THE BELL SYSTEM
TECHNICAL JOURNAL

Volume 48 | January 1969 | Number 1

Some Theorems on Properties of DC Equations of Nonlinear Networks

By I. W. SANDBERG and A. N. WILLSON, JR.

(Manuscript received July 3, 1968)

Several results are presented concerning the equation $F(x) + Ax = B$ (with $F(\cdot)$ a "diagonal" nonlinear mapping of real Euclidean n-space E^n into itself, and A a real $n \times n$ matrix) which plays a central role in the dc analysis of transistor networks. In particular, we give necessary and sufficient conditions on A such that the equation possesses a unique solution x for each real n-vector B and each strictly monotone increasing $F(\cdot)$ that maps E^n onto itself.

There are several direct circuit-theoretic implications of the results. For example, we show that if the short-circuit admittance matrix G of the linear portion of the dc model of a transistor network satisfies a certain dominance condition, then the network cannot be bistable. Therefore, a fundamental restriction on the G matrix of an interesting class of switching circuits is that it must violate the dominance condition.

I. INTRODUCTION

For each positive integer n let $\mathfrak{F}^n$ denote that collection of mappings of the real n-dimensional Euclidean space E^n onto itself, defined by: $F \in \mathfrak{F}^n$ if and only if there exist, for $i = 1, \cdots, n$, strictly monotone increasing functions f_i mapping E^1 onto E^1 such that, for each $x \equiv (x_1, \cdots, x_n)^t \in E^n$, $F(x) = (f_1(x_1), \cdots, f_n(x_n))^t$.

The main purpose of this paper is to report on some results concerning properties of the equation

$$F(x) + Ax = B, \tag{1}$$

where A is an $n \times n$ matrix of real numbers, F maps E^n into E^n, and $B \in E^n$. In particular, a condition to be satisfied by A is given which is both necessary and sufficient to guarantee that for each $F \in \mathfrak{F}^n$ and each $B \in E^n$ there exists a unique solution of equation (1).

We also study the problem of obtaining bounds on the solution of equation (1). These bounds show that (if $F \in \mathfrak{F}^n$ and our condition on A is satisfied) the solution depends continuously on B. The bounds are often of use in computing the solution by standard iteration methods such as the Newton-Raphson method. By appealing to a theorem of R. S. Palais it is shown that the bounds can also be used to obtain a theorem essentially the same as, but somewhat weaker than, our principal result.

Several results can be found in the literature which specify sufficient conditions for the existence of a unique solution of equation (1). For example, if A is positive semidefinite then a special case of a theorem of Ref. 1 guarantees the existence of a unique solution of equation (1) for all those $F \in \mathfrak{F}^n$ which have the property that the slope of each f_i is bounded from above and below by positive constants, and for all $B \in E^n$. This theorem also specifies that a certain iteration scheme will always converge to the solution.

A theorem of G. J. Minty[2], when applied to equation (1), also implies essentially the same result. The boundedness condition on the slopes of the functions f_i is not required by Minty's theorem. On the other hand, Minty's theorem does not provide a procedure for computing the solution of equation (1).

In Ref. 3 it is proved that a sufficient condition for the existence and uniqueness of a solution of equation (1) for all $F \in \mathfrak{F}^n$ and $B \in E^n$ is that A satisfy a weak row-sum dominance condition:

$$a_{ii} \geqq \sum_{j \neq i} |a_{ij}|, \qquad i = 1, \cdots, n.^*$$

Other information concerning the location and the computation of the solution is also given in Ref. 3.

The class of matrices satisfying the condition of our theorem (which is defined in Section III and denoted by P_0) includes all positive semidefinite matrices as well as all matrices which satisfy any one of several

* Appendix A contains a simpler proof of a similar result and a proof of a new related result. These results specify convergent algorithms for obtaining the solution.

dominance conditions. Many other matrices are included in P_0; and since the condition of our theorem is both a necessary and a sufficient one, we are assured that P_0 is the largest class of matrices A for which equation (1) has a unique solution for all $F \varepsilon \mathfrak{F}^n$ and all $B \varepsilon E^n$.

II. NONLINEAR NETWORKS

Equation (1) is often encountered in the study of nonlinear electrical networks. In the case of networks containing only resistors (that is, linear resistors with nonnegative resistance), dependent and independent sources, and two-terminal nonlinear resistors that are described by functions in $\mathfrak{F}^1$ (diodes, for example), this is rather obvious.[3] Even for networks which contain more general nonlinear devices, however, equation (1) can often provide a convenient characterization. For example, D. A. Calahan shows in his recent book that the transistor network of Fig. 1 may be described by the equation

$$\begin{pmatrix} I_{es}(e^{qV_r/kT} - 1) \\ I_{cs}(e^{qV_f/kT} - 1) \end{pmatrix} + \begin{bmatrix} 0.0225 & 0.309 \\ -0.168 & 0.494 \end{bmatrix} \begin{pmatrix} V_r \\ V_f \end{pmatrix} = \begin{pmatrix} 0.00177\,V_{cc} \\ -0.188\,V_{cc} \end{pmatrix}$$

if the Ebers–Moll model is used to represent the transistor. (See pp. 13ff of Ref. 4.) In this equation I_{es}, I_{cs}, q, k, T, and V_{cc} all represent fixed real parameters. It is quite trivial to apply the theory of this paper (in particular, Corollary 3 of Section IV) to Calahan's example and prove that this equation has a solution, the solution is unique, and the solution depends continuously on V_{cc}. We also show how bounds on the solution can be obtained.

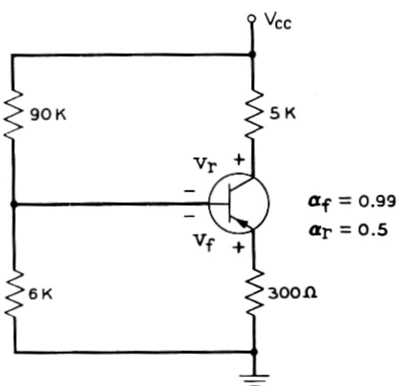

Fig. 1 — Biased transistor-stage.

More generally, it is frequently the case that networks which contain transistors, as well as the previously mentioned linear and nonlinear elements, may be described by the equation

$$TF(x) + Ax = B. \tag{2}$$

In this case, x is a vector whose components are the voltages across the nonlinear resistors and the transistor base-emitter and base-collector voltages. The $n \times n$ matrix A is the y-parameter matrix of the linear n-port network which is obtained by removing all nonlinear resistors and transistors and setting the value of each independent source to zero. The function $TF(x)$ describes the behavior of the nonlinear resistors and the transistors. It happens that the matrix T is nonsingular; therefore equation (2) can be put into the form of equation (1).

Networks which contains inductors and capacitors as well as the memoryless elements already mentioned are of course described by differential equations. Even the study of such networks, however, can often lead to the consideration of equations of the same type as equation (1). One usually finds the solution of such an equation is necessary, for example, when computing the solution of the differential equations by using some implicit numerical integration formula.

The problem of determining the equilibrium states of the abovementioned dynamic networks is one in which the consideration of equations of type (1) often arises in perhaps a more direct manner. In this regard, if it happens that equation (1) has a unique solution, then the network cannot possibly be bistable.

When the determination of equilibrium states of a transistor network leads first to the consideration of equation (2), then as a rather direct application of our existence and uniqueness theorem it follows that if the matrix A satisfies a weak column-sum dominance condition,

$$a_{ii} \geqq \sum_{j \neq i} | a_{ji} |, \qquad i = 1, \cdots, n,$$

then $T^{-1}A \in P_0$ and hence the network has exactly one equilibrium state. This result and related results which are proved in Section IV have the following interesting corollary: One cannot synthesize a bistable network which consists of resistors, inductors, capacitors, diodes, independent voltage and current sources, and one (Ebers–Moll modeled) transistor—or even an arbitrary number of (Ebers–Moll modeled) transistors with a common base connection.

The authors feel that in many respects the main contributions of this paper are in the techniques used to prove the results. For this reason, we have not chosen to summarize all of the results at the outset and relegate proofs to later sections. But rather, the results and the proofs will appear in the order in which they will best illustrate the techniques developed.

III. MATRICES OF CLASSES P AND P_0

The following notation will be used throughout the remainder of the paper: The origin in E^n will be denoted by θ. If D is a diagonal matrix then $D > 0$ ($D \geqq 0$) means that each element of D on the main diagonal is positive (nonnegative).

In Ref. 5 and Ref. 6 M. Fiedler and V. Pták define the classes of matrices denoted by P and P_0. They in fact prove that the following properties of a square matrix A are equivalent:

(*i*) All principal minors of A are positive.

(*ii*) For each vector $x \neq \theta$ there exists an index k such that $x_k y_k > 0$ where $y = Ax$.

(*iii*) For each vector $x \neq \theta$ there exists a diagonal matrix $D_x > 0$ such that the scalar product $\langle Ax, D_x x\rangle > 0$.

(*iv*) For each vector $x \neq \theta$ there exists a diagonal matrix $H_x \geqq 0$ such that $\langle Ax, H_x x\rangle > 0$.

(*v*) Every real eigenvalue of A, as well as of each principal submatrix of A, is positive.

The class of all matrices satisfying one of the above conditions is denoted by P. Fiedler and Pták prove that the following properties of a square matrix A are also equivalent:

(*i*) All principal minors of A are nonnegative.

(*ii*) For each vector $x \neq \theta$ there exists an index k such that $x_k \neq 0$ and $x_k y_k \geqq 0$ where $y = Ax$.

(*iii*) For each vector $x \neq \theta$ there exists a diagonal matrix $D_x \geqq 0$ such that $\langle x, D_x x\rangle > 0$ and $\langle Ax, D_x x\rangle \geqq 0$.

(*iv*) Every real eigenvalue of A, as well as of each principal submatrix of A, is nonnegative.

The class of all matrices satisfying one of the above conditions is denoted by P_0.

The following theorems follow directly from the above definitions.

Theorem 1. If $A \in P_0$ then for every diagonal matrix $\Delta \geqq 0$ ($\Delta > 0$), $\Delta + A \in P_0$ ($\Delta + A \in P$).

