Solvability and Consistency for Linear Equations and Inequalities
Author(s): H. W. Kuhn
Source: *The American Mathematical Monthly*, Vol. 63, No. 4 (Apr., 1956), pp. 217-232
Published by: Taylor & Francis, Ltd. on behalf of the Mathematical Association of America
Stable URL: https://www.jstor.org/stable/2310345
Accessed: 17-12-2019 09:11 UTC

SOLVABILITY AND CONSISTENCY FOR LINEAR EQUATIONS AND INEQUALITIES*

H. W. KUHN, Byrn Mawr College

1. Introduction. The term *consistent* is in common use in two contexts that appear to be quite different at first sight. It is often applied to systems of linear equations as a synonym for *solvable*. Thus, Dickson says in [3]: "We shall call two or more equations *consistent* if there exist values of the unknowns which satisfy all of the equations." Again, Bôcher writes in [1]: "The equations may have no solution, in which case they are said to be *inconsistent*." Elsewhere, it is applied by logicians to deductive systems as a synonym for *non-contradictory*. Thus, Tarski defines the term in [8]: "A deductive theory is called *consistent* or non-contradictory if no two asserted statements of this theory contradict each other."

Our preliminary purpose is to reconcile these two usages, agreeing informally that "solvable" means "satisfiable" and that "consistent" means "non-contradictory." The reconciliation is brought about by setting forth explicitly the definition of consistency that has been employed implicitly in ordinary treatments of linear equations. This definition is based on a non-effective characterization of the logical consequences of the system, and is almost trivially equivalent to solvability under much more general conditions than are considered here. Therefore, this equivalence adds little or nothing to our knowledge of the special subject of linear equations.

These considerations are in sharp contrast with another use of consistency, which considers only those consequences that can be derived by applying a finite number of algebraic operations to the system. We shall show that the ordinary criteria for the solvability of systems of linear equations follow directly from the latter notion when the only algebraic operation that is allowed is that of forming linear combinations. Whatever novelty there is in this approach consists in viewing the system of equations as a set of postulates added to an underlying logic that includes the laws of real numbers, and then investigating the "methods of proof" appropriate to it.

The main object of this paper is to extend this formulation to systems of linear inequalities. It is somewhat remarkable that the same theorems persist with only minor modifications, namely, with reasonable care in the type of linear combinations that are allowed, and with the use of a process of elimination designed for linear inequalities. The criteria for solvability that are proved in this manner form a basis for the modern disciplines of linear programming and game theory. They also admit important geometric interpretations. These and other applications will be treated in a sequel.

* A preliminary version of this paper was prepared while the author was a consultant to the National Bureau of Standards. The preparation of the present version was supported, in part, by the Social Science Research Council and, in part, by the Office of Naval Research Logistics Project, Princeton University.

2. Systems of linear equations. The system of linear equations

$$
\begin{aligned}
3x - 6y + 4z &= 1\\
-x + 2y - 2z &= 3\\
x - 2y + z &= 0
\end{aligned}
$$

does not have a solution. Despite its triviality, let us examine an argument that might be advanced as a proof of this statement.

Suppose the equations are multiplied by new unknowns, u, v, and w, respectively, and added. Represent this by listing the multipliers to the left of the equations and the sum below, thus:

$$
\begin{array}{llll}
u: & 3x & -6y & +4z = 1\\
v: & -x & +2y & -2z = 3\\
w: & x & -2y & +z = 0\\
\hline
\multicolumn{4}{l}{(3u - v + w)x + (-6u + 2v - 2w)y + (4u - 2v + w)z = u + 3v}
\end{array}
$$

If the original system were satisfied by numbers $\bar{x}$, $\bar{y}$, and $\bar{z}$, then the same operations would yield

$$(3u - v + w)\bar{x} + (-6u + 2v - 2w)\bar{y} + (4u - 2v + w)\bar{z} = u + 3v$$

and this would be a true statement for all values of u, v, and w. However, if the multipliers $u=1$, $v=1$, $w=-2$ are chosen, this says

$$0\bar{x} + 0\bar{y} + 0\bar{z} = 4,$$

a false statement about numbers.

We shall see later how such multipliers can be found by successively eliminating the unknowns. For the moment it is important to remark that they were chosen to satisfy

$$
\begin{aligned}
3u - v + w &= 0\\
-6u + 2v - 2w &= 0\\
4u - 2v + w &= 0\\
u + 3v \quad &\neq 0
\end{aligned}
$$

and thus establish the contradiction. This is, of course, the original system transposed in a special way.

So much for motivation; the object of this section is to state precise conditions for the solvability of general *systems S* of linear equations. The systems will be assumed to contain r inhomogeneous equations in n unknowns, $x_1, \cdots, x_n$, and hence will have the form

$$
\text{(S)} \qquad \begin{array}{c} c_{11}x_1 + \cdots + c_{1n}x_n = c_1 \\ \cdots\cdots\cdots\cdots \\ c_{r1}x_1 + \cdots + c_{rn}x_n = c_r. \end{array}
$$

The c_{kl} and c_k are given real numbers (for $k=1, \cdots, r$ and $l=1, \cdots, n$) and define the system S. The adjective "inhomogeneous" means that no assumption is made about the right hand members c_k.

An indexed set $\overline{X}$ of real numbers $(\bar{x}_1, \cdots, \bar{x}_n)$ is called a *solution* for S if all of the equations

$$
c_{k1}\bar{x}_1 + \cdots + c_{kn}\bar{x}_n = c_k \qquad (k = 1, \cdots, r)
$$

are true statements. The set of solutions for S is denoted by $\mathfrak{S}(S)$ and S is called *solvable* if $\mathfrak{S}(S)$ is not empty. An equation

$$
d_1x_1 + \cdots + d_nx_n = d
$$

in the unknowns $x_1, \cdots, x_n$ is called a (logical) *consequence* of S if $d_1\bar{x}_1 + \cdots + d_n\bar{x}_n = d$ is a true equation whenever $\overline{X} = (\bar{x}_1, \cdots, \bar{x}_n)$ is a member of $\mathfrak{S}(S)$, in symbols, if $\overline{X} \in \mathfrak{S}(S)$. A curious situation occurs if S is not solvable, Then the definition of a consequence places no restriction on an equation $d_1x_1 + \cdots + d_nx_n = d$ because there are no solutions $\bar{x}_1, \cdots, \bar{x}_n$ to check for equality. Hence we are forced to conclude that, if S is not solvable, then *every* equation in the unknowns $x_1, \cdots, x_n$ is a consequence of S.

