we shall assume that the known upper bound, b, is large in the sense of Theorem III, 2.11. We have seen that the heuristic solution to the Basic Problem (205) provides the factory with a good, simple solution procedure when M is known. However, the solution (or decision rule) used depends on the value of M falling into a certain interval. Furthermore, in certain instances, the solution requires a knowledge of the exact value of M to calculate the number of knapsacks to be filled with the appropriate j. Nevertheless, one can certainly envision applications where the number of knapsacks to be processed in a shift depends on any number of factors and is, for all intents and purposes, a random variable. We shall study a certain class of decision rules in such an operating environment.

Let M be a random variable with distribution F. We assume that the mean and variance are finite. The integer nature of M means, of course, that F is a discrete distribution over a countable probability space. As a consequence, we need not be concerned with difficulties involving the measure theoretic aspects of probability. Denote the mean and variance by μ and σ^2 respectively.

If we imagine that the factory can observe a realization of M prior to the beginning of a shift then, of course, the decision rule for the shift is that specified in Section 2. From this viewpoint, the expected return to a shift is easily calculated. However, this situation is essentially the same as the deterministic case in the sense that M is known before the decision rule is chosen. Of more interest to us here is the situation where M is not known until the end of the shift, but the factory must adopt decision rules which provide solutions for the first and succeeding knapsacks.

For the purpose of study, we shall assume that $j^* = \hat{j}^{(2)}$ in the underlying knapsack problem. This assumption is made solely in the interests of exposition as it covers the interesting cases while at the same time avoids much of the extra algebra and notation required to discuss the problem in full generality. In short, we lose nothing by the assumption in terms of insight. Furthermore, extensions to more complex cases are easily made. With this assumption, the results in Section 2 establish that we only need to consider three ranges of M:

1) $0 < M \leq [\frac{b}{j^*}]$

2) $[\frac{b}{j^*}] < M < [\frac{b}{\hat{j}}]$

3) $[\frac{b}{\hat{j}}] \leq M$

Recall that the solutions associated with these intervals are (M is an integer):

1) $X_{j^*} = M$, $X_j = 0$ otherwise. The associated value is $MZ(j^*)$.

2) $X_{j^*} = [\frac{b-M\hat{j}}{j^*-\hat{j}}]$, $X_{\hat{j}} = M - X_{j^*}$, $X_j = 0$ otherwise. The associated value is $[\frac{b-M\hat{j}}{j^*-\hat{j}}](Z(j^*)-Z(\hat{j})) + MZ(\hat{j})$

3) $X_{\hat{j}} = [\frac{b}{\hat{j}}]$, $X_0 = M - X_{\hat{j}}$, $X_j = 0$ otherwise. The associated value is $[\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) + MZ(0)$.

We note immediately that the factory cannot choose the decision rule associated with the second interval in advance of knowing M. That is, they cannot provide instructions to the stuffer on how many knapsacks to fill with j^* because that number depends on M in such a way that the value of M must be known in advance. This is not the case with the decision rules associated with the first and third intervals since the factory can choose either of those decision rules without knowing M. That this is the case is clear when we realize that these rules can be translated into instructions to the stuffer as follows:

"Start filling knapsacks with $\hat{j}$ (or j^*) until you either run out of knapsacks (i.e. reach M) or there are less than $\hat{j}(j^*)$ of the restricted filler types available. That is, when you have filled $[\frac{b}{\hat{j}}]$, $([\frac{b}{j^*}])$ knapsacks. Fill any remaining knapsacks with 0."

We shall not discuss the remainder problem, which occurs when $\frac{b}{\hat{j}} > [\frac{b}{\hat{j}}](\frac{b}{j^*} > [\frac{b}{j^*}])$, in the absence of any further assumptions, except to note that a positive remainder can only serve to improve your solution.

We shall call decision rules that state "fill knapsacks with j until you either run out of knapsacks or units of the restricted filler types" simple rules and refer to them as Rule j, denoted by $R(j)$. Such rules are characterized by the fact that they are made in advance of the shift and are carried out for the shift without change.

Certainly rules which permit changes during the sequence of knapsacks, either as a mixture of simple rules which the factory is committed to in advance of the shift or which permit changes based on increasing information on the value of M gained as knapsacks arrive (i. e. based on conditional probability ideas), may be better in an expected value sense than a simple rule. Nonetheless the simple rules have the advantage of including optimal rules for certain ranges of M, by which we mean rules which achieve the same result as our decision rule for the deterministic case. For example, $R(j^*)$ is truly optimal when the realization of M is less than or equal to b/j^* and thus is optimal with probability equal to the probability that M is less than or equal to b/j^*. Similarly, $R(\hat{j})$ is good and perhaps optimal with probability equal to the probability that M is greater than or equal to the greatest integer in $b/\hat{j}$.

In the interests of providing a start in the analysis of a stochastic version of the Basic Problem , we include a study of the behavior of simple rules in this section. We adopt as our criteria the expected value realized on a shift under decision rule j, which we shall denote as $E(R(j))$. We shall use the following notation: Let

$$A = \{x : 0 < x \leq [\frac{b}{j^*}]\}$$

$$B = \{x : [\frac{b}{j^*}] < x < [\frac{b}{\hat{j}}]\}$$

$$C = \{x : [\frac{b}{\hat{j}}] \leq x\}.$$

Then,

$$P(A) = \text{Probability that } M \in A, \quad \text{etc.}$$

Now, if $P(A) + P(B) = 0$ then certainly we would adopt $R(\hat{j})$ all the time. Similarly, if $P(B) + P(C) = 0$, we would adopt $R(j^*)$. Thus, we assume that $P(B) > 0$.

Consider the expected value of $R(\hat{j})$, $E(R(\hat{j}))$. Now

$$E(R(\hat{j})) = \int_0^\infty R(j)dF(t)$$

or

$$E(R(j)) = \int_0^{z-1} tZ(\hat{j})dF(t) + \int_z^\infty z(Z(\hat{j}) - Z(0))dF(t)$$

$$+ \int_z^\infty tZ(0)dF(t) ,$$

where $z = [\frac{b}{\hat{j}}]$. We can write

$$E(R(j)) = (Z(\hat{j}) - Z(0)) \int_0^{z-1} tdF(t) + [\frac{b}{\hat{j}}](Z(\hat{j}) - Z(0))P(C) + \mu Z(0).$$

Since the first term is not larger than $[\frac{b}{\hat{j}}](Z(\hat{j})-Z(0))(P(A)+P(B))$, we see that $E(R(\hat{j})) \leq [\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) + \mu Z(0)$. Therefore, if $\mu \geq [\frac{b}{\hat{j}}]$ and we were to adopt rule $\hat{j}$ based on the mean value of M, then our expected return is bounded above by the value we would obtain by setting $M = \mu$ in Section 2 of this chapter.

Similarly, we can show that

$$E(R(j^*)) = \int_0^x tZ(j^*)dF(t) + \int_{x+1}^{\infty} x(Z(j^*)-Z(0))dF(t)$$

$$+ \int_{x+1}^{\infty} tZ(0)dF(t), \qquad \text{where} \quad x = [\frac{b}{j^*}].$$

$$= (Z(j^*)-Z(0))\int_0^x tdF(t) + [\frac{b}{j^*}](Z(j^*)-Z(0))(1-P(A)) + \mu Z(0)$$

and that $E(R(j^*)) \leq [\frac{b}{j^*}](Z(j^*)-Z(0)) + \mu Z(0)$.

In general, we see, for any j,

$$E(R(j)) = (Z(j)-Z(0))\int_0^x tdF(t) + [\frac{b}{j}](Z(j)-Z(0))P(M > x) + \mu Z(0),$$

where $x = [\frac{b}{j}]$.

In theory, given F, one can find the j which maximizes $E(R(j))$, since the number of candidates is finite. Further, if we denote the expected value of M given that M is less than or equal to x as μ_x, we can simplify the notation by writing, when

$x = [\frac{b}{j}]$,

$$E(R(j)) = (Z(j)-Z(0))(\mu_x P(M \leq x)+xP(M > x)) + \mu Z(0)$$

or

$$E(R(j)) = (Z(j)-Z(0))(x+(\mu_x - x)P(M \leq x)) + \mu Z(0)$$

Alternately, setting $x = [\frac{b}{j}]$, we can write

$$E(R(j)) = (Z(j)-Z(0))([\frac{b}{j}]+P(M < \frac{b}{j})(\mu_{[\frac{b}{j}]} - [\frac{b}{j}]) + \mu Z(0)$$

Let us compare the two obvious candidates, $R(\hat{j})$ and $R(j^*)$. The value of rule $\hat{j}$, denoted $R(\hat{j})$, and the value $R(j^*)$ for the three intervals are:

	$R(j^*)$	$R(\hat{j})$
A	$MZ(j^*)$	$MZ(\hat{j})$
B	$[\frac{b}{j^*}](Z(j^*)-Z(0)) + MZ(0)$	$MZ(\hat{j})$
C	$[\frac{b}{j^*}](Z(j^*)-Z(0)) + MZ(0)$	$[\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) + MZ(0)$

We wish to know whether $E(R(j^*)-R(\hat{j})) \geq 0$ in order to choose between $R(j^*)$ or $R(\hat{j})$. Let $\delta = R(j^*) - R(\hat{j})$. Then, over the ranges indicated, we have:

$$\begin{array}{ll} & \delta \\ \hline \text{A:} & M(Z(j^*)-Z(\hat{j})) \\ \text{B:} & [\frac{b}{j^*}](Z(j^*)-Z(0)) - M(Z(\hat{j})-Z(0)) \\ \text{C:} & [\frac{b}{j^*}](Z(j^*)-Z(0)) - [\frac{b}{\hat{j}}](Z(\hat{j})-Z(0)) \end{array}$$

We note that for $M \in A$, $\delta > 0$ and for $M \in C$, $\delta < 0$ since we have assumed that b is large enough in the sense of Theorem 2.11. In any event, if $M \in C$ and $\delta = 0$, then it is clear that the problem has no content in the sense that $R(j^*)$ is certainly optimal. In effect, $j^* = \hat{j}$, which we assumed not to be the case. So, we have that $\delta < 0$ when $M \in C$. Therefore, it must be true that $\delta = 0$ for some $x \in B$. If we set $\delta = 0$ for some x, say $\bar{M}$, in B we have $[\frac{b}{j^*}](Z(j^*)-Z(0))-\bar{M}(Z(\hat{j})-Z(0)) = 0$ or $\bar{M} = [\frac{b}{j^*}] \frac{Z(j^*)-Z(0)}{Z(\hat{j})-Z(0)}$. If we recall Definition 2.1, we have the following:

<u>PROPOSITION 4.1.</u> $\delta = 0$ when $M = [\frac{b}{j^*}] \frac{j^*}{\hat{j}} \rho_{j^*}$.

Of course, there is no guarantee that $\bar{M}$ is an integer but it provides us with a convenient break point between $R(\hat{j})$ and $R(j^*)$ as shown in Figure 4.1. Also on the figure is the value of the optimal solution to the linear program (301). Of course, the figure is a continuous representation of what is in reality a discrete valued function.