Proof: Let $x \neq \theta$. Then, since $A \in P_0$, there exists an index k such that $x_k \neq 0$ and $x_k(Ax)_k \geqq 0$. Thus, $x_k(\Delta x + Ax)_k \geqq 0$ (>0). □

In particular, Theorem 1 implies that if $A \in P_0$ and $\Delta \geqq 0$ ($\Delta > 0$) then det $(\Delta + A) \geqq 0$ (>0).

Theorem 2. If $A \in P$ then $A^{-1} \in P$.

Proof: Suppose $A \in P$. Let $x \neq \theta$ be given and let $y = A^{-1}x$. $y \neq \theta$ since A^{-1} is nonsingular. Thus, there exists a diagonal matrix $D > 0$ such that $\langle Ay, Dy \rangle > 0$, which implies $\langle x, DA^{-1}x \rangle > 0$, or $\langle Dx, A^{-1}x \rangle > 0$, or $\langle A^{-1}x, Dx \rangle > 0$. That is, for every $x \neq \theta$ there exists $D > 0$ such that $\langle A^{-1}x, Dx \rangle > 0$. Hence $A^{-1} \in P$. □

Because of the similarity of the definitions of the classes of matrices P and P_0, one might conjecture that this proposition is also true: *If $A \in P_0$, and det $A \neq 0$, then $A^{-1} \in P_0$.* This conjecture is in fact true. Interestingly enough, however, its proof is not obtained as one might at first suspect, by simply modifying the proof of Theorem 2. Moreover, the proof of this conjecture does not even seem to follow directly from any of the above definitions of P_0. Rather, upon making the trivial observation that for every diagonal matrix $D > 0$, det $(A^{-1} + D) =$ det $(A^{-1}) \cdot$det $(D^{-1} + A) \cdot$det (D), the conjecture is easily seen to follow from the fact that det $(D + A) \neq 0$ for every diagonal $D > 0$ if and only if $A \in P_0$. This fact is a direct corollary to the proof of Theorem 3.

IV. EXISTENCE AND UNIQUENESS THEOREM

The following theorem is the principal result of this paper.

Theorem 3. If A is an $n \times n$ matrix then there exists a unique solution of equation (1) *for each $F \in \mathfrak{F}^n$ and for each $B \in E^n$ if and only if $A \in P_0$.*

Proof: *(if)* Let $A \in P_0$, $F \in \mathfrak{F}^n$, and $B \in E^n$. The solution of equation (1) is then unique (if it exists) since if x and y are both solutions then, using the strict monotonicity property of F, there exists a diagonal matrix $D > 0$ such that $F(x) - F(y) = D(x - y)$. But $[D + A](x - y) = \theta$ and, by Theorem 1, $D + A$ is nonsingular. This means that $x = y$.

We prove the existence of a solution of equation (1) by induction. For $k = 1, \cdots, n$, let

$$F_k(x) = \begin{bmatrix} f_1(x_1) \\ \vdots \\ f_k(x_k) \end{bmatrix}, \qquad A_k = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \cdot & \cdot & \cdot \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}, \qquad B_k = \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix}.$$

Clearly, $A_k \in P_0$, $F_k \in \mathfrak{F}^k$, and $B_k \in E^k$. Also, it is clear that there exists a unique solution of $F_1(x) + A_1x = B_1$ for each $F_1 \in \mathfrak{F}^1$ and for each $B_1 \in E^1$, and that this solution is a continuous function of b_1.

Assume that there exists a unique solution of $F_k(x)+A_kx=B_k$ for each $F_k \varepsilon \mathfrak{F}^k$, $B_k \varepsilon E^k$, and that this solution depends continuously on any scalar parameter η upon which B_k depends continuously. Let the matrices $A_{k,k+1}$ and $A_{k+1,k}$ be defined by

$$A_{k,k+1} = \begin{bmatrix} a_{1,k+1} \\ \vdots \\ a_{k,k+1} \end{bmatrix},$$

$$A_{k+1,k} = [a_{k+1,1} \cdots a_{k+1,k}].$$

Then, for every real number x_{k+1}, the equation

$$F_k(x) + A_kx + A_{k,k+1}x_{k+1} = B_k \tag{3}$$

has a (unique) solution which is a continuous function of x_{k+1} and of η. Let the components of this solution be denoted by $x_i = m_i(x_{k+1}, \eta)$, for $i = 1, \cdots, k$, and define the vector $M_k(x_{k+1}, \eta)$ by $M_k = (m_1, \cdots, m_k)^t$.

We now prove that the function

$$\varphi(x_{k+1}, \eta) \equiv A_{k+1,k}M_k(x_{k+1}, \eta) + a_{k+1,k+1}x_{k+1} - b_{k+1}(\eta)$$

is monotone increasing in x_{k+1}: Let $x_{k+1}^1, x_{k+1}^2 \varepsilon E^1$ with $x_{k+1}^1 < x_{k+1}^2$. Then, if M^1 (M^2) denotes the solution of equation (3) when $x_{k+1} = x_{k+1}^1$ (x_{k+1}^2), we have

$$F_k(M^2) - F_k(M^1) + A_k(M^2 - M^1) + A_{k,k+1}(x_{k+1}^2 - x_{k+1}^1) = \theta.$$

Because of the strict monotonicity of the function F_k, however, there exists a $k \times k$ diagonal matrix $\Delta > 0$ such that

$$F_k(M^2) - F_k(M^1) = \Delta(M^2 - M^1).$$

Hence,

$$M^2 - M^1 = -[\Delta + A_k]^{-1}A_{k,k+1}(x_{k+1}^2 - x_{k+1}^1).$$

Thus,

$$\varphi(x_{k+1}^2) - \varphi(x_{k+1}^1) = \{a_{k+1,k+1} - A_{k+1,k}[\Delta + A_k]^{-1}A_{k,k+1}\}(x_{k+1}^2 - x_{k+1}^1).$$

But then, from the easily verified relation

$$a_{k+1,k+1} - A_{k+1,k}[\Delta + A_k]^{-1}A_{k,k+1} = \frac{\det\left\{\begin{bmatrix} \Delta & \begin{matrix} 0 \\ \vdots \end{matrix} \\ 0 \cdots & 0 \end{bmatrix} + A_{k+1}\right\}}{\det(\Delta + A_k)}$$

and since

$$\det(\Delta + A_k) > 0, \qquad \det\left(\begin{bmatrix} \Delta & \begin{matrix} 0 \\ \vdots \end{matrix} \\ 0 \cdots & 0 \end{bmatrix} + A_{k+1}\right) \geqq 0 \text{ (Theorem 1)},$$

and $x_{k+1}^2 - x_{k+1}^1 > 0$, it follows that $\varphi(x_{k+1}^2) \geqq \varphi(x_{k+1}^1)$.

Now since φ is monotone increasing and, obviously, continuous in x_{k+1}, it follows that the left side of the equation

$$f_{k+1}(x_{k+1}) + \varphi(x_{k+1}) = 0 \tag{4}$$

is a strictly monotone increasing function mapping E^1 onto E^1, and hence equation (4) has a unique solution. If x_{k+1}^0 denotes this solution then

$$x^0 \equiv \begin{bmatrix} m_1(x_{k+1}^0) \\ \vdots \\ m_k(x_{k+1}^0) \\ x_{k+1}^0 \end{bmatrix}$$

is the (unique) solution of

$$F_{k+1}(x) + A_{k+1}x = B_{k+1} .$$

We must now prove that this solution is a continuous function of any scalar parameter η upon which B_{k+1} depends continuously. It suffices to prove that x_{k+1} depends continuously on η (see equation (3)). This may be done as follows:

Let x_{k+1}^0 be the solution of equation (4) corresponding to $\eta = \eta^0$. That is, let

$$f_{k+1}(x_{k+1}^0) + \varphi(x_{k+1}^0 , \eta^0) = 0,$$

and let $\epsilon > 0$ be given. Since f_{k+1} is a strictly monotone increasing mapping of E^1 onto E^1, so is f_{k+1}^{-1}, and hence f_{k+1}^{-1} is continuous. Hence, there exists $\delta' > 0$ such that if $| f_{k+1}(x_{k+1}^0) - f_{k+1}(x_{k+1}) | < \delta'$ then $| x_{k+1}^0 - x_{k+1} | < \epsilon$. Since φ is a continuous function of η, there exists $\delta > 0$ such that $| \eta^0 - \eta | < \delta$ implies $| \varphi(x_{k+1}^0 , \eta^0) - \varphi(x_{k+1}^0 , \eta) | < \delta'$. If $| \eta^0 - \eta | < \delta$, and

$$f_{k+1}(x_{k+1}) + \varphi(x_{k+1} , \eta) = 0,$$

then,

$$f_{k+1}(x_{k+1}^0) - f_{k+1}(x_{k+1}) + \varphi(x_{k+1}^0, \eta) - \varphi(x_{k+1}, \eta)$$
$$= -[\varphi(x_{k+1}^0, \eta^0) - \varphi(x_{k+1}^0, \eta)].$$

But since both f_{k+1} and φ are monotone increasing in x_{k+1},

$$(x_{k+1}^0 - x_{k+1})[f_{k+1}(x_{k+1}^0) - f_{k+1}(x_{k+1})] \geqq 0,$$

and

$$(x_{k+1}^0 - x_{k+1})[\varphi(x_{k+1}^0, \eta) - \varphi(x_{k+1}, \eta)] \geqq 0.$$

Therefore,

$$| (x_{k+1}^0 - x_{k+1})[f_{k+1}(x_{k+1}^0) - f_{k+1}(x_{k+1})] |$$
$$\leqq | (x_{k+1}^0 - x_{k+1})[\varphi(x_{k+1}^0, \eta^0) - \varphi(x_{k+1}^0, \eta)] |.$$

Now, if $x_{k+1}^0 = x_{k+1}$ then of course $| x_{k+1}^0 - x_{k+1} | < \epsilon$. Otherwise,

$$| f_{k+1}(x_{k+1}^0) - f_{k+1}(x_{k+1}) | \leqq | \varphi(x_{k+1}^0, \eta^0) - \varphi(x_{k+1}^0, \eta) |.$$

But then,

$$| f_{k+1}(x_{k+1}^0) - f_{k+1}(x_{k+1}) | < \delta',$$

and hence $| x_{k+1}^0 - x_{k+1} | < \epsilon$. Thus, x_{k+1} is a continuous function of η.

