Cowles Foundation Paper 648

Econometrica, Vol. 54, No. 3 (May, 1986), pp. 507-532

NEIGHBORHOOD SYSTEMS FOR PRODUCTION SETS WITH INDIVISIBILITIES[1]

BY HERBERT E. SCARF

A production set with indivisibilities is described by an activity analysis matrix with activity levels which can assume arbitrary integral values. A neighborhood system is an association with each integral vector of activity levels of a finite set of neighboring vectors. The neighborhood relation is assumed to be symmetric and translation invariant. Each such neighborhood system can be used to define a local maximum for the associated integer programs obtained by selecting a single commodity whose level is to be maximized subject to specified factor endowments of the remaining commodities. It is shown that each technology matrix (subject to mild regularity assumptions) has a unique, minimal neighborhood system for which a local maximum is global. The complexity of such minimal neighborhood systems is examined for several examples.

1 INTRODUCTION

I HAVE WRITTEN in this journal before on the subject of production sets with indivisibilities [**13, 14**]. In the present paper I would like to present some simplifications of the previous arguments, which I hope will make them more accessible to the general reader, and to describe some recent work on this topic by other authors and by myself.

Our inability to incorporate efficiencies of large-scale production and other forms of nonconvexity in a conceptual framework which possesses the generality of the Walrasian model, has long seemed to me to be a major deficiency of economic theory. When I first became aware of the game theoretic concept of the core it seemed to be ideally suited to the study of increasing returns to scale in production. If the production possibility set is a convex cone a competitive equilibrium will typically exist and be in the core. Moreover, when the number of consumers is large (and when all coalitions have access to the same production set), a feasible production and distribution plan, which is undominated by any coalition, will be close to a competitive equilibrium. When the production possibility set displays increasing returns to scale, the grand coalition, which can employ larger stocks of factors than those of a small coalition, is proportionately stronger than in the case of constant returns to scale. Outcomes proposed by the collection of all consumers would seem to have a lesser chance of being dominated by smaller coalitions, and as a consequence the existence of a nonempty core seems likely in the presence of increasing returns to scale.

This intuition is surely correct in the most elementary example in which production involves a single input, say labor, producing a single generalized output according to the production function $y=f(l)$, with $f(l)/l$ increasing in l.

[1] This paper, which is based on the Presidential Address of the Econometric Society delivered at Northwestern University and at Pisa in 1983, was supported by a grant from the National Science Foundation. I am very much indebted to Andrew Caplin, Philip White, and Ludo Van der Heyden for many stimulating conversations on the subject of this paper, and to one of the referees of the paper for his insightful comments

If the ith consumer has an initial endowment of labor given by l_i, and a utility function which only values output, then the allocation

$$y_i = \frac{f(l)}{l} \cdot l_i, \qquad \text{with} \quad l = \sum_i^n l_j,$$

is collectively feasible and undominated by any coalition.

Unfortunately the existence of a nonempty core is not guaranteed under more general conditions. Let Y be a production possibility set in m dimensional Euclidean space, with inputs represented by negative entries and outputs by positive entries of the typical production plan $y \in Y$. We shall require that Y be a closed set, contain the origin, and satisfy the customary free disposal assumption. Moreover, let Y be additive, in the sense that $y^1, y^2 \in Y$ implies $y^1 + y^2 \in Y$; and possess the property that $\{y \in Y | y \geq -\omega\}$ be bounded for any nonnegative vector ω; all very mild assumptions.

The ith consumer (for $i = 1, \ldots, n$) will have the utility function $u_i(x)$ and vector of initial holdings w^i. Let the productive knowledge available to an arbitrary coalition $S \subseteq N = \{1, 2, \ldots, n\}$ be described by the same set Y. The coalition may therefore achieve, by its own efforts, any utility vector u_i, for $i \in S$, with $u_i \leq u_i(x^i)$, for some $\{x^i\}$ and $y \in Y$ satisfying

$$\sum_{i \in S} x^i = y + \sum_{i \in S} w^i.$$

As usual a utility vector $u_1, \ldots, u_n$ is in the core if it is feasible for the grand coalition N, and if no coalition S can achieve an alternative utility vector which is strictly preferable for all of its members.

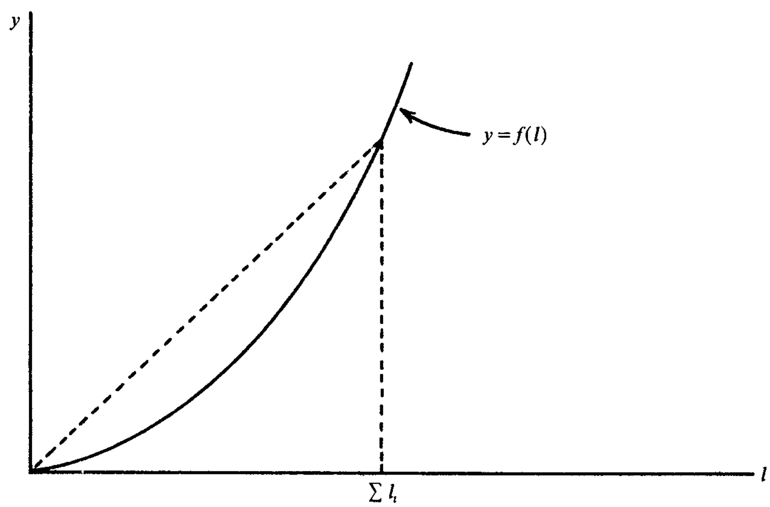

FIGURE 1.

The following rather surprising conclusion may be shown to be correct. If Y is *not* a convex cone, then it can be incorporated in an economy described by a given number of consumers n, a distribution of strictly positive initial assets $\{\omega^i\}_{i=1,\ldots,n}$, and an assignment of continuous, concave, and monotone utility functions u_i $(i=1,\ldots,n)$, each defined on the nonnegative orthant of R^m, for which *the core of the resulting economy is empty* (Scarf [**12**]).

I continue to find this result quite disturbing ever since I first came upon it over twenty years ago. It is true that the counterexamples which yield an empty core are somewhat unrealistic in that they involve utility functions which depend on all of the goods and services in the economy, and initial distributions of assets which are positive for all commodities. Certain restrictions on the generality of the economy which do yield a nonempty core have been found by Sharkey [**16**], Ichiishi [**7**], Quinzii [**11**], Ichiishi and Quinzii [**8**], and by myself [**12**]. But the conditions are by and large artificial and they do not yield a fully satisfactory cooperative solution to the general Walrasian model in the presence of increasing returns to scale.

Economies of scale are the major economic argument for the existence of large productive units, which by virtue of their size have an ability to influence the prices at which outputs are sold. The strategic selection of prices and outputs by economic agents may be analyzed by noncooperative game theory, an approach which has been taken by many authors during the last decade in reexamining theories of imperfect competition originally proposed some fifty years ago. But in spite of this substantial body of work (an excellent survey may be found in Hart [**5**]), a theory of imperfect competition with the range and generality of the Walrasian model is not yet available, even in the case in which economies of scale are absent.

The work of Brown and Heal [**1, 2**] is also concerned with increasing returns to scale, but their analysis makes no use of game theoretic considerations. Their contribution is to provide existence theorems for marginal cost pricing and average cost pricing equilibria—less ambitious, to be sure, than a game theoretic approach, but perhaps more satisfactory at this moment precisely because of its simplicity.

The general equilibrium model and its game theoretic counterparts place a heavy reliance on an *a priori* designation of consumer preferences, which are independent of the productive side of the economy. The efficiencies of large scale production, however, may be realized largely in the provision of new goods and services for which consumers have no measure of utility prior to their actual availability. To take only one of many obvious examples, the degree to which a typical consumer will substitute between the acquisition of computer services and other commodities has been drastically changed by the remarkable technical innovations in the computer industry during the post war period. The emptiness of the core in an economic model based on stable preferences for computer services may not be fully relevant in examining the consequences of increasing returns to scale in this industry.

Theorists will have various motivations for attempting to extend or modify the neo-classical paradigm so as to include nonconvex production possibility sets.

To my way of thinking, the most significant feature of the Walrasian model is its ability to evaluate the consequences—for a host of variables—of changes in economic policy or in the environment in which the economy finds itself. At the present level of development of economic theory, we simply do not have a corresponding ability to engage in this form of comparative analysis in the presence of increasing returns to scale in production.