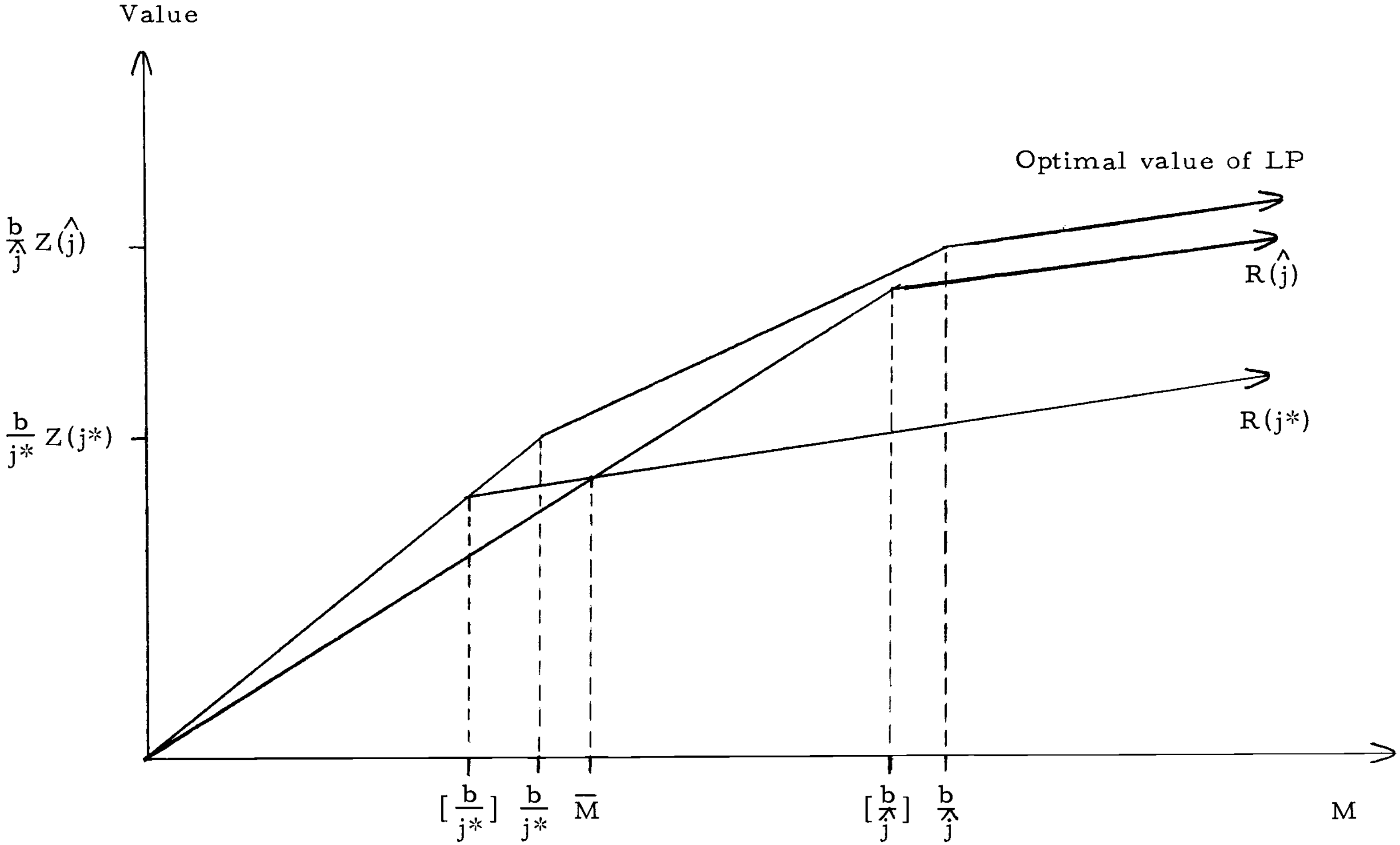


Figure III. 4. 1. Values of $R(\hat{j})$ and $R(j^*)$ versus M.

COROLLARY 1. If $P(M \leq \overline{M}) = 0$, then $E(\delta) < 0$.

PROOF. An immediate consequence of the fact that $\delta < 0$ for $M > \overline{M}$.

In general, if we set $x = [\frac{b}{j^*}]$, $z = [\frac{b}{\hat{j}}]$,

$$E(\delta) = \int_0^x (Z(j^*) - Z(\hat{j}))t\,dF(t) + \int_{x+1}^{z-1} [\frac{b}{j^*}](Z(j^*) - Z(0))dF(t)$$

$$- \int_{x+1}^{z-1} (Z(\hat{j}) - Z(0))t\,dF(t) + \int_z^{\infty} [\frac{b}{j^*}](Z(j^*) - Z(0))dF(t)$$

$$- \int_z^{\infty} [\frac{b}{\hat{j}}](Z(\hat{j}) - Z(0))dF(t).$$

If we add and subtract the quantity $\int_0^x (Z(\hat{j}) - Z(0))t\,dF(t)$ in this expression, we get

$$E(\delta) = (Z(j^*) - Z(0))\int_0^x t\,dF(t) - (Z(\hat{j}) - Z(0))\int_0^{z-1} t\,dF(t)$$

$$+ [\frac{b}{j^*}](Z(j^*) - Z(0))(P(B) + P(C)) - [\frac{b}{\hat{j}}](Z(\hat{j}) - Z(0))P(C)$$

One can manipulate this equation in a number of ways to derive conditions for $E(\delta)$ to be negative, none of which are any more instructive than this form.

Another question of interest is this, "How much does the simple rule cost us in the sense that because of the random nature of M we cannot realize the full potential of the decision rules of Section 2?" To study the question, we will compare the expected value of the rule j to the expected value of the heuristic solution, $\overline{IP}(M)$, which we could have realized had we known M in advance. We shall conduct the analysis specifically for rule $\hat{j}$ to illustrate the process. For the same intervals as before we have the solutions:

	$\overline{IP}$	$R(\hat{j})$
A.	$MZ(j^*)$	$MZ(\hat{j})$
B.	$\left[\frac{b-M\hat{j}}{j^*-\hat{j}}\right](Z(j^*)-Z(\hat{j})) + MZ(\hat{j})$	$MZ(\hat{j})$
C.	$\left[\frac{b}{\hat{j}}\right](Z(\hat{j})-Z(0)) + MZ(0)$	$\left[\frac{b}{\hat{j}}\right](Z(\hat{j})-Z(0)) + MZ(0)$

Let the loss function, $\Delta(\hat{j}) = \overline{IP} - R(\hat{j})$. Then, for the usual intervals, we have:

	$\Delta(\hat{j})$
A.	$M(Z(j^*)-Z(\hat{j}))$
B.	$\left[\frac{b-M\hat{j}}{j^*-\hat{j}}\right](Z(j^*)-Z(\hat{j}))$
C.	0

Figure 4.2 provides a graphic example of the loss functions, $\Delta(\hat{j})$ as well as $\Delta(j^*)$. This graph portrays typical loss functions as

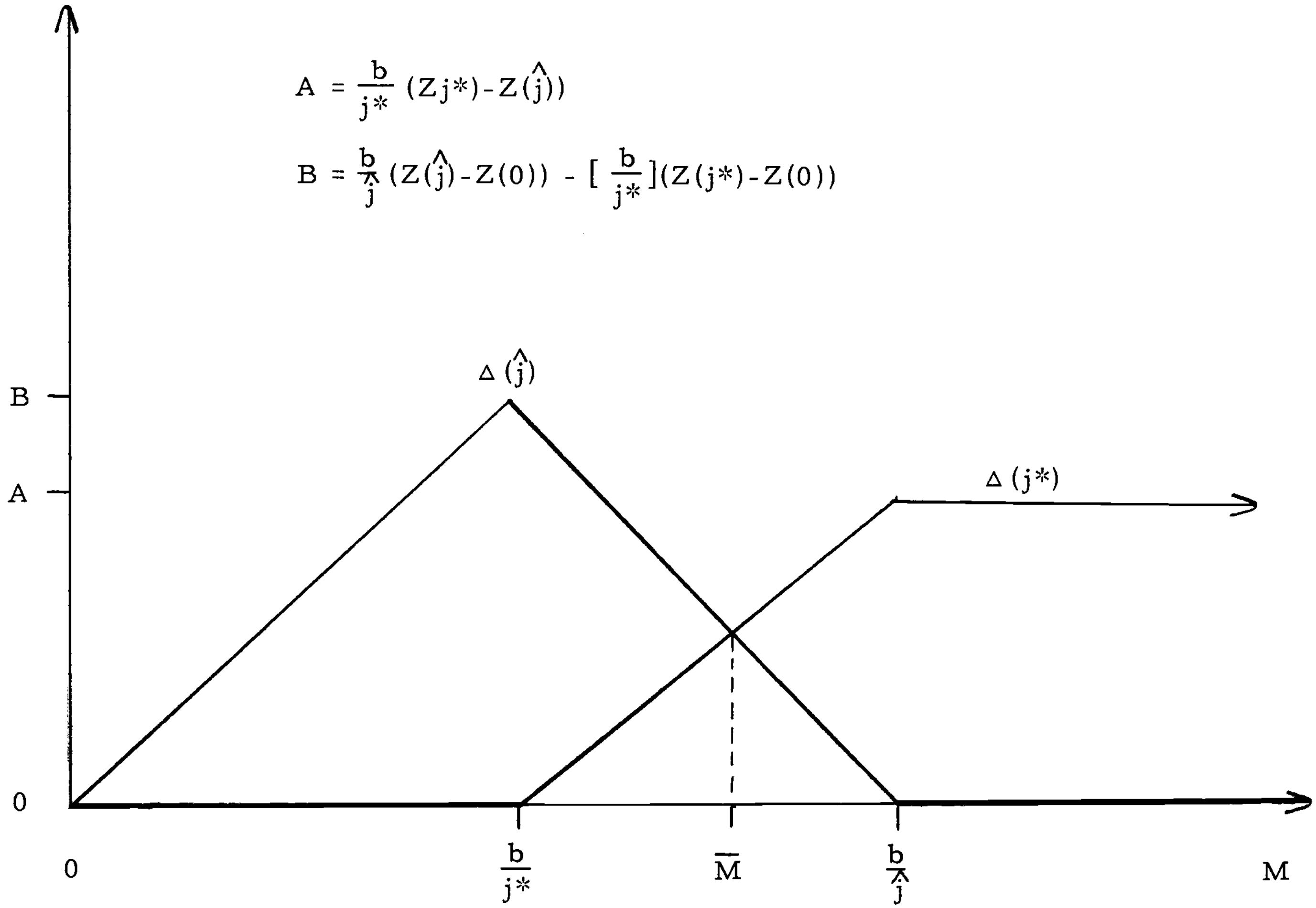


Figure 4.2. $\Delta(j^*)$ and $\Delta(\hat{j})$ vs. M.

continuous functions when they are not, but serves to illustrate their general behavior. We can write the expected value of the loss as

$$E(\Delta(\hat{j})) = (Z(j^*)-Z(\hat{j})) \int_0^x tdF(t) + \int_{x+1}^{z-1} [\frac{b-t\hat{j}}{j^*-\hat{j}}]dF(t)$$

where $x = [\frac{b}{j^*}]$ and $z = [\frac{b}{\hat{j}}]$. It is difficult to work with the greatest integer notation and so we shall place an upper bound on $E(\Delta)$ by noting that $[\frac{b-t\hat{j}}{j^*-\hat{j}}] \leq \frac{b-t\hat{j}}{j^*-\hat{j}}$. Then,

$$E(\Delta(\hat{j})) \leq (Z(j^*)-Z(\hat{j})) \int_0^x tdF(t) + \int_{x+1}^{z-1} (\frac{b-t\hat{j}}{j^*-\hat{j}})dF(t)$$

and the right hand term in brackets becomes

$$\int_0^x tdF(t) - (\frac{\hat{j}}{j^*-\hat{j}}) \int_{x+1}^{z-1} tdF(t) + \frac{b}{j^*-\hat{j}} P(B)$$

or

$$\frac{1}{j^*-\hat{j}} \quad (j^*-\hat{j}) \int_0^x tdF(t) - \hat{j} \int_{x+1}^{z-1} tdF(t) + bP(B) \quad .$$

Rearranging terms and substituting for μ_x as defined earlier, the term in brackets becomes

$$(\frac{1}{j^*-\hat{j}} (j^*\mu_x P(Z) - \hat{j}\mu_{z-1}(P(A)+P(B))+bP(B)).$$

Therefore,

$$E(\Delta) \leq \frac{Z(j^*)-Z(\hat{j})}{j^*-\hat{j}} \left(P(A)(j^*\mu_x - \hat{j}\mu_{z-1}) + P(B)(b-\hat{j}\mu_{z-1})\right)$$

When we apply the same analysis to the loss created by rule j^*, $\Delta(j^*)$, the following results:

INTERVAL	$\underline{\Delta(j^*)}$
A.	0
B.	$\left[\frac{b-M\hat{j}}{j^*-\hat{j}}\right](Z(j^*)-Z(\hat{j}))+M(Z(\hat{j})-Z(0)) - \left[\frac{b}{j^*}\right](Z(j^*)-Z(0))$
C.	$\left[\frac{b}{\hat{j}}\right](Z(\hat{j})-Z(0)) - \left[\frac{b}{j^*}\right](Z(j^*)-Z(0))$.

