# New Turnpike Theorems for the Unbounded Knapsack Problem

Ping Heidi Huang[1]* and Thomas L. Morin[2]

[1]Krannert School of Management, [2]School of Industrial Engineering

Purdue University, West Lafayette, Indiana 47907

**Abstract:**

We develop sharp bounds on turnpike theorems for the unbounded knapsack problem. Turnpike theorems specify when it is optimal to load at least one unit of the *best* item (i.e., the one with the highest "bang-for-buck" ratio) and, thus can be used for problem preprocessing. The successive application of the turnpike theorems can drastically reduce the size of the knapsack problems to be solved. Two of our theorems subsume known results as special cases. The third one is an entirely different result. We show that all three theorems specify *sharp bounds* in the sense that no smaller bounds can be found under the assumed conditions. We also prove that two of the bounds can be obtained in constant time. Computational results on randomly generated problems demonstrate the effectiveness of the turnpike theorems both in terms of how often they can be applied and the resulting reduction in the size of the knapsack problems.

**Keywords:**

Integer Programming, Knapsack Problem, Number Theory, Preprocessing, Turnpike Theorems

* Ping Heidi Huang, huang74@mgmt.purdue.edu, Krannert School of Management, Purdue University, 403 W. State St., West Lafayette, IN 47907, USA. Tel. 001-765-496-7115 Fax. 001-765-494-9658

## 1. Introduction

The *knapsack problem* is one of the most celebrated problems in operations research. The *unbounded knapsack problem* can be stated as follows. Given a knapsack (or backpack) with known weight capacity and an unlimited supply of items, each of which has known unit value and unit weight, how can one pack the knapsack with integral amounts of the items so as to maximize the value of the load carried?

It is well known that this problem is NP-hard (Garey and Johnson, 1990). There are a host of algorithms for the knapsack problem and its various variants (Kellerer *et al.*, 2004; Martello and Toth, 1990; Morin and Marsten, 1976; Zukerman *et al.*, 2001). However, it is oftentimes possible to drastically reduce the size of the knapsack problem to be solved even before applying one of these algorithms. This can be viewed as a form of problem pre-processing (Nemhauser, 2007). Specifically, one such way of cutting down the computational requirements of solving knapsack problems with large, but bounded weight capacities is to employ *turnpike theorems* (Garfinkel and Nemhauser, 1972; Gilmore and Gomory, 1966; Hu, 1969; Shapiro, 1970). If the items are indexed according to the non-increasing order of the "bang-for-buck" ratios of their unit values to unit weights, then for "large enough" weight capacities (i.e., "long enough" trips) it can be shown that it is optimal to load at least one unit of "best" item (to take the trip on the turnpike, as it were), which has the highest value-to-weight ratio. Turnpike theorems specify *lower bounds* on what constitutes a "large enough" weight capacity (right-hand-side) and their successive application can drastically reduce the right-hand-sides and the resulting computational requirements. For instance, a knapsack problem with several hundred items and a carrying capacity of 100,000 can be reduced to a problem with a carrying capacity of only 40 (see Example 4.4). The primary goal of the present paper is to provide sharper bounds and, thus, more effective turnpike theorems for the unbounded knapsack problem.

The plan of our paper is as follows. In section 2, we introduce notations and then develop the first turnpike theorem, which is shown to subsume a known result attributed to Hu (1969) by Garfinkel and Nemhauser (1972). Moreover, we prove that our bound can be obtained in constant time. We also prove that it is a *sharp bound* in the sense that in general, no smaller lower bound can be found under the assumed conditions. In section 3, we develop the second turnpike theorem and prove it also subsumes another known result from Garfinkel and Nemhauser (1972) as a corollary. The third turnpike theorem, which is an entirely different result, is given in section 4. Number theoretic results are used to prove our new turnpike theorems, and examples are given to show that our theorems are stronger than the known results. We also show that all three of our turnpike theorems specify *sharp bounds* in the sense that no smaller bounds can be found under the assumed conditions and prove that two of the bounds can be obtained in constant time. Computational experiments on randomly generated problems clearly demonstrate the effectiveness of the turnpike theorems not only in terms of when they can be applied but also in the resulting reduction in the size of the knapsack problems. Indeed, the three turnpike theorems yielded average reductions of the right-hand-sides of 99.97%, 98.10% and 99.93%, respectively. The paper concludes with a discussion in section 5.

## 2. Preliminaries and the First Turnpike Theorem

The unbounded knapsack problem can be stated as follows: given an unlimited number of *n* items, each of which has unit weight $a_j$ and unit value $c_j$, how can one fill a knapsack with weight capacity *b* in order to maximize the value of the load carried? More formally, one wants to find non-negative integers $\{x_1, x_2, \ldots, x_n\}$ in order to

$$\text{maximize} \quad \sum_{j=1}^{n} c_j x_j$$

$$\text{subject to} \quad \sum_{j=1}^{n} a_j x_j \le b \text{, and}$$

$$x_j \ge 0 \text{, integer ,}$$

where, as usual, we assume that all data are integral. As noted in the introduction, one way of reducing the computational requirements of solving knapsack problems with large, but bounded, right-hand-side weight capacities is to employ *turnpike theorems*. Other distinctly different ways include dominance results and periodicity properties (Kellerer *et al.*, 2004). If we index the items according to the non-increasing order of their "bang-for-buck" ratios, $v_j = c_j / a_j$ , so that $v_1 \ge v_2 \ge v_3 \ge \cdots$, then for "large enough" *b* we can prove that it is optimal to load at least one unit of "best" item 1 which has the highest value-to-weight ratio $v_1$, into the knapsack. Turnpike theorems specify a lower bound on what constitutes a "large enough" *b*. As noted in the

Introduction, successive application of such turnpike theorems can oftentimes drastically reduce the right-hand-side *b*, and the resulting computational requirements.