(*only if*) Suppose $A \notin P_0$. If $\det A < 0$ then for sufficiently small $\zeta > 0$, $\det (\zeta I + A) < 0$. For sufficiently large ζ, however,

$$\det (\zeta I + A) = \zeta^n \cdot \det \left(I + \frac{1}{\zeta} A\right) > 0.$$

Thus, since $\det (\zeta I + A)$ is a continuous function of ζ, there is some value of $\zeta > 0$ such that $\det (\zeta I + A) = 0$. For this value of ζ let $F(x) = \zeta I x$. Clearly, for this choice of $F \in \mathfrak{F}^n$, equation (1) cannot have a unique solution.

If $\det A \geqq 0$, but A has a negative principal minor, we can still find a diagonal matrix $\Delta > 0$ such that $\det (\Delta + A) = 0$; however, in this case Δ will not, in general, be simply the identity matrix multiplied by a positive constant ζ.

For some positive integer $k < n$ let A have a $k \times k$ principal minor which is negative and let

$$\Delta^{(1)} = \operatorname{diag} [\delta_1, \cdots, \delta_n].$$

Since the determinant of $\Delta + A$ is not altered if any two rows and then the corresponding pair of columns are interchanged we may, without loss of generality, assume that the matrix A is partitioned as

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix},$$

where A_1 is a $k \times k$ matrix with $\det A_1 < 0$. Let $\xi > 0$ be chosen so small that $\det (\xi I + A_1) < 0$, and let $\delta_1 = \cdots = \delta_k = \xi$. Now, if $\delta_{k+1} = \cdots = \delta_n = \zeta > 0$, then

$$\det (\Delta^{(1)} + A) = \det \begin{bmatrix} \xi I + A_1 & A_2 \\ A_3 & \zeta I + A_4 \end{bmatrix}$$

$$= \zeta^{n-k} \cdot \det \begin{bmatrix} \xi I + A_1 & A_2 \\ \frac{1}{\zeta} A_3 & I + \frac{1}{\zeta} A_4 \end{bmatrix}.$$

Thus, for $\zeta > 0$ chosen to be sufficiently large, $\det (\Delta^{(1)} + A) < 0$. ($\det (\Delta^{(1)} + A) \to \zeta^{n-k} \cdot \det (\xi I + A_1) < 0$ as $\zeta \to \infty$.) Now, if for $\eta > 0$, $\Delta^{(2)} = \eta I$, then it is clear that for η chosen sufficiently large, $\det [\Delta^{(2)} + A] = \eta^n \cdot \det (I + (1/\eta)A) > 0$. Thus, if

$$\Delta(\epsilon) = \epsilon \Delta^{(1)} + (1 - \epsilon)\Delta^{(2)},$$

it is clear, since $\det [\Delta(0) + A] > 0$ and $\det [\Delta(1) + A] < 0$ and since $\det [\Delta(\epsilon) + A]$ is a continuous function on $0 \leqq \epsilon \leqq 1$, that there is a value of $\epsilon > 0$ $(0 < \epsilon < 1)$ such that $\det [\Delta(\epsilon) + A] = 0$. For this value of ϵ, $\Delta(\epsilon) > 0$ is the required diagonal matrix. □

Notice that our proof shows that if $F \in \mathfrak{F}^n$ and $A \in P_0$, then the solution of equation (1) depends continuously on any scalar parameter upon which B depends continuously. The arguments of Section V show, under these assumptions on F and A, that the operator $(F + A)^{-1}$ is in fact a continuous map of E^n into itself.

In the proof of Theorem 3 we see that the uniqueness of the solution follows simply from the hypotheses that each f_i is strictly monotone increasing and that $A \in P_0$. The additional hypotheses that each f_i is continuous and maps E^1 *onto* E^1 are not necessary (continuity of each f_i is not explicitly hypothesized, but follows from the "monotonicity" and "onto" hypotheses). Hence, we have:

Corollary 1. If, for $i = 1, \cdots, n$, S_i *is a subset of* E^1, *and if* $S = S_1 \times \cdots \times S_n$, *and if* $F(x) = (f_1(x_1), \cdots, f_n(x_n))'$, *where each* f_i *maps* E^1 *into* E^1 *and is strictly monotone increasing on* S_i, *then if* $A \in P_0$ *and* $B \in E^n$, *there exists at most one solution of equation* (1) *in* S.

We now prove another interesting corollary of Theorem 3. We first define some additional notation.

For each positive integer n let $\mathcal{S}^n$ denote the collection of all subsets of E^n defined by: $S \varepsilon \mathcal{S}^n$ if and only if $S = S^1 \times \cdots \times S^n$ where, for $i = 1, \cdots, n$, $S^i \subset E^1$ and S^i has the same cardinality as E^1. For each $S \subset \mathcal{S}^n$ we define the collection $\mathfrak{F}^n(S)$ of functions mapping S onto E^n by: $F \varepsilon \mathfrak{F}^n(S)$ if and only if there exist, for $i = 1, \cdots, n$, strictly monotone increasing functions f_i mapping S^i onto E^1 such that for each $x \varepsilon S^n$, $F(x) = (f_1(x_1), \cdots, f_n(x_n))^t$.

Corollary 2. If A is an $n \times n$ matrix and the collection $\mathfrak{F}^n(S)$ is non-empty then there exists a unique solution of the equation

$$F_1(x) + AF_2(x) = B \tag{5}$$

for each $F_1 \varepsilon \mathfrak{F}^n(S)$, $F_2 \varepsilon \mathfrak{F}^n(S)$, and each $B \varepsilon E^n$ if and only if $A \varepsilon P_0$.

Proof: Since $F_2 \varepsilon \mathfrak{F}^n(S)$, $F_2^{-1} : E^n \to S$ exists and $F_1 \circ F_2^{-1} \varepsilon \mathfrak{F}^n$. Thus, there exists a unique solution of equation (5) if and only if there exists a unique solution of

$$F_1(F_2^{-1}(y)) + Ay = B. \quad \square$$

As special cases of Corollary 2 we have: there exists a unique solution of each of the equations

$$F_1(x) + AF_2(x) = B,$$

and

$$x + AF(x) = B,$$

for each F_1, F_2, $F \varepsilon \mathfrak{F}^n$ and each $B \varepsilon E^n$ if and only if $A \varepsilon P_0$.

In Theorem 3 (and Corollary 2) the hypothesis that each of the functions f_i is an onto mapping is quite necessary in order to guarantee the *existence* of a solution for each $A \varepsilon P_0$. In the following example all of the hypotheses of Theorem 3 except this one are satisfied:

$$e^{x_1} + x_1 - x_2 = 1$$

$$e^{x_2} - x_1 + x_2 = -2.$$

It is of course impossible for these equations to have a solution since, by adding both sides, we find that the solution would have to satisfy

$$e^{x_1} + e^{x_2} = -1,$$

which is absurd.

Even though the functions f_i are not "onto," it is still possible to specify sufficient conditions for the existence of a unique solution of equation (5) [and equation (1)] by strengthening the hypothesis on the matrix A—namely, by requiring that $A \in P$. This is the essence of Corollary 3. We first require additional notation.

With $\mathcal{S}^n$ defined as above, we define, for each $S \in \mathcal{S}^n$, the collection of functions $\mathfrak{F}_0^n(S)$ mapping S into E^n by: $F \in \mathfrak{F}_0^n(S)$ if and only if there exist, for $i = 1, \cdots, n$, monotone increasing functions f_i mapping S^i onto a connected set in E^1 such that, for each $x \in S$, $F(x) = (f_1(x_1), \cdots, f_n(x_n))^t$. When $S = E^n$ we denote $\mathfrak{F}_0^n(S)$ by $\mathfrak{F}_0^n$.

Corollary 3. If A is an $n \times n$ matrix then there exists a unique solution of equation (5) *for each $F_1 \in \mathfrak{F}_0^n(S)$, $F_2 \in \mathfrak{F}^n(S)$, or $F_1 \in \mathfrak{F}^n(S)$, $F_2 \in \mathfrak{F}_0^n(S)$, and for each $B \in E^n$, if $A \in P$.*

Proof: If $F_2 \in \mathfrak{F}^n(S)$, $F_2^{-1} : E^n \to S$ exists and $F_1 \circ F_2^{-1} \in \mathfrak{F}_0^n$. Thus, in this case, there exists a unique solution of equation (5) if there exists a unique solution of

$$F_1(F_2^{-1}(y)) + Ay = B. \tag{6}$$

Now, since $A \in P$, it follows from the fact that the determinant of a matrix is a continuous function of each of its elements, that there is a matrix $A^* \in P \subset P_0$ and an $\epsilon > 0$, such that $A = \epsilon I + A^*$. Hence, equation (6) is equivalent to

$$F(y) + A^*y = B, \tag{7}$$

where we have defined

$$F(y) \equiv F_1(F_2^{-1}(y)) + \epsilon Iy.$$

But, since $F_1 \circ F_2^{-1} \in \mathfrak{F}_0^n$ and $\epsilon I \in \mathfrak{F}^n$, it follows that $F \in \mathfrak{F}^n$. Therefore, since $A^* \in P_0$, equation (7) and hence equation (6) and hence equation (5) have unique solutions.

The case when $F_1 \in \mathfrak{F}^n(S)$ and $F_2 \in \mathfrak{F}_0^n(S)$ can be reduced to the case just considered by making the simple observations that, in this case, equation (5) has a unique solution if

$$A^{-1}F_1(x) + F_2(x) = A^{-1}B$$

has a unique solution, and $A \in P$ implies $A^{-1} \in P$ (Theorem 2). □

In Corollary 3 a *sufficient* condition is given for the existence of a unique solution to say equation (1) when the functions f_i which specify F are not necessarily mappings onto E^1. That the condition ($A \in P$) is not *necessary* is easily demonstrated by the counterexample: Let $F \in \mathfrak{F}_0^2$ and

$B \varepsilon E^2$; then the equations

$$f_1(x_1) - x_2 = b_1 , \text{ and } f_2(x_2) + x_1 = b_2$$

have a unique solution in spite of the fact that the matrix

$$A \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \not\varepsilon P.$$

This is true because the function $f_2(f_1(x_1) - b_1)$ is obviously a continuous monotone increasing function of x_1 , and hence the left side of the equation

$$f_2(f_1(x_1) - b_1) + x_1 = b_2 \tag{8}$$

is a strictly monotone increasing mapping of E^1 onto E^1. Thus equation (8) has a unique solution.