Therefore,

$$E(\Delta(j^*)) = (Z(j^*)-Z(\hat{j}))\int_{x+1}^{z-1}\left[\frac{b-t\hat{j}}{j^*-\hat{j}}\right]dF(t) + (Z(\hat{j})-Z(0))\int_{x+1}^{z-1} t\,dF(t)$$

$$- \left[\frac{b}{j^*}\right](Z(j^*)-Z(0))\int_{x+1}^{z-1} dF(t)$$

$$+ \left(\left[\frac{b}{\hat{j}}\right](Z(\hat{j})-Z(0)) - \left[\frac{b}{j^*}\right](Z(j^*)-Z(0))\right)\int_{z}^{\infty} dF(t)$$

where $x = \left[\frac{b}{j^*}\right]$ and $z = \left[\frac{b}{\hat{j}}\right]$. If we eliminate the greatest integer quantity under the first integral and consolidate, we get the following bound on the expected loss:

$$E(\Delta(j^*)) \leq \frac{(Z(j^*)-Z(\hat{j}))}{j^*-\hat{j}} bP(B) + \frac{j^*(Z(\hat{j})-Z(0)-\hat{j}(Z(j^*)-Z(0))}{j^*-\hat{j}} \int_{x+1}^{z-1} t dF(t)$$

$$+ [\frac{b}{\hat{j}}](Z(\hat{j})-Z(0))P(C) - [\frac{b}{j^*}](Z(j^*)-Z(0))(P(B)+P(C)).$$

We note finally that this upper bound is in fact a good estimate of the loss. If we denote this bound by K, and we note that $[\frac{b-t\hat{j}}{j^*-\hat{j}}] \geq \frac{b-t\hat{j}}{j^*-\hat{j}} - 1$, one can easily show that $K \geq E(\Delta(j^*)) \geq K - (Z(j^*)-Z(\hat{j}))P(B) > K - (Z(j^*)-Z(\hat{j}))$. As a consequence, we shall use the value of K as our conservative estimate of the expected loss. This expected loss can be interpreted as the price the factory pays for being content to take the knapsacks as they come rather than collecting all the knapsacks in the shift before the shift begins, thereby knowing the value of M.

EXAMPLE 4.1. To illustrate the above discussion we shall consider a specific example. It is not unreasonable to consider a situation where the factory knows bounds on the number of knapsacks to expect but the actual number realized is equally likely for any value within those bounds. That is, the distribution, F, is the discrete analog of the uniform distribution as follows:

$$P_x = P(M=x) = \frac{1}{b-a+1} \qquad x = a, a+1, \ldots, b$$

$$a > 0$$

$$= 0 \quad \text{otherwise}$$

We note that

$$\sum_{i=1}^{b} i = \frac{b(b+1)}{2}$$

so that

$$\sum_{i=a}^{b} i = \sum_{i=1}^{b} i - \sum_{i=1}^{a-1} i = \frac{b(b+1)+a(a-1)}{2}$$

and therefore

$$\mu = \int_0^\infty t dF(t) = \sum_{x=a}^{b} xP_x$$

or

$$\mu = \sum_{x=a}^{b} \frac{x}{b-a+1} = \frac{b^2+b-a^2+a}{2(b-a+1)} = \frac{b+a}{2} .$$

Also, if $a \leq c < d \leq b$, c, d integer,

$$P(c \leq x \leq d) = \int_c^d dF(t) = \sum_{x=c}^{d} P_x = \frac{d-c+1}{b-a+1}$$

and

$$\int_c^d t dF(t) = \sum_{x=c}^{d} xP_x = \frac{d(d+1)-c(c-1)}{2(b-a+1)} .$$

Let us modify slightly example III. 2. 2 in order to meet the assumptions of this section, that is so that $j^* = \hat{j}^{(2)}$. Consider the following underlying knapsack function:

j	$Z(j)$	$Z(j)-Z(0)$	$\frac{Z(j)-Z(0)}{j}$
0	915	0	0
1	950	35	35
2	1025	110	55
3	1029	114	38
4	1055	140	35
5	1155	200	40

Thus $\hat{j} = 2$, $j^* = 5 = \hat{j}^{(2)}$. As before, $b = 119$, which is certainly larger than $\frac{\hat{j}-1}{1-\rho} = \frac{1}{1-\rho} = \frac{55}{15}$ $(\rho = \frac{40}{55})$ and so b is "large enough." We have that $\overline{M} = [\frac{b}{j^*}] \frac{Z(j^*)-Z(0)}{Z(\hat{j})-Z(0)} = 41 \frac{9}{11}$. In order to have an interesting example, we shall require positive probability for values of M less than 42, in accordance with Corollary 1 of Proposition 4. 1. Therefore, we shall assume the following density function:

$$P_x = P(M=x) = \frac{1}{37} \qquad 23 \leq x \leq 59$$

$$= 0 \qquad \text{otherwise}$$

where we note that $[\frac{b}{j^*}] = 23$ and $[\frac{b}{\hat{j}}] = 59$. Further $\mu = 41$ and $P(A) = \frac{1}{37}$, $P(B) = \frac{35}{37}$ and $P(C) = \frac{1}{37}$. We calculate $E(R(\hat{j}))$ as

$$Z(\hat{j}) - Z(0) \sum_{x=0}^{58} xP_x + [\frac{b}{\hat{j}}](Z(\hat{j})-Z(0))P(C) + \mu Z(0)$$

which equals

$$\frac{110}{37} \sum_{x=23}^{58} x + \frac{59(110)}{37} + 41(915)$$

or 42025. If we proceed to do the same for all j, we have the following result:

j	E(R(j)) (to the nearest integer)
1	38,950
2	42,025
3	41,541
4	41,495
5	42,115

Consequently we would choose $R(j^*)$ despite the fact that it is optimal in the sense of this section only with probability $\frac{1}{37}$. This is the same probability that $R(\hat{j})$ is "optimal."

<u>EXAMPLE 4.2.</u> Let us take the same problem except to modify the distribution a little to see if the answer shifts to $R(\hat{j})$: Let the distribution F be characterized by this density function:

$$P_x = P(M=x) = \frac{1}{30} \qquad 29 \leq x \leq 58$$

$$= 0 \qquad \text{otherwise}$$

In this case, $\mu = 43.5$. Proceeding as above, we have

j	$E(R(j))$ (to the nearest integer)
1	41,325
2	44,588
3	44,040
4	43,863
5	44,403

Now we would choose rule $\hat{j}$ as the best simple rule, even though it is "optimal" with probability 0.

EXAMPLE 4.3. What is the expected value of the loss function for $R(j^*)$ in Example 4.1. We have define it as $E(\overline{IP}-R(j^*)) = E(\Delta(j^*))$ and we calculate a bound

$$E(\Delta j^*) \leq \frac{(1115-1025)}{3} 119P(B) + \frac{5(110)-2(200)}{3} \sum_{x=24}^{58} xP_x$$

$$+ 59(110)P(C) - 23(200)(P(B)+P(C))$$

or

$$E(\Delta(j^*)) \leq 1016$$

and we know that

$$E(\Delta(j^*)) \geq 1016 - (Z(j^*)-Z(\hat{j}))P(B) = 931.$$

Therefore,

$$931 \leq E(\Delta(j^*)) \leq 1016.$$

Using the upper bound as a conservative estimate, we see that the expected loss is $\frac{1016}{42115}$ x 100 percent = 2.41 percent of the expected value realized over the shift by rule j^*. Another way of looking at it is that if one observed the realization of M before filling any knapsacks, he could realize an expected value at least 931 greater than that using the best simple rule, i.e. realize an expected return greater than or equal to 43,046. Alternately, it costs the factory at least 931 per shift on the average to use rule j^* rather than go to the expense of determining the number of knapsacks in advance.

IV. THE COMPLEX KNAPSACK FACTORY PROBLEM

1. Introduction

The preceding chapter provided a good and simple solution to a problem consisting of a sequence of identical generic knapsack problems linked together by a single constraint on the permissible number of certain types of fillers used during the sequence. The next level of complexity in the knapsack factory problem is to remove the restriction that all the knapsack type problems in the sequence be identical. We call this next level the Complex Problem despite the fact that there are certainly more complicated variations.

The problem was formulated in Section II. 3 as follows:

Let K be the number of different knapsack types found in a shift. Let M_k be the number of knapsacks of type k processed during a shift, $k = 1, \ldots, K$. Let $M = \sum_{k=1}^{K} M_k$. Also, $Z_k(j)$ is the value of one knapsack of type k filled with j, which has the meaning described in Section II. 1. Let X_{kj} be the number of knapsacks of type k filled with j. Note that for a given k, j is an integer between 0 and $j_{max}(k)$, a maximum value which can differ for each knapsack type. The Complex Problem is:

$$(207) \qquad \text{maximize} \quad \sum_{k=1}^{K} \sum_{j=0}^{j_{max}(k)} X_{kj} Z_k(j)$$

$$\text{subject to} \quad \sum_{j=0}^{j_{max}(k)} X_{kj} = M_k, \quad k = 1, \ldots, K$$

and

$$\sum_{k=1}^{K} \sum_{j=1}^{j_{max}(k)} jX_{kj} \leq b \quad (\geq b)$$

$$X_{kj} \geq 0, \; X_{kj} \text{ integer for all } k \text{ and } j \text{ and } b > 0.$$

For each knapsack type k, $k = 1, \ldots, K$, we shall define, in exactly the same way as Definition II. 4. 1, the quantities $\hat{j}_k^{(1)}, \hat{j}_k^{(2)}, \ldots, \hat{j}_k^{(\bar{n})}$, which characterize the underlying generic knapsack function of type k. Thus, $\hat{j}_k^{(1)}$ is chosen so that

$$\frac{Z_k(\hat{j}^{(1)}) - Z_k(0)}{\hat{j}^{(1)}} \geq \frac{Z_k(j) - Z_k(0)}{j}$$

for all j, $0 \leq j \leq j_{max}(k)$, etc. As before, we shall drop the superscript $n = 1$ and merely write $\hat{j}_k$. Note that we have dropped the redundant subscript k in the interest of clarity, so that $Z_k(j) = Z_k(j_k)$. Furthermore, we shall restrict our analysis of the Complex Problem to the upper bound case $(\leq b)$ in the same interest. The translation of the results achieved to the lower bound case ($\geq b$) is straightforward given the discussion in Section III. 3 and will not be pursued here. In the upper bound case, a simple extension of the argument used in the Basic Problem will show that we can restrict our search for good or optimal solutions to those knapsacks of any type that are filled with j where $0 \leq j \leq j_k^*$, $k = 1, \ldots, K$. Recall that j_k^* is the parametric solution which corresponds to the optimal solution of the underlying knapsack problem of type k. Consequently, we will replace the index $j_{max}(k)$ with j_k^* wherever the former appears in the formulation (207).