In constructing an optimal solution of a given knapsack problem, we (inductively) want to decide if at least one unit of item 1 (recall that by assumption, item 1 has the highest value-to-weight ratio $v_1$) should be packed into the knapsack. Certainly if it is optimal to pack one unit of item 1, then there is a reduction of the (remaining) capacity, and thus, the problem is easier to solve. Before stating and proving our first theorem, we will state and prove a simple Lemma.

**Lemma 2.1** *Let q be a positive integer and s be a real number where* $s \geq q$, *then we have*

$$\frac{q}{q+1} s < \lfloor s \rfloor ,$$

*where* $\lfloor s \rfloor$ = *largest integer* $\leq s$.

**Proof.** Let $\lfloor s \rfloor = y$, it follows from the assumption $s \geq q$ that $y \geq q$. Let $s - y = t$ with $0 \leq t < 1$. Thus, we have

$$\frac{q}{q+1} s = \frac{q}{q+1} y + \frac{q}{q+1} t = (1 - \frac{1}{q+1}) y + \frac{q}{q+1} t$$

$$= (y - \frac{y}{q+1}) + \frac{qt}{q+1}$$ and since $y \geq q$, we have

$$\leq (y - \frac{q}{q+1}) + \frac{qt}{q+1} = y - \frac{q(1-t)}{q+1}$$ and by definition $t < 1$, we have

$$< y = \lfloor s \rfloor ,$$

or $\frac{q}{q+1} s < \lfloor s \rfloor$, which completes the proof of the Lemma 2.1. ■

Note that the interval [0, 1) can be written as the disjoint union of intervals of the form $\left[ \frac{q-1}{q}, \frac{q}{q+1} \right)$, i.e.,

$$[0, 1) = \bigcup_{q=1}^{\infty} \left[ \frac{q-1}{q}, \frac{q}{q+1} \right)$$

and any real number $r$ with $0 \le r < 1$ falls precisely into one of the subintervals above. We have the following theorem,

**Theorem 2.2** *Assume that* $\frac{c_1}{a_1} > \frac{c_2}{a_2}$ (i.e., $v_1 > v_2$), *and let the positive integer q be uniquely determined by the following relation*

$$\frac{q-1}{q} \le \frac{v_2}{v_1} < \frac{q}{q+1} \quad .$$

*Then there is a weight* $h_I = qa_1$ *such that if* $b \ge h_I$, *it is optimal to pack at least one unit of item 1 into the knapsack.*

**Proof.** We have $\frac{c_1}{a_1} > \frac{c_2}{a_2} \ge \frac{c_3}{a_3} \ge \cdots$ (i.e., $v_1 > v_2 \ge v_3 \ge \cdots$). Clearly, if no units of item 1 are loaded, then the total value of the objective function is $z \le bv_2$ which leads to $z < b\left(\frac{q}{q+1}v_1\right)$, due to the assumption that $\frac{v_2}{v_1} < \frac{q}{q+1}$. Therefore, we have

$$z < b\frac{q}{q+1}v_1 \quad . \qquad (2.1)$$

We want to show that any solution without item 1 loaded can not be optimal.

We claim that

$$b\frac{q}{q+1}v_1 < \left\lfloor \frac{b}{a_1} \right\rfloor c_1 \quad . \qquad (2.2)$$

Note that the right-hand-side of inequality (2.2) is the value of filling the knapsack with as many units of item 1 as possible (which is certainly allowed). The above claim will establish our theorem.

Now let us prove the claim. By definition, $v_j = c_j / a_j$, thus

$$\frac{v_1}{c_1} = \frac{1}{a_1} \quad .$$

Therefore, inequality (2.2) is equivalent to the following

$$\left(\frac{q}{q+1}\right)\frac{b}{a_1} < \left\lfloor \frac{b}{a_1} \right\rfloor ,$$

and the above inequality follows at once from Lemma 2.1 by taking $s = \frac{b}{a_1}$ (Note that $b \geq qa_1$ in the assumption, so we have $\frac{b}{a_1} \geq q$. Thus, Lemma 2.1 can be employed). Therefore, inequality (2.2) holds.

Combining (1) and (2), we have

$$z < \left\lfloor \frac{b}{a_1} \right\rfloor c_1$$

The right-hand-side is equal to the objective function value of a way of loading item 1. So the original solution can not be optimal. Therefore, an optimal solution includes at least one unit of item 1. ■

Notice that Theorem 2.2 subsumes a known result, attributed to Hu (1969) by Garfinkel and Nemhauser (1972) as a special case that is stated in the following corollary.