When the economy is in competitive equilibrium with the production side described by a convex cone, the question, for example, of whether a newly discovered activity may be used to provide an improvement in the utility of each consumer, has a remarkably simple answer. Under mild technical assumptions a necessary and sufficient condition that such an improvement be possible is that the new activity make a positive profit at the old equilibrium prices. This conclusion has an important counterpart in the simplex method for the solution of linear programming problems, in which prices are used to test whether a feasible solution to a system of linear inequalities is indeed optimal. Given a basic feasible solution, a vector of prices is determined which yields a profit of zero for each activity in use. A necessary and sufficient condition that this solution be optimal is that the remaining activities make a profit which is less than or equal to zero.

If, in the competitive model, a new activity is discovered which can only be used at *integral levels*, its profitability at the equilibrium prices is no longer sufficient to guarantee higher utility levels for all consumers. The analogous conclusion for integer programming problems, in which all activity levels are required to be *integral*, is that there need not exist prices which permit us to conclude that a feasible solution is indeed optimal.

Consider the following example of an integer program:

$$\max -4h_1 - 3h_2$$

$$2h_1 + h_2 \geqslant 3,$$

$$h_1, \quad h_2 \geqslant 0.$$

The constraint set is given in Figure 2, with the objective function denoted by a dashed line. If the integrality condition were relaxed the solution of the corresponding linear program would occur at the point $(3/2, 0)$. If the prices associated with the objective function and the inequality are 1 and 2, respectively, then the profit of the activity in use is zero, and the profit of the second activity is -1. The integer programming solution is at the point $(1, 1)$, and there are no prices which yield a zero profit for both activities.

Production sets with indivisibilities, and the related integer programming problems that arise when an endowment of factors is specified, represent the most extreme form of nonconvexities in production. Such production sets capture some of the major features which give rise to the efficiencies of large scale production: set up costs which must be laid out prior to the use of a particular activity, and the construction of large, indivisible pieces of machinery whose employment is economically merited only for high levels of output. My own research has been to concentrate on integer programming problems, and to replace

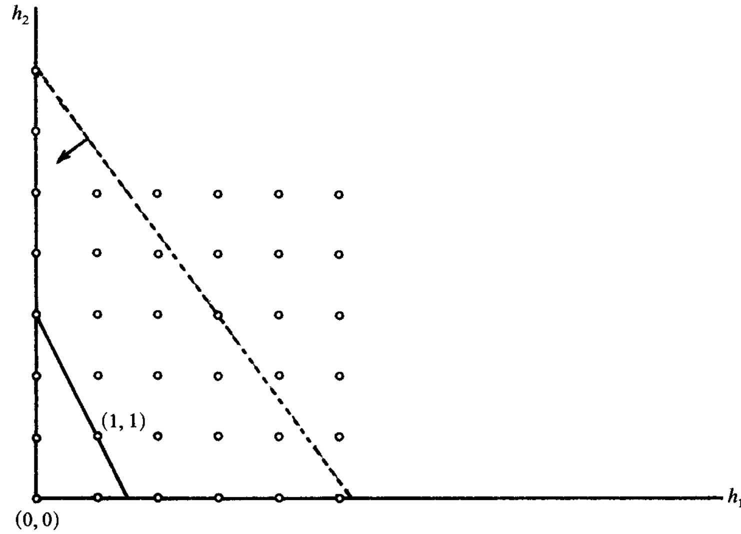

FIGURE 2.

the pricing argument for optimality by a search through neighboring lattice points which are, in a specific sense, close to a given feasible activity vector. Instead of prices we shall be concerned with quantity information in verifying that a feasible plan is optimal.

Let the general integer program have the form

$$\begin{aligned} \max\ & a_{01}h_1+\cdots+a_{0n}h_n \\ & a_{11}h_1+\cdots+a_{1n}h_n \geqslant b_1, \\ & \quad\vdots \\ & a_{m1}h_1+\cdots+a_{mn}h_n \geqslant b_m, \end{aligned}$$

and $h=(h_1,\ldots,h_n)$ integral. Nonnegativity inequalities, to the extent that they appear, will be assumed to be incorporated in the constraints of the problem. By a *neighborhood system* we mean an assignment to each integral vector $h=(h_1,h_2,\ldots,h_n)$ of a *finite* set of integral vectors $N(h)$, satisfying the following two conditions:

CONDITION 1: $N(h)=h+N(0)$.

CONDITION 2: *If* $k\in N(h)$, *then* $h\in N(k)$.

The first condition states that the set of neighbors associated with different lattice points are simply translates of each other, and the second condition that the neighborhood relation is symmetric.

A local maximum, with respect to a given neighborhood system, is a feasible integral vector, all of whose neighbors are either infeasible, or yield an inferior value of the objective function. A local maximum for a particular neighborhood system need not be a global maximum to the integer program. We will shortly show, however, under mild conditions on the technology matrix A, *that a unique, smallest neighborhood system will always exist, for which a local maximum is indeed global.* The minimal neighborhood system will depend on the technology matrix alone and not on the factor endowment.

The demonstration that such a minimal neighborhood system exists is quite elementary and is best illustrated by a specific example, such as that of Figure 2, in which the technology matrix is given by

$$\begin{bmatrix} -4 & -3 \\ 1 & 0 \\ 0 & 1 \\ 2 & 1 \end{bmatrix}.$$

We shall show that the neighbors of the origin for this example are given by the six lattice points of Figure 3; the neighbors of other points are obtained by translation.

In order to verify, for example, that the point $(1, -2)$ is a neighbor of the origin in the minimal neighborhood system, it is sufficient to show that there is a specification of $b = (b_0, b_1, b_2, b_3)'$ so that the only integral vectors satisfying the inequalities

$$\begin{aligned} -4h_1 - 3h_2 &\geq b_0, \\ h_1 \qquad &\geq b_1, \\ h_2 &\geq b_2, \\ 2h_1 + h_2 &\geq b_3, \end{aligned}$$

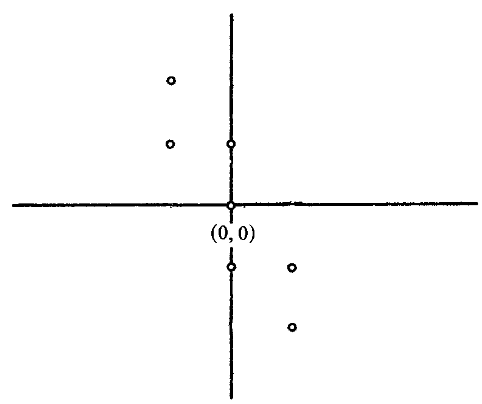

FIGURE 3.

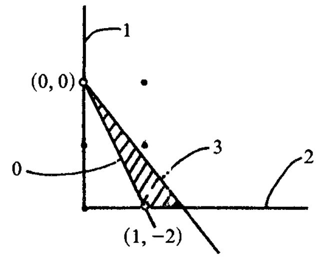

FIGURE 4.

are the points (0, 0) and (1, −2) themselves. For if this were so, the constraints of the integer program

$$\begin{aligned} &\max -4h_1 - 3h_2 \\ &h_1 \geqslant b_1, \\ &h_2 \geqslant b_2, \\ &2h_1 + h_2 \geqslant b_3, \end{aligned}$$

would be satisfied by (0, 0) and (1, −2) (with the latter point giving the higher value of the objective function). No other lattice point would both satisfy the constraints and yield a value of the objective function greater than 0. If (1, −2) were not in the neighborhood system, (0, 0) would, therefore, be judged incorrectly to be optimal.

To see whether there is a specification of the inequalities for which (0, 0) and (1, −2) are the only feasible lattice points, we ask whether the smallest convex body—obtained by parallel movements of the inequalities—which contains these two lattice points, will contain any other lattice points. This smallest convex body is the set in Figure 4. The absence of other lattice points from this body verifies that (1, −2) is a neighbor of the origin in the minimal neighborhood system.

Figure 5 displays the minimal convex bodies containing (0, 0) and (−1, 1), (0, −1), respectively. Since both of these bodies are free of other lattice points, (−1, 1) and (0, −1) must each be neighbors of the origin in the minimal neighborhood system. In Figure 6 we see that (2, −3) is not a neighbor of the origin.

FIGURE 5.

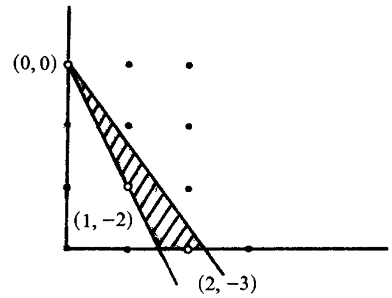

FIGURE 6.