Without loss of generality, we shall order the knapsack types so that

$$\frac{Z_k(\hat{j}) - Z_k(0)}{\hat{j}_k} \geq \frac{Z_{k+1}(\hat{j}) - Z_{k+1}(0)}{\hat{j}_{k+1}}$$

for $k = 1, \ldots, K$. In the event of a tie, we shall require that $\hat{j}_k \geq \hat{j}_{k+1}$. If $\hat{j}_k = \hat{j}_{k+1}$ break the tie arbitrarily.

Further, to reduce our notational requirements, we shall adopt the following convention:

$$\Delta(k, n) = \frac{Z_k(\hat{j}^{(n)}) - Z_k(\hat{j}^{(n-1)})}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} .$$

We note that the index n can range from 0 to $\bar{n}$ where, for k fixed, $\bar{n}$ is defined by Definition II. 4. 1. Of course, $\bar{n}$ may differ for different k, but it is always the case that $\hat{j}_k^{(\bar{n})} = j_k^*$.

We assume hereafter that $\sum_{k=1}^{K} j_k^* M_k > b$. Otherwise, the problem is uninteresting in that the factory constraint is not binding and the solution is apparent:

Set

$$X_{kj^*} = M_k \qquad k = 1, \ldots, K$$

$$X_{kj} = 0 \qquad \text{otherwise.}$$

Finally, we note that the case where each knapsack in the shift is different but its type is known is handled in this formulation by setting $M_k = 1$ for all k. By the type being known we mean that the underlying knapsack function for each knapsack is known prior to the shift (i. e. prior to solving (207)).

2. The Linear Programming Analogy

As usual, we shall exploit the special form of the linear programming relaxation of (207) to arrive at a simple procedure to obtain an optimal solution to the LP. Then, we shall easily obtain an algorithm which provides us with a good and sometimes optimal solution to (207). We note that the solution procedure does not require access to a computer linear programming package and can, in fact, be done "on the back of an envelope."

Consider this problem:

$$\text{(401)} \qquad \text{maximize} \quad \sum_{k=1}^{K} \sum_{j=0}^{j_k^*} X_{kj} Z_k(j)$$

$$\text{subject to} \quad \sum_{j=0}^{j_k^*} X_{kj} = M_k, \qquad k = 1, \ldots, K$$

$$\text{and} \quad \sum_{k=1}^{K} \sum_{j=1}^{j_k^*} j X_{kj} \le b, \qquad b > 0$$

$$X_{kj} \ge 0, \; 1 \le k \le K, \; 0 \le j \le j_k^* .$$

This is a linear program with $K+1$ constraints where the right hand side is integer.

Since we have assumed that $\sum_{k=1}^{K} j_k^* M_k > b$, it should be clear by the same argument we pursued in the Basic Problem that a slack variable in the last (factory) constraint will be zero in an optimal solution. Thus, we can easily put (401) in standard form by making the factory constraint an equality. By standard form, we mean that one can write (401) in matrix notation as:

$$\begin{array}{ll} \text{maximize} & CX \\ \text{subject to} & AX = d, \quad d \geq 0 \\ \text{and} & X \geq 0 . \end{array}$$

In this instance,

$$X = \begin{bmatrix} X_{10} \\ X_{11} \\ \cdots \\ X_{Kj^*_K} \end{bmatrix} \qquad C = \begin{bmatrix} Z_1(0) \\ Z_1(1) \\ \cdots \\ Z_K(j^*) \end{bmatrix} \qquad d = \begin{bmatrix} M_1 \\ \cdots \\ M_K \\ b \end{bmatrix}$$

and A has the following form,

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 1 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & & & \cdots & & & & & \cdots & & & \cdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & j^*_1 & 0 & 1 & 2 & \cdots & 0 & 1 & \cdots & j^*_K \end{bmatrix}$$

and has dimension $K+1$ by $\sum_{k=1}^{K} (j^*_k+1)$. Certainly $K+1 < \sum_{k=1}^{K} (j^*_k+1)$ and so any basis matrix, B, has dimension $K+1$ by $K+1$. Furthermore, as a direct consequence of the block diagonal structure of the first K rows of A and the fact that all entries are either 0 or

1 in the first K rows, it is also clear that the first K rows of B each have to have at least one positive element (equal to one). In fact, $K-1$ of the first K rows must have exactly one positive element and one row must have exactly two ones, all the other entries must be zero. Furthermore, each column of B must have exactly one element equal to one and all the rest are zero, if we ignore the last element in each column (i.e. the $K+1$st row). The last row of B is made up of components drawn from the set of feasible j's. That is, the last element in the k^{th} column can be an integer between 0 and j_k^*.

Therefore, B must look like

$$\begin{array}{cc}
\begin{array}{ccccccccc} & & & & \overbrace{}^{\ell^{th}\text{ column}} & & \phantom{j_{\ell+1}} & & \end{array} & \\
\left[\begin{array}{ccccccccc}
1 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 0 & 1 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\
j_1 & j_2 & j_3 & \cdots & j_\ell & j'_\ell & j_{\ell+1} & \cdots & j_K
\end{array}\right] &
\begin{array}{l} \\ \\ \\ \\ \} \ \ell^{th} \text{ row} \\ \\ \\ \\ \\ \end{array}
\end{array}$$

where $0 \leq j_1 \leq j_1^*, \ldots, 0 \leq j_K \leq j_K^*$ and, without loss of generality, we can assume $0 \leq j_\ell < j'_\ell \leq j_\ell^*$.

It is this very tractable structure for B that enables us to pursue the analysis and derive a solution algorithm.

We require the inverse of B, denoted B^{-1}. To obtain the inverse we shall perform elementary row and column operations on B to obtain a matrix, $\hat{B}$, whose inverse is easily derived. Then recalling that row transformations on B translate into column transformations on B^{-1}, we can return from $\hat{B}^{-1}$ to B^{-1}. First, we note that as long as we recall which variable a column is associated with, the order of the columns in a basis matrix for an LP is not relevant. Accordingly, we shall move the ℓ^{th} and $\ell+1^{st}$ columns to the end of the matrix and otherwise preserve the order, that is, if b_k is the k^{th} column of B, we now have $B = (b_1, b_2, \ldots, b_{\ell-1}, b_{\ell+2}, \ldots, b_K, b_{K+1}, b_\ell, b_{\ell+1})$. Next, we shall insert the ℓ^{th} row (the row with two 1's) between the last and second to last row of $\hat{B}$, otherwise preserving the order. Then we can partition $\hat{B}$ as

$$\hat{B} = \begin{bmatrix} I & 0 \\ C & D \end{bmatrix}$$

where I is a $K-1$ by $K-1$ identity matrix, 0 is a $K-1$ by 2 matrix all of whose elements are zero. We see that

$$C = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ j_1 & j_2 & j_3 & \cdots & j_K \end{bmatrix}$$

and has dimension 2 by K-1 and

$$D = \begin{bmatrix} 1 & 1 \\ j_\ell & j'_\ell \end{bmatrix}$$

where D is 2 by 2. Consequently, if we partition $\hat{B}^{-1}$ in the same way and write it as:

$$\hat{B}^{-1} = \begin{bmatrix} E & F \\ G & H \end{bmatrix}$$

we have that

$$EI + FC = I$$

$$DF = 0$$

$$GI + HC = 0$$

$$DH = I$$

Therefore,

$$H = D^{-1}$$

$$G = -D^{-1}C$$

$$F = 0$$

$$E = I.$$

Now D^{-1} is easily written,

$$D^{-1} = \frac{1}{(j'_\ell - j_\ell)} \begin{bmatrix} j'_\ell & -1 \\ -j_\ell & 1 \end{bmatrix}$$

and so

$$G = \frac{1}{j'_\ell - j_\ell}\begin{bmatrix} +j_1 & +j_2 & \cdots & +j_K \\ -j_1 & -j_2 & \cdots & -j_K \end{bmatrix}$$

Thus

$$\hat{B}^{-1} = \frac{1}{d}\begin{bmatrix} d & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & d & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & d & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & d & 0 & 0 \\ j_1 & j_2 & j_3 & \cdots & j_K & j'_\ell & -1 \\ -j_1 & -j_2 & -j_3 & \cdots & -j_K & -j_\ell & 1 \end{bmatrix}$$

where $d = j'_\ell - j_\ell > 0$.

As long as we recall the transformations that have been made, it is as easy to use $\hat{B}^{-1}$ as B^{-1} in the analysis. Nonetheless, we see that we can also write, using the appropriate transformations:

$$B^{-1} = \frac{1}{d}\begin{array}{c} \qquad\qquad\qquad \ell^{th}\text{ column} \\ \left[\begin{array}{ccccccccc} d & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & d & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & d & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ j_1 & j_2 & j_3 & \cdots & j'_\ell & j_{\ell+1} & \cdots & j_K & -1 \\ -j_1 & -j_2 & -j_3 & \cdots & -j_\ell & -j_{\ell+1} & \cdots & -j_K & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & d & 0 \end{array}\right] \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \} \ell^{th}\text{ row} \\ \\ \\ \\ \\ \end{array}$$

Again, $d = \frac{1}{j'_\ell - j_\ell}$.

A non-trivial consequence of the form of the inverse of any basis described above, in concert with the fact that the right hand side, d, is integer, is that at least K-1 of the variables in any solution to (401), and thus in the optimal solution, are integer valued. Therefore, the optimal solution to (401) is a good start on the solution to (207).

The preceding discussion provides us with the necessary tools for working through the simplex procedure to solve (401). The reader can easily perceive that doing so would be excessively tedious, if for no other reason than the notational morass. We shall not do so. Instead, subsequent sections of this chapter will be devoted to a discussion of an extension to the greedy solution to the Basic Problem, which we can consider as an initial basic feasible solution. This will be followed by a discussion of a solution algorithm and then a demonstration of the optimality of the solution generated by that algorithm.

3. The Complex Greedy Solution

We showed in the Basic Problem that when $M\hat{j} \geq b$, the optimal solution to its linear program relaxation, (301), was given by $X_{\hat{j}} = \frac{b}{\hat{j}}$, $X_0 = M - \frac{b}{\hat{j}}$, $X_j = 0$, otherwise. We shall define an analogous solution in the Complex Case and show that this extension is not necessarily optimal. Necessary and sufficient conditions for such a solution to be optimal will then be derived. The analysis of this case will provide valuable insight and motivation for the solution to (401).

The analogy to the condition $M\hat{j} \geq b$ in the Basic Problem is that $\sum_{k=1}^{K} \hat{j}_k M_k \geq b$. Accordingly, we assume this to be the case.