Our first definition of consistency will be based on the idea of consequence and the form of the patently false equation, $0\bar{x}+0\bar{y}+0\bar{z}=4$, derived in the example above. Namely, a system S is said to be *inconsistent* if some equation

$$
0x_1 + \cdots + 0x_n = d, \qquad \text{with } d \neq 0,
$$

is a consequence of S; otherwise it is called *consistent*. (Thus, the contradictory assertions used in this definition are $0=d$ and $d \neq 0$.)

The first theorem will establish the logical connection between solvability and this notion of consistency. Its statement explains immediately why these terms have often been confused (or used interchangeably) in textbook discussions of linear equations. The triviality of its proof reveals that it is a theorem without special content for linear equations.

THEOREM 1. *A system S is solvable if and only if it is consistent.*

Proof. If S is solvable, choose an $\overline{X} = (\bar{x}_1, \cdots, \bar{x}_n) \in \mathfrak{S}(S)$. Then $0\bar{x}_1 + \cdots + 0\bar{x}_n$ is equal to zero by the rules of operating with real numbers and hence $0\bar{x}_1 + \cdots + 0\bar{x}_n = d$ is not a true equation for any $d \neq 0$. Therefore, no $0x_1 + \cdots + 0x_n = d$, with $d \neq 0$, is a consequence of S, and S is consistent.

On the other hand, if S is not solvable, then every equation in $x_1, \cdots, x_n$ is a consequence of S. In particular, $0x_1 + \cdots + 0x_n = 1$ is a consequence of S, and S is inconsistent.

The significant content of the theory of linear equations comes not from the fact that an absurd consequence can always be exhibited when S is not solvable, but from the *form* of this equation and from *how* it can be obtained. To this end, form a scheme analogous to that used in the example, with multipliers appearing at the left of the system and the sum appearing below:

$$\begin{array}{rl} w_1: & c_{11}x_1 + \cdots + c_{1n}x_n = c_1 \\ w_2: & c_{21}x_1 + \cdots + c_{2n}x_n = c_2 \\ & \cdots\cdots\cdots\cdots \\ w_r: & \underline{c_{r1}x_1 + \cdots + c_{rn}x_n = c_r} \\ & d_1x_1 + \cdots + d_nx_n = d. \end{array}$$

The coefficients of the sum are easily read off; they are:

$$\begin{aligned} d_1 &= w_1c_{11} + \cdots + w_rc_{r1} \\ &\cdots\cdots\cdots\cdots \\ d_n &= w_1c_{1n} + \cdots + w_rc_{rn} \\ d &= w_1c_1 + \cdots + w_rc_r. \end{aligned}$$

An equation,

$$d_1x_1 + \cdots + d_nx_n = d,$$

that is formed in this manner, is called a *linear combination* of the equations of S. The multipliers $w_1, \cdots, w_r$ are called the *coefficients* of the linear combination.

With this set of definitions, the central theorem on the solvability of systems of inhomogeneous linear equations is:

THEOREM 2. (a) *Every linear combination of the equations of S is a consequence of S.* (b) *If S is solvable, then every consequence of S is a linear combination of the equations of S. If S is not solvable, then an equation*

$$0x_1 + \cdots + 0x_n = d, \qquad \textit{with } d \neq 0,$$

is a linear combination of the equations of S.

Several informal remarks may help to explain the content and intent of this theorem. First, it characterizes the consequences of *solvable* systems S as being exactly the linear combinations. Since "consequence" is a logical notion and "linear combination" is algebraic, this is quite remarkable. Often we may obtain the information that some statement holds whenever other statements are true; however, it is seldom that this forces such a narrow and explicit connection. Secondly, this theorem says that it is always possible to demonstrate the inconsistency of an unsolvable system by exhibiting as a linear combination an equation that is false for all values of the unknowns. Thus both parts of the theorem yield information through the medium of linear combinations.

The statements of Theorem 2 (b) can be strengthened to constructive assertions, that is, there is a finite computational procedure that will find the multipliers $w_1, \cdots, w_r$ in the linear combinations (and hence the solution to the transposed system) for every case. This procedure is based on the familiar technique for eliminating an unknown from the system.

THEOREM 3. *The process of elimination either yields a solution for S or exhibits an equation*

$$0x_1 + \cdots + 0x_n = d, \qquad \text{with } d \neq 0,$$

as a linear combination of the equations of S.

Although the technique of elimination is known to every algebra student in practice, it will be described here for the sake of comparison with the generalizations to follow. (Actually, "elimination" is something of a misnomer; for formal reasons, we will not eliminate an unknown, but will make all of its coefficients equal to zero.) It is applied only to systems S in which some coefficient c_{kl} is different from zero. For notational convenience, assume that $c_{11} \neq 0$, renumbering unknowns and equations if necessary. Then define a new system S' of r equations in the unknowns $x_1, \cdots, x_n$ by:

$$(S')\quad \begin{aligned} &0x_1 + 0x_2 + \cdots + 0x_n = 0 \\ &0x_1 + \left(c_{22} - \frac{c_{21}c_{12}}{c_{11}}\right)x_2 + \cdots + \left(c_{2n} - \frac{c_{21}c_{1n}}{c_{11}}\right)x_n = c_2 - \frac{c_{21}}{c_{11}}c_1 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &0x_1 + \left(c_{r2} - \frac{c_{r1}c_{12}}{c_{11}}\right)x_2 + \cdots + \left(c_{rn} - \frac{c_{r1}c_{1n}}{c_{11}}\right)x_n = c_r - \frac{c_{r1}}{c_{11}}c_1. \end{aligned}$$

The formation of this system follows an obvious rule; it is obtained by subtracting c_{k1}/c_{11} times the first equation from the kth equation for $k=1, \cdots, r$. Thus the coefficient of x_1 in the kth sum equation is $c_{k1} - c_{k1}c_{11}/c_{11} = 0$ for $k = 1, \cdots, r$. The seemingly redundant first equation, $0x_1 + \cdots + 0x_n = 0$, is retained to avoid the logical complications involved in considering void systems. The proof of Theorem 3 is then based on the following three assertions:

1° *Every equation of S′ is a linear combination of the equations of S.*
2° *If S′ is solvable, then S is solvable.*
3° *More unknowns have all zero coefficients in S′ than in S.*

The details of the proof are straight-forward and will not be given. Theorem 2 (b) follows immediately from Theorem 3; however, the details of this derivation are also omitted because we shall prove a more general statement for inequalities in Section 3. At this stage, presenting the proofs might obscure the simplicity of the logical structure of the results. To emphasize this structure, the theorems will be reproduced in a condensed form to arm the reader for the section that follows.