V. BOUNDED SOLUTIONS AND RELATED PROBLEMS

For many systems whose behavior is described by an equation having the form of equation (1), the vector B may be regarded as the system's input and the vector x may be regarded as the system's response, or output. Thus, if a sequence $B^1, B^2, B^3, \cdots$ of input vectors for the system is given, the corresponding sequence $x^1, x^2, x^3, \cdots$ of output vectors is specified by equation (1). An important property that such systems might have is that of producing a bounded sequence of output vectors for each bounded sequence of input vectors; that is, the property that whenever an input sequence $B^1, B^2, B^3, \cdots$ is contained in some bounded region of E^n, then the corresponding output sequence $x^1, x^2, x^3, \cdots$ (exists and) also is contained in some bounded region of E^n. By considering matrices A which are not members of P_0 , it is easy to demonstrate that all equations having the form of equation (1) do not have this property. For example, if $f(x) \equiv x + e^x$ ($f \varepsilon \mathfrak{F}^1$), then the sequence of solutions of the equation $f(x) + (-1)x = b$ is unbounded, even though the sequence $b = 1, \frac{1}{2}, \frac{1}{3}, \cdots$ of inputs is bounded. The fact that one *must* resort to matrices A which are not in P_0 , and the fact that by choosing *any* $A \not\varepsilon P_0$, an example of the above kind can be constructed by an appropriate choice of $F \varepsilon \mathfrak{F}^n$, follows from our next theorem.

Theorem 4. If A is an $n \times n$ matrix then $A \varepsilon P_0$ if and only if for each $F \varepsilon \mathfrak{F}^n$ and each unbounded sequence of points $x^1, x^2, x^3, \cdots$ in E^n, the corresponding sequence $B^1, B^2, B^3, \cdots$ ($B^k = F(x^k) + Ax^k$, $k = 1, 2, 3, \cdots$) is unbounded.

Proof: (*if*) If $A \notin P_0$ then, as shown in the "only if" part of the proof of Theorem 3, there exists a diagonal matrix $D > 0$ such that $D + A$ is singular. Hence, there exists some point $p \in E^n$, $p \neq \theta$, such that $Dp + Ap = \theta$. Let p_j, the j-th component of p, be nonzero. Let the diagonal elements of the matrix D be denoted by $d_1, \cdots, d_n$ and let the mapping $F \in \mathfrak{F}^n$ be defined by

$$f_i(x_i) = \begin{cases} d_i x_i, & \text{for} \quad i \neq j, \\ d_i x_i + e^{x_i}, & \text{for} \quad i = j. \end{cases}$$

If $p_j < 0$ let $\epsilon = 1$, if $p_j > 0$ let $\epsilon = -1$. Consider the unbounded sequence $x^1, x^2, x^3, \cdots$ defined by $x^k = k \cdot \epsilon \cdot p$, for $k = 1, 2, 3, \cdots$. The members of the corresponding sequence $B^1, B^2, B^3, \cdots$ are $B^k = (0, \cdots, 0, e^{k\epsilon p_j}, 0, \cdots, 0)^t$, $k = 1, 2, 3, \cdots$, where the j-th element of each B^k is nonzero. Since for $k = 1, 2, 3, \cdots$, $k \epsilon p_j < 0$, the sequence $B^1, B^2, B^3, \cdots$ is bounded.

(*only if*) Our proof of the "only if" part of Theorem 4 consists of proving Theorem 5 which is referred to later for another purpose. □

Theorem 5. Let $F \equiv (f_1(\cdot), \cdots, f_n(\cdot))^t \in \mathfrak{F}^n$, $A \in P_0$, *and, for* $i = 1, \cdots, n$, $\alpha_i \leqq \beta_i$ *be given. There exist, for* $i = 1, \cdots, n$, *real numbers* $\gamma_i \leqq \delta_i$ *such that for any* $B \equiv (b_1, \cdots, b_n)^t \in E^n$ *with* $\alpha_i \leqq b_i \leqq \beta_i$ *for* $i = 1, \cdots, n$, *if* x *satisfies equation* (1) *then* $\gamma_i \leqq x_i \leqq \delta_i$ *for* $i = 1, \cdots, n$.

Proof of Theorem 5: We first prove a useful lemma.

Lemma 1. Let f *be a strictly monotone increasing mapping of* E^1 *onto itself. Let* x, b, α, β *be real numbers such that* $xf(x) \leqq xb$ *with* $\alpha \leqq b \leqq \beta$. *Then* $\gamma \leqq x \leqq \delta$, *where* $\gamma = \min \{f^{-1}(\alpha), 0\}$ *and* $\delta = \max \{f^{-1}(\beta), 0\}$.

Proof: Let $\alpha \leqq b \leqq \beta$ and define $\gamma = \min \{f^{-1}(\alpha), 0\}$ and $\delta = \max \{f^{-1}(\beta), 0\}$. Let x satisfy $xf(x) \leqq xb$. Then $x(f(x) - b) \leqq 0$. Clearly, $\gamma \leqq 0 \leqq \delta$ and hence if $x = 0$ then $\gamma \leqq x \leqq \delta$. If $x > 0$ then $f(x) \leqq b \leqq \beta$ which implies $x \leqq f^{-1}(\beta) \leqq \delta$ and hence $\gamma \leqq 0 < x \leqq \delta$. If $x < 0$, then $f(x) \geqq b \geqq \alpha$ which implies $x \geqq f^{-1}(\alpha) \geqq \gamma$ and hence $\gamma \leqq x < 0 \leqq \delta$. □

(*Proof of Theorem 5*) Since $A \in P_0$ there exists $k_1 \in \{1, \cdots, n\}$ such that $x_{k_1}(Ax)_{k_1} \geqq 0$ and hence,

$$x_{k_1} b_{k_1} = x_{k_1} f_{k_1}(x_{k_1}) + x_{k_1}(Ax)_{k_1} \geqq x_{k_1} f_{k_1}(x_{k_1}).$$

Thus, by Lemma 1, there exist $\gamma_{k_1}^{(1)} = \gamma_{k_1}^{(1)}(f_{k_1}, \alpha_{k_1})$ and $\delta_{k_1}^{(1)} = \delta_{k_1}^{(1)}(f_{k_1}, \beta_{k_1})$

such that $\gamma_{k_1}^{(1)} \leqq x_{k_1} \leqq \delta_{k_1}^{(1)}$. Now if F_{n-1} denotes the mapping of E^{n-1} onto E^{n-1} defined by

$$F_{n-1} \equiv (f_1(\cdot), \cdots, f_{k_1-1}(\cdot), f_{k_1+1}(\cdot), \cdots, f_n(\cdot))^t,$$

if A_{n-1} denotes the $(n-1) \times (n-1)$ matrix obtained from A by deleting the k_1-st row and column (note that $A_{n-1} \varepsilon P_0$), if

$$a_{n-1} = (a_{1.k_1}, \cdots, a_{k_1-1.k_1}, a_{k_1+1.k_1}, \cdots, a_{n.k_1})^t,$$

and if

$$B_{n-1} = (b_1, \cdots, b_{k_1-1}, b_{k_1+1}, \cdots, b_n)^t,$$

then

$$F_{n-1}(x) + A_{n-1}x = B_{n-1} - a_{n-1}x_{k_1}.\text{*}$$

Since $A_{n-1} \varepsilon P_0$, there is a $k_2 \varepsilon \{1, \cdots, k_1-1, k_1+1, \cdots, n\}$ such that $x_{k_2}(A_{n-1}x)_{k_2} \geqq 0$ and hence, as before,

$$x_{k_2}(b_{k_2} - a_{k_2.k_1}x_{k_1}) \geqq x_{k_2}f_{k_2}(x_{k_2}).$$

But, if $a_{k_2.k_1} \leqq 0$, then

$$\alpha_{k_2} - a_{k_2.k_1}\gamma_{k_1}^{(1)} \leqq b_{k_2} - a_{k_2.k_1}x_{k_1} \leqq \beta_{k_2} - a_{k_2.k_1}\delta_{k_1}^{(1)},$$

and if $a_{k_2.k_1} > 0$, then

$$\alpha_{k_2} - a_{k_2.k_1}\delta_{k_1}^{(1)} \leqq b_{k_2} - a_{k_2.k_1}x_{k_1} \leqq \beta_{k_2} - a_{k_2.k_1}\gamma_{k_1}^{(1)}.$$

Therefore, by Lemma 1, there is a $\gamma_{k_2}^{(1)} = \gamma_{k_2}^{(1)}(f_{k_2}, \alpha_{k_2} - a_{k_2.k_1}\gamma_{k_1}^{(1)})$ and $\delta_{k_2}^{(1)} = \delta_{k_2}^{(1)}(f_{k_2}, \beta_{k_2} - a_{k_2.k_1}\delta_{k_1}^{(1)})$ such that $\gamma_{k_2}^{(1)} \leqq x_{k_2} \leqq \delta_{k_2}^{(1)}$ if $a_{k_2.k_1} \leqq 0$, and similarly for $a_{k_2.k_1} > 0$.