<u>DEFINITION 3.1.</u> Let $\bar{k}$ be an integer such that $0 \leq \bar{k} \leq K$, $\sum_{k=1}^{\bar{k}-1} \hat{j}_k M_k < b$ and $\sum_{k=1}^{\bar{k}} \hat{j}_k M_k \geq b$. By our assumption above, such a $\bar{k}$ exists.

<u>DEFINITION 3.2.</u> The Complex Greedy Solution (CGS) is:

$$X_{k\hat{j}} = M_k \,, \qquad k \leq \bar{k}-1$$

$$X_{k\hat{j}} = \frac{b - \sum_{k=1}^{\bar{k}-1} \hat{j}_k M_k}{\hat{j}_k} \,, \qquad k = \bar{k}$$

$$X_{k0} = M_k - X_{k\hat{j}} \,, \qquad k = \bar{k}$$

$$X_{k0} = M_k \,, \qquad k > \bar{k}$$

We have defined this solution in such a way that it is feasible in the standard form of (401) under the current assumption. The value of this solution can be written:

$$Z = \sum_{k=1}^{\bar{k}-1} M_k Z_k(\hat{j}) + \sum_{k=\bar{k}}^{K} M_k Z_k(0) + X_{\bar{k}\hat{j}}(Z_{\bar{k}}(\hat{j}) - Z_{\bar{k}}(0))$$

or

$$Z = \sum_{k=1}^{\bar{k}-1} M_k (Z_k(\hat{j}) - Z_k(0)) + X_{\bar{k}\hat{j}}(Z_{\bar{k}}(\hat{j}) - Z_{\bar{k}}(0)) + \sum_{k=1}^{K} M_k Z_k(0) .$$

We are interested in whether or not this solution is optimal. In order to answer the question, we shall consider the normal simplex procedure commencing with the CGS as the basic feasible solution. The associated basis is

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & & & & \cdots & & & & \cdots & \\ 0 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & & & & \cdots & & & & \cdots & \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \\ \hat{j}_1 & \hat{j}_2 & \hat{j}_3 & \cdots & \hat{j}_{\bar{k}} & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \\ \}\ (k^{th}\ \text{row}) \\ \}\ (k+1^{st}\ \text{row}) \\ \\ \\ \\ \end{array}$$

By the process discussed in the preceding section, we can easily find

$$B^{-1} = \frac{1}{d}\begin{bmatrix} d & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & d & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & d & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{j}_1 & \hat{j}_2 & \hat{j}_3 & \cdots & 0 & 0 & \cdots & 0 & -1 \\ -\hat{j}_1 & -\hat{j}_2 & -\hat{j}_3 & \cdots & -\hat{j}_{\bar{k}} & 0 & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & d & 0 \end{bmatrix}$$

where $d = -\hat{j}_{\bar{k}}$. We have

$$C_B = (Z_1(\hat{j}), Z_2(\hat{j}), \ldots, Z_{\bar{k}}(\hat{j}), Z_{\bar{k}}(0), Z_{\bar{k}+1}(0), \ldots, Z_K(0)).$$

Furthermore, if we designate the column of the A matrix corresponding to X_{kj} as a_{kj}, we note that $a_{kj} = e_k + je_{K+1}$, $k = 1, \ldots, K$ and $j = 0, \ldots, j_k^*$ where e_j is the j^{th} column of the $K+1^{st}$ order identity matrix. To check the optimality of this solution, we compute the optimality criterion, π_{kj}, and determine if $\pi_{kj} \geq 0$ for all k and j. We can write $\pi_{kj} = \pi'_{kj} - Z_k(j)$, where

$$\pi'_{kj} = C_B B^{-1} a_{kj} = C_B B^{-1}(e_k + je_{K+1}).$$

We will proceed to do so by cases.

Case 1: $k \geq \bar{k}+1$

$$\pi_{kj} = j\left(\frac{Z_{\bar{k}}(\hat{j}) - Z_{\bar{k}}(0)}{\hat{j}_{\bar{k}}}\right) - (Z_k(j) - Z_k(0)).$$

We know by the way we have ordered k and by the definition of $\hat{j}$, that

$$\Delta(\bar{k}, 1) \geq \Delta(k, 1) \geq \frac{Z_k(j) - Z_k(0)}{j}$$

for all $k \geq \bar{k}+1$ and for all j, $0 \leq j \leq j_k^*$. This is equivalent to stating that $\pi_{kj} \geq 0$ for all $k \geq \bar{k}+1$ and for all j, $0 \leq j \leq j_k^*$.

Case 2: $k = \bar{k}$

$$\pi_{\bar{k}j} = j\left(\frac{Z_{\bar{k}}(\hat{j}) - Z_{\bar{k}}(0)}{\hat{j}_{\bar{k}}}\right) - (Z_{\bar{k}}(j) - Z_{\bar{k}}(0))$$

$$= j\Delta(\bar{k}, 1) - (Z_{\bar{k}}(j) - Z_{\bar{k}}(0)).$$

By the definition of $\hat{j}_{\bar{k}}$, we have that $\Delta(\bar{k}, 1) \geq \dfrac{Z_{\bar{k}}(j) - Z_{\bar{k}}(0)}{j}$ for all j, $0 \leq j \leq j_{\bar{k}}^*$, which implies $\pi_{\bar{k}j} \geq 0$ for all j, $0 \leq j \leq j_{\bar{k}}^*$.

Case 3: $k < \bar{k}$.

$$\pi_{kj} = Z_k(\hat{j}) - Z_k(j) - (\hat{j}_k - j)\Delta(\bar{k}, 1)$$

within this case, we must consider the following subcases (recall that

$j_k = \hat{j}_k$ implies $\pi_{kj} = 0$ for $k \leq \bar{k}$)

a. $j = 0$.

We can easily see that $\pi_{k0} \geq 0$ since we know that $\Delta(k, 1) \geq \Delta(\bar{k}, 1)$, by the way we ordered k.

b. $0 < j < \hat{j}_k$.

From Corollary 2 to Lemma II. 4. 1, where we set $A = \Delta(\bar{k}, 1)$, we can see that $\dfrac{Z_k(\hat{j}) - Z_k(j)}{\hat{j}_k - j} \geq \Delta(\bar{k}, 1)$ or equivalently $\pi_{kj} \geq 0$ for all j, $0 < j < \hat{j}$. By now, we may have lulled the reader into thinking we have found the optimal solution to (401). Unfortunately, we have not necessarily done so.

c. $\hat{j} < j \leq j^*$.

If it were true that $\pi_{kj} \geq 0$ in this instance, then we would have the equivalent condition that $\dfrac{Z_k(\hat{j}) - Z_k(j)}{\hat{j} - j} \leq \Delta(\bar{k}, 1)$. We know from Lemma II. 4. 1 that

$$\Delta(k, 1) > \frac{Z_k(\hat{j}) - Z_k(j)}{\hat{j}_k - j},$$

and also that $\Delta(k, 1) \geq \Delta(\bar{k}, 1)$ for $k < \bar{k}$, but these two conditions do not guarantee that $\pi_{kj} \geq 0$. That is, it is possible to have the situation where

$$\Delta(k, 1) > \frac{Z_k(\hat{j}) - Z_k(j)}{\hat{j}_k - j} > \Delta(\bar{k}, 1)$$

which would imply that $\pi_{kj} < 0$.

First, we note that if there exists a $k' < \bar{k}$ such that $\pi'_{kj} < 0$, then we must have $\Delta(k', 2) > \Delta(\bar{k}, 1)$ since

$$\Delta(k', 2) \geq \frac{Z_k(\hat{j}) - Z_k(j)}{\hat{j}_k - j}$$

for all $j > \hat{j}$ by definition. Therefore, we can restrict our search for non-basic variables as candidates to enter the basis to those X_{kj}'s such that $k < \bar{k}$ and $j = \hat{j}_k^{(2)}$.

Second, we have proved the following:

<u>THEOREM 3.1.</u> If $\sum_{k=1}^{K} \hat{j}_k M_k \geq b$, then the Complex Greedy Solution is optimal in (401) if and only if there does not exist a $k < \bar{k}$ such that $\Delta(k, 2) > \Delta(\bar{k}, 1)$.

<u>COROLLARY 1.</u> If $\hat{j}_k = j_k^*$ and $k < \bar{k}$, then $\pi_{kj} \geq 0$ for all j, $0 \leq j \leq j^*$.

<u>PROOF.</u> $\hat{j}_k = j_k^*$ necessarily means there does not exist a $j > j_k^*$.

Of course, if for each $k < \bar{k}$ we have that $Z_k(\cdot)$ is discretely quasiconvex over the set $\{0, \ldots, j^*\}$, then by Lemma II. 4. 4 and Corollary 1 above we know that the CGS is optimal.

COROLLARY 2. If $k < \bar{k}$ and $\Delta(k, 1) = \Delta(\bar{k}, 1)$ then $\pi_{kj} \geq 0$ for all j, $0 \leq j \leq j_k^*$.

PROOF. Obvious.

The conditions specified in Theorem 3.1 are easy to check and provide us with a means of determining quickly whether or not the CGS is optimal in the LP. If it is, we can derive a good integer solution from the CGS since we immediately have at least K-1 integer basic variables. If $X_{\bar{k}\hat{j}}$ is integer, then so is $X_{\bar{k}0}$ and the CGS is optimal in (207) directly. If $X_{\bar{k}\hat{j}}$ is non integer, we can write a feasible integer solution as follows:

$$
\begin{array}{ll}
X_{k\hat{j}} = M_k & k \leq \bar{k}-1 \\
X_{\bar{k}\hat{j}} = \left\lfloor \dfrac{b - \sum\limits_{k=1}^{\bar{k}-1} \hat{j}_k M_k}{\hat{j}_{\bar{k}}} \right\rfloor & \\
X_{\bar{k}0} = M_{\bar{k}} - X_{\bar{k}\hat{j}} & \\
X_{k0} = M_k & k > \bar{k} \\
X_{kj} = 0 & \text{otherwise.}
\end{array}
$$

If we compare the value of this solution to the value of the CGS, which in this case is an upper bound on the value of the solution to

(207), we are assured that the difference between our good solution and the solution to (207) is strictly bounded above by

$(b' \bmod \hat{j}_{\bar{k}})\Delta(\bar{k}, 1)$, where $b' = b - \sum_{k=1}^{\bar{k}-1} \hat{j}_k M_k$, which in turn is strictly bounded above by $Z_{\bar{k}}(\hat{j}) - Z_{\bar{k}}(0)$, a quantity which does not depend on b or M_k for any k.

4. A Solution Algorithm

The preceding section addressed a special case of (207), that where we have $\sum_{k=1}^{K} \hat{j}_k M_k \geq b$. Nonetheless, the discussion was illustrative of the nature of the problem in general and provides some motivation behind our solution algorithm for problem (401).

The essence of the solution procedure is to choose those feasible solutions which realize the most return to an incremental increase in the number of fillers in the special class. In this light, $\Delta(k, 1)$ represents the maximum return per unit of j over and above $j = 0$ in knapsack type k. Similarly, $\Delta(k, 2)$ is the maximum return per unit of j given that you are going to increase the number of constrained fillers beyond $\hat{j}_k$ and so on. This would be a simple matter of rank ordering the various $\Delta(k, n)$'s except that on the one hand we only have M_k knapsacks of a given type and on the other hand we can only choose one j_k to fill every knapsack of a given type for $K-1$ of the knapsack types. Of course, the particular j chosen can change from one type to another. Furthermore, we must have exactly two different solutions spread among the knapsacks of the one remaining type.