The core of this section is contained in the diagram:

$$
\begin{array}{r|ccccc|c}
\text{solution} \; \rightarrow & & & & & & \\
\text{contradiction} \; \downarrow & x_1 & x_2 & \cdots & x_n & & \\
\hline
w_1 & c_{11} & c_{12} & \cdots & c_{1n} & & = c_1 \\
w_2 & c_{21} & c_{22} & \cdots & c_{2n} & & = c_2 \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
w_r & c_{r1} & c_{2r} & \cdots & c_{rn} & & = c_r \\
\hline
 & = 0 & = 0 & \cdots & = 0 & & \neq 0
\end{array}
$$

It presents a schematic expression of the statement that exactly one of the two systems

$$
(S) \qquad \begin{array}{c}
c_{11}x_1 + \cdots + c_{1n}x_n = c_1 \\
\cdots\cdots\cdots\cdots \\
c_{r1}x_1 + \cdots + c_{rn}x_n = c_r
\end{array}
$$

and

$$
(T) \qquad \begin{array}{c}
w_1c_{11} + \cdots + w_rc_{r1} = 0 \\
\cdots\cdots\cdots\cdots \\
w_1c_{1n} + \cdots + w_rc_{rn} = 0 \\
w_1c_1 \;\; + \cdots + w_rc_r \;\; \neq 0
\end{array}
$$

is solvable. A solution to T means that S is inconsistent. Successive elimination of $x_1, \cdots, x_n$ leads either to a solution for S or to a solution for T.

The relation of solvability to consistency for linear equations should now be clear. Any confusion is caused by two different uses of the word consistency. The first (which we have called "consistency") might be called "consistency with respect to consequences" (following Church [2]) and is the notion that is used interchangeably with "solvability" because Theorem 1 is valid. The second might be called "consistency with respect to linear combinations" and leads to the following formulation of the results: Forming a linear combination is a rule of inference (*i.e.*, never leads from true statements to a false statement) for systems of linear equations. If a system S is solvable then every consequence can be derived in this manner. If S is inconsistent with respect to consequences, then it is inconsistent with respect to linear combinations. Elimination is an effective procedure for deciding whether a system is solvable or inconsistent (in either sense).

3. Systems of linear inequalities. The object of this section is to establish conditions for the solvability of general systems S of linear inequalities. We shall pattern our definitions upon those given for equations and, happily, the same theorems will be found to hold with only slight and obvious modifications.

The systems S under consideration are assumed to contain $p+q>0$ inequalities in the unknowns $x_1, \cdots, x_n$; of these $p \geqq 0$ are assumed to be strict. Thus, they can be written

$$(S)\qquad \begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &> a_i \qquad (i = 1, \cdots, p) \\ b_{j1}x_1 + \cdots + b_{jn}x_n &\geqq b_j \qquad (j = 1, \cdots, q) \end{aligned}$$

where the a_{ik}, a_i, b_{jk}, and b_j $(i=1, \cdots, p; j=1, \cdots, q; k=1, \cdots, n)$ are given real numbers that define the system S. Again, no assumption is made concerning the right hand members a_i and b_j, and the inequalities are called inhomogeneous.

An indexed set $\overline{X}$ of real numbers $(\bar{x}_1, \cdots, \bar{x}_n)$ is called a *solution* for S if all of the inequalities

$$\begin{aligned} a_{i1}\bar{x}_1 + \cdots + a_{in}\bar{x}_n &> a_i \qquad (i = 1, \cdots, p) \\ b_{j1}\bar{x}_1 + \cdots + b_{jn}\bar{x}_n &\geqq b_j \qquad (j = 1, \cdots, q) \end{aligned}$$

are true statements. The set of solutions for S is denoted by $\mathfrak{S}(S)$, and is possibly empty. An inequality,

$$d_1x_1 + \cdots + d_nx_n \mathfrak{R} d,$$

where $\mathfrak{R}$ is one of the relations $>$ or $\geqq$, in the unknowns $x_1, \cdots, x_n$ is called a *consequence* of S if $d_1\bar{x}_1 + \cdots + d_n\bar{x}_n \mathfrak{R} d$ is a true inequality wherever $\bar{x} = (\bar{x}_1, \cdots, \bar{x}_n)$ is a member of $\mathfrak{S}(S)$. If S is not solvable then *every* inequality in the unknowns $x_1, \cdots, x_n$ is a consequence of S. A system S is said to be *inconsistent* if the inequality

$$0x_1 + \cdots + 0x_n > 0$$

is a consequence of S; otherwise, it is called *consistent*. The choice of this "standard" contradiction will be explained below.

The logical connection between solvability and consistency holds without change; to emphasize the strict parallel, this theorem and those following it will be given roman numerals corresponding to their counterparts for equations.

THEOREM I. *A system S is solvable if and only if it is consistent.*

Proof. If S is solvable, choose any $\overline{X} = (\bar{x}_1, \cdots, \bar{x}_n) \in \mathfrak{S}(S)$. Then $0\bar{x}_1 + \cdots + 0\bar{x}_n$ is equal to zero by the rules for operating with real numbers and hence $0\bar{x}_1 + \cdots + 0\bar{x}_n > 0$ is not a true inequality. Hence $0x_1 + \cdots + 0x_n > 0$ is not a consequence of S and S is consistent.