**Corollary 2.3** *Let us assume that* $\frac{c_1}{a_1} > \frac{c_2}{a_2}$ (i.e., $v_1 > v_2$), *then there is a weight* $h_{I'} = \frac{c_1}{v_1 - v_2}$ *such that if* $b \geq h_{I'}$, *it is optimal to pack at least one unit of item 1 into the knapsack.*

**Proof.** Let us use the notation of the preceding theorem. Note that there is a unique positive integer $q$ that is the maximal possible and satisfies the following inequality,

$$\frac{q-1}{q} \leq \frac{v_2}{v_1} \quad \Leftrightarrow \quad v_1 - \frac{1}{q}v_1 \leq v_2 \quad \Leftrightarrow \quad v_1 - v_2 \leq \frac{1}{q}v_1 \quad \Leftrightarrow \quad \frac{c_1}{v_1 - v_2} \geq \frac{qc_1}{v_1} = qa_1$$

Therefore, $b \geq \frac{c_1}{v_1 - v_2}$ implies that $b \geq qa_1$. Thus, the corollary follows from Theorem 2.2. ■

**Remark 2.4:** It follows from the proof of the corollary that $h_{I'} = \frac{c_1}{v_1 - v_2} \geq qa_1 = h_I$. ■

Here is an interesting issue, as pointed out by one of the anonymous referees, of finding the number $q$ in the statement of the preceding theorem. If we consider the sequence $0 < \frac{1}{2} < \frac{2}{3} < \frac{3}{4} < \cdots$ and try to fit $\frac{v_2}{v_1} < 1$ into the segment $\left[\frac{q-1}{q}, \frac{q}{q+1}\right)$, i.e., we compare $\frac{v_2}{v_1}$ with $\frac{n}{n+1}$ one by one for $1 \le n < a_2 c_1$ (since $\frac{v_2}{v_1} = \frac{c_2/a_2}{c_1/a_1} = \frac{a_1 c_2}{a_2 c_1} \le \frac{a_2 c_1 - 1}{a_2 c_1}$), then the process would require a linear time in the size of $a_2 c_1$, and we would end up in a situation with a tighter bound but a longer time (to find the bound). However, the appropriate $q$ can be calculated in at most two steps if we use the following lemma and, thus, our bound can be found in constant time instead of linear time.

**Lemma 2.5** *Let $0 < r < 1$ be any real number. There exists a unique expression such that $r = \frac{1}{n_1 + r_1}$ with $n_1 \ge 1$ being an integer and $0 \le r_1 < 1$. We have $r \in \left[\frac{q-1}{q}, \frac{q}{q+1}\right)$, where the positive integer $q$ is determined by*

*(i) if $n_1 = 1$: we must have $0 < r_1 < 1$. Thus, there exists a unique expression such that*

$$r_1 = \frac{1}{n_2 + r_2} \text{ with } n_2 \ge 1 \text{ being an integer and } 0 \le r_2 < 1, \text{ then } q = n_2 + 1;$$

*(ii) if $n_1 = 2$ and $r_1 = 0$, then $q = 2$;*

*(iii) all other cases, $q = 1$.*

**Proof.** Clearly, $\frac{1}{r} > 1$ can be uniquely expressed as $\frac{1}{r} = n_1 + r_1$ with $n_1 \ge 1$ (integer) and $0 \le r_1 < 1$. Now consider case (i) $n_1 = 1$. If $r_1 = 0$, then $r = \frac{1}{n_1 + r_1} = 1$ contradicting to our assumption. Therefore, we have $0 < r_1 < 1$. Again, $\frac{1}{r_1} > 1$ can be uniquely expressed as $\frac{1}{r_1} = n_2 + r_2$ with $n_2 \ge 1$ (integer) and $0 \le r_2 < 1$. This implies $n_2 \le \frac{1}{r_1} < n_2 + 1$, so we have $\frac{1}{n_2+1} < r_1 \le \frac{1}{n_2}$ . Going back to $r$ with $n_1 = 1$, we have $r = \frac{1}{1+r_1} \ge \frac{1}{1+\frac{1}{n_2}} = \frac{n_2}{n_2+1}$ and

$r = \frac{1}{1+r_1} < \frac{1}{1+\frac{1}{n_2+1}} = \frac{n_2+1}{n_2+2}$. Thus, we obtain $\frac{n_2}{n_2+1} \le r < \frac{n_2+1}{n_2+2}$. Therefore, $q = n_2 + 1$.

Consider case (ii) $n_1 = 2$ and $r_1 = 0$. We have $r = \frac{1}{n_1+r_1} = \frac{1}{2}$ , $r \in \left[\frac{q-1}{q}, \frac{q}{q+1}\right)$, thus $q = 2$.