Notice first that $(2, -3)$ yields a higher value of the objective function $-4h_1 - 3h_2$ than does $(0, 0)$. If the right-hand side of the integer program is specified so that $(0, 0)$ is a feasible solution, then the feasibility of $(2, -3)$ will certainly imply that $(0, 0)$ is not an optimal solution. But the feasibility of both $(0, 0)$ and $(2, -3)$ implies the feasibility of a lattice point we have already verified to be a neighbor of the origin, $(1, -2)$, and this lattice point also yields a higher value of the objective than does $(0, 0)$. Therefore $(2, -3)$ need not be examined in order to decide on the optimality of $(0, 0)$. This argument can be used to show that no lattice points other than those appearing in Figure 3 are neighbors of the origin.

For a general technology matrix A, two lattice points k and k' will be neighbors of each other if the smallest convex body—obtained by parallel movements of the linear inequalities—which contains k and k' contains no other lattice points. Conditions on A which are somewhat stronger than necessary to guarantee the existence of a unique, minimal, finite neighborhood system are the following:

ASSUMPTION 1.1: *For each* $b = (b_0, b_1, \ldots, b_m)'$ *the set of integral vectors* h *for which* $Ah \geqslant b$ *is finite.*

NON-DEGENERACY ASSUMPTION 1.2: *For each* $i = 0, 1, \ldots, m$, *the origin is the only integral vector satisfying* $\sum_{j=1}^{n} a_{ij} h_j = 0$.

Assumption 1.2, which is not valid in our previous example, can be obtained by perturbing the coefficients of the matrix A, or by a lexicographic tie-breaking rule to order the components of $y = Ah$ as h ranges over the integers in R^n.

Given these two assumptions the proof of existence of the unique, minimal neighborhood system is quite straightforward. For each integral vector k, different from 0, define

$$S_k = \left\{ x \in R^n \;\middle|\; \sum_{j=1}^{n} a_{ij} x_j \geqslant \min\left(0, \sum_{j=1}^{n} a_{ij} k_j\right), \text{for } i = 0, \ldots, m \right\}.$$

S_k is the smallest convex set obtained by parallel movements of the inequalities which contains both 0 and k. Then define a partial ordering of the nonzero lattice points by $k \preccurlyeq k'$ if $S_k \subseteq S_{k'}$. From Assumption 1.1, each lattice point is preceded

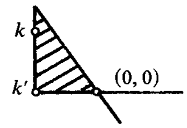

FIGURE 7.

by a finite number of lattice points in this ordering, and from Assumption 1.2 $k \leqslant k'$ and $k' \leqslant k$ imply $k = k'$. The lattice points which are minimal with respect to this ordering form the neighbors of the origin in the minimal neighborhood system for which a local maximum is global.[2]

If the entries in the technology matrix are perturbed slightly the minimal neighborhood system will be unchanged unless Assumption 1.2 is violated. Figure 7 illustrates such a case (based on a matrix with three rows and two columns) in which *either* of the lattice points k or k' can be included in a minimal neighborhood of the origin. This arbitrariness can be removed by the lexicographic tie-breaking rule. Finiteness of the minimal neighborhood system is also guaranteed by the lexicographic rule [**13**], or, as Philip White has shown in his thesis [**17**], by the assumption that all $n \times n$ minors of A have rank n, in conjunction with Assumption 1.2.

It is important to remark that considerations of degeneracy can be avoided completely if we are willing to settle for a neighborhood system which is not necessarily minimal. When degeneracy is present the ordering of lattice points given above may not be simple: there may be pairs k and k' for which $k \leqslant k'$ and $k' \leqslant k$. A lattice point k is *minimal* with respect to this ordering if any $k' \leqslant k$ must also satisfy $k \leqslant k'$. It is then easy to see that, as long as Assumption 1.1 is satisfied, the set of lattice points which are minimal with respect to this ordering will form a neighborhood of the origin in the following sense: Let the right-hand side $(b_1, b_2, \ldots, b_m)'$ be selected so that 0 is feasible. If there is any lattice point, different from 0, which is also feasible and the value of whose objective function is nonnegative, then at least one of the lattice points in this neighborhood of the origin will have the same property.

Precise knowledge of the neighborhood system associated with a matrix A would yield an obvious algorithm for integer programs based on a specification of the factor endowment b. When a feasible lattice point is given, its neighbors are examined to see whether one of them satisfies the constraints and yields a higher value of the objective function. If there is such a neighbor we move to it and iterate; if not the current feasible lattice point is optimal.

In the next section I will present some specific examples of neighborhood systems associated with certain classes of integer programming problems. As we shall see, the cardinality of the set of neighbors of the origin may be quite high when compared to the magnitude of the coefficients in the matrix A. This cardinality is the most immediate measure of the complexity of an algorithm

[2] This argument is due to Andrew Caplin and Philip White.

based on iterated neighborhood searches. It specifies the number of alternative lattice points that are required to be examined in order to verify the optimality of a particular feasible solution.

On the other hand, the set of neighbors of the origin may have sufficient structure so that each of its members need not be considered independently. For example, the determination of whether a particular linear inequality is satisfied by some member of a finite set of lattice points may be simplified considerably if the set can be decomposed into a small number of linear segments. Each of our examples will display structural properties of this type permitting the development of an algorithm which proceeds much more rapidly than might be imagined by simply counting the number of neighbors of the origin. If the approach taken in this paper is to lead to efficient algorithms for integer programming it will be necessary to extend these structural simplifications from our examples to the general problem.

A possible conjecture concerning the structure of minimal neighborhood systems may be phrased using the language of complexity theory (an excellent introduction to this very significant topic may be found in Garey and Johnson [**4**]). Assume, to be specific, that the entries in the technology matrix A are themselves integral. A measure of the complexity of A is given by the number of binary bits required to store all of the entries of A (including their sign) in a computer; a quantity which may be shown to be

$$S=\sum_m \sum_n (1+\lceil \log_2 (1+|a_{ij}|)\rceil),$$

with $\lceil x \rceil$ defined as the least integer greater than or equal to x. An algorithm for solving integer programs based on A is then said to be *polynomial* if the number of iterations of its basic steps—or more precisely its running time—is bounded from above by a polynomial function of S (to be accurate S should be increased by the number of bits required to store the vector of factor endowments specifying the particular program). If the algorithm is not polynomial, it may lead to "exponential" searches which rapidly become infeasible to carry out.

Using techniques of the Geometry of Numbers, Lenstra [**10**] has demonstrated the remarkable result that there is a polynomial algorithm for all integer programs in which the number of variables is fixed in advance. It is easy to show, by means of examples [**14**], that the cardinality of the set of neighbors of the origin is definitely *not* polynomial in the data, but Lenstra's result suggests the very interesting and difficult conjecture that the minimal neighborhood system has sufficient structure so that it can be described in a polynomial fashion.

I feel that such a result, if correct, might be capable of interpretation in terms of the internal organization of a large firm whose production possibility set involves significant indivisibilities. Such a firm, faced with the necessity of revising its decisions because of changing factor endowments, may construct a mode of organization—an algorithm, if you will—in which the examination and testing of alternative plans can be carried out in the most efficient way. But at the present time, I am far from being able to provide a convincing argument which relates the structure of neighborhood systems to the administrative arrangements that might be undertaken by a large industrial enterprise.

2 EXAMPLES OF NEIGHBORHOOD SYSTEMS

In the present section I shall present several examples of technology matrices whose minimal neighborhood systems can easily be described. In each of these examples the neighborhood structure can be exploited so as to yield rapid algorithms for the corresponding integer programming problems.

A. *A Special Leontief Matrix*

Let the technology matrix consist of $m+1$ rows and n columns. Row zero will have all of its entries negative, and the remaining m rows will be partitioned into n subsets $I_1, I_2, \ldots, I_n$ with the property that $a_{ij} > 0$ if $i \in I_j$ and $a_{ij} < 0$ otherwise. We make the special assumption that $\sum_j a_{ij} > 0$ for rows $1, \ldots, m$. This last condition is restrictive; because of the assumption of integral activity levels the multiplication of the columns of A by arbitrary scalars is not permitted. As we shall see, the neighbors of the origin consist of all nonzero vectors $h = (h_1, h_2, \ldots, h_n)$ whose entries are zero or ones, or the negatives of these vectors. The argument begins by establishing the following well known lemma.

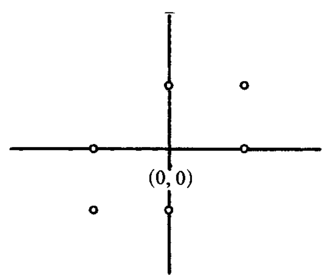

FIGURE 8.