The above process may be repeated successively until the n pairs of real numbers $\gamma_{k_i}^{(1)}, \delta_{k_i}^{(1)}, (i = 1, \cdots, n)$ have been obtained. Thus, for any given B with $\alpha_i \leqq b_i \leqq \beta_i$ for $i = 1, \cdots, n$, the components of the solution x of equation (1) will be bounded by these pairs of numbers, provided it is known at each step which coordinate k_i to choose. The appropriate coordinate choice, however, will in general depend on the particular solution x which is associated with the given B. For different input vectors B the appropriate choice will in general be different. Therefore, in order to obtain bounds on x which are valid for all B with $\alpha_i \leqq b_i \leqq \beta_i$ $(i = 1, \cdots, n)$ we must consider each of the $n!$ permutations of the coordinates $\{1, \cdots, n\}$ and, for each one, generate the set of bounds $\{\gamma_{k_i}^{(\nu)}, \delta_{k_i}^{(\nu)}: i = 1, \cdots, n\}$ for $\nu = 1, \cdots, n!$. We then define $\gamma_i =$

* In this equation x is understood to be $(x_1, \cdots, x_{k_1-1}, x_{k_1+1}, \cdots, x_n)^t$.

min $\{\gamma_i^{(\nu)}: \nu = 1, \cdots, n!\}$ and $\delta_i = \max \{\delta_i^{(\nu)}: \nu = 1, \cdots, n!\}$ for $i = 1, \cdots, n$. Then, for each B with $\alpha_i \leqq b_i \leqq \beta_i$ for $i = 1, \cdots, n$, we have that $\gamma_i \leqq x_i \leqq \delta_i$ for $i = 1, \cdots, n$, since at least one of the sets of bounds $\{\gamma_{k_i}^{(\nu)}, \delta_{k_i}^{(\nu)}: i = 1, \cdots, n\}$ must always apply. □

If the matrix A of Theorem 5 satisfies a stronger condition than $A \varepsilon P_0$ (that is, if A satisfies a weak row-sum dominance condition),

$$a_{ii} \geqq \sum_{j \neq i} | a_{ij} |, \qquad \text{for} \quad i = 1, \cdots, n,$$

it is possible to use a method that requires much less computational effort than that of Theorem 5 to compute the vectors γ and δ whose components bound the corresponding components of the solution of equation (1). This method of computing the bounds, a straightforward generalization of an idea presented in Ref. 3, is explained in Appendix B.

From Theorems 3 and 4 we now have the result: *Every bounded input sequence* $B^1, B^2, B^3, \cdots$ *is mapped by equation* (1) *into a bounded output sequence* $x^1, x^2, x^3, \cdots$, *for each* $F \varepsilon \mathfrak{F}^n$, *if and only if* $A \varepsilon P_0$.

In the proof of Theorem 5, the number of real numbers $\gamma_{k_i}^{(\nu)}, \delta_{k_i}^{(\nu)}$ which must be computed, in order to determine bounds for x, is $2n \times (n!)$. At the expense of obtaining poorer bounds it is easy to reduce this number to $2n^2$. Suppose we compute, at the first step, the $2n$ numbers $\gamma_1^{(1)}, \delta_1^{(1)}, \cdots, \gamma_n^{(1)}, \delta_n^{(1)}$ and set $\lambda_1 = \min \{\gamma_1^{(1)}, \cdots, \gamma_n^{(1)}\}$, $\mu_1 = \max \{\delta_1^{(1)}, \cdots, \delta_n^{(1)}\}$. Then, for each B with $\alpha_i \leqq b_i \leqq \beta_i$ for $i = 1, \cdots, n$, one of the components of the corresponding x will be bounded by λ_1 (from below) and μ_1 (from above). We next compute the $2n$ numbers $\gamma_i^{(2)} = \gamma_i^{(2)}(f_i, \alpha_i - p_i^{(1)})$, $\delta_i^{(2)} = \delta_i^{(2)}(f_i, \beta_i - q_i^{(1)})$, where $p_i^{(1)} = \max \{a_{ij}\lambda_1, a_{ij}\mu_1 : j \neq i\}$, $q_i^{(1)} = \min \{a_{ij}\lambda_1, a_{ij}\mu_1 : j \neq i\}$, and denote the smallest $\gamma_i^{(2)}$ by λ_2 and the largest $\delta_i^{(2)}$ by μ_2. Then we have bounds which apply for two of the components of the x which corresponds to any B with $\alpha_i \leqq b_i \leqq \beta_i$ for $i = 1, \cdots, n$. By computing $\gamma_i^{(3)} = \gamma_i^{(3)}(f_i, \alpha_i - p_i^{(1)} - p_i^{(2)})$, $\delta_i^{(3)} = \delta_i^{(3)}(f_i, \beta_i - q_i^{(1)} - q_i^{(2)})$, etc., the above process may be continued to obtain the numbers $\lambda_1, \cdots, \lambda_n$, $\mu_1, \cdots, \mu_n$. Each component of the x corresponding to any B with $\alpha_i \leqq b_i \leqq \beta_i$ for $i = 1, \cdots, n$ will be bounded by $\lambda = \min \{\lambda_1, \cdots, \lambda_n\}$ (from below) and $\mu = \max \{\mu_1, \cdots, \mu_n\}$ (from above).

A matter that is closely related to the proofs of the above theorems on the boundedness of solutions of equation (1) is that of proving: *For each* $F \varepsilon \mathfrak{F}^n$ *and each* $A \varepsilon P_0$ *the solution* x *of equation* (1) *is a continuous function of the vector* B. It is obvious that it will suffice to prove that for each $F \varepsilon \mathfrak{F}^n$ with $F(\theta) = \theta$, and for each $A \varepsilon P_0$, the solution x of equation (1) is continuous in B at $B = \theta$. We then note that if f

satisfies the hypotheses of Lemma 1 and, in addition, if $f(0) = 0$ then, due to the continuity of f^{-1}, for every $\epsilon > 0$ there exists $\zeta > 0$ such that if α, β in Lemma 1 satisfy $-\zeta < \alpha \leqq b \leqq \beta < \zeta$ then γ, δ in Lemma 1 satisfy $-\epsilon < \gamma \leqq x \leqq \delta < \epsilon$. This observation may be used to incorporate a simple "ϵ-δ argument" into the steps of the previous paragraph to show that when $F(\theta) = \theta$ then for arbitrary $\epsilon > 0$, one can determine $\zeta > 0$ such that $\| B \| < \zeta$ implies $\| x \| < \epsilon$.

At this point we return to the matter of the existence and uniqueness of solutions of equation (1). We state first a theorem of R. S. Palais (Ref. 7—see also the Appendix of Ref. 8) which shows the connection between the concepts of existence and uniqueness of solutions and the boundedness of solutions.

Palais' Theorem. Let $f_1, \cdots, f_n$ *be* n *continuously differentiable real valued functions of* n *real variables. Necessary and sufficient conditions that the mapping* $f : E^n \to E^n$ *defined by* $f(x) = (f_1(x), \cdots, f_n(x))^t$ *be a diffeomorphism of* E^n *onto itself are*:

(*i*) $\det [\partial f_i/\partial x_j]$ *never vanishes.*
(*ii*) $\lim_{\|x\| \to \infty} \| f(x) \| = \infty$.

Palais' Theorem may be used to prove a result which is almost equivalent to our Theorem 3, that is:

Theorem 6. If A *is an* $n \times n$ *matrix then there exists a unique solution of equation* (1) *for each* $F \equiv (f_1(x_1), \cdots, f_n(x_n))^t$ *with continuously differentiable, strictly monotone increasing functions* f_i *which map* E^1 *onto itself, and whose slopes are everywhere positive, and for each* $B \in E^n$, *if and only if* $A \in P_0$.

A proof of Theorem 6 which is independent of our Theorem 3 is easy to construct: For all $A \in P_0$, the rather trivial Theorem 1 guarantees that condition (*i*) of Palais' Theorem is satisfied, and Theorem 5 guarantees that condition (*ii*) is satisfied. If $A \notin P_0$ then a choice of F such as is specified in the "*if*" part of the proof of Theorem 4 provides a case in which condition (*ii*) of Palais' Theorem is violated.

VI. SUFFICIENT CONDITIONS FOR $A \in P_0$ OR P

For a given matrix A, it is not in general an easy task to determine whether or not A satisfies any one of the four equivalent conditions of Fiedler and Pták which are given in Section III and which serve to define the class of matrices P_0 (or the conditions which define P). This is particularly true when the order of A is large. For this reason, we now give several conditions which are sufficient to insure that a matrix A is in P_0 or P (and which are not so difficult to verify).

Suppose it were known that every eigenvalue of A as well as every eigenvalue of each principal submatrix of A had a nonnegative (positive) real part. Then this would guarantee that $A \varepsilon P_0$ (P). This is the main idea involved in the following theorem.

Theorem 7. If any one of the following inequalities is satisfied by the elements a_{ij} of the matrix A, for all $i = 1, \cdots, n$, then $A \varepsilon P_0$.

$$(i) \qquad a_{ii} \geqq (\sum_{j \neq i} | a_{ij} |)^{\alpha} (\sum_{k \neq i} | a_{ki} |)^{1-\alpha}, \qquad 0 \leqq \alpha \leqq 1;$$

$$(ii) \qquad a_{ii} \geqq \alpha_i^{1/q} (\sum_{j \neq i} | a_{ij} |^p)^{1/p}, \qquad p \geqq 1, \qquad p^{-1} + q^{-1} = 1,$$

α_i positive numbers satisfying $\sum_{i=1}^{n} (1 + \alpha_i)^{-1} \leqq 1$;

$$(iii) \qquad a_{ii} \geqq \alpha \max_{j \neq i} | a_{ij} |, \ \alpha \text{ positive satisfying}$$

$$\sum_{i=1}^{n} \{ \sum_{j \neq i} | a_{ij} | (\max_{j \neq i} | a_{ij} |)^{-1} \} \leqq \alpha(1 + \alpha), \qquad (0/0 = 0).$$

If any one of the above inequalities with $\geqq$ replaced by $>$ is satisfied for $i = 1, \cdots, n$, then $A \varepsilon P$.

Proof: If the right-hand side of any of the above inequalities is denoted by the nonnegative number r_i then it is well known that all of the eigenvalues of the matrix A are contained in the union $\bigcup \{C_i : i = 1, \cdots, n\}$ of the disks $C_i = \{z: | z - a_{ii} | \leqq r_i\}$.[9] But the condition $a_{ii} \geqq (>) \ r_i$ guarantees that if $z \varepsilon C_i$ then $\text{Re}(z) \geqq (>) \ 0$. Thus, each of the eigenvalues of the matrix A has a nonnegative (positive) real part. The same is true of each eigenvalue of every principal submatrix of A, for if one of the above inequalities is satisfied by the elements of A it is also satisfied by the elements of any principal submatrix. □

VII. COMPUTATION OF THE SOLUTION

At present, the authors know of no single computational algorithm which is guaranteed to yield the solution of equation (1) for all $F \varepsilon \mathfrak{F}^n$, $A \varepsilon P_0$, $B \varepsilon E^n$. However, there are several ways that the solution may be computed for large classes of such equations.