Now consider case (iii), *if* $n_1 = 2$ and $r_1 > 0$, then $r = \frac{1}{n_1+r_1} < \frac{1}{2}$, thus $q = 1$;

*if* $n_1 > 2$, then $r = \frac{1}{n_1+r_1} < \frac{1}{2}$, we also obtain $q = 1$. ■

**Remark 2.6:** In Theorem 2.2, $0 < \frac{v_2}{v_1} < 1$. Thus, Lemma 2.5 can be applied to find *q*. Since $\frac{v_2}{v_1} = \frac{c_2/a_2}{c_1/a_1} = \frac{a_1c_2}{a_2c_1}$, we can simply use *long division* (or *continuous fraction*) to calculate $n_1$ and $n_2$ (if $r_1 \ne 0$). The exact solutions for $n_1$ and $n_2$ satisfy $a_2c_1 = n_1(a_1c_2) + w_1 = a_1c_2(n_1 + r_1)$ and $a_1c_2 = n_2w_1 + w_2 = w_1(n_2 + r_2)$, where $0 \le w_1 < a_1c_2$ and $0 \le w_2 < w_1$. ■

**Example 2.7:** Our theorem is stronger than the corollary. For consider an example with

$v_1 = 200$ and $v_2 = 99$.

Hence $h_{I'} = \frac{c_1}{v_1 - v_2} = \frac{v_1}{v_1 - v_2} a_1 = \frac{200}{101} a_1 (\approx 2a_1)$,

while $0 \le \frac{v_2}{v_1} = \frac{99}{200} < \frac{1}{2}$, we have $q = 1$, thus $h_I = qa_1 = a_1$.

To achieve an optimal solution, Corollary 2.3 states that if $b \ge \frac{200}{101} a_1$, then we may pack at least one unit of item 1 whereas Theorem 2.2 states that we may do so with just $b \ge a_1$. Therefore, for arbitrary *b*, we may fill the knapsack by a suitable number of units of item 1 until the remaining capacity is less than $a_1$ (by our theorem) or less than $\frac{200}{101}a_1$ (by the corollary). Certainly it is more advantageous to use our theorem. ■

In fact, $h_I = qa_1$ is a *sharp bound* in the sense that no smaller bound can be found under the conditions specified in Theorem 2.2 as the following proposition shows.

**Proposition 2.8** *Given the assumptions of Theorems 2.2, there are examples such that if* $b = h_I - 1$*, then it is not optimal to pack item 1 into the knapsack.*

**Proof.** Consider three types of items with $v_1 = 2q+1$, $v_2 = 2q-1$ and $v_3 = \frac{1}{2}$, where $q$ is any positive integer. Their corresponding unit weights are $a_1 > 4q$, $a_2 = qa_1 - 1$ and $a_3 = 1$. It is easy to verify that $v_1 > v_2$ and $\frac{q-1}{q} \le \frac{v_2}{v_1} < \frac{q}{q+1}$. Hence, we have $h_I = qa_1$ as shown in Theorem 2.2. Now let $b = qa_1 - 1$. Apparently we can load one unit of item 2 into the knapsack, and the objective function value is

$$z^* = a_2 v_2 = (qa_1 - 1)(2q - 1) = 2a_1 q^2 - a_1 q - 2q + 1.$$

We claim that for any other loading, the resulting value is strictly less than $z^*$. Suppose we load one unit of item 1, then the remaining capacity is $b - a_1 = (qa_1 - 1) - a_1 < qa_1 - 1 = a_2$. Thus, no item 2 can be loaded (i.e., only item 1 and item 3 may be considered). We assume that altogether there are $k$ units of item 1 loaded (it must have $k < q$ since $b = qa_1 - 1$), and the rest is filled up by a suitable number of units of item 3. The resulting value is $z = ka_1 v_1 + (b - ka_1)v_3$ $= ka_1(2q+1) + (qa_1 - 1 - ka_1)\frac{1}{2} = 2ka_1 q + \frac{1}{2}ka_1 + \frac{1}{2}qa_1 - \frac{1}{2}$, and it reaches the maximum when $k = q - 1$. Therefore, we have

$$\begin{aligned} z &\le 2(q-1)a_1 q + \frac{1}{2}(q-1)a_1 + \frac{1}{2}qa_1 - \frac{1}{2} = 2a_1 q^2 - a_1 q - \frac{1}{2}(a_1 + 1) \\ &< 2a_1 q^2 - a_1 q - \frac{1}{2}(4q+1) < 2a_1 q^2 - a_1 q - 2q + 1 = z^* \end{aligned}$$

Hence, it is not optimal to pack item 1 into the knapsack. Thus, one can not improve on $h_I$. ■

Computational experiments were conducted as suggested by one of the referees, in order to ascertain just how effective our turnpike theorems are. Specifically, how often could they be applied and what was the average percentage reduction in the size of the resulting knapsack problem. Toward that end we used MATLAB 7.0 (The MathWorks Inc.) to randomly generate both $c_i$ and $a_i$ for $i = 1, \cdots, 500$ from a uniform distribution of integers from 1 to 1,000. Following Gabrel and Minoux (2002) the knapsack capacity, *b*, was randomly generated from a uniform distribution of integers where the range was $\left[ \frac{1}{3}\sum_{i=1}^{500} a_i,\ \frac{2}{3}\sum_{i=1}^{500} a_i \right]$. After running dozens 10,000 randomly generated instances, the typical results are as follows: we ran into 4 out of 10,000 cases where $v_1 = v_2$, i.e., the sufficient conditions of Theorem 2.2 were not met and among all cases, $b \geq h_I$. Therefore, our turnpike theorem could be applied in 99.96% of the cases. The average reduction in the right-hand-side, $\frac{b - h_I}{b}$ , $= 99.97\%$.