LEMMA 2.1: *Let the vector $b = (b_1, b_2, \ldots, b_m)'$ be selected so that the $m+1$ inequalities $Ah \geq b$ are satisfied by the two vectors k and k'. Then the same inequalities are satisfied by $k'' = (\min(k_1, k_1'), \ldots, \min(k_n, k_n'))$.*

Notice first of all that $\sum_j a_{ij}k_j \geq b_i$, for $i \in I_l$, implies that $a_{il}k_l \geq b_i - \sum_{j \neq l} a_{ij}k_j \geq b_i - \sum_{j\neq l} a_{ij} \min(k_j, k_j')$, since $a_{ij} < 0$ for $j \neq l$. Replacing k by k' yields a similar inequality so that $a_{il} \min(k_l, k_l') \geq b_i - \sum_{j \neq l} a_{ij} \min(k_j, k_j')$. This verifies that inequalities $1, 2, \ldots, m$ are satisfied by k''. The zeroth inequality is certainly satisfied by k'', establishing Lemma 2.1.

This conclusion may be described geometrically by saying that the smallest convex set obtained by parallel movements of the inequalities, which contains k and k', will also contain k''. But that immediately implies that the two lattice points k and k' will not be neighbors of each other unless one of them is, in fact, equal to k'', i.e. unless $k \leq k'$ or $k' \leq k$. It follows that a neighbor of the origin must have all of its coordinates greater than or equal to zero, or less than or equal to zero.

Now let $k = (k_1, \ldots, k_n) \geq 0$ be a neighbor of the origin; we shall demonstrate that $k_j \leq 1$ for all j, as a consequence of the assumption $\sum_j a_{ij} > 0$ for rows $1, \ldots, m$. By this assumption the vector $e = (1, 1, \ldots, 1)$ will satisfy $\sum_j a_{ij} e_j > 0 \geq \min(0, \sum_j a_{ij} k_j)$ for $i = 1, \ldots, m$, and from the arguments of Lemma 2.1, it follows that $\sum_j a_{ij} \min(k_j, e_j) \geq \min(0, \sum_j a_{ij} k_j)$ for these values of i. Since $\sum_j a_{0j} \min(k_j, e_j) \geq \sum_j a_{0j} k_j \geq \min(0, \sum_j a_{0j} k_j)$ it follows that $\min(k, e)$ must be in the smallest convex set (obtained by translating the $m+1$ inequalities) which contains 0 and k. This eliminates the possibility that a neighbor of the origin has some coordinates greater than one.

The possible nonnegative neighbors of the origin are therefore vectors $e^S = (e_1^S, e_2^S, \ldots, e_n^S)$ with $e_j^S = 1$ if $j \in S$ and $e_j^S = 0$ if $j \notin S$, with S an arbitrary subset of $(1, 2, \ldots, n)$. To show that such a vector is indeed a neighbor of the origin it is sufficient to show that no other vector $e^{S'}$ is in the smallest convex body, obtained by translations of the inequalities, which contains 0 and e^S. This is equivalent to showing that for any other subset S' the inequalities

$$\sum_j a_{ij} e_j^{S'} \geq \min\left[0, \sum_j a_{ij} e_j^S\right], \quad \text{for} \quad i = 0, \ldots, m$$

are not all satisfied. But $\sum_j a_{ij} e_j^S$ is positive for all $i \in I_l$ if $l \in S$ and negative otherwise. The above inequalities for $i = 1, \ldots, m$ therefore imply that $S' \supset S$ and this is inconsistent with $\sum_j a_{0j} e_j^{S'} \geq \sum_j a_{0j} e_j^S$. We have therefore demonstrated the following theorem.

THEOREM 2.2: *The neighbors of the origin consist precisely of those vectors whose coordinates are* 0's *and* 1's *and their negatives.*

Consider the integer program

$$\max \sum_j a_{0j} h_j \qquad \text{subject to}$$

$$\sum_j a_{ij} h_j \geq b_i \qquad (i = 1, 2, \ldots, m),$$

with $b_1, b_2, \ldots, b_m$ a specific right-hand side. A direct application of neighborhood searches may lead to an extremely slow algorithm if the initial feasible solution is far from the optimum. As we shall see, there is an alternative algorithm which reaches the optimum solution in a number of steps which is bounded above by some function of n, independently of the entries in the matrix A. Let us begin by solving the associated *linear* program

$$\max \sum_j a_{0j} x_j \quad \text{subject to}$$

$$\sum_j a_{ij} x_j \geq b_i, \quad \text{for} \quad i = 1, 2, \ldots, m,$$

with optimal solution $x = x^*$. For simplicity of notation we shall assume that $0 < x_j^* \leq 1$ for $j = 1, \dots, n$; this can be achieved by a suitable translation of the right-hand side.

Our first observation is that the optimal solution to the integer program must then satisfy $h_j \leq \max_{l \neq j} h_l + 1$. Suppose, to the contrary, that $(h_1, h_2, \dots, h_n)$ is the optimal solution and that $h_1 \geq h_j + 2$ for $j = 2, \dots, n$. But then the neighbor $(h_1 - 1, h_2, \dots, h_n)$ will certainly satisfy all of the inequalities in rows $I_2, I_3, \dots, I_n$. In order to show that the inequalities in I_1 are also satisfied by $(h_1 - 1, h_2, \dots, h_n)$ we begin by observing that the vector $(x_1^* + h_1 - 2, x_2^* + h_1 - 2, \dots, x_n^* + h_1 - 2)$ is feasible, since $(x_1^*, \dots, x_n^*)$ is feasible, $h_1 \geq 2$, and $\sum_j a_{ij} > 0$. But $(h_1, h_2, \dots, h_n)$ is also feasible, and we may therefore use Lemma 2.1 to conclude that $(x_1^* + h_1 - 2, h_2, \dots, h_n) = \min((h_1, h_2, \dots, h_n), (x_1^* + h_1 - 2, x_2^* + h_1 - 2, \dots, x_n^* + h_1 - 2))$ is feasible. Since $x_1^* + h_1 - 2 \leq h_1 - 1$, we see immediately that the inequalities in I_1 are also satisfied by $(h_1 - 1, h_2, \dots, h_n)$ and h is therefore not optimal.

It follows from this argument that the optimal solution must have a pair of coordinates, say h_j and h_k, whose difference is either +1, 0, or −1. If this pair were known the two variables h_j and h_k could be replaced by a single variable t with one of the following conditions holding:

$$h_j = t, \qquad h_k = t;$$

$$h_j = t, \qquad h_k = t + 1;$$

$$h_j = t + 1, \qquad h_k = t.$$

The original integer program would then be reduced to three integer programs involving $n - 1$ rather than n variables. Moreover the new technology matrix would be obtained by adding together columns j and k, and would therefore satisfy precisely the same conditions as the original problem. In order to initiate the next stage of the algorithm, each of the three new *linear* programs would be solved.

Of course, the precise pair j and k is not known in advance. But there are $n(n-1)/2$ such pairs; when all of these are considered at each step, we obtain an algorithm whose running time is bounded above by a function of n, independently of the remaining data.

The neighborhood system may be considerably more complex when the crucial assumption $\sum_j a_{ij} > 0$ for $j = 1, \dots, m$ is relaxed. In this more general case there will be a smallest positive integer vector $h_1^*, \dots, h_n^*$ satisfying $\sum a_{ij} h_j^* > 0$, and the nonnegative neighbors of the origin will be a subset of $\{h \mid 0 \leq h \leq h^*\}$. But substantially more analysis seems necessary in order to develop an algorithm which is polynomial in the data.

B. *The Transportation Problem*

As our second example we shall consider the classical Hitchcock-Koopmans transportation problem. We are given m sources of a single homogeneous commodity and n destinations. The unit cost of shipping from source i to destination

j is a positive number c_{ij}. The available supply at source i is denoted by s_i, for $i = 1, \ldots, m$, and the demand at destination j by d_j, for $j = 1, \ldots, n$; it is typically assumed that $\sum_i s_i \geq \sum_j d_j$.

Let $x_{ij} \geq 0$ be the amount shipped from source i to destination j. Such a shipping plan will be feasible—consistent with the available supplies and the demands—if $\sum_j x_{ij} \leq s_i$ and $\sum_i x_{ij} \geq d_j$. The goal of the transportation problem is to select a feasible shipping plan which minimizes the total cost $\sum_i \sum_j c_{ij} x_{ij}$. As such the problem is an ordinary linear program involving $m \cdot n$ variables x_{ij} and $m+n$ linear inequalities in addition to the nonnegativity requirements on the variables themselves. It can be solved in a most expeditious fashion by the simplex method.