The algorithm described below is designed to cascade down the incremental returns (the $\Delta(k,n)$'s) in decreasing order while preserving feasibility until exactly b units of the restricted filler types are used up.

<u>CASCADING ALGORITHM 4.1.</u> <u>STEP 1.</u> Calculate $\Delta(k,n)$ for all k and for all n, $1 \leq n \leq \bar{n}(k)$. Continue.

<u>STEP 2.</u> If there is a k' such that $j^*_{k'} = 0$, by definition we set $\Delta(k', 1) = 0$. Continue.

<u>STEP 3.</u> Reorder the variables so that $\Delta(k, 1) \geq \Delta(k+1, 1)$ for all k, $1 \leq k \leq K$. In case of a tie, choose so that $\hat{j}_k > \hat{j}_{k+1}$. If $\Delta(k, 1) = \Delta(k+1, 1)$ and $\hat{j}_k = \hat{j}_{k+1}$ then make an arbitrary choice. Continue.

<u>STEP 4.</u> Set $X_{k0} = M_k$ for all k, set $X_{kj} = 0$ otherwise. Continue.

<u>STEP 5.</u> If $\sum_{k=1}^{K} j^*_k M_k > b$ go to Step 6. Otherwise, set $X_{kj^*} = M_k$ for all k, $X_{kj} = 0$ otherwise. Go to Step 11.

<u>STEP 6.</u> If $\sum_{k=1}^{K} \hat{j}_k M_k < b$ go to Step 7. Otherwise,

determine $\bar{k}$ by Definition 3.1 and see if there is any $k < \bar{k}$ such that $\Delta(k, 2) > \Delta(\bar{k}, 1)$. If so, go to Step 7. Otherwise, the CGS is optimal, so set

$$X_{k\hat{j}} = M_k \qquad\qquad \text{for} \quad k < \bar{k},$$

$$X_{\bar{k}\hat{j}} = \frac{b - \sum_{k=1}^{\bar{k}-1} \hat{j}_k M_k}{\hat{j}_{\bar{k}}}$$

$$X_{\bar{k}0} = M_{\bar{k}} - X_{\bar{k}\hat{j}}$$

$$X_{kj} = 0 \qquad\qquad \text{otherwise.}$$

Go to Step 11.

STEP 7. Arrange all $\Delta(k, n)$ in descending order without regard to k or n. In the event of a tie, choose the one with largest k first. Thus, we can associate with any $\Delta(k, n)$ a unique integer i, where i ranges from 1 to $\sum_{k=1}^{K} \bar{n}_k$ and we know that $\Delta_i(k, n) \geq \Delta_{i+1}(k', n')$ for any k, n, k' or n'.

NOTE: In the event of a tie where $\Delta(k, n) = \Delta(k', n')$, there exists the possibility of alternative optimal solutions, depending on the value of b. Nonetheless, we are actually indifferent to the method of breaking a tie.

PROPOSITION 4.1. $\Delta_1(k,n) = \Delta(1,1)$. That is, the $\max_k \max_n \Delta(k,n) = \Delta(1,1)$.

PROOF. By Lemma II.4.2, for fixed k, $\Delta(k,1) \geq \Delta(k,n)$ for all n, $1 \leq n \leq \bar{n}$. By the way we have ordered k, $\Delta(1,1) \geq \Delta(k,1)$ for all $k \geq 1$.

STEP 8. Set $i = 1$. There is a unique X_{kj} associated with every $\Delta(k,n)$, namely $X_{k\hat{j}(n)}$. We shall call the X_{kj} associated with $\Delta_i(k,n)$, $X(i)$. There is a specific value of M_k associated with every $\Delta(k,n)$ by the value of k. We shall denote the M_k connected to $\Delta_i(k,n)$ by $M(i)$. Set $X(1) = M(1)$. We know that this is equivalent to setting $X_{1\hat{j}} = M_1$ by Proposition 4.1. Set $X_{10} = 0$. Set $T = \hat{j}_1 M_1$. Continue.

STEP 9. Set $i = i+1$. With every i there is associated only one k and one $\hat{j}^{(n)}$, which we call $k(i)$ and $j(i)$. Does $k(i) = k(j)$ for any $j < i$? If not, set $X(i) = M(i)$ and set $X_{k(i),0} = 0$.

If so, set $X(i) = M(i)$ and set $X(j) = 0$. Note that $k(i) = k(j)$ means that the $\Delta(k,n)$ uniquely associated with i has the same value for k as the $\Delta(k',n')$ associated with j and $n = n'+1$. Also, if $k(i) = k(j)$ for some $j < i$, we must already have set $X_{k(i),0} = 0$. At this stage of the algorithm, we are ensuring that only one variable associated with knapsack type k has the value M_k. Set $T(i) = \sum_{m=1}^{i} j(i)X(i)$. Note that $T(i) = \sum_{k=1}^{K} \sum_{j=1}^{j_k^*} jX_{kj}$ at each stage of the algorithm. Continue.

STEP 10. Is $T < b$? If so, return to Step 9. If not, set

$$X(i) = \frac{b - \sum_{m=1}^{i-1} j(m)X(m)}{j(i)},$$

i.e. $X(i) = \frac{b-T(i-1)}{j(i)}$. We know that $X(i) = X_{k(i),\hat{j}(n)}$ for some $n \geq 1$ since these are the only candidates we are considering. Set $X_{k(i),\hat{j}(n-1)} = M(i) - X(i)$. Continue.

STEP 11. You have generated the optimal solution to (401). Calculate the value directly as $\sum_{k=1}^{K} \sum_{j=0}^{j_k^*} X_{kj}Z_k(j)$. STOP.

Before we proceed to discuss the optimality of the solution generated by this procedure, an example will help to illustrate what is going on.

EXAMPLE 4.1. Let $K = 7$ and let each knapsack type have the following parameterized generic knapsack functions:

k =	1	2	3	4	5	6	7
$Z_k(0)$	802	915	747	102	114	116	2411*
$Z_k(1)$	863	950	790	103	117	118	1842
$Z_k(2)$	847	1025	784	122*	121*	122	1372
$Z_k(3)$	873	1027	791	115	111	123	681
$Z_k(4)$	881	1091*	762	-	85	115	-
$Z_k(5)$	895*	-	802	-	-	124*	-
$Z_k(6)$	-	-	804*	-	-	-	-
$Z_k(7)$	-	-	-	-	-	-	-

where we have indicated the optimal value with * and indicated an infeasible solution, i.e. $j > j_{max}$, by - . Calculating $\Delta(k, n)$ for all k and n leads to the following data:

k =	1	2	3	4	5	6	7
$\Delta(k, 1)$	61	55	43	$\overline{10}$	$\overline{3.5}$	3	$\overline{0}$
$\Delta(k, 2)$	$\overline{8}$	$\overline{33}$	3	-	-	1	-
$\Delta(k, 3)$	-	-	$\overline{2}$	-	-	$\overline{0.5}$	-

We have indicated $\Delta(k, \bar{n})$ by a bar over the appropriate value. We initially ordered the variables in the correct order for our algorithm to avoid confusion. We have the following data,

$k =$	1	2	3	4	5	6	7
$M_k =$	8	6	5	12	4	10	3
$j^*_k =$	5	4	6	2	2	5	0
$\hat{j}_k =$	1	2	1	2	2	2	0

from which we can calculate the following data needed for the algorithm:

i	$\Delta_i(k, n)$	$M(i)$	$X(i)$	$k(i)$	$j(i)$
1	$61 = \Delta(1, 1)$	8	X_{11}	1	1
2	$55 = \Delta(2, 1)$	6	X_{22}	2	2
3	$43 = \Delta(3, 1)$	5	X_{31}	3	1
4	$33 = \Delta(2, 2)$	6	X_{24}	2	4
5	$10 = \Delta(4, 1)$	12	X_{42}	4	2
6	$8 = \Delta(1, 2)$	8	X_{15}	1	5
7	$3.5 = \Delta(5, 1)$	4	X_{52}	5	2
8	$3 = \Delta(6, 1)$	10	X_{62}	6	2
9	$3 = \Delta(3, 2)$	5	X_{35}	3	5
10	$2 = \Delta(3, 3)$	5	X_{36}	3	6
11	$1 = \Delta(6, 2)$	10	X_{63}	6	3
12	$0.5 = \Delta(6, 3)$	10	X_{65}	6	5
13	$0 = \Delta(7, 0)$	3	X_{70}	7	0

Now, suppose $b = 101$. Let us proceed through the algorithm. Steps 1, 2 and 3 have been accomplished. Note that $j_7^* = 0$ so we defined $\Delta(7,1) = 0$.

STEP 4. We set $X_{10} = 8$, $X_{20} = 6$, $X_{30} = 5$, $X_{40} = 12$, $X_{50} = 4$, $X_{60} = 10$, $X_{70} = 3$, $X_{kj} = 0$ otherwise.

STEP 5. $\sum_{k=1}^{7} j_k^* M_k = 176 > 101 = b.$

STEP 6. $\sum_{k=1}^{7} \hat{j}_k M_k = 77 < 101 = b.$

STEP 7. This was accomplished in the previous data table. Note that we broke the tie between $\Delta(6,1)$ and $\Delta(3,2)$ according to our rule, which is purely arbitrary.

STEP 8. $i = 1$. Set $X_{11} = 8$, $X_{10} = 0$, otherwise no change. $T = 8$.

STEP 9. $i = 2$, $k(2) \neq k(1)$ so set $X(2) = X_{22} = 6$, $X_{20} = 0$ no change otherwise. $T = 20$.

STEP 10. $T = 20 < 101$. Let $i = 3$. $k(3) \neq k(2), k(1)$, so set $X(3) = X_{31} = 5$, $X_{30} = 0$, no change otherwise. $T = 25 < b$. Set $i = 4$. Now, $k(4) = k(2)$, so set $X(4) = X_{24} = 6$ and set $X(2) = X_{22} = 0$, no change otherwise. $T = 37 < b$. Set $i = 5$. $k(5) \neq k(4), k(3), k(2), k(1)$ so we set $X(5) = X_{42} = 12$, $X_{40} = 0$, no change otherwise. $T = 61 < b$. Set $i = 6$. $k(6) = k(1)$. Therefore we set $X(6) = X_{15} = 8$ and $X(1) = X_{11} = 0$, no change otherwise. $T = 94 < b$. Set $i = 7$. $k(7) \neq k(6), k(5), k(4), k(3), k(2), k(1)$, so we set $X(7) = X_{52} = 4$, $X_{50} = 0$ no change otherwise. However $T = 102 > b$. So we set $X(7) = X_{52} = \frac{101-94}{2} = 3.5$ and set $X_{50} = 4 - 3.5 = 0.5$. We are done. The optimal solution is $X_{15} = 8$, $X_{24} = 6$, $X_{31} = 5$, $X_{42} = 12$, $X_{52} = 3.50$, $X_{50} = 0.50$, $X_{60} = 10$, $X_{70} = 3$, $X_{kj} = 0$ otherwise. The value of this solution is 27993.50.