On the other hand, if S is not solvable, then every inequality in $x_1, \cdots, x_n$

is a consequence of S. In particular, $0x_1 + \cdots + 0x_n > 0$ is a consequence of S and S is inconsistent.

The significant content of the theory of linear inequalities is found again in the form of the consequences that can be derived from S and the manner of their derivation, namely, as linear combinations. We must exercise some care in forming these, but no more than is indicated by the following rules for manipulating inequalities.

RULE 1. Any inequality (strict $(>)$ or ordinary $(\geqq)$) still holds if it is multiplied by a *positive* number throughout.

RULE 2. Strict inequality $(>)$ holds for the sum of two similarly directed inequalities if and only if strict inequality holds in at least one of the summands. Ordinary inequality $(\geqq)$ can always be asserted for the sum.

RULE 3. Strict inequality $(>)$ always implies ordinary inequality $(\geqq)$.

These three rules make possible a precise description of the nature of the linear combinations that will be used. Rule 1 suggests that the multipliers should be restricted to be *non-negative* (a zero multiplier means that that inequality is not being used). Rule 2 says that the relation holding for the linear combination can be $>$ only if some strict inequality has a positive coefficient. Rules 2 and 3 say that the relation $\geqq$ can always be asserted for the linear combination. Following these precepts, form a multiplier scheme patterned on the scheme used for equations, with *non-negative* multipliers at the left and the sum below:

$$\begin{array}{rl}
u_0 \geqq 0: & 0x_1 + \cdots + \quad 0x_n > -1 \\
u_1 \geqq 0: & a_{11}x_1 + \cdots + a_{1n}x_n > a_1 \\
\cdots & \cdots\cdots\cdots\cdots\cdots \\
u_p \geqq 0: & a_{p1}x_1 + \cdots + a_{pn}x_n > a_p \\
v_1 \geqq 0: & b_{11}x_1 + \cdots + b_{1n}x_n \geqq b_1 \\
\cdots & \cdots\cdots\cdots\cdots\cdots \\
v_q \geqq 0: & b_{q1}x_1 + \cdots + b_{qn}x_n \geqq b_q \\
\hline
 & d_1x_1 + \cdots + d_nx_n \mathcal{R} d
\end{array}.$$

The insertion of the first line is related to the choice of a standard inconsistent inequality. The coefficients of the sum are easily calculated to be:

$$\begin{aligned}
d_1 &= u_1a_{11} + \cdots + u_pa_{p1} + v_1b_{11} + \cdots + v_qb_{q1} \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
d_n &= u_1a_{1n} + \cdots + u_pa_{pn} + v_1b_{1n} + \cdots + v_qb_{qn} \\
d &= -u_0 + u_1a_1 + \cdots + u_pa_p + v_1b_1 + \cdots + v_qb_q.
\end{aligned}$$

An inequality,

$$d_1x_1 + \cdots + d_nx_n \mathcal{R} d,$$

that is formed in this manner from a system S, is called a *legal linear combination* of the inequalities of S provided that $\mathfrak{R}$ is $\geqq$, *or* $\mathfrak{R}$ is $>$ and some u_i is positive ($i=0, 1, \cdots, p$). The reader should notice that this definition has been framed so as to conform to Rules 1, 2, and 3 and thus to insure the fact that every legal linear combination is a consequence. Another important fact to be noted is that every inequality in the multiplier scheme above can be proved to be a legal linear combination by choosing its multiplier equal to one and all other multipliers equal to zero. In particular, the definitions make $0x_1+ \cdots +0x_n>-1$ a legal linear combination of every system.

With these definitions, the central theorem on the solvability of systems of inhomogeneous linear inequalities is:

THEOREM II. (a) *Every legal linear combination of the inequalities of S is a consequence of S.* (b) *If S is solvable then every consequence of S is a legal linear combination of the relations of S. If S is not solvable, then the inequality*

$$0x_1 + \cdots + 0x_n > 0$$

is a legal linear combination of the relations of S.

Proof. Theorem II (a) follows immediately from the definitions and Rules 1, 2, and 3. As before, the statements of Theorem II (b) will be strengthened to constructive assertions; that is, elimination is an effective procedure for finding the multipliers $u_0, u_1, \cdots, u_p, v_1, \cdots, v_q$ in the legal linear combinations for every case. Precisely, the proof of Theorem II (b) will be based on

THEOREM III. *The process of elimination either yields a solution for S or exhibits the inequality* $0x_1+ \cdots +0x_n>0$ *as a legal linear combination of the inequalities of S.*

Proof. The process of elimination is applied only to systems S in which some coefficient a_{il} or b_{jl} is different from zero. Assume that $l=1$, renumbering unknowns if necessary, and set about eliminating x_1 from the system. Separate the inequalities of the system into three classes (Classes I, II, III) according to whether the coefficient of x_1 is positive, negative, or zero. These classes will be distinguished notationally by the number of primes on the indices. Define a new system S' of inequalities in the unknowns $x_1, \cdots, x_n$ by:

$$0x_1 + \qquad 0x_2 + \cdots + \qquad 0x_n > -1$$

$$(S') \qquad 0x_1 + \left(\frac{a_{i'2}}{a_{i'1}} - \frac{a_{i''2}}{a_{i''1}}\right) x_2 + \cdots + \left(\frac{a_{i'n}}{a_{i'1}} - \frac{a_{i''n}}{a_{i''1}}\right) x_n > \frac{a_{i'}}{a_{i'1}} - \frac{a_{i''}}{a_{i''1}}$$

for *all* pairs i' and i'' with $a_{i'1}>0$ and $a_{i''1}<0$,

$$0x_1 + \left(\frac{a_{i'2}}{a_{i'1}'} - \frac{b_{j''2}}{b_{j''1}}\right) x_2 + \cdots + \left(\frac{a_{i'n}}{a_{i'1}} - \frac{b_{j''n}}{b_{j''1}}\right) x_n > \frac{a_{i'}}{a_{i'1}} - \frac{b_{j''}}{b_{j''1}}$$

for *all* pairs i' and j'' with $a_{i'1}>0$ and $b_{j''1}<0$, (S' *continues on next page.*)