The problem becomes an integer program when the variables x_{ij} are required to be integral. But even in this form the problem is elementary since it may be shown that the optimal solution to the linear program will yield variables x_{ij} which are integral if the supplies s_i and the demands d_j are themselves integral (Dantzig [3]).

Our purpose will be to exhibit the neighborhood system for the integral transportation problem and to show that an accelerated neighborhood search is identical to the simplex method. For this problem the technology matrix A will consist of $1+n+m+m \cdot n$ rows and $m \cdot n$ columns. The matrix will map a vector of integral activity levels $\{h_{ij}\}$ into a production plan

$$
\begin{aligned}
y &= -\sum_i \sum_j c_{ij} h_{ij}, \\
t_1 &= \sum_i h_{i1}, \\
&\vdots \\
t_n &= \sum_i h_{in}, \\
u_1 &= -\sum_j h_{1j}, \\
&\vdots \\
u_m &= -\sum_j h_{mj}, \\
h_{11} &= h_{11}, \\
&\vdots \\
h_{mn} &= h_{mn},
\end{aligned}
$$

with y the negative of the shipping cost, t_j the total quantity shipped to destination j, and $-u_i$ the quantity shipped from source i.

It is useful to introduce the graph G associated with the transportation problem which has as its set of nodes, the m sources and n destinations, and as its arcs, all simple paths from a source to a destination. A cycle of the graph is a sequence of arcs $(n_1, n_2), (n_2, n_3), \ldots, (n_k, n_1)$ each of which connects a source and a destination, beginning and ending with the same node.

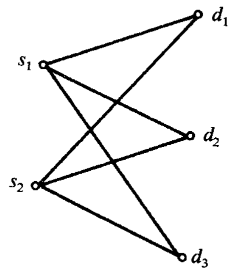

FIGURE 9.

The transportation problem is degenerate in the sense used in Section 1 of this paper and our neighborhood system will not necessarily be minimal. We shall assume that $c_{ij} > 0$, $c_{ij_1} \neq c_{ij_2}$, for $j_1 \neq j_2$ and $\sum c_{ij} e_{ij} \neq 0$ around any cycle.

A typical neighbor of the origin will be an integral vector $\{h_{ij}\}$ whose coordinates may be of mixed sign. Such a vector will be a neighbor of the origin if there is no other integral vector $\{h'_{ij}\}$ with $\min(0, y') \geqslant \min(0, y)$, $\min(0, t'_j) \geqslant \min(0, t_j)$ for $j = 1, \ldots, n$, $\min(0, u'_i) \geqslant \min(0, u_i)$ for $i = 1, \ldots, m$, and $\min(0, h'_{ij}) \geqslant \min(0, h_{ij})$ for all i and j, unless all of these inequalities are equalities. Let us begin our analysis by considering a potential neighbor of the origin for which one of the t_j's, say t_1, is different from zero. Since the negative of a neighbor is also a neighbor there is no loss of generality in assuming that $t_1 > 0$. But then one or more of the coordinates h_{i1} must also be strictly positive, say $h_{i^*1} > 0$. The integral vector $\{h'_{ij}\}$ defined to be

$$h'_{ij} = h_{ij} \quad \text{for all } (i, j) \neq (i^*, 1),$$

$$h'_{i^*1} = h_{i^*1} - 1,$$

will yield a production plan $\{y', t', u', h'\}$ which differs from $\{y, t, u, h\}$ in only four coordinates:

$$y' = y + c_{i^*1},$$

$$t'_1 = t_1 - 1,$$

$$u'_{i^*} = u_{i^*} + 1,$$

$$h'_{i^*1} = h_{i^*1} - 1.$$

Since t'_1 and h'_{i^*1} are both >0, this new production plan will certainly be in the smallest convex body containing 0 and h, obtained by translating the inequalities of the transportation problem: $h' \leqslant h$, and of course it is not correct that $h \leqslant h'$. For h to be a neighbor, h' must therefore be the zero vector, so that h is a plan which ships a single unit from source i^* to destination 1. All other neighbors of the origin must satisfy $t_j = 0$ for $j = 1, \ldots, n$.

As the next step in our analysis we let $\{h_{ij}\}$ be a neighbor of the origin for which $u_i=\sum_j h_{ij}=0$ and $t_j=\sum_i h_{ij}=0$ for all i and j. Assume, for specificity, that $h_{11}>0$. It follows that $h_{i1}<0$ for some i, say i_1. As a consequence $h_{i_1j}>0$ for some j, say j_1. By continuing in this fashion we shall find a cycle along which the corresponding h_{ij}'s alternate in sign. We define $\{e_{ij}\}$ to be $+1$ if (i,j) is in the cycle and $h_{ij}>0$, to be -1 if (i,j) is in the cycle and $h_{ij}<0$, and to be 0 otherwise. The production plan $\{e_{ij}\}$, illustrated in Figure 10, consists of shipping a single unit of the commodity around the arcs of a cycle.

Let us assume that the cost $\sum\sum c_{ij}e_{ij}$ associated with this elementary shipping plan is positive. If we then define $h'_{ij}=h_{ij}-e_{ij}$, it will be true that $y'>y$, $t'_j=0$, and $u'_i=0$ for all (i,j), that $h'_{ij}\geq h_{ij}$ for $h_{ij}\leq 0$, and that $h'_{ij}\geq 0$ for $h_{ij}>0$. It follows that h' is in the smallest convex body containing 0 and h, obtained by translating the inequalities of the transportation problem; $h'\leqslant h$, and it is not correct that $h\leqslant h'$. h will therefore be a neighbor of the origin only if it is equal to the elementary shipping plan e. If, on the other hand, the cost $\sum\sum c_{ij}e_{ij}<0$, consideration of the plan $h'=-h+e$ will also show that h is equal to e. We conclude that a neighbor of the origin for which u_i and t_j are all 0 is an elementary plan consisting of shipping a single unit in some direction around a cycle.

To complete our analysis we consider those neighbors of the origin for which $t_j=0$, for $j=1,\ldots,n$, and for which one of the u_i, say u_1, is negative. Since $u_1<0$, it follows that there is some j, say j_1, for which $h_{1j_1}>0$, and as a consequence there will be some i, say i_1 for which $h_{i_1j_1}<0$. There is no loss in generality in assuming that $u_{i_1}>0$, since if this were not so the argument could be continued to find $h_{i_1j_2}>0$ and $h_{i_2j_2}<0$, etc., and must terminate with a positive u, since a cycle is ruled out by precisely those arguments which have just been employed.

If $c_{1j_1}>c_{i_1j_1}$, the plan

$$h'_{ij}=h_{ij}\quad \text{for all } (i,j)\neq(1,j_1) \text{ or } (i_1,j_1),$$
$$h'_{1j_1}=h_{1j_1}-1,$$
$$h'_{i_1j_1}=h_{i_1j_1}+1,$$

is in the smallest translated convex body containing 0 and h, and we conclude that $h_{1j_1}=1$, $h_{i_1j_1}=-1$ and is otherwise 0. If $c_{1j_1}<c_{i_1j_1}$, $-h'$ will lie in the smallest

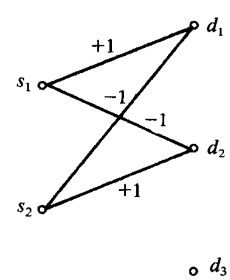

FIGURE 10.

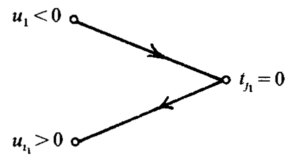

FIGURE 11.

translated convex body containing 0 and $-h$, and the same conclusion follows. We have therefore demonstrated the following theorem.

THEOREM 2.3: *The neighbors of the origin for the transportation problem are of three types: shipping a single unit from a source to a destination (or its negative), shipping a single unit around a cycle, and replacing the shipment of a single unit from a particular source to a destination by the shipment of a single unit from a different source to the same destination.*

Now let h_{ij} be a feasible solution to the transportation problem with a given set of supplies s_i and demands d_j; we shall assume that $\sum s_i = \sum d_j$. (We shall also make the standard nondegeneracy assumption of the transportation problem that the sum of the supplies is different from the sum of the demands for any proper subset of nodes.) In this case the neighbors of a feasible plan h obtained by shipping one more (or less) unit along a given arc will not be feasible, nor will those obtained by shipping one more unit along a given arc and one less from a different source to the same destination. We need only concern ourselves with neighbors that arise by shipping one unit around a cycle.

