The Role of Pivoting in Proving
Some Fundamental Theorems of Linear Algebra

Emil Klafszky
Mathematical Institute
T. U. Miskolc
3515. Miskolc Egyetemváros
Pf.1
Hungary

and

Tamás Terlaky
Department of Operations Research
Eötvös University
1088. Budapest
Múzeum krt.6-8
Hungary

Submitted by Richard A. Brualdi

ABSTRACT

We give a constructive proof for some classical theorems of linear algebra (Steinitz, rank of matrices, Rouche-Kronecker-Capelli, Farkas, Wey, Minkowski). The construction is based on pivoting, and its most important part is the pivot selection rule, which is based on Bland's least index resolution.

1. INTRODUCTION

Several authors have given constructive proofs for key theorems of linear algebra, especially for alternative theorems of linear inequalities [1, 4, 8, 16]. One of the referees kindly called our attention to Nef's [11] book, in which the author also uses tableaux and pivots in his "exchange method" approach to linear algebra.

655 Avenue of the Americas, New York, NY 10010 0024-3795/91/$3.50
This paper contains a new approach to some classical theorems of linear algebra (Steinitz, matrix rank, Rouche-Kronecker-Capelli, Farkas, Weyl, Minkowski). The constructive proofs are based on pivoting. Defining pivoting in a more general way—using generating tableaux—has made it possible to give a new proof for the Steinitz theorem as well.

Our pivot selection strategies are based essentially on Bland's [2] minimal index rule. The famous theorems of Farkas, Weyl, and Minkowski are proved by using pivot tableaux. Theorem 4.1 is essentially a new, very simple form of the alternative theorem of linear inequalities, and its proof is a pretty application of the minimal index rule. One can apply this theorem and its proof to combinatorial structures (for example to oriented matroids) as well (Klafszky and Terlaky [9]). The algorithms presented are mostly not efficient computationally (see e.g. Roos [13] for an exponential example), but they are surprisingly simple.

We will use the symbols 0, +, −, Φ, ∈ introduced by Balinski and Tucker [1], which denote zero, positive, negative, nonnegative, and nonpositive numbers respectively. On the other hand Gale's [7] notation will be used, so matrices and vectors are denoted by capital and small Latin letters, and their components are denoted by the corresponding Greek letters. Index sets are denoted by I and J (with proper subscripts), and the cardinality of an index set J is denoted by ∥J∥.

2. PIVOTING

Pivoting and some of its consequences (Steinitz's theorem) are presented in this section. Our considerations and the definition of pivoting are a little more general than those known from the literature.

Let \{a_1, \ldots, a_n\} ⊂ R^m be arbitrary real vectors. Denote by J = \{1, \ldots, n\} the index set of vectors \{a_1, \ldots, a_n\}. A vector b ∈ R^m is said to be a linear combination of the vectors \{a_j : j ∈ J\} if there are numbers \{λ_1, \ldots, λ_n\} ⊂ R such that

\[ b = \sum_{j=1}^{n} a_jλ_j. \]

**Definition 2.1.** A set of vectors \{a_i : i ∈ I_G ⊂ J\} is called a generating system if for all \(j ∈ \hat{I}_G = J − J_G\), the vector \(a_j\) is a linear combination of the vectors \{a_i : i ∈ J_G\}. (A generating system is often referred to as a “spanning set.”)
For simplicity generating systems will be referred to by their index set $J_G$. Let $J_G \subseteq J$ be a generating system, and choose the coefficients $\tau_{ij}$ so that

$$a_j = \sum_{i \in J_G} \tau_{ij} a_i \quad \text{for all} \quad j \in J_G,$$

that is, $\tau_{ij}$ is the coefficient of the generating vector $a_i$ for $i \in J_G$ in the $J_G$ representation of the vector $a_j$ for $j \in J_G$. We can collect these coefficients in a tableau.

Notice that in general a lot of different systems of the coefficients $\tau_{ij}$ may exist for a given generating system $J_G$. Since representation of vectors in generating systems may not be unique, an infinite number of essentially different systems (tableaux) may occur (not just permutations of columns and rows).

If in position $(r, s)$ the coefficient $\tau_{rs} \neq 0$, then the generating vector $a_r$ can be replaced by the nongenerating vector $a_s$. The new coefficients can be derived as follows:

$$\tau_{ij}' = \tau_{ij} - \frac{\tau_{ij} \tau_{is}}{\tau_{rs}}, \quad i \in J_G', \quad i \neq s, \quad j \in J_G', \quad j \neq r, \quad (1)$$

$$\tau_{sj}' = \frac{\tau_{ij}}{\tau_{rs}}, \quad j \in J_G', \quad j \neq r, \quad (2)$$

$$\tau_{ir}' = -\frac{\tau_{is}}{\tau_{rs}}, \quad i \in J_G', \quad i \neq s, \quad (3)$$

$$\tau_{sr}' = \frac{1}{\tau_{rs}}. \quad (4)$$
Recall that $\tau_{ij}$ is the coefficient of the generating vector $a_i$ in the actual representation of the vector $a_j$. The new system of vectors $J' = (J - \{r\}) \cup \{s\}$ is again a generating system $[J' = (J - \{s\}) \cup \{r\}]$. The element $\tau_{rs}$ is called the pivot element; the transformation (1), (2), (3), (4) is called pivoting. The tableaux of Figure 1 are called generating or pivot tableaux.