## 3. The Second Turnpike Theorem

Let *S* be a feasible solution to the knapsack problem. Assume $S = \{j_1, j_2, ..., j_m\}$ consists of *m* units of various items, each is denoted by $j_t$. The values of $j_t$ s don't have to be unique, i.e., the same type of item may have more than one occurrence in *S*. For instance, $S = \{1, 2, 1, 1, 3\}$ indicates {item 1, item 2, item 1, item1, item 3} . Now let $S_i = \{j_1, j_2, ..., j_i\}$ consists of first *i* units of *S*, where $i \leq m$ . Thus, $S_i$ denotes a partial solution with total weight $A_i = \sum_{t=1}^{i} a_{j_t}$ . Finally, let $r_i$ be the *residue* of $A_i (\text{mod } a_1)$, where $0 \leq r_i < a_1$ and $i \leq m$ ( Hardy and Wright, 1938).

First we have the obvious result.

**Proposition 3.1** *If* $b \equiv 0 (\text{mod } a_1)$*, i.e., b is divisible by* $a_1$*, then it is optimal to pack* $b/a_1$ *units of item 1 and no units of item 2,…, item n.* ■

**Remark 3.2:** If $S'$ is any partial solution of an optimal solution, and the weight of $S' \equiv 0 (\text{mod } a_1)$, then $S'$ consists entirely of units of item 1. ■

Next for the following two lemmas, assume the contrary that $S$ is an optimal solution with no item 1 included. Also, recall that $r_i \equiv A_i (\text{mod } a_1)$.

**Lemma 3.3** *Let* $\frac{c_1}{a_1} > \frac{c_2}{a_2} \geq \frac{c_3}{a_3} \geq \cdots$ (i.e., $v_1 > v_2 \geq v_3 \geq \cdots$). *If S is an optimal solution with no item 1 included, then* $r_i \neq 0$, *for* $i = 1, \cdots, m$.

**Proof**. If $r_i = 0$, i.e., $A_i \equiv 0 (\text{mod } a_1)$, then it implies that the total weight of $S_i$ is divisible by $a_1$. Thus $S_i$ can be replaced entirely by units of item 1, and the total value of $S$ will increase strictly (as $S_i$ was a partial solution of $S$). This violates the optimality of the original solution and completes the proof. ■

**Lemma 3.4** *Let* $\frac{c_1}{a_1} > \frac{c_2}{a_2} \geq \frac{c_3}{a_3} \geq \cdots$ (i.e., $v_1 > v_2 \geq v_3 \geq \cdots$). *If S is an optimal solution with no item 1 included, then* $r_i \neq r_k$ *for* $i \neq k$.

**Proof**. WLOG assume that $k > i$. If $r_i = r_k$, then $A_k - A_i \equiv 0 (\text{mod } a_1)$. This implies that the weight difference between partial solutions $S_k$ and $S_i$ is divisible by $a_1$. Thus, the resulting partial solution $S_k \backslash S_i$ can be replaced by a suitable number of units of item 1 (see Remark 3.2), and the total value of $S$ will increase strictly. This violates the optimality of the original solution and completes the proof. ■

**Theorem 3.5** *Let* $L = \max_{j \geq 2} \{a_j\}$. *Then there is a weight* $h_{II} = (a_1 - 1)L$ *such that if* $v_1 > v_2$, $b > h_{II}$, *then it is optimal to pack at least one unit of item 1 into the knapsack.*

**Proof.** We have $v_1 > v_2 \geq v_3 \geq \cdots$. Assume the contrary that there is an optimal solution $S$ without item 1 included. There are two cases and several sub-cases.

*Case 1:* Suppose that $a_1 - 1 > m$, where $m$ is the number of units in $S$. We have $A_m \leq mL$. Thus, the remaining weight capacity $b' = b - A_m > (a_1 - 1)L - mL \geq L$. Therefore, we may pack at least one more unit from item 2,..., item n into the knapsack. Thus, the original solution can not be optimal.

*Case 2:* Now consider $a_1 - 1 \leq m$, there are two sub-cases: $a_1 - 1 < m$ and $a_1 - 1 = m$.

*Sub-case 1:* $a_1 - 1 < m$. Recall that $r_i \equiv A_i (\text{mod } a_1),\ i = 1, \cdots, m.$ There are $m$ many $r_i$ s. Furthermore, it follows from two preceding lemmas that these $r_i$ s can not be zero and should all be distinct. Thus, we conclude that there exist $m$ distinct non-zero residues, where $m > a_1 - 1$. Since there are at most $a_1 - 1$ non-zero possibilities in a complete residue system $(\text{mod } a_1)$ (Serre, 1973), this is, therefore, an absurd conclusion.

*Sub-case 2:* $a_1 - 1 = m$. The remaining weight capacity $b' = b - A_m > (a_1 - 1)L - mL = 0.$ We thus have $b' \geq 1$ (since $b' > 0$ and $b'$ is an integer). Furthermore we have two possibilities: $b' \geq a_1$ and $1 \leq b' < a_1$.