Let us suppose that the feasible solution is positive for a set of arcs which contain a cycle and let e be the plan which ships one unit around that cycle, with the orientation selected in such a way that $\sum c_{ij}e_{ij} < 0$. Then $h + e$ will remain feasible and yield an improvement in the objective function. But this neighborhood search can easily be accelerated by adding the largest integral multiple of

FIGURE 12.

e to h which retains feasibility; in so doing we shall obtain an improved feasible solution containing fewer positive arcs than the previous solution.

Repeated application of the same procedure will yield an improved solution for which the set of positive arcs forms a *tree*: a subgraph of G containing no cycles. And given the nondegeneracy assumption of the transportation problem, this tree is maximal in the sense that adding any arc to the tree will yield a subgraph containing precisely one cycle. The reader familiar with linear programming will know that at this point we have reached a *feasible basis* for the transportation problem.

In order to see whether $h+e$ remains feasible, where e is a plan which ships a single unit around any particular cycle, we need only select an arc which is not in the tree. If $h+e$ ships a positive amount along this arc and remains feasible, the cycle—and its orientation—will be uniquely determined. Once the cycle is known we evaluate the cost $\sum c_{ij}e_{ij}$. If this cost is positive this particular neighbor should not be selected, since it results in a worsening of the objective function. If the cost is negative, we add to h the largest possible multiple of e which retains feasibility, yielding a new tree containing the new arc and with one of the arcs of the original tree deleted.

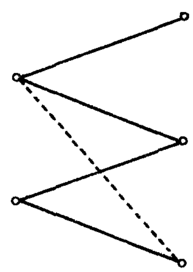

FIGURE 13.

This movement from one feasible basis to an adjacent one with an improvement in the objective function is precisely a pivot step of the ordinary simplex method. As we see an accelerated neighborhood search is identical to the simplex method for the transportation problem. A similar conclusion is undoubtedly correct for other types of network flow problems as well.

C. *The Knapsack Problem with Two Variables*

The problem of minimizing a linear function of nonnegative integral variables subject to a single linear inequality is known as the knapsack problem. It is an extremely difficult problem and undoubtedly does not possess an algorithm which is polynomial in the data when the number of variables is considered as well as the coefficients of the two linear inequalities. In this section we shall provide a complete description of the neighborhood structure when the number of variables is equal to two and verify that an accelerated neighborhood search does lead to a polynomial algorithm.

A polynomial algorithm for the two variable knapsack problem was first provided by Hirschberg and Wong [**4**]. This was followed by a generalization to a larger class of integer programs involving two variables by Kannan [**7**], and extended to the general two variable problem by myself [**12**].[3] The analysis of the neighborhood structure for the two variable knapsack problem can be applied with minor modifications to the general integer program with two variables; the knapsack problem has been selected in order to simplify the exposition.

The technology matrix A for the knapsack problem has the form

$$\begin{bmatrix} - & - \\ 1 & 0 \\ 0 & 1 \\ + & + \end{bmatrix}.$$

The middle two rows of A express the nonnegativity constraints, the top row the objective function (we are minimizing a linear function with positive coordinates), and the last row states that a positive linear function of the two variables is greater than or equal to a preassigned right-hand side. We adopt the lexicographic tie breaking rule which stimulates that a 0 coordinate in the vector $y = Ah$ is to be considered positive if and only if the vector y is lexicographically positive. With this convention the 0's in A are interpreted as being negative.

The analysis of the neighborhood system is greatly facilitated by the following elementary observation.

LEMMA 2.4: *Let the matrix A have the sign pattern*

$$\begin{bmatrix} - & - \\ + & - \\ - & + \\ + & + \end{bmatrix}.$$

Then $(1, 0)$, $(0, 1)$, *and* $(1, 1)$ *are neighbors of the origin.*

When the four linear inequalities, corresponding to the four rows of A, are placed on the vertices of the unit square, as in Figure 14, the resulting convex body will contain no other lattice points. Therefore the smallest convex body—obtained by translating the four linear inequalities—containing the origin and any one of the other three lattice points will be free of additional lattice points. This demonstrates Lemma 2.4.

Now let us assume that the difference between the two columns of A has the same sign pattern as one of the columns; for example that $a^1 - a^2$ has the same sign pattern as a^1, where a^j denotes column j. But then Lemma 2.4 can be applied

[3] As previously mentioned, Lenstra has demonstrated the existence of a polynomial algorithm when the number of variables is fixed at any value

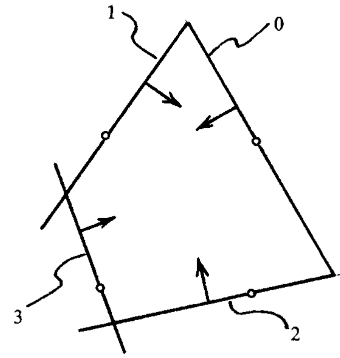

FIGURE 14.

to the matrix B given by

$$\begin{bmatrix} a_{01}-a_{02} & a_{02} \\ 1 & 0 \\ -1 & 1 \\ a_{31}-a_{32} & a_{32} \end{bmatrix},$$

and we conclude that $a^1 - a^2 = B \cdot \binom{1}{0}$ is a neighbor of the origin for the matrix B. But as h and h' range over the lattice points in the plane, $B \cdot h'$ and $A \cdot h$ generate the same set of vectors, and it follows that $(1, -1)$ is a neighbor of the origin for the original problem.

This process of subtraction may be continued. If k_1 is the largest integer such that $a^1 - k_1 a^2$ has the same sign pattern as a^1, then $(1,-1), (1,-2), \ldots, (1, -k_1)$ will all be neighbors of the origin. Letting $a^3 = a^1 - k_1 a^2$, the result of these repeated subtractions will be a matrix

$$A^2 = \overset{\begin{matrix} a^3 & \quad a^2 \end{matrix}}{\begin{bmatrix} a_{01}-k_1 a_{02} & a_{02} \\ 1 & 0 \\ -k_1 & 1 \\ a_{31}-k_1 a_{32} & a_{32} \end{bmatrix}}$$

whose sign pattern is identical to that of the original matrix.

It will no longer be correct that $a^3 - a^2$ has the same sign pattern as a^3, but it is possible that $a^2 - a^3$ has the same sign pattern as a^2. If this is so we let k_2 be the largest integer such that $a^4 = a^2 - k_2 a^3$ has the same sign pattern as a^2. This produces a new linear sequence of neighbors $(-1, 1+k_1)$, $(-2, 1+2k_1), \ldots,$ $(-k_2, 1+k_2k_1)$, and a final matrix A^3 whose columns are given by a^3, a^4.

The process moves through a sequence of matrices $A^1 = A, A^2, A^3, \ldots, A^{n+1}$ each having the same sign pattern as A. A^j has columns (a^{j+1}, a^j) if j is even and (a^j, a^{j+1}) if j is odd, where $a^{j+1} = a^{j-1} - k_{j-1} a^j$; k_{j-1} is the largest integer such that a^{j+1} and a^{j-1} have the same sign pattern. The process terminates at the matrix A^{n+1} which has the property that neither column difference has the same sign pattern as one of the columns. In moving from A^j to A^{j+1} a new set of k_j neighbors is determined:

$$(a_1^j - a_1^{j+1},\, a_2^j - a_2^{j+1}), \ldots, (a_1^j - k_j a_1^{j+1},\, a_2^j - k_j a_2^{j+1}) = (a_1^{j+2}, a_2^{j+2}).$$

As is quite apparent, this set of neighbors, which we denote by S_j, is linear.

I shall reserve for the next section the proof that all of the neighbors of the origin—aside from sign—are obtained by this procedure. Assuming this conclusion to be correct, the neighbors of the origin are partitioned into $S_0 = \{(1,0), (0,1), (1,1)\}$ and n linear subsets—an observation which permits a considerable acceleration of the repeated neighborhood search for an optimal solution.

An example may be useful at this point. Let the knapsack problem be

$$\min 329\, h_1 + 103\, h_2 \qquad \text{subject to}$$

$$195\, h_1 + 61\, h_2 \geq b$$

with h_j nonnegative integers.