If, for example, the matrix A satisfies either a weak row-sum or weak column-sum dominance condition (inequality (i) of Theorem 7 with either $\alpha = 1$ or $\alpha = 0$) and if $F \in \mathfrak{F}^n$ with, roughly speaking, the slopes of each f_i bounded from below by some positive constant, then it can be shown (see Appendix A) that an algorithm for computing the solution can be obtained by the use of Banach's contraction-mapping fixed point theorem.

If the matrix A is positive semidefinite then, as mentioned in Section I, the existence of a unique solution of equation (1) for all $F \in \mathfrak{F}^n$ follows from the earlier work of Sandberg and Minty. If, in addition, there exists for $i = 1, \cdots, n$, positive constants α_i and β_i such that

$$\alpha_i \leqq \frac{f_i(u) - f_i(v)}{u - v} \leqq \beta_i$$

for all $u \neq v$, then Sandberg's iteration scheme (also resulting from an application of the contraction-mapping fixed point theorem) can be used to compute the solution.[1] In this regard, if the techniques of Section V are first used to obtain bounds on the location of the solution then one could modify equation (1) by changing the nature of the functions f_i outside the domain in which the solution is known to lie (but still keeping the f_i strictly monotone increasing from E^1 onto E^1) and obtain a new equation which has the same solution as the original equation. By doing this, the functions f_i in the new equation might be made to satisfy the above inequalities in cases where this was impossible for the original f_i. Also, even if these inequalities could be satisfied for the original equation, larger values of α_i and smaller values of β_i might be used for the modified equation. This can result in a more rapidly converging iteration process (see Section VII of Ref. 3). Similarly, the bounds can be used to improve the performance of other iteration schemes.

In case $A \in P_0$ is not positive semidefinite, it might be that there exist diagonal matrices $\Delta_1, \Delta_2 > 0$ such that $\Delta_1 A \Delta_2$ is positive semidefinite. If such matrices can be found, then Sandberg's iteration scheme could be used to compute the solution of the equation

$$\Delta_1 F(\Delta_2 x) + \Delta_1 A \Delta_2 x = \Delta_1 B,$$

from which the solution of equation (1) may be obtained directly. Unfortunately, it is not the case that such $\Delta_1, \Delta_2 > 0$ exist for all $A \in P_0$. For example, it is quite easily verified that for

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix},$$

even though $A \varepsilon P_0$, the matrix $\Delta_1 A \Delta_2$ is not positive semidefinite for any choice of $\Delta_1, \Delta_2 > 0$.

It is easily verified, however, that appropriate $\Delta_1, \Delta_2 > 0$ can be found for all 2×2 matrices $A \varepsilon P_0$ except those for which

(*i*) $a_{11}a_{22} = 0$,

and

(*ii*) $a_{12}a_{21} = 0$,

and

(*iii*) either $a_{12} \neq 0$ or $a_{21} \neq 0$.

In particular, for all nonsingular 2×2 matrices $A \varepsilon P_0$, appropriate Δ_1, Δ_2 ($\Delta_2 = I$) can be found. Thus, Sandberg's iteration scheme could be used, for example, to compute the solution of the example problem of Section II which was taken from Calahan's book.

VIII. APPLICATION TO EQUATIONS FOR TRANSISTOR NETWORKS

In this section some of the above theory is applied to the equations which describe the behavior of electrical networks containing transistors. By the word transistor we refer to the three-terminal device whose equivalent circuit is shown in Fig. 2.* Considering the transistor as a nonlinear two-port network, the following equations which express the port currents in terms of the port voltages follow immediately from inspection of Fig. 2:

$$\begin{pmatrix} i_1 \\ i_2 \end{pmatrix} = \begin{bmatrix} 1 & -\alpha_{12} \\ -\alpha_{21} & 1 \end{bmatrix} \begin{pmatrix} f_1(v_1) \\ f_2(v_2) \end{pmatrix}.$$

We assume, as is the case for the usual large-signal model of a physical transistor, that $0 < \alpha_{12} < 1$, $0 < \alpha_{21} < 1$, and that both of the functions f_1 and f_2 are continuous and strictly monotone increasing. The character of the functions f_1 and f_2 which describe the transistor's nonlinear conductances will depend on whether the transistor is designated as NPN or PNP. We shall, however, have no occasion to distinguish between these two cases in what is to follow.

Suppose an electrical network is synthesized by connecting together,

* In some respects this equivalent circuit is an *ideal* model of a transistor. Nevertheless, since this model is often used in the design and the computer analysis of transistor networks, consideration of it is important. The presence of series resistance at the base, emitter, and collector terminals of a transistor will be considered by the authors in another paper.

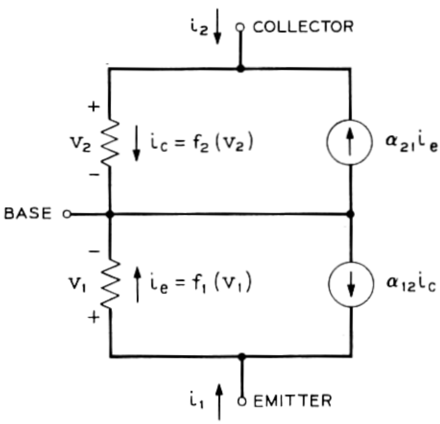

Fig. 2 — The equivalent circuit of a transistor.

in an arbitrary manner, any (finite) number of transistors, resistors (that is, linear resistors with nonnegative resistance), voltage sources, current sources, and nonlinear resistors which are described by strictly monotone increasing conductance functions (and which we shall henceforth refer to as "diodes"). Suppose the network contains n transistors and d diodes. For $k = 1, \cdots, n$, let x_{2k-1}, x_{2k}, y_{2k-1}, and y_{2k} denote the voltage and current variables v_1, v_2, i_1, and i_2, respectively, for the k-th transistor. For $k = 1, \cdots, d$, let x_{2n+k} and y_{2n+k} denote the voltage across, and the current through, the kth diode. Let these variables be related by $y_{2n+k} = f_{2n+k}(x_{2n+k})$. Then, if $x = (x_1, \cdots, x_{2n+d})^t$ and $y = (y_1, \cdots, y_{2n+d})^t$, we have

$$y = TF(x), \tag{9}$$

where $T = \operatorname{diag}(T_1, T_2)$, with T_1 a block diagonal matrix with n 2×2 diagonal blocks of the form

$$\begin{bmatrix} 1 & -\alpha_{12}^{(k)} \\ -\alpha_{21}^{(k)} & 1 \end{bmatrix},$$

and T_2 a $d \times d$ identity matrix. The nonlinear function F has the form $F(x) \equiv (f_1(x_1), \cdots, f_{2n+d}(x_{2n+d}))^t$.

Consider now the $(2n+d)$-port network of resistors and independent sources which is formed from the original network by removing the transistors and diodes. If the y-parameter matrix G of this $(2n+d)$-port exists then we have the additional equation relating the vectors

x and y:

$$y = -Gx + u, \tag{10}$$

where u is some vector of constants which is, in general, nonzero since sources are present in the $(2n+d)$-port.

Combining equations (9) and (10) we obtain

$$TF(x) + Gx = u. \tag{11}$$

Now T is a nonsingular matrix and hence, if equation (11) is multiplied by T^{-1}, we obtain an equation having the form of equation (1). If the matrix $T^{-1}G \in P_0$ then, by Corollary 1, there exists at most one set of transistor and diode voltages satisfying equation (11). Moreover, if each of the nonlinear functions describing the transistors and diodes in our network maps E^1 *onto* E^1, or if $T^{-1}G \in P$, then Theorem 3, or Corollary 3, guarantees the *existence* of a unique solution of equation (11).

We have been careful to distinguish between the case when our theory guarantees only the uniqueness of a solution and the case when it guarantees both the solution's existence and its uniqueness for the following reason: In the analysis of transistor networks the nonlinear functions which are used to describe diodes or to describe the nonlinear conductances in the equivalent circuit of a transistor are often taken to be of the form

$$f(x) = I_0(e^{\lambda x} - 1),$$

where I_0 and λ are constants. The range of such a function is not the entire real line. Presumably, therefore, one can construct transistor networks having the property that if functions of the above type are used in a transistor's equivalent circuit then the network admits no solution. We now give a simple example of such a network. We wish to emphasize, though, that even for these networks whose equations may sometimes have no solution, our theory still guarantees that if $T^{-1}G \in P_0$ and if a solution of equation (11) exists, then it is unique.

Consider the network of Fig. 3. For this network, equation (11) becomes

$$\begin{bmatrix} 1 & -\alpha_{12} \\ -\alpha_{21} & 1 \end{bmatrix} \begin{bmatrix} f_1(v_1) \\ f_2(v_2) \end{bmatrix} + \begin{bmatrix} g & -g \\ -g & g \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} I_a \\ I_b \end{bmatrix}.$$

Suppose $\alpha_{12} = 0.5$, $\alpha_{21} = 0.9$, and $g = 5.5$ mhos. Then, the above equation is equivalent to

$$\begin{bmatrix} f_1(v_1) \\ f_2(v_2) \end{bmatrix} + \begin{bmatrix} 5 & -5 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 20 & 10 \\ 18 & 20 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \end{bmatrix}.$$

Hence, v_1 and v_2 must satisfy

$$f_1(v_1) + 5f_2(v_2) = 10(I_a + I_b).$$

If we now assume that the transistor's nonlinear conductances are described by the functions

$$f_1(v_1) = -I_e(e^{-\lambda_e v_1} - 1),$$
$$f_2(v_2) = -I_c(e^{-\lambda_c v_2} - 1),$$

where the parameters I_e, I_c, λ_e, and λ_c are each positive, then for all v_1, v_2 we have

$$f_1(v_1) + 5f_2(v_2) < I_e + 5I_c.$$

Hence, if the values of the independent current sources of Fig. 3 are chosen such that

$$I_a + I_b \geqq \tfrac{1}{10}I_e + \tfrac{1}{2}I_c,$$

then the equation for this network has no solution.

Let us now consider the problem of determining whether or not, for a given network, the matrices T and G in equation (11) satisfy the condition $T^{-1}G \in P_0$ (or $T^{-1}G \in P$). (The existence of many transistor bistable circuits assures us that this condition is not always satisfied.)