$$(S' \textit{ cont.}) \quad 0x_1 + \left(\frac{b_{j'2}}{b_{j'1}} - \frac{a_{i''2}}{a_{i''1}}\right) x_2 + \cdots + \left(\frac{b_{j'n}}{b_{j'1}} - \frac{a_{i''n}}{a_{i''1}}\right) x_n > \frac{b_{j'}}{b_{j'1}} - \frac{a_{i''}}{a_{i''1}}$$

for *all* pairs j' and i'' with $b_{j'1} > 0$ and $a_{i''1} < 0$,

$$0x_1 + \left(\frac{b_{j'2}}{b_{j'1}} - \frac{b_{j''2}}{b_{j''1}}\right) x_2 + \cdots + \left(\frac{b_{j'n}}{b_{j'1}} - \frac{b_{j''n}}{b_{j''1}}\right) x_n \geqq \frac{b_{j'}}{b_{j'1}} - \frac{b_{j''}}{b_{j''1}}$$

for *all* pairs j' and j'' with $b_{j'1} > 0$ and $b_{j''1} < 0$,

$$0x_1 + \qquad a_{i'''2}x_2 + \cdots + \qquad a_{i'''n}x_n > a_{i'''}$$

for *all* i''' with $a_{i'''1} = 0$,

$$0x_1 + \qquad b_{j'''2}x_2 + \cdots + \qquad b_{j'''n}x_n \geqq b_{j'''}$$

for *all* j''' with $b_{j'''1} = 0$.

Although it may seem formidable at first glance, the formation of the system S' follows a simple rule. Namely, the coefficient of x_1 in each inequality in Class I is made $+1$ through the use of the *positive* multipliers $1/a_{i'1}$ and $1/b_{j'1}$, while the coefficient of x_1 in each inequality of Class II is made -1 through the use of the *positive* multipliers $-1/a_{i''1}$ and $-1/b_{j''1}$. By Rule 1, the inequalities are all preserved. Then, *all* pairs of inequalities, one from Class I and one from Class II are added and the inequality type of the sum is assigned by Rule 2. The inequalities of Class III are copied, unchanged; the first inequality, $0x_1 + \cdots + 0x_n > -1$, is inserted in S' to avoid the logical complications involved in considering the empty system that would arise if all of the inequalities of S were in Class I or II. This characterization of the formation of S' proves:

1° *Every inequality of S' is a legal linear combination of the inequalities of S, and hence is a consequence of S.*

Another manner of viewing the formation of S' is revealed by isolating x_1 on one side of all of the inequalities of Classes I and II, then pairing the results as follows:

$$\frac{a_{i''}}{a_{i''1}} - \frac{a_{i''2}}{a_{i''1}} x_2 - \cdots - \frac{a_{i''n}}{a_{i''1}} x_n > x_1 > \frac{a_{i'}}{a_{i'1}} - \frac{a_{i'2}}{a_{i'1}} x_2 - \cdots - \frac{a_{i'n}}{a_{i'1}} x_n$$

for *all* pairs i' and i'' with $a_{i'1} > 0$ and $a_{i''1} < 0$,

$$\frac{b_{j''}}{b_{j''1}} - \frac{b_{j''2}}{b_{j''1}} x_2 - \cdots - \frac{b_{j''n}}{b_{j''1}} x_n \geqq x_1 > \frac{a_{i'}}{a_{i'1}} - \frac{a_{i'2}}{a_{i'1}} x_2 - \cdots - \frac{a_{i'n}}{a_{i'1}} x_n$$

for *all* pairs i' and j'' with $a_{i'1} > 0$ and $b_{j''1} < 0$,

$$\frac{a_{i''}}{a_{i''1}} - \frac{a_{i''2}}{a_{i''1}} x_2 - \cdots - \frac{a_{i''n}}{a_{i''1}} x_n > x_1 \geqq \frac{b_{j'}}{b_{j'1}} - \frac{b_{j'2}}{b_{j'1}} x_2 - \cdots - \frac{b_{j'n}}{b_{j'1}} x_n$$

for *all* pairs j' and i'' with $b_{j'1}>0$ and $a_{i''1}<0$,

$$\frac{b_{j''}}{b_{j''1}} - \frac{b_{j''2}}{b_{j''1}}x_2 - \cdots - \frac{b_{j''n}}{b_{j''1}}x_n \geqq x_1 \geqq \frac{b_{j'}}{b_{j'1}} - \frac{b_{j'2}}{b_{j'1}}x_2 - \cdots - \frac{b_{j'n}}{b_{j'1}}x_n$$

for *all* pairs j' and j'' with $b_{j'1}>0$ and $b_{j''1}<0$.

(Again, the reader should not ignore the possibility that either Class I or Class II might be empty, in which case there would be no such pairs; he should check that each assertion holds in this case, too.) If x_1 is dropped from each of the pairs and the result arranged with unknowns on the left and constants on the right, and if the inequalities of Class III and $0x_1+\cdots+0x_n>-1$ are adjoined to the system, the result is exactly S'.

Suppose $\bar{x}_2, \cdots, \bar{x}_n$ solve S' (note that x_1 can be given any value in this system). Then, arranging S' as in the previous paragraph, the inequalities that are not in S are:

$$(i', i''):\quad \frac{a_{i''}}{a_{i''1}} - \frac{a_{i''2}}{a_{i''1}}\bar{x}_2 - \cdots - \frac{a_{i''n}}{a_{i''1}}\bar{x}_n > \frac{a_{i'}}{a_{i'1}} - \frac{a_{i'2}}{a_{i'1}}\bar{x}_2 - \cdots - \frac{a_{i'n}}{a_{i'1}}\bar{x}_n$$

$$(i', j''):\quad \frac{b_{j''}}{b_{j''1}} - \frac{b_{j''2}}{b_{j''1}}\bar{x}_2 - \cdots - \frac{b_{j''n}}{b_{j''1}}\bar{x}_n > \frac{a_{i'}}{a_{i'1}} - \frac{a_{i'2}}{a_{i'1}}\bar{x}_2 - \cdots - \frac{a_{i'n}}{a_{i'1}}\bar{x}_n$$

$$(j', i''):\quad \frac{a_{i''}}{a_{i''1}} - \frac{a_{i''2}}{a_{i''1}}\bar{x}_2 - \cdots - \frac{a_{i''n}}{a_{i''1}}\bar{x}_n > \frac{b_{j'}}{b_{j'1}} - \frac{b_{j'2}}{b_{j'1}}\bar{x}_2 - \cdots - \frac{b_{j'n}}{b_{j'1}}\bar{x}_n$$

$$(j', j''):\quad \frac{b_{j''}}{b_{j''1}} - \frac{b_{j''2}}{b_{j''1}}\bar{x}_2 - \cdots - \frac{b_{j''n}}{b_{j''1}}\bar{x}_n \geqq \frac{b_{j'}}{b_{j'1}} - \frac{b_{j'2}}{b_{j'1}}\bar{x}_2 - \cdots - \frac{b_{j'n}}{b_{j'1}}\bar{x}_n.$$