**Theorem 2.1.** The system of vectors $J_G'$ is a generating system of $J$, and the coefficients given by the formulae (1), (2), (3), and (4) give a generating tableau.

**Proof.** Elementary; it is left to the reader. 

**Remark.** The tableau can be completed by unit vectors associated to the generating system (i.e., the generating system trivially generates itself). The resulting tableau is called the complete tableau, while the original one is called a dictionary [3].

Linear independence of vectors plays a fundamental role in linear algebra.

**Definition 2.2.** A set of vectors $\{a_j : j \in J_F\}$, $J_F \subset J$, is called independent if it does not contain the zero vector and there is no vector $a_r$, $r \in J_F$, that can be generated as a linear combination of the vectors $\{a_j : j \in J_F \text{ and } j \neq r\}$. Independent sets will be referred to by their index set $J_F$.

Pivoting can be used easily to prove the relation of generating and independent systems. This is Steinitz’s famous theorem.

**Corollary 2.1 (Steinitz theorem [6, 10]).** If $J_F \subset J$ is an independent and $J_G \subset J$ is a generating system, then

$$\|J_F\| < \|J_G\|.$$  

**Proof.** Let us denote $J = J_F \cup J_G$ and suppose, to the contrary, that $\|J_F\| > \|J_G\|$. Choose a dictionary as shown in Figure 2. Make a pivot operation on position $(i, j)$ if $i \in J_G \cap \hat{J}_F$, $j \in J_F \cap \hat{J}_G$, and $\tau_{ij} \neq 0$. Repeat this as many times as possible. Finally the tableau and generating system $J_G'$ shown in Figure 3 are obtained. Here $\tau_{ii} = 0$ for $i \in J_G' \cap \hat{J}_F$, $j \in J_F \cap \hat{J}_G$ by the construction. Furthermore $J_F \cap \hat{J}_{G'} \neq \emptyset$, since we have assumed that $\|J_F\| > \|J_G\|$.

If $J_F \cap J_G' = \emptyset$, then all vectors of $J_F$ are identically zero vectors, which is a contradiction. Otherwise the vectors $a_j$ for $j \in J_F \cap \hat{J}_{G'}$ are linear combina-
tions of the vectors \( \{ a_j : j \in J_F \cap J_G \} \), which also contradicts the linear independence of \( J_F \).

In our further considerations the independent generating systems play an important role.

**DEFINITION 2.3.** A set of vectors \( \{ a_j : j \in J_B \} \), \( J_B \subset J \), is called a *basis* if it is linearly independent and generating.

An obvious consequence of the Steinitz theorem is that bases have the same cardinality. (In fact this is equivalent to the Steinitz theorem.) So the following terminology can be used.

**DEFINITION 2.4.** The cardinality of a basis of the set of vectors \( \{ a_1, \ldots, a_n \} \) is called the rank of the vector set, and it is denoted by \( \text{rank}(a_1, \ldots, a_n) \).
REMARK. A generating tableau is called a basic tableau if its generating system is a basis.

**Lemma 2.1.** A basis uniquely determines its basic tableau, and if two arbitrary bases are given, then there is a sequence of pivots which transforms one of the basic tableaux to the other.

**Proof.** Elementary, well known [3, 5].

**Definition 2.5.** The set \( L(a_1, \ldots, a_n) = \{ b : b = \sum_{j=1}^{n} a_j \lambda_j \text{ for all } \lambda_j \in R, j = 1, \ldots, n \} \) is called a linear subspace of \( R^n \). (Vectors \( a_1, \ldots, a_n \in R^n \) are called a generating system of this subspace.)

Well-known elementary properties of linear subspaces and pivoting (how to perform it) will be used without proving them. So for convenience the following notation will be used: \( \text{rank}(a_1, \ldots, a_n) = \text{rank} L(a_1, \ldots, a_n) \).

It can be seen easily that the linear subspace of the row vectors of the complete tableau remains unchanged during pivot operations, since they can also be performed backwards. Furthermore, if \( J'' \subset J \), then the subspaces \( L(t_i'' : i \in J''_B) \) and \( L(t_i''' : i \in J''_B) \) are equal. (See Figure 4.)