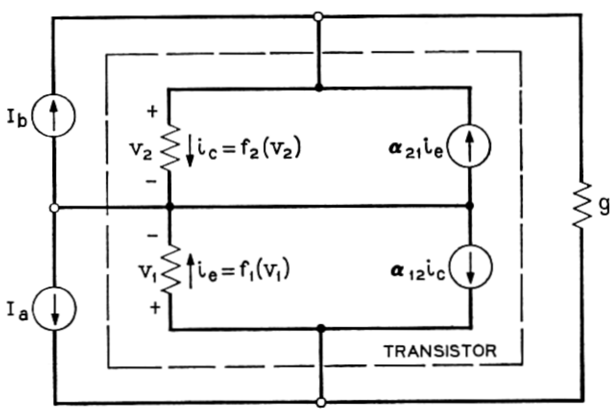

Fig. 3 — A transistor network whose equations may have no solution.

There is a large class of networks for which this condition is satisfied, and for which a simple inspection of the G matrix suffices to identify a member of the class.

Since the matrix T satisfies a strong column-sum dominance condition, that is, since

$$t_{ii} > \sum_{j \neq i} | t_{ji} | \qquad \text{for} \quad i = 1, \cdots, 2n + d,$$

the following theorem guarantees that if the matrix G also satisfies a strong column-sum dominance condition, then $T^{-1}G \varepsilon P$, and that if the matrix G satisfies a weak column-sum dominance condition,

$$g_{ii} \geqq \sum_{j \neq i} | g_{ji} |,$$

then $T^{-1}G \varepsilon P_0$ and, hence, the above conclusions concerning the existence and the uniqueness of a solution follow.

Theorem 8. If the square matrix A satisfies a strong column-sum dominance condition and if the square matrix B satisfies a weak (strong) column-sum dominance condition, then $A^{-1}B \varepsilon P_0$ (P).

Proof: Suppose $A^{-1}B \not\varepsilon P_0$. Then, by the main result of the "only if" part of the proof of Theorem 3, there exists some diagonal matrix $D > 0$ such that $\det (D + A^{-1}B) = 0$. But $\det (D + A^{-1}B) = \det (A^{-1}) \cdot \det (AD + B)$, and $\det (A^{-1}) \neq 0$. Likewise, $\det (AD + B) \neq 0$ since $AD + B$ satisfies a strong column-sum dominance condition. Hence, $A^{-1}B \varepsilon P_0$.

With B strongly column-sum dominant, let $\delta > 0$ be such that $B - \delta A$ also possesses the strong dominance property. Suppose that $A^{-1}B - \delta I \not\varepsilon P_0$. Then, as above, there is a $D > 0$ such that $A^{-1}B - \delta I + D = A^{-1}[B - \delta A + AD]$ is singular, which is a contradiction. Therefore $A^{-1}B - \delta I \varepsilon P_0$, and, by Theorem 1, $A^{-1}B \varepsilon P$. □

IX. COMMON-BASE TRANSISTOR NETWORKS

We now consider a special class of the networks which are comprised of transistors, resistors, diodes, and independent sources. We consider the class of all such networks for which there is a single node (called ground) to which the base terminal of each transistor is connected. Let us first consider a subclass of this class of networks; that is, let us temporarily assume that no diodes are present. For all networks in this subclass it is easily verified that when the G matrix for equation (11) exists, then it satisfies the above weak column-sum dominance condition and hence, by Theorem 8, $T^{-1}G \varepsilon P_0$. This fact is made evident if we consider the network of resistors which is described by G (that is, the linear multiport to which the transistors are connected, with all sources removed) and first simplify this network by using the star-mesh transformation to remove all internal nodes. Of course for many networks of this sublcass G is strongly column-sum dominant, in which case $T^{-1}G \varepsilon P$.

It is clear that the networks for which the G matrix fails to exist are exactly those networks in which either one or more of the collector or emitter terminals are connected, through the resistor network, directly to ground (that is, through a branch having infinite conductance), or else two (or more) of the transistors' collector or emitter terminals are connected directly together (through a branch of the resistor network having infinite conductance). These direct connections can exist in the resistor network either because of corresponding short-circuits in the original linear multiport, or because of corresponding connections involving branches which contain only ideal voltage sources.

If one assumes that each transistor in the network has a nonzero series resistance associated with both its emitter and its collector terminals (this assumption certainly being consistent with physical reality) then one need not be concerned about the possibility of the nonexistence of the G matrix since the situations mentioned in the previous paragraph cannot occur. We now show, however, that one need not rely upon this assumption in order to prove the uniqueness of the solution of the equations which describe the networks that we are considering.

We have observed that the matrix G will not exist if and only if the linear multiport has fewer independent port voltages than it has ports. In this case we modify the nonlinear multiport in such a manner that we can break some of the connections to the linear multiport so that it then possesses a G matrix and hence can be described by an equation having the form of equation (10). The modifications to the nonlinear multiport which are called for are obviously the addition of voltage sources between certain nodes, the values of these sources being the same as those of the voltage sources connecting the corresponding nodes in the linear multiport. This simple concept is illustrated in Fig. 4. Here, the network of Fig. 4a, containing a linear 6-port, has been replaced by the "equivalent" network of Fig. 4b containing a linear 3-port. Although the G matrix of the 6-port does not

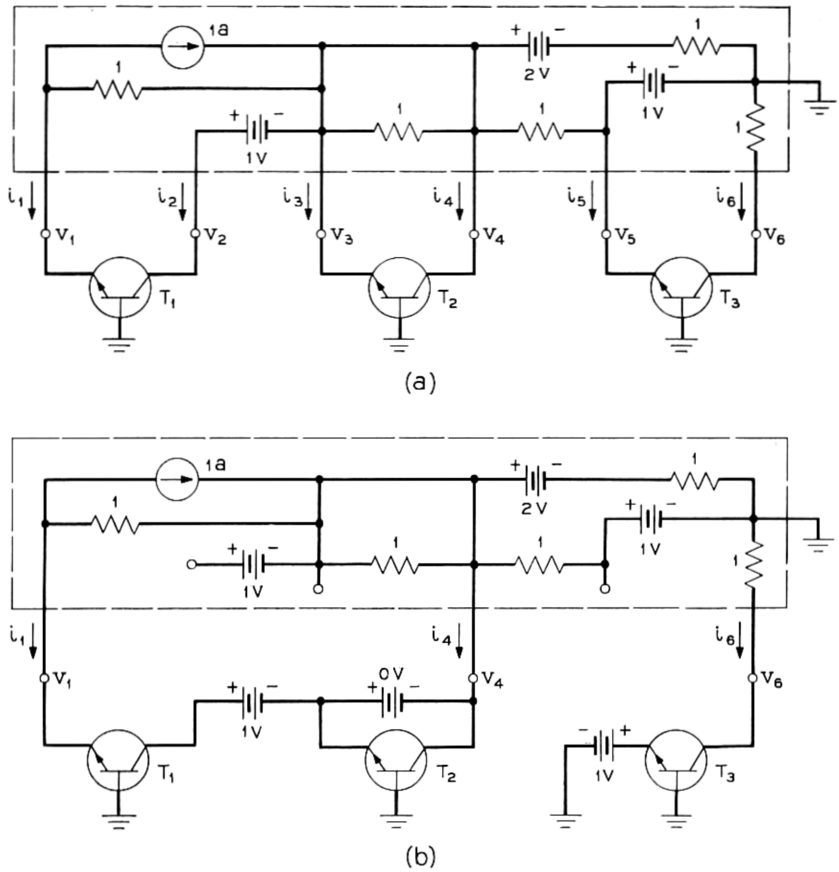

Fig. 4 — Example of a grounded-base transistor network.

exist, it does exist for the 3-port which can be described by

$$\begin{bmatrix} i_1 \\ i_4 \\ i_6 \end{bmatrix} = -\begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_4 \\ v_6 \end{bmatrix} + \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}.$$

We have shown that the above artifice allows an equation having the form of equation (10) to always be written to describe the linear multiport contained in our network. We now show that an equation like equation (9) can be written to describe the nonlinear part of our modified network. The equation which we obtain is of the form $y = PTF(P^t x + C)$ with P an $m \times 2n$ matrix $(m < 2n)$ and C a $2n$-vector.

Consider the equation which describes the nonlinear part of a common-base transistor network before any of the above-mentioned modifications (that is, the addition of voltage sources) are made. This equation has the form of equation (9) with T being a $2n \times 2n$ block diagonal matrix (recall that n is the number of transistors present). Let us consider the effect on this equation of the modification of the network by adding voltage sources, one at a time. There are two different ways of adding voltage sources that must be considered.

Suppose a voltage source of voltage E is connected between nodes j and k (with plus reference at node j), and suppose the connections between node j and the linear multiport are then open-circuited. This situation is illustrated in Fig. 5. Using the notation indicated in this figure, we have

$$
\begin{aligned}
i' &= TF(v), \\
i_\nu &= i'_\nu \qquad \text{for} \quad \nu \neq j, k, \\
i_j &= 0, \\
i_k &= i'_j + i'_k\,, \\
v_j &= v_k + E.
\end{aligned}
$$

Let us now define the vectors v^* and i^* to be the $(2n-1)$-vectors obtained from v and i, respectively, by deleting the v_j and i_j elements. Then, if $F^*(v^*)$ is the $2n$-vector obtained from $F(v)$ by replacing the

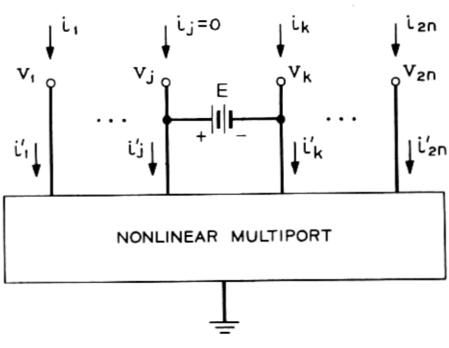

Fig. 5 — Typical modification of the nonlinear multiport network.

argument v_j by $v_k + E$, we then have that

$$i^* = T^*F^*(v^*), \tag{12}$$

where the $(2n-1) \times 2n$ matrix T^* is obtained from T by adding the j-th row to the k-th row and then deleting the j-th row. Observe that $T^*F^*(v^*)$ can be written as $QTF(Q^t v^* + R)$ in which the j-th element of the $2n$-vector R is E, all other elements of R are zero, and Q is obtained from the identity matrix of order $2n$ by adding the j-th row to the k-th row and then deleting the j-th row.