*Sub-sub-case 1*: $b' \geq a_1$. Then we would be able to add at least one unit of item 1 and thus, the original solution can not be optimal.

*Sub-sub-case 2:* $1 \leq b' < a_1$. Then $0 < a_1 - b' \leq a_1 - 1$, this is the same as $1 \leq a_1 - b' < a_1$. Thus, $a_1 - b'$ is a residue of a complete residue system $(\text{mod } a_1)$. On the other hand, we have shown that there exist $m$ distinct non-zero residues, here $m = a_1 - 1$ which is the maximal non-zero possibilities in the complete residue system $(\text{mod } a_1)$. Therefore, $a_1 - b'$ must be one of these $r_i$ s. Thus, there exists a $k$ such that $r_k = a_1 - b'$, which means $r_k + b' \equiv 0 \ (\text{mod } a_1)$. This says that the sum of $S_k$'s weight and the remaining weight is divisible by $a_1$. So if we combine the partial solution $S_k$ and the remaining capacity $b'$ together, we could replace it entirely with units of item 1. Hence the total value of *S* will increase strictly and thus, the original solution can not be optimal. ■

The next corollary follows from Theorem 3.5.

**Corollary 3.6** *Let* $L' = \max_{j \geq 1} \{a_j\}$. *Then there is a weight* $h_{II'} = (a_1 - 1)L'$ *such that if* $v_1 > v_2$, $b > h_{II'}$, *then it is optimal to pack at least one unit of item 1 into the knapsack.*

**Proof.** Follows at once from Theorem 3.5 since $L' \geq L$. ■

The following corollary is a known result. The reader is referred to problem 25 on p.247 of Garfinkel and Nemhauser (1972).

**Corollary 3.7** *Let* $L' = \max_{j\geq 1}\{a_j\}$. *Then there is a weight* $h_{II''} = a_1(L'+1)$ *such that if* $v_1 > v_2$, $b > h_{II''}$, *then it is optimal to pack at least one unit of item 1 into the knapsack.*

**Proof.** Follows at once from Theorem 3.5 since $a_1 > a_1 - 1$ and $L'+1 > L' \geq L$. ■

**Example 3.8:** Our theorem is stronger than the corollaries. Consider an example with

$$a_1 = 10,\ a_2 = 2 \text{ and } a_3 = 1.\ \text{Also, } v_1 > v_2 \geq v_3.$$

Then $L = \max_{j\geq 2}\{a_j\} = 2$, and $L' = \max_{j\geq 1}\{a_j\} = 10$. Corollary 3.7 states that if $b > 110$ then it is optimal for us to pack at least one unit of item 1 into the knapsack. Corollary 3.6 states that we may do so if $b > 90$. However, Theorem 3.5 states that we should include item 1 to achieve optimal with just $b > 18$. Certainly it is advantageous to use our theorem.

Also notice that neither corollary is applicable when $b = 89$. Using Theorem 3.5, one immediately gets an obvious optimal solution by loading *8* units of item 1, *4* units of item 2 and *1* unit of item 3. ■

We may further show that $h_{II} = (a_1 - 1)L$ is a *sharp bound* in the sense that no smaller bound can be found under the conditions specified in Theorem 3.5.

**Proposition 3.9** *Given assumptions as in Theorems 3.5. There are examples such that if* $b = h_{II}$, *then it is not optimal to pack item 1 into the knapsack.*

**Proof.** Consider items with $a_2 = a_3 = \cdots = 1$, then $L = \max_{j\geq 2}\{a_j\} = 1$. If $b = h_{II}$, then we have $b = a_1 - 1$. Clearly item 1 can not be included in any optimal solution, and this completes our proof. ■

Again, we performed computational experiments in order to determine just how effective our second turnpike theorem was. MATLAB 7.0 was used to randomly generate integral values for $c_i$ and $a_i$, $i = 1, \cdots, 500$ from a uniform distribution of integers from 1 to 1,000 and, following Gabrel and Minoux (2002), the knapsack capacity, *b*, was chosen to be an integer randomly generated in the range $\left[\frac{1}{3}\sum_{i=1}^{500} a_i,\ \frac{2}{3}\sum_{i=1}^{500} a_i\right]$. After computing dozens 10,000 randomly generated instances, we ran into 5 out of 10,000 cases where $v_1 = v_2$, i.e., the sufficient conditions of Theorem 3.5 were not met and among all cases, $b > h_{II}$. Therefore, our theorem could be applied in 99.95% of the cases, and the average reduction $\frac{b - h_{II}}{b} = 98.10\%$.

## 4. The Third Turnpike Theorem

In Theorem 2.2, we examined the ratio of "bang-for-buck" ratios $\frac{v_2}{v_1}$. Exploring further along the path, it may also be advantageous to examine the ratio of weights $\frac{a_2}{a_1}$. When $\frac{a_2}{a_1} < 1$, we have the following new theorem.