The sequence of matrices A^1, A^2, A^3, A^4 is then given by

$$\begin{bmatrix} -329 & -103 \\ 1 & 0 \\ 0 & 1 \\ 195 & 61 \end{bmatrix}, \quad \begin{bmatrix} -20 & -103 \\ 1 & 0 \\ -3 & 1 \\ 12 & 61 \end{bmatrix}, \quad \begin{bmatrix} -20 & -3 \\ 1 & -5 \\ -3 & 16 \\ 12 & 1 \end{bmatrix}, \quad \begin{bmatrix} -2 & -3 \\ 31 & -5 \\ -99 & 16 \\ 6 & 1 \end{bmatrix},$$

with neighbors $(1,0)$, $(0,1)$, $(1,1)$ and

$(1,-1)$	$(-1,4)$	$(6,-19)$
$(1,-2)$	$(-2,7)$	$(11,-35)$
$(1,-3)$	$(-3,10)$	$(16,-51)$
	$(-4,13)$	$(21,-67)$
	$(-5,16)$	$(26,-83)$
		$(31,-99)$.

In S_j, for $j \geq 1$, the neighbors are $h^j(t) = (a_1^j - t a_1^{j+1}, a_2^j - t a_2^{j+1})$ with $t = 1, 2, \ldots, k_j$. If j is odd then $h_1^j(t) > 0$ and is monotonically increasing, $h_2^j(t) < 0$ and is monotonically decreasing; moreover $h_1^j(k_j) < h_1^{j+2}(1)$ and $h_2^j(k_j) > h_2^{j+2}(1)$. The sign patterns and inequalities are reversed if j is even. The values of the constraint function at this sequence of neighbors are given by

$$a_3 \cdot h^j(t) = a_3^j - t a_3^{j+1}$$

for $t = 1, \ldots, k_j$, with $a_3^j > a_3^{j+1} > 0$. The sequence $a_3 \cdot h^j(t)$ is positive and decreasing from a_3^j when $t = 0$ to a_3^{j+2} when $t = k_j$. It follows that $a_3 \cdot h^l(t) \geq a_3^{l+2} \geq a_3^{j+1}$ for $t = 1, 2, \ldots, k_l$ and $l \leq j-1$.

Now let h be a feasible solution to the knapsack problem for a particular value of the right-hand side b, i.e. $h \geq 0$ and $a_3 \cdot h \geq b$. As we shall see, a single division is required to determine whether one of the neighbors of the origin in S_j, for $j \geq 1$, can be subtracted from h (addition of one of these neighbors will worsen the objective function) and still retain feasibility.

LEMMA 2.5: *Let h be a feasible solution, and define t^* to be the smallest positive integer such that*

$$a_3 \cdot (h - h^j(t)) = a_3 \cdot h - a_3^j + t a_3^{j+1} \geq b.$$

If $h - h^j(t)$ is feasible for some $t = 1, 2, \ldots, k_j$, then $h - h^j(t^*)$ is feasible. (If $t^* > k_j$ then $h - h^j(t)$ is infeasible for all $t = 1, 2, \ldots, k_j$.)

The proof of the Lemma is quite direct. If $h - h^j(t^*)$ is not feasible, then $h_i - h_i^j(t^*) < 0$ for $i = 1$ or 2. Assume, to be specific, that j is odd, so that $h_1^j(t) > 0$ and increasing in t. This implies that $h_1 - h_1^j(t^*) < 0$ and therefore $h_1 - h_1^j(t) < 0$ for $t > t^*$. On the other hand, for $t < t^*$ we have $a_3 \cdot (h - h^j(t)) = a_3 \cdot h - a_3^j + t a_3^{j+1} < b$. This demonstrates Lemma 2.5.

We see that at most n computations (plus some trivial ones associated with subtracting $(1, 0)$, $(0, 1)$ and $(1, 1)$ from h) are required to test for the optimality of a given feasible vector h. But even more can be said. Let h be an arbitrary feasible vector such that neither $h - (1, 0)$ nor $h - (0, 1)$ are feasible. Then at most $2n$ steps are required to pass from h to the optimal solution if at each step we subtract the largest multiple of a neighbor of the origin from a feasible vector so as to retain feasibility.

To establish this assertion, let j be the smallest integer (assume that $j \geq 2$) such that S_j contains a neighbor $h^j(t)$ with $h - h^j(t)$ feasible. By Lemma 2.5 it will be correct that $h' = h - h^j(t^*)$ is feasible with t^* the smallest positive integer such that $a_3 \cdot (h - h^j(t)) = a_3 \cdot h - a_3^j + t a_3^{j+1} \geq b$. Consider first the case in which $t^* > 1$. In this case $a_3 \cdot h' = a_3 \cdot h - a_3^j + t^* a_3^{j+1} < b + a_3^{j+1}$. Since $a_3 \cdot h^l(t) \geq a_3^{j+1}$ for $l = 1, 2, \ldots, j-1$ and any $t = 1, 2, \ldots, k_l$, it follows that the only neighbors of the origin which can be subtracted from h' with feasibility retained must lie in $S_j, S_{j+1}, \ldots, S_n$. Moreover, since $a_3 \cdot h^j(t) > a_3^{j+1}$ for $t = 1, 2, \ldots, k_j - 1$ the only possible neighbor in S_j that can be subtracted is the last one, $h^j(k_j)$. If we subtract the largest possible multiple of $h^j(k_j)$ which retains feasibility then no further subtractions are available in any of the sets $S_1, \ldots, S_j$.

To complete the argument we consider the case in which $t^* = 1$. But then I claim that

$$a_3 \cdot (h - h^j(1)) = a_3 \cdot h - a_3^j + a_3^{j+1} < a_3^{j+1} + b.$$

If this were not so then $a_3 \cdot h \geq b + a_3^j$. Since $a_3 \cdot h^{j-2}(k_{j-2}) = a_3^j$ it follows that $a_3 \cdot (h - h^{j-2}(k_{j-2})) \geq b$. From the fact that the positive coordinate of $h^{j-2}(k_{j-2})$ is less than the same coordinate of $h^j(1)$, we see that $h - h^{j-2}(k_{j-2})$ is feasible, contradicting the assumption that S_j is the first set containing a neighbor which can be subtracted from h. It follows that even in the case in which $t^* = 1$, $a_3 \cdot h' = a_3 \cdot (h - h^j(1)) < b + a_3^j$ and the only subsequent subtraction in S_j is $h^j(k_j)$. Again we subtract the largest possible multiple of $h^j(k_j)$ which retains feasibility, and are left with the situation that all subsequent subtractions are from $S_{j+1}, \ldots, S_n$. This demonstrates that the optimal solution may be found in no more than $2n$ steps.

Consider the problem

$$\min 329\, h_1 + 103\, h_2 \qquad \text{subject to}$$
$$195\, h_1 + \;\, 61\, h_2 \geq 6999$$
$$\text{with } h_j \text{ nonnegative integers,}$$

whose neighbors have been determined above. The feasible solution $h = (28, 26)$ yields a value of the constraint of 7046. The only neighbor in S_1 which can be subtracted is $(1, -3)$. Subtracting this neighbor three times yields $h' = (25, 35)$ with a value of the constraint function given by 7010. The only neighbor in S_2 which can then be subtracted is $(-5, 16)$. Subtracting this neighbor twice yields $h'' = (35, 3)$ with a value of the constraint given by 7008. No subsequent subtraction is possible and we have reached the optimal solution.

The number of binary bits required to store a positive integer a is $f(a) = \lceil \log_2 (a+1) \rceil$, with $\lceil x \rceil$ the smallest integer greater than or equal to x. The function satisfies the identity $f(2a) = f(a) + 1$.

It is elementary to argue that n, the number of linear subsets into which the neighborhood system is partitioned, is no larger than $f(a_3^1) + f(a_3^2)$, the number of bits required to store the coefficients of the constraint function. To see this, define $T_j = f(a_3^j) + f(a_3^{j+1})$. Then

$$T_j - T_{j+1} = f(a_3^j) - f(a_3^{j+2}) \geq 1,$$

since $a_3^j \geq a_3^{j+1} \geq a_3^{j+2}$, and $a_3^j = a_3^{j+2} + k_j a_3^{j+1} \geq a_3^{j+2} + a_3^{j+1} \geq 2a_3^{j+2}$. It follows that T_j decreases by at least unity for each linear subset of neighbors; the number of such subsets is therefore not larger than T_1. In the terminology of complexity theory the accelerated neighborhood search leads to a linear algorithm for the knapsack problem. As I have previously mentioned, virtually identical arguments can be applied to the general integer programming problem with two variables.

D. *The Totality of Neighbors for the Knapsack Problem*

We shall demonstrate that the procedure of Section 2C finds all of the neighbors of the origin for the two variable knapsack problem. It is easy to see that the only neighbors in the nonnegative quadrant are $(1, 0)$, $(0, 1)$ and $(1, 1)$; other than the negatives of these three, all other neighbors are in the second or fourth quadrant.