The problem of completing $\bar{x}_2, \cdots, \bar{x}_n$ to a solution for S by a proper choice of $\bar{x}_1$ is clearly that of fitting $\bar{x}_1$ into these inequalities as required by the system immediately above them. With an eye to picking out those inequalities of S that restrict the choice of $\bar{x}_1$ most severely, define

$$\alpha' = \max_{i'} \left\{\frac{a_{i'}}{a_{i'1}} - \frac{a_{i'2}}{a_{i'1}}\bar{x}_2 - \cdots - \frac{a_{i'n}}{a_{i'n}}\bar{x}_n\right\},$$

$$\beta' = \max_{j'} \left\{\frac{b_{j'}}{b_{j'1}} - \frac{b_{j'2}}{b_{j'1}}\bar{x}_2 - \cdots - \frac{b_{j'n}}{b_{j'1}}\bar{x}_n\right\},$$

$$\alpha'' = \min_{i''} \left\{\frac{a_{i''}}{a_{i''1}} - \frac{a_{i''2}}{a_{i''1}}\bar{x}_2 - \cdots - \frac{a_{i''n}}{a_{i''1}}\bar{x}_n\right\},$$

$$\beta'' = \min_{j''} \left\{\frac{b_{j''}}{b_{j''1}} - \frac{b_{j''2}}{b_{j''1}}x_2 - \cdots - \frac{b_{j''n}}{b_{j''1}}x_n\right\}.$$

(We shall adopt the convention that a maximum over an empty set is $-\infty$ while a minimum over an empty set is $+\infty$. Thus, if there are no inequalities in Class I, $\alpha'=-\infty$, *etc.*) It is important to realize these maxima and minima are actually achieved for a proper choice of the indices i', j', i'', and j'' if the sets concerned are non-empty, and so among the inequalities immediately above one can find:

$$\alpha''>\alpha', \qquad \beta''>\alpha', \qquad \alpha''>\beta', \quad \text{and} \quad \beta''\geqq\beta',$$

(where the inequalities are trivially true if the sets are empty and the convention is applied).

If $\bar{x}_1>\alpha'$, $\bar{x}_1\geqq\beta'$, $\bar{x}_1<\alpha''$, *and* $\bar{x}_1\leqq\beta''$, *then* $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n$ *is a solution for* S.

Proof. Since the inequalities in S of Class III are also in S', only those of Classes I and II need be checked. For these,

$$\begin{aligned} a_{i'1}\bar{x}_1+a_{i'2}\bar{x}_2+\cdots+a_{i'n}\bar{x}_n &> a_{i'1}\alpha'+a_{i'2}\bar{x}_2+\cdots+a_{i'n}\bar{x}_n \\ &\geqq (a_{i'}-a_{i'2}\bar{x}_2-\cdots-a_{i'n}\bar{x}_n)+(a_{i'2}\bar{x}_2+\cdots+a_{i'n}\bar{x}_n)=a_{i'} \end{aligned}$$

for all i' with $a_{i'1}>0$,

$$\begin{aligned} a_{i''1}\bar{x}_1+a_{i''2}\bar{x}_2+\cdots+a_{i''n}\bar{x}_n &> a_{i''1}\alpha''+a_{i''2}\bar{x}_2+\cdots+a_{i''n}\bar{x}_n \\ &\geqq (a_{i''}-a_{i''2}\bar{x}_2-\cdots-a_{i''n}\bar{x}_n)+(a_{i''2}\bar{x}_2+\cdots+a_{i''n}\bar{x}_n)=a_{i''} \end{aligned}$$

for all i'' with $a_{i''1}<0$,

$$\begin{aligned} b_{j'1}\bar{x}_1+b_{j'2}\bar{x}_2+\cdots+b_{j'n}\bar{x}_n &> b_{j'1}\beta'+b_{j'2}\bar{x}_2+\cdots+b_{j'n}\bar{x}_n \\ &\geqq (b_{j'}-b_{j'2}\bar{x}_2-\cdots-b_{j'n}\bar{x}_n)+(b_{j'2}\bar{x}_2+\cdots+b_{j'n}\bar{x}_n)=b_{j'} \end{aligned}$$

for all j' with $b_{j'1}>0$,

$$\begin{aligned} b_{j''1}\bar{x}_1+b_{j''2}\bar{x}_2+\cdots+b_{j''n}\bar{x}_n &\geqq b_{j''1}\beta''+b_{j''2}\bar{x}_2+\cdots+b_{j''n}\bar{x}_n \\ &\geqq (b_{j''}-b_{j''2}\bar{x}_2-\cdots-b_{j''n}\bar{x}_n)+(b_{j''2}\bar{x}_2+\cdots+b_{j''n}\bar{x}_n)=b_{j''} \end{aligned}$$

for all j'' with $b_{j''1}<0$. This proves that S is solved by $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n$.

To choose an $\bar{x}_1$ satisfying these conditions, consider the four exhaustive cases, corresponding to the four possible orderings of α' with β' and α'' with β'', separately:

CASE 1. $\beta'\leqq\alpha'$ and $\alpha''\leqq\beta''$.
Choose $\bar{x}_1$ so that $\alpha'<\bar{x}_1<\alpha''$.

CASE 2. $\beta'\leqq\alpha'$ and $\beta''<\alpha''$.
Choose $\bar{x}_1$ so that $\alpha'<\bar{x}_1<\beta''$.