Denote

\[
t^{(i)} = \left( \tau^{(i)}_j \right) = \begin{cases} 
\tau_{ij} & \text{if } j \in J_B, \\
1 & \text{if } j = i, \\
0 & \text{otherwise}
\end{cases}
\]

and

\[
t_{(j)} = \left( \tau_{(j)i} \right) = \begin{cases} 
\tau_{ij} & \text{if } j \in J_B, \\
-1 & \text{if } i = j, \\
0 & \text{otherwise}
\end{cases}
\]

![Figure 4](image-url)
As an illustration, let us consider the following simple numerical example:

\[
\begin{align*}
    a_1 &= \begin{pmatrix} 1 \\ 2 \end{pmatrix}, & a_2 &= \begin{pmatrix} 3 \\ 4 \end{pmatrix}, & a_3 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, & a_4 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix},
    J_B &= \{4, 3\}.
\end{align*}
\]

Then the tableau is

\[
\begin{array}{cccc}
    & a_1 & a_2 & a_3 & a_4 \\
    a_4 & 1 & 3 & 0 & 1 \\
    a_3 & 2 & 4 & 1 & 0
\end{array}
\]

In this case \( t_{(1)} = (-1, 0, 2, 1), \) \( t_{(2)} = (0, -1, 4, 3), \) \( t_{(3)} = (2, 4, 1, 0), \) \( t_{(4)} = (1, 3, 0, 1). \)

The matrix rank theorem can be proved by using the abovementioned properties of pivot tableaux.

**Theorem 2.2 (Matrix rank).** If \( A = (a_1, \ldots, a_n) = (a^{(1)}, \ldots, a^{(m)}) \) is an arbitrary matrix, then \( \text{rank}(a_1, \ldots, a_n) = \text{rank}(a^{(1)}, \ldots, a^{(m)}), \) where columns and rows of \( A \) are denoted respectively by \( a_j \) and \( a^{(i)} \).

**Proof.** Consider the matrix \( A \) as a dictionary associated with the vectors \( a_1, \ldots, a_n, e_1, \ldots, e_m \), where unit vectors are the basis vectors. By pivoting, let us bring as many \(^*\)\(^*\)\(^*\)\(^*\) vectors into the basis as possible. When it is not possible to pivot more such vectors into the basis, then we have the tableau shown in Figure 5. (The index sets \( J, J_B \) and \( I, I_B \) contain the corresponding indices of the vectors \( a_j \) and \( e_i \) respectively.)

According to the above properties, \( \mathcal{L}(a^{(1)}, \ldots, a^{(m)}) = \mathcal{L}(t^{u(j)}: j \in J_B), \)

\[
\begin{array}{c|c|c|c}
    J_B & J & I \\
    \hline
    J_B & \cdots & L^{(j)} & \text{0} \\
    \hline
    I_B & 0 & 0 & 1 & \cdots & 1
\end{array}
\]

**Fig. 5.**
since the rows associated to \( i \in I_B \) are zero rows. So \( \text{rank} \{ t^{u(j)} : j \in I_B \} = \| J_B \| \), because these rows are independent. So we have \( \text{rank}(a^{(1)}, \ldots, a^{(m)}) = \text{rank}(d^{(1)}, \ldots, d^{(m)}) = \text{rank} \{ t^{u(j)} : j \in I_B \} = \| J_B \| \).

On the other hand, the set of vectors \( \{ a_j : j \in I_B \} \) is a basis of the set \( \{ a_1, \ldots, a_n \} \) (see Figure 5), which by definition says that \( \text{rank}(a_1, \ldots, a_n) = \| J_B \| \). Our proof is complete.

Finally in this section we show some orthogonality properties of the pivot tableau. The following definitions are necessary.

**Definition 2.6.** The *inner product* of two vectors \( a, c \in \mathbb{R}^m \) is defined by

\[
ac = \sum_{j=1}^{m} a_j c_j.
\]

**Definition 2.7.** Vectors \( a, c \in \mathbb{R}^m \) are called *orthogonal* if \( ac = 0 \).

If a vector is orthogonal to a generating system of a subspace, then it is orthogonal to all vectors of the subspace. In this case we say that the vector is orthogonal to the subspace.

**Theorem 2.3** (Orthogonality). If two arbitrary bases \( J_B, J_{B^*} \) of the vectors \( (a_1, \ldots, a_n) \) and their dictionaries are given, then \( t^{(i)} t^{(j)} = 0 \) holds for all \( i \in J_B \) and \( j \notin J_{B^*} \).

**Proof.** Well known [2, 15].

The following well-known, simple consequence of the orthogonality theorem will be often used.