To illustrate a case where the Complex Greedy solution is optimal, consider the following example:

<u>EXAMPLE 4.2.</u> Everything remains the same, except $b = 24$. Then in Step 6, we have $\sum_{k=1}^{7} \hat{j}_k M_k = 77 > 24$. $\bar{k} = 3$ in this case and it is true that there is no $k < 3$ such that $\Delta(k, 2) > \Delta(\bar{k}, 1) = \Delta(3, 1)$. Therefore, the Complex Greedy Solution, to wit: $X_{11} = 8$, $X_{22} = 6$, $X_{31} = 4$, $X_{30} = 1$, $X_{40} = 12$, $X_{50} = 4$, $X_{60} = 10$, $X_{70} = 3$, $X_{kj} = 0$ otherwise, is optimal with value 27034.0.

We shall now show that the solution generated by the cascading algorithm is in fact optimal in (401). We require some definitions and preliminary results. It is clear that the solution given when we terminate in step 5 is optimal in (401) and in fact, in (207). If we terminate with the solution we found in step 6, the Complex Greedy

Solution, Theorem 3.1 establishes its optimality in (401). Consequently we only need to concern ourselves with the solution we arrive at by following steps 7 through 10 until we terminate. Let i be the value of the counting index i in the algorithm when we terminate. There is a unique $\Delta(k,n)$ associated with that i, as well as a specified k which we have called k(i). Let the n in the unique $\Delta_i(k,n)$ be denoted as n(i). It is clear from the algorithm itself that there are K+1 variables identified in the solution, of which at least K and at most K+1 are positive. Furthermore, feasibility is preserved since we ensure that $\sum_{j=0}^{j_k^*} X_{kj} = M_k$ for all k and we end up with $T = \sum_{k=1}^{K} \sum_{j=1}^{j_k^*} jX_{kj} = b$. Therefore, we can consider the algorithm's solution a basic feasible solution in the LP.

It is also true that the knapsack type which contributed two variables to the basic feasible solution generated by the algorithm is type k(i). All other knapsack types contribute one positive variable to the solution. The two variables from knapsack type k(i) are either both positive or one is zero. Furthermore, any variable in the solution must be of the form $X_{k\hat{j}(n)}$ for given k, where $n \geq 0$. Recall that $\hat{j}(0)$ is defined to be 0. This is the case simply because we do not consider any other variables in the solution. We shall mean, in $\Delta(k,n)$ or $\hat{j}_k^{(n)}$, the particular n chosen by the algorithm for each k upon termination. This notational convenience may get confusing unless the reader keeps in mind that the n for one k may not equal the n associated with another.

Let $\ell = k(i)$. We note, for any $k < \ell$,

1) $\hat{j}_k^{(n)} \neq 0$ i.e. $n > 0$ and

2) $\Delta(k,n) \geq \Delta(k(i), n(i)) = \Delta(\ell, n)$.

The first claim is clear since $k < \ell$ means that $\Delta(k,1) \geq \Delta(\ell,1) \geq \Delta(\ell,n)$ and therefore we have at least worked our way past $\hat{j}_k$ by the time we arrive at $\Delta(\ell,n)$. Part 2 follows when we note that all that is meant by $\Delta(k,n)$ is that this is the particular $\Delta(k,n)$ associated with the variable the algorithm has brought into the basic feasible solution at a positive level. If it were true that $\Delta(k,n) < \Delta(\ell,n)$ we simply would not have come to it as we worked down the list of $\Delta(k,n)$'s. Note also that $X_{k\hat{j}^{(n)}} = M_k$ for $k < \ell$. We can express this argument as the proof of

PROPOSITION 4.2. For $k < \ell$, it must be true that $\Delta(k,n+1) \leq \Delta(\ell,n)$.

PROOF. This is really a definition, since if $\Delta(k,n+1) > \Delta(\ell,n)$, it must be true that the variable $X_{k\hat{j}^{(n')}}$ is in the solution at a positive level and $n' = n+1$. But we defined $n' = n$.

We also note that the two variables in the solution from knapsack type ℓ are necessarily $X_{\ell \hat{j}(n)}$ and $X_{\ell \hat{j}(n-1)}$. That is, whenever $\hat{j}(n) = j(i)$, the other variables has $j = \hat{j}(n-1)$.

Now, when $k > \ell$, we either have $\hat{j}_k(n) = 0$, i.e. $X_{k0} = M_k$ in the final solution, or we have $\hat{j}_k(n) > 0$.

PROPOSITION 4.3. If, for $k > \ell$, $\hat{j}_k(n) = 0$, then $\Delta(\ell, n) > \Delta(k, 1)$.

PROOF. Suppose $\Delta(k, 1) \geq \Delta(\ell, n)$. Then $\Delta(k, 1)$ would appear on the ordered list before $\Delta(\ell, n)$ even in the case of equality by our tie breaking rule. Therefore, we would have brought in $X_{k\hat{j}} = M_k$, which contradicts the fact that $\hat{j}_k(n) = 0$ implies $X_{k0} = M_k$.

PROPOSITION 4.4. If, for $k > \ell$, $\hat{j}_k(n) > 0$ then $\Delta(k, n) \geq \Delta(\ell, n) > \Delta(k, n+1)$.

PROOF. Again, this is simply what we have defined the n in $\Delta(k, n)$ to mean: that value of n such that $X_{k\hat{j}(n)} = M_k$ in the final solution.

The preceding rather convoluted discussion has prepared us to prove the optimality of the solution derived by the algorithm of this section.

<u>THEOREM 4.1.</u> The solution generated by the cascading algorithm, that is

$$X_{k\hat{j}^{(n)}} = M_k\,, \qquad k \neq \ell$$

$$X_{\ell\hat{j}^{(n)}} = b - \sum_{k\neq\ell} \hat{j}_k^{(n)} M_k$$

$$X_{\ell\hat{j}^{(n-1)}} = M_\ell - X_{\ell\hat{j}^{(n)}}$$

$$X_{kj} = 0 \qquad \text{otherwise,}$$

is optimal in the linear program (401).

<u>PROOF.</u> Consider the basis generated by this solution:

$$B = \begin{array}{cc} \begin{matrix} & & & & \ell^{\text{th}}\text{ column} & & & & \end{matrix} & \\ \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \cdots & & & \cdots & \cdots & & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & 1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \cdots & & & \cdots & \cdots & & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \\ \hat{j}_1^{(n)} & \hat{j}_2^{(n)} & \hat{j}_3^{(n)} & \cdots & \hat{j}_\ell^{(n-1)} & \hat{j}_\ell^{(n)} & \cdots & \hat{j}_{K-1}^{(n)} & \hat{j}_K^{(n)} \end{bmatrix} & \begin{matrix} \\ \\ \\ \\ \} \ \ell^{\text{th}}\text{ row} \\ \\ \\ \\ \\ \end{matrix} \end{array}$$

which has inverse

$$B^{-1} = \frac{1}{d}\begin{array}{c} \begin{array}{ccccccccc} & & & & \overset{\ell^{th}\text{ column}}{\frown} & & & & \end{array} \\ \left[\begin{array}{ccccccccc} d & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & d & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & d & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \cdots & & & & \cdots & & & \cdots & \\ \hat{j}^{(n)}_1 & \hat{j}^{(n)}_2 & \hat{j}^{(n)}_3 & \cdots & \hat{j}^{(n)}_\ell & \hat{j}^{(n)}_{\ell+1} & \cdots & \hat{j}^{(n)}_K & -1 \\ -\hat{j}^{(n)}_1 & -\hat{j}^{(n)}_2 & -\hat{j}^{(n)}_3 & \cdots & -\hat{j}^{(n-1)}_\ell & -\hat{j}^{(n)}_{\ell+1} & \cdots & -\hat{j}^{(n)}_K & 1 \\ \cdots & \cdots & \cdots & & \cdots & & & \cdots & \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & d & 0 \end{array}\right] \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \} \ \ell^{th}\text{ row} \\ \\ \\ \\ \\ \end{array}$$

where $d = \hat{j}^{(n)}_\ell - \hat{j}^{(n-1)}_\ell > 0$. We have that

$$C_B = (Z_1(\hat{j}^{(n)}), \ldots, Z_\ell(\hat{j}^{(n-1)}), Z_\ell(\hat{j}^{(n)}), Z_{\ell+1}(\hat{j}^{(n)}), \ldots, Z_K(\hat{j}^{(n)}))$$

and as usual the kj^{th} column of the A matrix, $a_{kj} = e_k + je_{K+1}$. Therefore, for $k \neq \ell$,

$$B^{-1}a_{kj} = e_k + \left(\frac{\hat{j}^{(n)}_k - j}{d}\right)e_\ell + \left(\frac{j - \hat{j}^{(n)}_k}{d}\right)e_{\ell+1}\,.$$

Consequently $\pi_{kj} = C_B B^{-1} a_{kj} - Z_k(j)$ or

$$\pi_{kj} = Z_k(\hat{j}^{(n)}) - Z_k(j) + (j - \hat{j}^{(n)}_k)\Delta(\ell, n).$$

1) Suppose $\hat{j}^{(n)}_k = 0$ and $\pi_{kj} < 0$. That implies that

$\Delta(\ell, n) < \dfrac{Z_k(j) - Z_k(0)}{j}$ which in turn implies $\Delta(\ell, n) \leq \Delta(k, 1)$. Proposition 4.2 and Proposition 4.3 show that this is a contradiction for any k. Therefore $\pi_{k\ell} \geq 0$.

2) Suppose $\hat{j}_k^{(n)} > 0$. a) Then, for $j < \hat{j}_k^{(n)}$, $\pi_{kj} < 0$ implies that

$$\frac{Z_k(\hat{j}^{(n)}) - Z_k(j)}{\hat{j}_k^{(n)} - j_k} < \Delta(\ell, n).$$

Now, we know from Lemma II. 4.2 and its corollary that $\dfrac{Z_k(\hat{j}^{(n)}) - Z_k(j)}{\hat{j}_k^{(n)} - j_k} \geq \Delta(k, n)$ but we have by the definition of our notation that $\Delta(k, n) \geq \Delta(\ell, n)$ if $k < \ell$ and Proposition 4.4 says that $\Delta(k, n) \geq \Delta(\ell, n)$ if, for $k > \ell$, $\hat{j}_k^{(n)} > 0$ as we have assumed. Therefore $\pi_{kj} \geq 0$. b) Suppose that $j = \hat{j}_k^{(n)}$, we know that $\pi_{kj} = 0$ in this case. c) Assume $j > \hat{j}_k^{(n)}$. $\pi_k < 0$ implies that

$$\frac{Z_k(j) - Z_k(\hat{j}^{(n)})}{j_k - \hat{j}_k^{(n)}} > \Delta(\ell, n).$$

Suppose there were a $j > j_k^{(n)}$ such that $\pi_k < 0$. Then certainly

$$\Delta(k, n+1) = \frac{Z_k(\hat{j}^{(n+1)}) - Z_k(\hat{j}^{(n)})}{\hat{j}_k^{(n+1)} - \hat{j}_k^{(n)}} \geq \frac{Z_k(j) - Z_k(\hat{j}^{(n)})}{j - \hat{j}_k^{(n)}} > \Delta(\ell, n)$$

but this contradicts Proposition 4.2. Therefore, $\pi_{kj} \geq 0$.