**Theorem 4.1** *Assume that $v_1 > v_2$ and $a_1 > a_2$. Let the positive integer $q$ be uniquely determined by the following relation*

$$\frac{q-1}{q} \le \frac{v_2}{v_1} < \frac{q}{q+1} \quad .$$

*Let $k = \left\lfloor \frac{a_2}{a_1} q \right\rfloor + 1$. Then there is a weight $h_{III} = ka_1$ such that if $b \ge h_{III}$, it is optimal to pack at least one unit of item 1 into the knapsack.*

**Proof.** We have $v_1 > v_2 \ge v_3 \ge ...$. Assume the contrary that there is an optimal solution with no item 1 included and with objective function value *z*. Then, clearly $z \le bv_2$.

It follows from our condition that $b - ka_1 \ge 0$. Now let $y$ be the residue of $(b - ka_1)(\text{mod } a_2)$ Then $b = ka_1 + pa_2 + y$ for some non-negative integer *p*. Thus, we have

$$z \le bv_2 = (ka_1 + pa_2 + y)v_2$$

We claim that

$$bv_2 = (ka_1 + pa_2 + y)v_2 < ka_1v_1 + pa_2v_2 (= kc_1 + pc_2) \qquad (4.1)$$

In other words, another packing with *k* units of item 1 and *p* units of item 2 produces a higher objective function value than *z*. This would violate the optimality of the original solution and complete the proof.

Now let us prove our claim. Notice that inequality (4.1) can be simplified to

$$(ka_1 + y)v_2 < ka_1v_1 \qquad (4.2)$$

Since $y$ is a residue $\pmod{a_2}$, we know that $y < a_2$, so we may replace inequality (4.2) by the following stronger inequality

$$(ka_1 + a_2)v_2 < ka_1v_1 \qquad (4.3)$$

We shall prove the above inequality (4.3).

It follows from our assumption that

$$\frac{v_2}{v_1} < \frac{q}{q+1} \quad \Rightarrow \quad \frac{v_1}{v_2} > \frac{q+1}{q} = 1 + \frac{1}{q}$$

$$k = \left\lfloor \frac{a_2}{a_1} q \right\rfloor + 1 \quad \Rightarrow \quad \frac{a_2}{a_1} q < k \quad \Rightarrow \quad \frac{1}{q} > \frac{a_2}{ka_1}$$

The above two relations imply

$$\frac{v_1}{v_2} > 1 + \frac{a_2}{ka_1} \quad \Rightarrow \quad ka_1v_1 > (ka_1 + a_2)v_2$$

This establishes inequality (4.3). Hence we may deduce inequalities (4.2) and (4.1), and our proof is completed. ■

**Remark 4.2:** Recall that $q$ can be determined in constant time by Lemma 2.5 and Remark 2.6. Now that $k = \left\lfloor \frac{a_2}{a_1} q \right\rfloor + 1$, so $k$ can also be obtained in constant time. ■

**Remark 4.3:** Since $a_2 < a_1$, it is clear that $k \le q$. ■

**Example 4.4:** Theorem 4.1 is stronger than Theorem 2.2. whenever $a_1 > a_2$ and $k < q$. Consider an example with

$$v_1 = 100,\ v_2 = 91,\ v_i \ge v_{i+1} \ \text{ for } \ i = 2, \cdots, 199, \ \text{ and}$$

$$a_1 = 120,\ \ a_2 = 10,\ \ a_i : \text{positive integers, for } i = 3, \cdots, 200.$$

Using Lemma 2.5, $r = \frac{v_2}{v_1} = \frac{91}{100} = \frac{1}{1+\frac{9}{91}}$, thus $n_1 = 1$ and $r_1 = \frac{9}{91}$.

Since $n_1 = 1$, we go one more step with $r_1 = \frac{9}{91} = \frac{1}{10+\frac{1}{9}}$, thus $n_2 = 10$ and $r_2 = \frac{1}{9}$.

Therefore, $q = n_2 + 1 = 11$, we have $h_I = qa_1 = 1320$.

Let $k = \left\lfloor \frac{a_2}{a_1} q \right\rfloor + 1 = \left\lfloor \frac{10}{120} * 11 \right\rfloor + 1 = 1$, then $h_{III} = ka_1 = 120$.

Theorem 2.2 states that if $b \geq 1320$, then it is optimal to pack at least one unit of item 1 whereas Theorem 4.1 states that we may do so with just $b \geq 120$. Certainly it is advantageous to use Theorem 4.1 whenever $a_1 > a_2$ and $k < q$. Specifically, consider a knapsack problem with several hundred items and a carrying capacity of 100,000. Since 100,000 = 833 x 120 + 40, using Theorem 4.1 we may go ahead with *833* units of item 1. The remaining capacity is thus reduced to a merely 40! Now it is obvious that an optimal solution is loading *833* units of item 1 and *4* units of item 2. Here we get an optimal solution without invoking any algorithm. ■

The following proposition illustrates that $h_{III} = ka_1$ is a *sharp bound* under conditions given in Theorem 4.1.