**Corollary 2.2.** If the vectors \( a_1, \ldots, a_n, e_1, \ldots, e_m \in \mathbb{R}^m \) and \( J_B \subseteq J = \{1, \ldots, n\}, I_B \subseteq I = \{1, \ldots, m\} \), where \( J_B \cup I_B \) is a base of \( \mathbb{R}^m \), are given, then \( \tau_{ij} = y^{(i)} a_j \) for all \( j \notin J_B \) and \( i \in I_B \cup J_B \).

See Figure 6.

Now we can turn to the examination of the system of linear equations.

3. **Systems of Linear Equations**

Let \( \{ a_1, \ldots, a_n, b \} \subseteq \mathbb{R}^m \) be arbitrary vectors, and let \( \{ e_1, \ldots, e_m \} \subseteq \mathbb{R}^m \) be the standard basis of \( \mathbb{R}^m \). Denote by \( J = \{1, \ldots, n\} \) the index set of the
vectors \{a_1, \ldots, a_n\}, and by \(I = \{1, \ldots, m\}\) the index set of the vectors \{e_1, \ldots, e_m\}.

**Theorem 3.1.** For the above (arbitrarily) given vectors \(\{b, a_1, \ldots, a_n, e_1, \ldots, e_m\}\) exactly one of the dictionaries shown in Figure 7 exists.

**Proof.** Using orthogonality properties it is easy to see that both cannot hold simultaneously. To show that at least one holds let us start with the basis \{e_1, \ldots, e_m\} (Figure 8). Let us change as many vectors \(e_i\) of the basis with nonbasic vectors \(a_j\) as possible. Finally we obtain the dictionary shown in
Figure 9. If the hatched part of Figure 9 contains only zeros, then the first case applies. If it contains a nonzero element, then pivoting on this (nonzero) position, the vector $b$ enters the basis (some vector $e_i$ leaves) and we get the second case. □

This theorem—the proof of which gives an algorithm—implies the following Rouche-Kronecker-Capelli theorem.

Corollary 3.1 (Rouche, Kronecker, and Capelli [6]). Exactly one of the following alternatives holds:

(a) $Ax = b$ has a solution, or
(b) $yA = 0, yb = 1$ has a solution.

Proof. Apply the previous theorem for the column vectors of the matrix $A$. If case (a) holds, then vector $b$ is a linear combination of some columns of $A$, i.e., we have a solution for $Ax = b$. If case (b) holds, then the tableau is the
one shown in Figure 10. By orthogonality \( y \) satisfies (b), so our proof is complete.

**Remark.** The tableaux tell us in both cases the general solution of the equation systems. In case (a) the linear combinations of vectors \( t'_j \) for \( j \in J_B' \) give the general solution of the homogeneous system \( Ax = 0 \). In case (b) the linear combinations of vectors \( y^{(i)} \) for \( i \in I_B'' \) give the general solution of the homogeneous system \( yA = 0, \ yb = 0 \).

---

**The Geometric Interpretation of Linear Equation Systems**

**Definition 3.1.** The set \( \mathcal{L}(a_1, \ldots, a_n) = \{ y : ya_j = 0, j = 1, \ldots, n \} \) is called the **orthogonal complement** of the linear subspace \( \mathcal{L}(a_1, \ldots, a_n) = \{ b : b = \sum_{j=1}^{n} a_j \lambda_j, \ \lambda_j \in \mathbb{R}, \ j = 1, \ldots, n \} \).

It is obvious that \( by = 0 \) for arbitrary \( b \in \mathcal{L} \) and \( y \in \mathcal{L}^\perp \).

Three famous theorems of orthogonal complements are now discussed as a consequence of Theorem 3.1. We know that simple proofs exist [6, 10] for these theorems; we prove them here to motivate the corresponding proofs for linear inequalities. Most of the key ideas have already been presented here.

**Corollary 3.2.** Given any set of vectors \( \{a_1, \ldots, a_n\} \subset \mathbb{R}^m \), there is a set of vectors \( \{ y_1, \ldots, y_k \} \subset \mathbb{R}^m \) such that:

(a) \( \mathcal{L}(a_1, \ldots, a_n) = \mathcal{L}^\perp(y_1, \ldots, y_k) \),
(b) \( \mathcal{L}^\perp(a_1, \ldots, a_n) = \mathcal{L}(y_1, \ldots, y_k) \),
(c) \( \mathcal{L}^{++}(a_1, \ldots, a_n) = \mathcal{L}(a_1, \ldots, a_n) \).
Proof. Consider a basic tableau (Figure 11) where \( \|J_B\| \) is maximal, that is, the number of vectors \( e_i \) is minimal in the basis. Let \( \{ y_1, \ldots, y_k \} \) be the set \( \{ y(i) : i \in I_B \} \), or let it be the single zero vector if \( I_B = \emptyset \). So by at most \( m \) pivots the set \( \{ y_1, \ldots, y_k \} \) can be constructed.