CASE 3. $\alpha'<\beta'$ and $\alpha''\leqq\beta''$.
Choose $\bar{x}_1$ so that $\beta'<\bar{x}_1<\alpha''$.

CASE 4. $\alpha'<\beta'$ and $\beta''<\alpha''$.
Choose $\bar{x}_1$ so that $\beta'\leqq\bar{x}_1\leqq\beta''$.

This proves the crucial fact:

2° *If* S' *is solvable, then* S *is solvable.*

Suppose that the process of elimination is applied several, say h, times to yield successively the systems S, S', S'', $\cdots$, $S^{(h)}$. It is clear that 1° and 2° still hold, with $S^{(h)}$ replacing S'. The only detail that needs verifying, namely, that legal linear combinations of legal linear combinations are legal linear combinations, is obvious and tedious to write out. The termination of the process is insured by a third property of S' relative to S.

3° *More unknowns have all zero coefficients in S' than in S.*

This is clear, since any unknown with all zero coefficients in S still has all zero coefficients in S' while x_1 had a non-zero coefficient in S (some a_{i1} or b_{j1} was assumed non-zero) and has all zero coefficients in S'.

Property 3° insures that the process of elimination must end after a finite number of steps (no more than n) with a system $S^{(h)}$ in which all of the unknowns have all zero coefficients. If all the right hand members of the strict (or ordinary) inequalities are negative (or non-positive) then any set of numbers $\overline{X}=(\bar{x}_1, \cdots, \bar{x}_n)$ solves $S^{(h)}$ and hence S is solvable by h applications of 2°. Otherwise, $0x_1+\cdots+0x_n>d$ with some $d\geqq 0$, or $0x_1+\cdots+0x_n\geqq d$ with some $d>0$ appears in $S^{(h)}$. In this case, the proof is completed by writing down the two multiplier schemes:

$$\begin{array}{rl} 1\colon & 0x_1+\cdots+0x_n > d, \quad \text{with any } d\geqq 0 \\ d\colon & 0x_1+\cdots+0x_n > -1 \\ \hline & 0x_1+\cdots+0x_n > 0 \end{array}$$

$$\begin{array}{rl} 1\colon & 0x_1+\cdots+0x_n \geqq d, \quad \text{with any } d > 0 \\ d\colon & 0x_1+\cdots+0x_n > -1 \\ \hline & 0x_1+\cdots+0x_n > 0 \end{array}$$

(The reader should verify that the result is a legal linear combination in each case.) Hence, $0x_1+\cdots+0x_n>0$ is a legal linear combination of $S^{(h)}$ and can be exhibited as a legal linear combination of the inequalities of S (and possibly $0x_1+\cdots+0x_n>-1$) by at most $h+1$ applications of 1°. This completes the proof of Theorem III.

Before returning to the proof of Theorem II (b), these results will be summarized in a multiplier scheme. This diagram expresses the fact that *exactly* one of the two systems listed there admits a solution. A solution to T means that S is inconsistent. Successive elimination of $x_1, \cdots, x_n$ leads either to a solution for S or to a solution for T.

solution → / contradiction ↓		x_1	x_2	$\cdots$	x_n	
	$u_0 \geqq 0$	0	0	$\cdots$	0	> -1
not	$u_1 \geqq 0$	a_{11}	a_{12}	$\cdots$	a_{1n}	$> a_1$
all	$u_2 \geqq 0$	a_{21}	a_{22}	$\cdots$	a_{2n}	$> a_2$
zero	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	$u_p \geqq 0$	a_{p1}	a_{p2}	$\cdots$	a_{pn}	$> a_p$
	$v_1 \geqq 0$	b_{11}	b_{12}	$\cdots$	b_{1n}	$\geqq b_1$
	$v_2 \geqq 0$	b_{21}	b_{22}	$\cdots$	b_{2n}	$\geqq b_2$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	$v_q \geqq 0$	b_{q1}	b_{q2}	$\cdots$	b_{qn}	$\geqq b_q$
		$= 0$	$= 0$	$\cdots$	$= 0$	$= 0$

The system S is defined by

$$a_{i1}x_1 + \cdots + a_{in}x_n > a_i \qquad (i = 1, \cdots, p)$$
$$b_{j1}x_1 + \cdots + b_{jn}x_n \geqq b_j \qquad (j = 1, \cdots, q).$$

The system T is defined by

$$u_1a_{1l} + \cdots + u_pa_{pl} + v_1b_{1l} + \cdots + v_qb_{ql} = 0 \qquad (l = 1, \cdots, n)$$
$$-u_0 + u_1a_2 + \cdots + u_pa_p + v_1b_1 + \cdots + v_qb_q = 0$$

with all of the unknowns $u_0, u_1, \cdots, u_p, v_1, \cdots, v_q$ non-negative and not all of the $u_0, u_1, \cdots, u_p$ equal to zero.

Proof of Theorem II(b). The case for systems without solutions is stated directly in Theorem III and so it is only necessary to show that, for solvable systems, every consequence is a legal linear combination. Assume that the strict inequality $d_1x_1 + \cdots + d_nx_n > d$ is a consequence of S. Thus, if $\overline{X} = (\bar{x}_1, \cdots, \bar{x}_n)$ is any solution for S, then $d_1\bar{x}_1 + \cdots + d_n\bar{x}_n$ is a number that is *larger* than d. This means the *insolvability* of the system U (to save rewriting the system, multipliers have been included):

$$(U)\qquad \begin{aligned} u_0 \geqq 0:&\quad 0x_1 + \cdots + 0x_n > -1 \\ u_i \geqq 0:&\quad a_{i1}x_1 + \cdots + a_{in}x_n > a_i \qquad (i = 1, \cdots, p) \\ v_j \geqq 0:&\quad b_{j1}x_1 + \cdots + b_{jn}x_n \geqq b_j \qquad (j = 1, \cdots, q) \\ t \geqq 0:&\quad -d_1x_1 - \cdots - d_nx_n \geqq -d. \end{aligned}$$