3) The only case we have left is for $k = \ell$. In this case,

$$B^{-1} a_{\ell j} = \left(\frac{\hat{j}_\ell^{(n)} - j}{d}\right) e_\ell + \left(\frac{j - \hat{j}_\ell^{(n-1)}}{d}\right) e_{\ell+1}$$

and so

$$\pi_{\ell j} = (\hat{j}_\ell^{(n)} - j) \frac{Z_\ell(\hat{j}^{(n)})}{d} + (j - \hat{j}_\ell^{(n-1)}) \frac{Z_\ell(\hat{j}^{(n-1)})}{d} - Z_\ell(j)$$

If we let $\hat{j}_\ell^{(n)} - j = \hat{j}_\ell^{(n)} - \hat{j}_\ell^{(n-1)} - (j - \hat{j}_\ell^{(n-1)})$ and rearrange terms and multiply through by d, we have

$$d\pi_{\ell j} = (\hat{j}_\ell^{(n)} - \hat{j}_\ell^{(n-1)})(Z_\ell(\hat{j}^{(n-1)}) - Z_\ell(j)) + (j - \hat{j}_\ell^{(n-1)})(Z_\ell(\hat{j}^{(n)}) - Z_\ell(\hat{j}^{(n-1)}))$$

Suppose $j < \hat{j}_\ell^{(n-1)}$. $\pi_{\ell j} < 0$ implies $d\pi_{\ell j} < 0$ since $d > 0$ and so we have

$$\frac{Z_\ell(\hat{j}^{(n-1)}) - Z_\ell(j)}{\hat{j}_\ell^{(n-1)} - j} < \frac{Z_\ell(\hat{j}^{(n)}) - Z_\ell(\hat{j}^{(n-1)})}{\hat{j}^{(n)} - \hat{j}^{(n-1)}} = \Delta(\ell, n)$$

but this contradicts Corollary 1 to Lemma II.4.2. So $\pi_{\ell j} \geq 0$ for $j < \hat{j}_\ell^{(n-1)}$. Suppose $j > \hat{j}_\ell^{(n-1)}$, then $\pi_{\ell j} < 0$ implies that

$$\Delta(\ell, n) < \frac{Z_\ell(j^{(n-1)}) - Z_\ell(j)}{j_\ell^{(n-1)} - j}$$

which contradicts the definition of $\Delta(\ell, n)$. Therefore, $\pi_{\ell j} \geq 0$ for $j > \hat{j}_\ell^{(n-1)}$.

Since $\pi_{\ell \hat{j}^{(n-1)}} = 0$, we are done. That is, we have shown that $\pi_{kj} \geq 0$ for all k, $1 \leq k \leq K$ and all j, $0 \leq j \leq j_k^*$ and therefore the feasible solution due to the algorithm is optimal.

The algorithm, which we have shown leads to an optimal solution, has an interesting interpretation. It says, for a given b, that there exists a number which we can consider a cutoff value, call it C, such that if $\Delta(k, n)$ is greater than or equal to C, then we can consider $X_{k\hat{j}^{(n)}}$ as a candidate to have a positive value in the optimal solution. Alternately, if $\Delta(k, n)$ is less than C, then we will set $X_{k\hat{j}^{(n)}} = 0$ in the optimal solution. Furthermore, for a given k, the variable which will appear in the optimal solution at a positive level corresponds to the largest n such that $\Delta(k, n) \geq C$. Of course, for that knapsack type k which has two solutions in the basis for the optimal solution, it is the variables associated with the largest two n's such that $\Delta(k, n) \geq C$. In this case, we know that the largest n of the pair is such that $\Delta(k, n) = C$.

We note, as the reader familiar with the theory of linear programming may have surmised, the value of C is precisely the value of the variable associated with the factory constraint in the optimal solution to the dual problem to (401).

One can state this as follows. There exists a value, C, such that if, for any knapsack type, you choose the largest number of fillers of restricted type whose marginal return to an incremental increase from $\hat{j}^{(n-1)}$ to $\hat{j}^{(n)}$ per unit of restricted filler, $\Delta(k, n)$ is greater than C, then that is the optimal thing to do.

One can see that the cut off value, C, is the optimal value of the dual variable associated with the factory constraint by noting that it is the last element in the vector of dual variables. Its value is calculated by computing $C_B B^{-1}$ where B^{-1} is the inverse of the basis of the primal problem and that, at optimality,

$$C_B = (Z_1(\hat{j}^{(n)}, \ldots, Z_\ell(\hat{j}^{(n-1)}), Z_\ell(\hat{j}^{(n)}), Z_{\ell+1}(\hat{j}^{(n)}), \ldots, Z_K(\hat{j}^{(n)}))$$

in the notation of Theorem 4. 1.

Solving (401) is equivalent to determining the value of C before the shift, giving us a dynamic decision rule: "Fill the knapsack with the largest value of j whose incremental marginal return is greater than C. If none, fill the knapsack with 0." Of course, C is a decreasing function of b and can be calculated directly only by knowing M and M_k for all k.

Finally, we point out that the order of receipt of the knapsacks in the Complex Problem is no more important here than in the Basic Problem once we have generated the solution. All we need to know of a specific knapsack is its type to determine the proper solution. Of course, for the ℓ^{th} type, we need to keep track of how many we have already filled with each of the two allowable solutions.

5. The Heuristic Solution

The algorithm described in Section 4 provides an optimal solution to the linear program (401). We can proceed in the same way as the Basic Problem to transform that solution into a solution to the Complex Factory Problem (207). The task is made easy by the fact that at least $K-1$ of the non-zero valued variables in the optimal solution to (401) are positive integers and so we need do nothing to them. All but 2 of the others are equal to zero and also need not be changed. The remaining two are either both integer valued, in which case the solution to (401) is also optimal in (207) or they are both non-integers. In this event, both variables are associated with the same knapsack type, call it type ℓ, and we proceed to modify them in exactly the same way as the Basic case.

Specifically, we have

$$X_{\ell \hat{j}^{(n)}} = \frac{b - \sum_{m=1}^{\ell-1} j(m)X(m)}{\hat{j}_\ell^{(n)}} ,$$

where $\hat{j}_\ell^{(n)}$, $j(m)$ and $X(m)$ are defined in algorithm 4. 1. In addition, $X_{\ell \hat{j}^{(n-1)}} = M_k - X_{\ell \hat{j}^{(n)}}$. By definition, $\hat{j}_\ell^{(n)} > \hat{j}_\ell^{(n-1)}$ so that if we round $X_{\ell \hat{j}^{(n)}}$ down (and thereby round $X_{\ell \hat{j}^{(n-1)}}$ up) to the nearest integer we shall generate an integer solution which maintains feasibility. That is, if we set

$$X_{\ell \hat{j}^{(n)}} = \left\lfloor \frac{b - \sum_{m=1}^{\ell-1} j(m)X(m)}{\hat{j}_\ell^{(n)}} \right\rfloor$$

then, the difference between the value of the solution to (401) and the value of our solution to (207) can be written as $r_\ell (Z_\ell (\hat{j}^{(n)}) - Z_\ell (\hat{j}^{(n-1)})$ where

$$r_\ell = \frac{b - \sum_{m=1}^{\ell-1} j(m)X(m)}{\hat{j}_\ell^{(n)}} - \left\lfloor \frac{b - \sum_{m=1}^{\ell-1} j(m)X(m)}{\hat{j}_\ell^{(n)}} \right\rfloor .$$

If we abbreviate the numerator in the remainder expression above as b', we can write the remainder

$$r_\ell = \frac{b' \bmod \hat{j}_\ell^{(n)}}{\hat{j}_\ell^{(n)}},$$

a number strictly less than 1. Therefore, the different between our solution and the optimal solution to (401) is strictly less than $Z_\ell(\hat{j}^{(n)}) - Z_\ell(\hat{j}^{(n-1)})$ which in turn is less than or equal to $Z_\ell(j^*) - Z_\ell(0)$. The determination of this value depends upon knowing the value of ℓ, which can only be specified after all the parameters of the problem are established. Nonetheless, we can establish a strict upper bound on the potential loss for any problem. Define this loss as $LP - \overline{IP}$, where LP is the value of the solution to (401) and $\overline{IP}$ is our solution to (207).

<u>PROPOSITION 5.1.</u> The potential loss

$$LP - \overline{IP} < \max_k j_k^* \Delta(1, 1).$$

<u>PROOF.</u> On the previous page, we wrote the loss as $r_\ell(Z_\ell(\hat{j}^{(n)}) - Z_\ell(\hat{j}^{(n-1)}))$ which is equivalent to $b' \bmod \hat{j}_\ell^{(n)} \Delta(\ell, n)$. Now $\Delta(\ell, n) \leq \Delta(1, 1)$ and $b' \bmod \hat{j}_\ell^{(n)}$ is an integer strictly less than $\hat{j}_\ell^{(n)}$. $\hat{j}_\ell^{(n)} \leq j_\ell^*$ and, of course, $\max_k j_k^* \geq j_\ell^*$.

This is a very conservative upper bound which is, nonetheless, independent of M, M_k and b. Consequently, as the number of knapsacks in a shift increases the percentage of potential loss goes to zero. Another limit which we can establish with a little more work is given for any problem by

PROPOSITION 5.2. $\quad LP - \overline{IP} < \max_k \max_n \{\hat{j}_k^{(n)} \Delta(k,n)\}.$

PROOF. This bound is obtained by observing that the actual potential loss in any particular problem is strictly bounded by $\hat{j}_\ell^{(n)} \Delta(\ell, n)$ for some ℓ and some n.

We will demonstrate the simplicity of the procedure in our previous example.

EXAMPLE 5.1. Find a good and simple integer solution to the problem specified in Example 4.1. The optimal solution to (401) in that case was $X_{15} = 8$, $X_{24} = 6$, $X_{31} = 5$, $X_{42} = 12$, $X_{52} = 3.5$, $X_{50} = 0.5$, $X_{60} = 0$, $X_{70} = 0$, $X_{kj} = 0$ otherwise with a value of 27993.5. Our solution to (207) is therefore the same except $X_{52} = 3$, $X_{50} = 1$. This solution has a value of 27990. The difference is $r_\ell(Z_\ell(\hat{j}^{(n)}) - Z_\ell(\hat{j}^{(n-1)})) = r_5(Z_5(2) - Z_5(0))$. In this case,

$$r_5 = \frac{b' \bmod \hat{j}_5}{\hat{j}_5} = \frac{(101-94) \bmod 2}{2} = 0.5$$

and $Z_5(2) - Z_5(0) = 121 - 114 = 7$, giving us a total difference of 3.5. This difference is 0.012 percent of the possible optimal value. The upper bound given by Proposition 5.1 is 305 which is very conservative. The bound specified by Proposition 5.2 is 132 and corresponds to $k = 3$, $n = 1$. This bound is valid for any M, M_k $k = 1, \ldots, K$ and any b, for the set of knapsack types in the example.

BIBLIOGRAPHY

[1] Garfinkel, R.S. and Nemhauser, G. L. (1972), Integer Programming, John Wiley & Sons, New York.

[2] Gilmore, P.C. and Gomory, R.E. (1966), "The Theory and Computation of Knapsack Functions," Operations Research, 14, 1045-1054.

[3] Hadley, G. (1962), Linear Programming, Addison-Wesley, Reading, Massachusetts.

[4] Hu, T.C. and Lenard, M. L. (1976), "Optimality of a Heuristic Solution for a Class of Knapsack Problems," Operations Research, 24, 193-196.

[5] Magazine, M. J. , Nemhauser, G. L. and Trotter, L.E. (1975), "When the Greedy Solution Solves a Class of Knapsack Problems," Operations Research, 23, 207-217.

[6] Mangasarian, O. L. (1969), Non Linear Programming, McGraw-Hill, New York.

[7] Salkin, H.M. and de Kluyver, C.A. (1975), "The Knapsack Problem: a Survey," Naval Research Logistics Quarterly, 22, 127-144.