(a): If \( b \in \mathcal{L}(a_1, \ldots, a_n) \), then \( b = \sum_{j=1}^{n} \lambda_j a_j \). Since \( y_i a_j = 0 \) for \( i = 1, \ldots, k \), \( j = 1, \ldots, n \), we have \( y_i b = \sum_{j=1}^{n} (y_i a_j) \lambda_j = 0 \), that is, \( b \in \mathcal{L}^+(y_1, \ldots, y_k) \). On the other hand, if \( b \not\in \mathcal{L}(a_1, \ldots, a_n) \), then (see Fig. 11) \( \xi_i \neq 0 \) for an \( i \in I_B \), that is, \( y_i b \neq 0 \), and so \( b \not\in \mathcal{L}^+(y_1, \ldots, y_k) \).

(b): If \( y \in \mathcal{L}(y_1, \ldots, y_k) \), then \( y = \sum_{i=1}^{k} \lambda_i u_i \). Since \( y_i a_j = 0 \) for \( i = 1, \ldots, k \), \( j = 1, \ldots, n \), we have \( y a_j = \sum_{i=1}^{k} (y_i a_j) = 0 \) for \( j = 1, \ldots, n \), that is, \( y \in \mathcal{L}^+(a_1, \ldots, a_n) \). On the other hand, if \( y \not\in \mathcal{L}(y_1, \ldots, y_k) \), that is, \( \sum_{i=1}^{k} \lambda_i y_i = y \) is inconsistent, then by Corollary 3.1 we have a vector \( b \in R^m \) such that \( y_i b = 0 \) for \( i = 1, \ldots, k \) and \( y b = 1 \). So by (a) we have \( b \in \mathcal{L}^+(y_1, \ldots, y_k) = \mathcal{L}(a_1, \ldots, a_n) \), but \( y \not\in \mathcal{L}^+(a_1, \ldots, a_n) \).

(c): Let us apply first (a), then (b), as follows: \( \mathcal{L}(a_1, \ldots, a_n) = [\mathcal{L}(y_1, \ldots, y_k)]^+ = [\mathcal{L}^+(a_1, \ldots, a_n)]^+ \). Our proof is complete.

Remark. The following statement is an immediate consequence of part (b): For any subspace \( \mathcal{L} \subset R^m \) we have rank \( \mathcal{L} = \text{rank} \mathcal{L}^+ = m \).

Linear inequality systems are examined in the last section of this paper.

4. LINEAR INEQUALITY SYSTEMS

Let \( \{ a_1, \ldots, a_n, b \} \subset R^m \) be vectors, and let \( \{ e_1, \ldots, e_m \} \subset R^m \) be the standard base of \( R^m \). Denote \( J = \{ 1, \ldots, n \} \) the index set of vectors.
\{a_1, \ldots, a_n\}$, and by $I = \{1, \ldots, m\}$ the index set of the vectors $\{e_1, \ldots, e_m\}$.

We can state the following theorem (type of Farkas [12]).

**Theorem 4.1.** For the above given vectors exactly one of the following alternatives holds:

(a) There exists a basis $(J_B', I_B)$ as shown in Figure 12(a).

(b) There exist a basis $(J_B^\prime, I_B^\prime)$ as shown in Figure 12(b).

**Proof.** Using orthogonality properties it is easy to see that both cannot hold simultaneously. To show that at least one holds, let us consider first the alternatives of Theorem 3.1. If the second situation holds there, then also the second alternative holds here (in a slightly stronger form, as required). If the first case of Theorem 3.1 holds, then we have the tableau of Figure 13.

Let us perform the following

**Algorithm (α).**

I. If $b$ is a nonbasic element and the hatched part of Figure 13 contains only nonnegative coefficients, then, the case (a) of Theorem 4.1 applies. If there are negative coefficients in this area, then denote $r = \min\{i : i \in J_B$ and $a_i$ has a negative coefficient in the column of $b\}$. Make a pivot step: $a_r$ leaves, $b$ enters the basis.

II. If $b$ is a basic element and its row contains only nonpositive coefficients in the nonbasic columns, then case (b) applies.
If there are positive coefficients in this row, then denote $s = \min\{j : j \in \overline{J}_B$ and $a_j$ has a positive coefficient in the row of $b\}$. Make a pivot step: $a_s$ enters, $b$ leaves the basis.

The above algorithm stops at either of the required two cases, so to prove the theorem we have just to show that this algorithm is finite. Since there are a finite number of distinct bases, one has to prove only that cycling cannot occur.