By Theorem III, there exist non-negative multipliers $\bar{u}_0, \bar{u}_1, \cdots, \bar{u}_p, \bar{v}_1, \cdots, \bar{v}_q, \bar{t}$ with not all of $\bar{u}_0, \bar{u}_1, \cdots, \bar{u}_p$ equal to zero, such that

$$\bar{x}_l: \qquad \bar{u}_1 a_{1l} + \cdots + \bar{u}_p a_{pl} + \bar{v}_1 b_{1l} + \cdots + \bar{v}_q b_{ql} - \bar{t}d = 0 \quad (l = 1, \cdots, n)$$
$$-1: \quad -\bar{u}_0 + \bar{u}_1 a_1 + \cdots + \bar{u}_p a_p + \bar{v}_1 b_1 + \cdots + \bar{v}_q b_q - \bar{t}d = 0.$$

If $\bar{t}$ were known to be positive, then all of these equations could be divided by it, establishing $d_1x_1 + \cdots + d_nx_n > d$ as a legal linear combination of the inequalities of S. Recall that S is assumed to be solvable and let $\bar{x}_1, \cdots, \bar{x}_n$ be a solution. These have been written as multipliers at the left of the last system, using -1 to multiply the last equation. Multiplying and adding,

$$\bar{u}_0 + \sum_i \bar{u}_i(a_{i1}\bar{x}_1 + \cdots + a_{in}\bar{x}_n - a_i) + \sum_j \bar{v}_j(b_{j1}\bar{x}_1 + \cdots + b_{jn}\bar{x}_n - b_j)$$
$$= \bar{t}(d_1\bar{x}_1 + \cdots + d_n\bar{x}_n - d).$$

Since the left side of this equation is positive, the right side is also, and hence $\bar{t} > 0$. Hence $d_1x_1 + \cdots + d_nx_n > d$ is a legal linear combination of the inequalities of S with multipliers $\bar{u}_0/\bar{t}, \bar{u}_1/\bar{t}, \cdots, \bar{u}_p/\bar{t}, \bar{v}_1/\bar{t}, \cdots, \bar{v}_q/\bar{t}$. If the consequence is an ordinary inequality $d_1x_1 + \cdots + d_nx_n \geqq d$, a few changes must be made in the proof. The inequality $-d_1x_1 - \cdots - d_nx_n \geqq -d$ in U should be replaced by $-d_1x_1 - \cdots - d_nx_n > -d$. Using the same notation for multipliers, the same transposed system of equations is obtained, with not all of $\bar{u}_0, \bar{u}_1, \cdots, \bar{u}_p$, *and* $\bar{t}$ equal to zero. If $\bar{t}$ is assumed zero, we are in the previous case (with not all of $\bar{u}_0, \bar{u}_1, \cdots, \bar{u}_p$ equal to zero) and can conclude $\bar{t} > 0$ as before, a contradiction. Hence $\bar{t}$ is positive and $d_1x_1 + \cdots + d_nx_n \geqq d$ is a legal linear combination of the inequalities of S with multipliers $\bar{u}_0/\bar{t}, \bar{u}_1/\bar{t}, \cdots, \bar{u}_p/\bar{t}, \bar{v}_1/\bar{t}, \cdots, \bar{v}_q/\bar{t}$. This completes the proof of Theorem II (b).

4. An infinite example. Of course, the results of Section 3 can be paraphrased in logical terms exactly in the form of the conclusion to Section 2. At this point the reader may be led to the erroneous belief that consistency with respect to consequences is always equivalent to consistency with respect to linear combinations for linear systems. The following (infinite) system shows that this is not so.

$$(S) \qquad x > -1/n \qquad (n = 1, 2, 3, \cdots).$$

Clearly $x \geqq 0$ is a consequence of this system, yet is not a linear combination of any number of inequalities of S. We infer that

$$(S_1) \qquad -x > 0, \quad x > -1/n \qquad (n = 1, 2, 3, \cdots)$$

is inconsistent with respect to consequences but is consistent with respect to linear combinations. The natural way out of this dilemma is to add new rules of inference based on limiting operations.

5. Acknowledgments and historical remarks. The author is indebted to T. S. Motzkin, who first suggested that his "transposition theorem" [6] might be viewed as asserting the disjoint alternatives of solvability or contradiction *via* linear combination. He also called attention to a remark of Fourier [5] that elimination was a natural method for solving linear inequalities. However, the

author takes full responsibility for the logical consequences of these ideas.

Elimination, as a method for solving linear inequalities, has been used for theoretical purposes by a number of authors, notably by Dines [4]. However, his treatment seems to conceal rather than emphasize the parallel with equations.

The distinction between logical consequences and provable consequences for logistic systems was first made by Tarski [7]. Our definition of the logical consequences of a system of linear equations or inequalities is in an obvious way parallel to Tarski's definition of logical consequence, but the two are not the same. They might be compared in the following manner. The logical consequences, in the sense of this paper, of a system of equations or inequalities are the same as the logical consequences, in Tarski's sense, of the system together with some categorical system of postulates for the real numbers—provided that the unknowns in the equations or inequalities, before considering the consequences in Tarski's sense, are first replaced by new symbols, which play the role of primitive constants, and which do not appear elsewhere. The author is indebted to A. Church for unraveling the relation between the two ideas; the general metatheorem parallel to Theorems 1 and I, and valid for functional calculi of all orders is stated by Church in [2].

References

[1] Bôcher, M. Introduction to Higher Algebra, New York, Macmillan, 1915.

[2] Church, A. Introduction to Mathematical Logic, Princeton University Press, in press.

[3] Dickson, L. E. New First Course in the Theory of Equations, New York, John Wiley and Sons, 1939.

[4] Dines, Lloyd L. Systems of linear inequalities, Annals of Mathematics, 2nd Series, vol. 20, 1918–19, pp. 191–199.

[5] Fourier, J. J.-B. Oeuvres II, Paris, 1890.

[6] Motzkin, T. S. Beiträge zur Theorie der linearen Ungleichungen (Dissertation, Basel, 1933) Jerusalem, 1936.

[7] Tarski, A. Über den Begriff der logischen Folgerung, Actes du Congrès International de Philosophie Scientifique VII, Paris, 1936.

[8] Tarski, A. Introduction to Logic and to the Methodology of Deductive Sciences, New York, Oxford University Press, 1941.