**Proposition 4.5** *Given assumptions as in Theorems 4.1. There are examples such that if* $b = h_{III} - 1$*, then it is not optimal to pack item 1 into the knapsack.*

**Proof.** Consider items with $v_1 = 2q+1$, $v_2 = 2q-1$, where $q$ is any positive integer. We also assume that $a_1 > a_2 q$. Clearly, $v_1 > v_2$ and $a_1 > a_2$. It is easy to verify that $\frac{q-1}{q} \leq \frac{v_2}{v_1} < \frac{q}{q+1}$.

Hence, we have $h_{III} = ka_1 = a_1$ (here, $k = \left\lfloor \frac{a_2}{a_1} q \right\rfloor + 1 = 1$ since $\frac{a_2 q}{a_1} < 1$). If $b = h_{III} - 1 = a_1 - 1$, then item 1 can not be included in any optimal solution, and this completes our proof. ■

We also performed computational experiments on our third turnpike theorem. MATLAB 7.0 was used to randomly generate integral values of $c_i$ and $a_i$ for $i = 1, \cdots, 500$ from a uniform distribution of integers from 1 to 1,000 and, following Gabrel and Minoux (2002), the knapsack capacity, *b*, was chosen to be an integer randomly generated in the range $\left[\frac{1}{3}\sum_{i=1}^{500} a_i,\ \frac{2}{3}\sum_{i=1}^{500} a_i\right]$. After computing dozens 10,000 randomly generated instances, we ran into 8,667 out of 10,000 cases where either $v_1 = v_2$ or $a_1 \le a_2$, i.e., the sufficient conditions of Theorem 4.1 were not met but among all cases, $b > h_{III}$. Therefore, although our theorem could only be applied in 13.33% of the cases, we observed that $\frac{h_{III}}{h_I} = 65.89\%$ in average. This implies that if the sufficient conditions of Theorem 4.1 are met, then using Theorem 4.1 may achieve further reduction than using Theorem 2.2. Out of those cases where the third turnpike theorem could be applied, the average reduction by Theorem 4.1 was $\frac{b - h_{III}}{b} = 99.93\%$ while the average reduction by Theorem 2.2 was $\frac{b - h_I}{b} = 99.88\%$.

## 5. Discussion

This paper developed sharp bounds under various conditions on turnpike theorems for the unbounded knapsack problem. Turnpike theorems specify when it is optimal to load at least one unit of the *best* item (i.e., the one with the highest "bang-for-buck" ratio) and their successive application can drastically reduce the size of the knapsack problems. We also show that all three of our turnpike theorems specify *sharp bounds* in the sense that no smaller bounds can be found under the assumed conditions and prove that two of our bounds can be obtained in constant time. Finally, computational results on randomly generated problems clearly demonstrated the effectiveness of our turnpike theorems. Indeed, the three turnpike theorems yielded average reductions in the knapsack capacities $\frac{b-h}{b}$ of 99.97%, 98.10% and 99.93%, respectively.

There are at least three extensions that we intend to pursue. The first involves considering both the first as well as the second best items. The resulting turnpike theorems could then be based upon results from the famous Frobenius problem (Ramírez-Alfonsín, 2005). The second extension was suggested by one of the referees and involves conducting a comprehensive computational study to see if the turnpike theorems actually reduce the total computing time when compared with solving the knapsack problems without them. The third topic involves extending the turnpike theorems to the multidimensional knapsack problem.

**Acknowledgments:** It is indeed a pleasure to thank Kwei Tang and the two anonymous referees for their numerous insightful comments, corrections and suggestions that significantly improved the paper.

**References**

-Gabrel V, Minoux M. A Scheme for Exact Separation of Extended Cover Inequalities and Application to Multidimensional Knapsack Problems. Operations Research Letters 2002; 30; 252-264.

-Garey M, Johnson D. Computers and Intractability: A Guide to the Theory of NP Completeness. W. H. Freeman and Co.: New York; 1990.

-Garfinkel R, Nemhauser G. Integer Programming. John Wiley & Sons: New York; 1972.

-Gilmore P, Gomory R. The Theory and Computation of Knapsack Functions. Operations Research 1966; 14; 1045-1074.

-Kellerer H, Pferschy U, Pisinger D. Knapsack Problems. Springer: Berlin; 2004.

-Hardy G, Wright E. An Introduction to the Theory of Numbers. Oxford University Press, New York; 1938.

-Hu T. Integer Programming and Network Flows. Addison-Wesley: Reading, MA; 1969.

-Martello S, Toth P. Knapsack Problems: Algorithms and Computer Implementation. John Wiley & Sons: New York; 1990.

-Morin T, Marsten R. An Algorithm for Nonlinear Knapsack Problems. Management Science 1976; 22; 1147-1158.

-Nemhauser G. Need and Potential for Real-Time Mixed-Integer Programming. Operations Research Management Science Today 2007; 34; 21-22.

-Ramírez-Alfonsín J. The Diophantine Frobenius Problem. Oxford University Press; 2005.

-Serre J-P. A Course in Arithmetic. Springer: Berlin; 1973.

-Shapiro J. Turnpike Theorem for Integer Programming Problem. Operations Research 1970; 18; 432-440.

-Zukerman M, Jia L, Neame T, Woeginger G. A Polynomial Solvable Special Case of the Unbounded Knapsack Problem. Operations Research Letters 2001; 29; 13-16.