**Lemma 4.1.** Algorithm $(\alpha)$ is finite.

*Proof.* Suppose to the contrary that cycling occurs. Denote $J^0 = \{j : a_j$ leaves the basis through the cycle$\}$. Note that vector $a_j$ enters the basis as well if $j \in J^0$ and vectors out of $J^0$ remain basic or nonbasic variables through the entire cycle.

Denote $p = \max\{j : j \in J^0\}$. Consider the two cases where $a_p$ leaves and enters the base. These situations are as shown in Figure 14. Leaving out the variables which are not in $J^0$ and denoting $a_p := -a_p$, we get the same tableaux as in Figure 12. We have seen that those two cases cannot hold simultaneously. Our proof is complete; the theorem is proved as well. 

A very important consequence of this theorem is the famous Farkas theorem. An algorithm also follows (to generate the alternatives) from our constructive proof.

**Corollary 4.1** (Farkas theorem [4, 5, 7]). Exactly one of the following
alternatives holds:

(a) the system $Ax = b$, $x \geq 0$ has a solution, or
(b) the system $yA \leq 0$, $yb = 1$ has a solution.

Proof. Let us apply the theorem for the column vectors $\{a_1, \ldots, a_n\}$ of the matrix $A$. If the first case holds [Figure 12(a)], then the vector $b$ is presented as a nonnegative linear combination of the vectors $\{a_1, \ldots, a_n\}$, that is, the system $Ax = b$, $x \geq 0$, is solved.

If the second case holds, then the tableau is as shown in Figure 15. By the orthogonality properties the vector $y$ of this tableau satisfies $yA \leq 0$, $yb = 1$. Our proof is complete.

Remark 1. The Farkas theorem is frequently used in the following form

\begin{align*}
J'_{B} & \quad J'_{B} & \quad J'_{B} & \quad J'_{B} \\
\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} & \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}
\end{align*}
(Gordan): Exactly one of the following alternatives holds:

(a) the system \( Ax = 0, cx = -1, x \geq 0 \), has a solution, or
(b) the system \( yA \leq c \) has a solution.

In this case Corollary 4.1 can be applied to the following vectors:

\[
a_1 = \begin{pmatrix} a_1 \\ -\gamma_1 \end{pmatrix}, \ldots, \quad a_n = \begin{pmatrix} a_n \\ -\gamma_n \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix},
\]

\[
e_1 = \begin{pmatrix} e_1 \\ 0 \end{pmatrix}, \ldots, \quad e_m = \begin{pmatrix} e_m \\ 0 \end{pmatrix}, \quad e_{m+1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\]

To prove the statement one has to notice only that the last equation of \( Ax = b \), gives the equality \( cx = -1 \), and the last coordinate of the vector \( y = (y, \eta) \) is \( \eta = 1 \).

**Remark 2.** Theorem 4.1 can also be stated in the following equivalent form: Exactly one of the following alternatives holds:

(a) There exists a basis \((J_B^B, I_B)\) as shown in Figure 16(a).
(b) There exist a basis \((J_B^B, I_B)\) as shown in Figure 16(b).

The first case is the same as it was in Theorem 4.1. From the second one, by a single pivot operation, we get the second case of Theorem 4.1. On the other hand, again a single pivot is necessary to get this form.

Note that a direct proof is also possible by a slight modification of (a)
Algorithm (α) ($b$ is always a nonbasic vector). Developing this algorithm is left to the reader. The modified algorithm is a special case of Terlaky's [14, 15] criss-cross method, and it is identical to the dual simplex method with Bland's [2] rule in a linear program where the coefficients of the objective function are identically zero, so that all the pivot operations are degenerate in the terminology of linear programming.

**Remark 3.** The following is equivalent to Theorem 4.1 as well: Exactly one of the following alternatives holds:

(a) There exist a basis ($J_B, b, I_B$) as shown in Figure 17(a).
(b) There exist a basis ($J_B'', b, I_B$) as shown in Figure 17(b).

The second case is the same as it was in Theorem 4.1. From the first one, by a single pivot operation case (a) of Theorem 4.1 is obtained. Conversely, again a single pivot is needed.

As in the previous case, a direct algorithmic proof is possible here as well ($b$ is kept in the basis). The resulting algorithm is the primal simplex algorithm with the minimal index rule, and so it is a special case of the criss-cross method.

**Remark 4.** It is easy to see that the Farkas theorem (and its Gordan form) follows from the alternative theorems formulated in Remarks 2 and 3. By equivalence this observation is obvious: one can get the solution vectors directly (without making any additional pivoting) from the final tableaux. Finding these solution vectors is left to the reader.
**Geometric Interpretation of Linear Inequality Systems**

**Definition 4.1.** The set of vectors \( \mathcal{C}(a_1, \ldots, a_n) = \{ b : b = \sum_{j=1}^{n} a_j \lambda_j, \quad 0 \leq \lambda_j, \quad j = 1, \ldots, n \} \) is called a *finite cone*, and the set \( \mathcal{C}^+(a_1, \ldots, a_n) = \{ y : y a_j \leq 0 \text{ for all } j = 1, \ldots, n \} \) is called the polar of the finite cone \( \mathcal{C}(a_1, \ldots, a_n) \).