Let $h = (q_1, -q_2)$ be a neighbor in the fourth quadrant, with q_1, q_2 nonnegative integers, and consider the smallest convex body—obtained by translating the four inequalities—containing h and the origin. In Figure 15 I have made a particular assumption—which will be maintained throughout the argument—about the way in which the linear inequalities relate to these two points; the small number of alternative assumptions can be dealt with by analogous arguments.

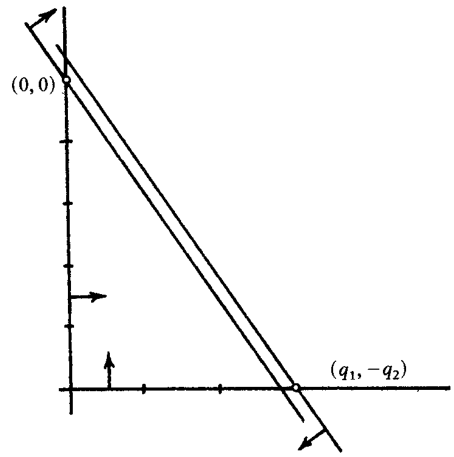

FIGURE 15.

Figure 15 implies that

$$Ah = \begin{bmatrix} - & - \\ 1 & 0 \\ 0 & 1 \\ + & + \end{bmatrix} \begin{pmatrix} q_1 \\ -q_2 \end{pmatrix} = \begin{bmatrix} - \\ + \\ - \\ + \end{bmatrix}.$$

CASE 1: $q_2 > q_1$. We have

$$0 < q_1 a_3^1 - q_2 a_3^2 = q_1(a_3^1 - a_3^2) - (q_2 - q_1)a_3^2, \qquad \text{and}$$
$$0 > q_1 a_0^1 - q_2 a_0^2 = q_1(a_0^1 - a_0^2) - (q_2 - q_1)a_0^2.$$

The first of these inequalities implies $a_3^1 - a_3^2 > 0$ and the second $a_0^1 - a_0^2 < 0$. It follows that the matrix A' with columns $(a^1 - a^2, a^2)$ has the same sign pattern as A, and that $(q_1, -q_2 + q_1)$ represents the same neighbor of the origin when A is replaced by A'.

CASE 2: $q_2 < q_1$.
In this case we write

$$q_1 a^1 - q_2 a^2 = (q_1 - q_2)a^1 - q_2(a^2 - a^1)$$

so that $(q_1 - q_2, -q_2)$ represents the original neighbor of the origin when A is replaced by A', the matrix with columns $(a^1, a^2 - a^1)$. To verify that A' is the matrix arising in the computations of Section 2C, we need to verify that $a_3^2 - a_3^1 > 0$ and $a_0^2 - a_0^1 < 0$. The first of these inequalities arises from the fact that $(1, -1)$ does not satisfy inequality 3 when it is placed through the origin, and the second from the observation that $(q_1 - 1, -q_2 + 1)$ does not satisfy inequality 0 when it is placed through the point $(q_1, -q_2)$.

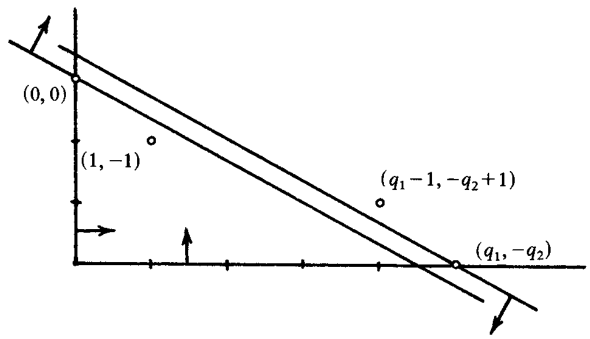

FIGURE 16.

In either case when the smaller of the two numbers is subtracted from the larger we obtain a new representation of the same neighbor with A replaced by A'. Since the pair of integers q_1, q_2 are relatively prime, a continuation of the process will ultimately represent the original neighbor using weights $(1, 0)$ or $(0, -1)$ with A replaced by a matrix arising in the computation of Section 2C. We have therefore verified that the neighbors of Section 2C are—aside from sign—all of the neighbors of the two variable knapsack problem.

E. *Integer Programs with Three Variables and Three Inequalities*

In this section I will briefly describe some recent work on neighborhood systems associated with a technology matrix A with four rows and three columns. The matrix is assumed to satisfy Assumption 1.1.

A matrix U is defined to be unimodular if it has integral entries and a determinant of ± 1. If U is unimodular the transformation $h' = U^{-1}h$ maps the set of lattice points in three space onto itself. Since $Ah = AUU^{-1}h = AUh'$ it follows that the two technology matrices A and AU generate equivalent integer programs, and we are free to select the unimodular matrix U in a way that simplifies the description of the neighborhood system.

In **[15]**, I have shown that there is a unimodular transformation U, with the property that all of the neighbors of $(0, 0, 0)$, associated with $A' = AU$, lie on one of the three planes $h_1 = -1$, $h_1 = 0$, $h_1 = +1$. This implies that in testing for the optimality of a feasible solution $h = (h_1, h_2, h_3)$ to the integer program

$$\max a_0' \cdot h \quad \text{subject to}$$

$$a_i' \cdot h \geq b_i \qquad (i = 1, 2, 3)$$

$$\text{and } h \text{ integral,}$$

it is sufficient to solve the three two variable integer programs obtained by fixing the first variable at $h_1 - 1$, h_1, $h_1 + 1$. Each one of these problems has, of course, a polynomial algorithm by the arguments of the previous example.

In fact, something more can be demonstrated. When the two variable integer program, obtained by fixing the first variable at any level, is solved, the information available at the solution is sufficient to tell whether the optimal solution to the three variable problem has a higher or lower value of its first coordinate. This is a very surprising conclusion which, unfortunately, does not seem capable of generalization to problems with more than three variables.

Yale University

Manuscript received November, 1984; final revision received August, 1985.

REFERENCES

[1] BROWN, DONALD J., AND GEOFFREY M. HEAL: "Marginal vs. Average Pricing in the Presence of a Public Monopoly," *American Economic Review*, 73(1983), 189-193.
[2] ———: "The Optimality of Regulated Pricing: A General Equilibrium Analysis," *Advances in Equilibrium Theory*. Berlin: Springer-Verlag, 1985.
[3] DANTZIG, GEORGE B.: *Linear Programming and Extensions*. Princeton: Princeton University Press, 1963.
[4] GAREY, MICHAEL R., AND DAVID S. JOHNSON: *Computers and Intractability*. San Francisco: W. H. Freeman, 1979.
[5] HART, OLIVER D.: "Imperfect Competition in General Equilibrium," London School of Economics, 1982.
[6] HIRSCHBERG, D. S., AND C. K. WONG: "A Polynomial Time Algorithm for the Knapsack Problem in Two Variables," *Journal of the Association of Computing Machinery*, 23(1976), 147-154.
[7] ICHIISHI, T.: "A Social Coalitional Existence Lemma," *Econometrica*, 49(1981), 369-377.
[8] ICHIISHI, T., AND M. QUINZII: "Decentralization for the Core of a Production Economy with Increasing Returns," Ecole Polytechnique, 1981.
[9] KANNAN, ROVINDRAN: "A Polynomial Algorithm for the Two Variable Integer Programming Problem," *Journal of the Association of Computing Machinery*, 27(1980), 118-122.
[10] LENSTRA, H. W., JR., "Integer Programming with a Fixed Number of Variables," *Mathematics of Operational Research*, 8(1983), 538-548.
[11] QUINZII, M.: "An Existence Theorem for the Core of a Production Economy with Increasing Returns," *Journal of Economic Theory*, 28(1982), 32-50.
[12] SCARF, HERBERT E.: "Note of the Core of a Productive Economy," mimeo, Stanford University, 1963.
[13] ———: "Production Sets with Indivisibilities, Part I: Generalities," *Econometrica*, 49(1981), 1-32.
[14] ———: "Production Sets with Indivisibilities, Part II: The Case of Two Activities," *Econometrica*, 49(1981), 395-423.
[15] ———: "Integral Polyhedra in Three Space," *Mathematics of Operational Research*, 10(1985), 403-438.
[16] SHARKEY, W. W.: "Existence of a Core When There Are Increasing Returns," *Econometrica*, 47(1979), 869-876.
[17] WHITE, PHILIP: "Discrete Activity Analysis," Ph.D. Thesis, Yale University, 1983.