It is obvious that for arbitrary vectors \( b \in \mathcal{C} \) and \( y \in \mathcal{C}^+ \) the relation \( by \leq 0 \) holds. Now we will present three famous theorems on finite cones as consequences of Theorem 4.1 [3–5, 7].

**Corollary 4.2.** If arbitrary vectors \( \{a_1, \ldots, a_n\} \subset \mathbb{R}^m \) are given, then there exist vectors \( \{y_1, \ldots, y_k\} \subset \mathbb{R}^m \) such that:

(a) \( \mathcal{C}(a_1, \ldots, a_n) = \mathcal{C}^+(y_1, \ldots, y_k) \) (Weyl),
(b) \( \mathcal{C}^+(a_1, \ldots, a_n) = \mathcal{C}(y_1, \ldots, y_k) \) (Minkowski),
(c) \( \mathcal{C}^{+\times}(a_1, \ldots, a_n) = \mathcal{C}(a_1, \ldots, a_n) \) (Farkas).

**Proof.** Consider all the basic tableaux where \( \|J_B\| \) is maximal, that is, the number of the vectors \( e_i \) is minimal in the basis, and there is an \( r \in J_B \) satisfying \( \tau_{rj} \leq 0 \) if \( j \in \bar{J}_B \) (Figure 18). Denote

\[
\{ y_1, \ldots, y_k \} = \{ \pm y^{(i)} : i \in I_B \} \cup \{ -y^{(r)} : \tau_{rj} \geq 0 \text{ for all } j \text{ in some tableau} \},
\]

or let it be the single zero vector if both vector sets on the right hand side are empty.

Note that \( \mathcal{L} = \mathcal{C}(\pm y^{(i)} : i \in I_B) \) is a subspace, \( \mathcal{L} = \mathcal{C}\{ -y^{(r)} : \tau_{rj} \geq 0 \text{ for all } j \text{ in some tableau} \} \) is a pointed cone (it does not contain any line), and \( \mathcal{C}\{ y_1, \ldots, y_k \} = \mathcal{C} + \mathcal{L} \).

![Figure 18](for arbitrary b)
(a): If \( b \in \mathcal{C}(a_1, \ldots, a_n) \), then \( b = \sum_{j=1}^{m} \lambda_j a_j \) for some \( \lambda_j \geq 0, j = 1, \ldots, m \).

Since \( y_i a_j \leq 0 \) for \( i = 1, \ldots, k, j = 1, \ldots, n \) we have \( y_i b = \sum_{j=1}^{n} (y_i a_j) \lambda_j \leq 0 \), that is, \( b \in \mathcal{C}^+(y_1, \ldots, y_k) \). On the other hand, if \( b \notin \mathcal{C}(a_1, \ldots, a_n) \), then (see Remark 2) \( y_i b > 0 \) for an \( i \in \{1, \ldots, k\} \), that is, \( b \notin \mathcal{C}^+(y_1, \ldots, y_k) \).

(b): If \( y \in \mathcal{C}(y_1, \ldots, y_k) \), then \( y = \sum_{i=1}^{k} \lambda_i y_i \), where \( \lambda_i \geq 0, i = 1, \ldots, k \).

Since \( y_i a_j \leq 0 \) for \( i = 1, \ldots, k, j = 1, \ldots, n \), we have \( y a_j = \sum_{i=1}^{n} (y_i a_j) \lambda_i \leq 0 \) for \( j = 1, \ldots, n, \lambda_i \geq 0, i = 1, \ldots, k \), that is, \( y \notin \mathcal{C}^+(a_1, \ldots, a_n) \). On the other hand, if \( y \notin \mathcal{C}(y_1, \ldots, y_k) \), then \( \sum_{i=1}^{k} \lambda_i y_i = y, \lambda_i \geq 0, i = 1, \ldots, k \) is inconsistent, then by Corollary 4.1 we have a vector \( b \in \mathbb{R}^m \) such that \( y_i b \leq 0 \) for \( i = 1, \ldots, k \) and \( yb = 1 \). So by (a) we have \( b \in \mathcal{C}^+(y_1, \ldots, y_k) = \mathcal{C}(a_1, \ldots, a_n) \), but \( yb \neq 1 \) implies that \( y \notin \mathcal{C}^+(a_1, \ldots, a_n) \).

(c): Let us apply first (a), then (b) as follows: \( \mathcal{C}(a_1, \ldots, a_n) = \mathcal{C}(y_1, \ldots, y_k) = \mathcal{C}^+(a_1, \ldots, a_n) \).

Our proof is complete.

Finally we show that the vectors \( \{y_1, \ldots, y_k\} \) [given by (5)] not only generate the cone \( \mathcal{C}(y_1, \ldots, y_k) = \mathcal{C}^+(a_1, \ldots, a_n) \), but this vector set is an extremal generating system of the cone \( \mathcal{C}(y_1, \ldots, y_k) \).

**Definition 4.2.** The set of vectors \( \{y_1, \ldots, y_k\} \) is an **extremal generating system** of the cone \( \mathcal{C} \) if \( \mathcal{C}(y_1, \ldots, y_k) = \mathcal{C} \) and for all \( j = 1, \ldots, k \) the system \( y_j = \sum_{i=1}^{k} \lambda_i y_i, \lambda_i \geq 0, i = 1, \ldots, k, i \neq j, \) is inconsistent.

In general \( \{y_1, \ldots, y_t\} \) is an extremal generating system of the cone \( \mathcal{C} \) if its **reduced** system \( \{y_1, \ldots, y_k\} \) is an extremal one, where reducing means that we leave out the vectors which are parallel to some other vector. (If \( y_i = \lambda y_j \), then we leave out \( y_j \), since if \( y_i \) is extremal then \( y_j \) is extremal as well.)

**Theorem 4.2.** The vector set \( \{y_1, \ldots, y_k\} \) is an extremal generating system of the cone \( \mathcal{C}(y_1, \ldots, y_k) = \mathcal{C}^+(a_1, \ldots, a_n) \).

**Proof.** According to the definition it is a generating system. By construction, the vectors \( \mathcal{L} = \mathcal{C}\{ \pm y^{(i)} : i \in I_n \} \) are extremals, since \( \{y^{(i)} : i \in I_n \} \) is a basis of the subspace \( \mathcal{L} \) and \( \mathcal{L} \cap \mathcal{C} = \emptyset \).

So we have to prove only that the set \( \{-y^{(r)} : \) the \( r \)th row of the tableau is nonnegative} = \( \{y_0, y_1, \ldots, y_t\} \) is extremal in \( \mathcal{C} \). We may assume that \( y_i \neq \lambda y_j, \) \( i, j = 0, 1, \ldots, t, i \neq j, \lambda > 0 \). Assume, to the contrary, that the set \( \{y_0, y_1, \ldots, y_t\} \) is not extremal. For convenience we may assume that \( y_0 \) is a
nonextremal element, that is, there are numbers $\lambda_1, \ldots, \lambda_t > 0$ such that

$$\sum_{i=1}^{t} \lambda_i y_i = y_0, \quad i = 1, \ldots, t. \quad (6)$$

The alternative pair of the system (6) is (see Corollary 4.1)

$$zy_i \leq 0, \quad i = 1, \ldots, t,$$

$$zy_0 > 0. \quad (7)$$

Now we show that (7) has a solution. Let $J_{B^0}$ be the base associated with $y_0$, that is, $a_j y_0 = 0$ for $j \in J_{B^0}$, $j \neq r$, and $y_0 a_r = -1$ for any $r$. Denote $z = \sum_{r \neq j \in J_{B^0}} a_j - \varepsilon a_r$, where $\varepsilon > 0$ is constant. Then

$$zy_0 = \sum_{r \neq j \in J_{B^0}} a_j y_0 - \varepsilon a_r y_0 = 0 - \varepsilon (-1) > 0$$

and

$$zy_i = \sum_{r \neq j \in J_{B^0}} a_j y_i - \varepsilon a_r y_i,$$

where

$$\sum_{r \neq j \in J_{B^0}} a_j y_i < 0,$$

since it is nonpositive, and if it were zero then $y_i$ would be parallel to $y_0$, contradicting our assumption. So $zy_i \leq 0$ if $0 < \varepsilon < \min \{(y_i \Sigma r \neq j \in J_{B^0} a_j)/(-y_i a_r): y_i a_r > 0, i = 1, \ldots, t\}$. That is, $z$ solves (7). This is a contradiction, so (6) is inconsistent, that is, \{y_0, y_1, \ldots, y_k\} is an extremal generating system. Our proof is complete.

Finally we note that the extremal elements of the original cone $\mathcal{C}(a_1, \ldots, a_n)$ can be obtained from pivot tableaux as well. A vector $a_r$ is an extremal element of the cone $\mathcal{C}$ if there is a pivot tableau $T(B)$ where $a_r$ is a basic vector and its row contains nonpositive coefficients in the nonbasic
columns. (See Figure 19.) The justification for this statement is left to the reader.

REFERENCES

3 V. Chvatal, Linear Programming, Freeman, 1983.

*Received 11 May 1987; final manuscript accepted 11 July 1990*