In case a voltage source of voltage E is connected between node j and ground (with the plus reference at node j) and all connections between node j and the linear multiport are open-circuited, then we can again form equation (12) from equation (9) by simply replacing v_j by E wherever it appears in the argument of F, to form $F^*(v^*)$, and deleting the j-th row of the matrix T, to form T^*. In this case $T^*F^*(v^*)$ can be written as $QTF(Q^t v^* + R)$ in which R is as defined earlier, but in this case Q is obtained from the identity matrix of order $2n$ by simply deleting the j-th row.

The above processes can be applied repeatedly to account for the addition of an arbitrary number of voltage sources to the nonlinear multiport. The resulting equation which describes the multiport will have the form

$$\begin{aligned} y &= Q_p \cdots Q_2 Q_1 TF(Q_1^t Q_2^t \cdots Q_p^t x + C) \\ &\equiv \tilde{T}\tilde{F}(x) \end{aligned}$$

with C some constant $2n$-vector and each of the Q_i obtained from the identity matrix of the appropriate order in one of the two ways described above.

Consider equation (9) in which T is a square matrix. Due to the strict monotonicity of each component function of F, the mapping $TF(x)$ has the following property: If p, q are arbitrary $2n$-vectors then there is a diagonal matrix $D > 0$ such that

$$TF(p) - TF(q) = TD(p - q), \tag{13}$$

and furthermore, the matrix TD is strongly column-sum dominant (since T is strongly column-sum dominant). We now wish to show that a similar fact is true in the more general case.

With m the number of rows of Q_p, let p and q denote arbitrary m-vectors. Then since there is a diagonal $D > 0$ such that

$$F(Q_1^t Q_2^t \cdots Q_p^t p + C) - F(Q_1^t Q_2^t \cdots Q_p^t q + C) = DQ_1^t Q_2^t \cdots Q_p^t (p - q),$$

we have

$$\tilde{T}\tilde{F}(p) - \tilde{T}\tilde{F}(q) = Q_p \cdots Q_2Q_1TDQ_1^tQ_2^t \cdots Q_p^t(p - q).$$

The fact that $Q_p \cdots Q_2Q_1TDQ_1^tQ_2^t \cdots Q_p^t$ is strongly column-sum dominant follows from the very easily verified proposition that the product $Q_kMQ_k^t$ $(k = 1, 2, \cdots, p)$ possesses that property whenever M does.

Therefore if x^1 and x^2 denote two solutions of the "generalized equation (11)," then $[\tilde{T}\tilde{D} + G](x^1 - x^2) = \theta$ in which

$$\tilde{T} = Q_p \cdots Q_2Q_1T \qquad \text{and} \qquad \tilde{D} = DQ_1^tQ_2^t \cdots Q_p^t .$$

But $\tilde{T}\tilde{D}$, and hence $\tilde{T}\tilde{D} + G$, is strongly column-sum dominant and hence, nonsingular. This implies that $x^1 = x^2$.

We have now shown that in any network constructed from resistors, independent sources, and transistors having a common-base connection, the transistors' base-emitter and base-collector voltages are unique. It is a trivial matter to show that the same result applies when diodes are also allowed to be present in the network.

Suppose the result was not true for some network containing at least one diode. Then there would be two different sets of voltages and currents which satisfy Kirchoff's laws. Thus for each diode in the network there would be two (not necessarily distinct) pairs of points $(v_d^{(1)}, i_d^{(1)})$, $(v_d^{(2)}, i_d^{(2)})$ at which the diode is biased, corresponding to each solution. Letting f denote the strictly monotone increasing function which characterizes the diode we have $i_d^{(1)} = f(v_d^{(1)})$ and $i_d^{(2)} = f(v_d^{(2)})$. But then, suppose the diode is replaced by the series combination of a resistor r and a voltage source E whose values are chosen so that the line $i_d = (1/r)v_d - E/r$ passes through the points $(v_d^{(1)}, i_d^{(1)})$ and $(v_d^{(2)}, i_d^{(2)})$. (Due to the strict monotonicity of f, this can certainly be done with some positive choice of r.) Performing the above type of substitution for each diode in the network, we obtain a new network of the type already considered. This new network would possess two different sets of transistor base-emitter and base-collector voltages (the same as before). This contradicts our previous result, and hence the previous result must apply, even when diodes are present in the network.

To determine the equilibrium solutions of the differential equations which describe a network containing inductors and capacitors as well as the elements mentioned above, one must determine the solutions of a dc equation for a network of the above class. Thus, in summary, what we have shown is: *One cannot synthesize a bistable network which consists of resistors, inductors, capacitors, diodes, independent* *voltage and current sources, and an arbitrary number of (Fig. 2) transistors having a common base connection (or, in particular, only one transistor).*

X. ACKNOWLEDGMENT

The authors would like to acknowledge the helpful conversations with their colleague H. C. So.

APPENDIX A

Algorithms for Computing Solutions of Equation (1)

In this appendix two algorithms for computing the solution of equation (1) are presented. It is proved that one of the algorithms will always converge to the solution of equation (1) if the matrix A satisfies either a weak row-sum or column-sum dominance condition (inequality (*i*) of Theorem 7 with either $\alpha = 1$ or $\alpha = 0$) and if, roughly speaking, the slopes of each f_i are bounded from below by some positive constant. In each case the proof of convergence relies upon Banach's contraction-mapping fixed point theorem, and therefore also represents an independent proof of the existence and uniqueness of a solution of equation (1) for the conditions stated above.

The following notation will be used: For fixed $F \in \mathfrak{F}^n$, $B \in E^n$, let $f(x) \equiv F(x) - B$; also, if A is a given $n \times n$ matrix with elements a_{ij}, we define the diagonal matrix D by $D = \text{diag}\,[a_{11}, a_{22}, \cdots, a_{nn}]$, and let $\Delta = A - D$.

Theorem A. If the $n \times n$ *matrix* A *satisfies*

$$a_{ii} \geqq \sum_{j \neq i} | a_{ij} |, \qquad \text{for} \quad i = 1, \cdots, n,$$

and if $F \in \mathfrak{F}^n$, $B \in E^n$, *and if there exists some* $\epsilon > 0$ *such that for each* $\alpha, \beta \in E^1$, $\epsilon \,| \alpha - \beta | \leqq | f_i(\alpha) - f_i(\beta) |$ *for* $i = 1, \cdots, n$, *then equation* (1) *possesses a unique solution, and if* x^0 *is an arbitrary point in* E^n, *the sequence* $x^0, x^1, x^2, \cdots$ *defined by*

$$x^{k+1} = (f + D)^{-1}(-\Delta)x^k$$

converges to the solution.

Proof: Equation (1) may be rewritten as

$$f(x) + Dx + \Delta x = \theta.$$

Hence, if the operation $T: E^n \rightarrow E^n$ is defined by $T = (f + D)^{-1}(-\Delta)$, then the solution of the equation $x = Tx$ is identical to the solution of equation (1). We now prove that the sequence $x^0, x^1, x^2, \cdots$ converges to this solution by proving that T is a contraction.

Let x and y be arbitrary points in E^n and let $g = Tx$, $h = Ty$. Then, $f(g) + Dg = -\Delta x$ and $f(h) + Dh = -\Delta y$. Thus, for $i = 1, \cdots, n$,

$$f_i(g_i) - b_i + a_{ii}g_i = -(\Delta x)_i ,$$

and

$$f_i(h_i) - b_i + a_{ii}h_i = -(\Delta y)_i .$$

Subtracting, we obtain

$$f_i(g_i) - f_i(h_i) + a_{ii}(g_i - h_i) = (\Delta y)_i - (\Delta x)_i .$$

Since f_i is strictly monotone increasing, we have

$$| f_i(g_i) - f_i(h_i) | + a_{ii} | g_i - h_i | = | (\Delta x)_i - (\Delta y)_i |,$$

and hence, since $\epsilon + a_{ii} > 0$,

$$| g_i - h_i | \leqq \frac{1}{a_{ii} + \epsilon} | (\Delta x)_i - (\Delta y)_i |.$$

Now,

$$\begin{aligned} | (\Delta x)_i - (\Delta y)_i | &= | \sum_{j \neq i} a_{ij}(x_j - y_j) | \\ &\leqq \sum_{j \neq i} (| a_{ij} | \cdot | x_j - y_j |) \\ &\leqq (\sum_{j \neq i} | a_{ij} |) \cdot \max_j | x_j - y_j |. \end{aligned}$$

Thus, defining the metric ρ on E^n by $\rho(x, y) = \max_j | x_j - y_j |$, we have, for $i = 1, \cdots, n$,

$$| g_i - h_i | \leqq \frac{1}{a_{ii} + \epsilon} (\sum_{j \neq i} | a_{ij} |) \cdot \rho(x, y).$$

But, since $0 \leqq \sum_{j \neq i} | a_{ij} | < a_{ii} + \epsilon$, there exists K, $0 \leqq K < 1$, such that $| g_i - h_i | \leqq K \cdot \rho(x, y)$ for $i = 1, \cdots, n$, and in particular, $\rho(Tx, Ty) = \max_i | g_i - h_i | \leqq K \cdot \rho(x, y)$. Hence T is a contraction. □

Theorem B. If the $n \times n$ *matrix* A *satisfies*

$$a_{ii} \geqq \sum_{j \neq i} | a_{ji} |, \qquad \text{for} \quad i = 1, \cdots, n,$$

and if $F \in \mathfrak{F}^n$, $B \in E^n$, *and if there exists some* $\epsilon > 0$ *such that for each*

$\alpha, \beta \in E^1$, $\epsilon \mid \alpha - \beta \mid \leqq \mid f_i(\alpha) - f_i(\beta) \mid$ for $i = 1, \cdots, n$, then equation (1) possesses a unique solution, and if z^0 is an arbitrary point in E^n, the sequence $z^0, z^1, z^2, \cdots$ defined by

$$z^{k+1} = -\Delta(f + D)^{-1}z^k$$

converges to some point z^* and the solution of equation (1) *is given by*

$$x^* = (f + D)^{-1}z^*.$$

Proof: As in Theorem A, the solution of equation (1) is also the solution of $x = (f + D)^{-1}(-\Delta)x$. For each $x \in E^n$, let $z = (f + D)x$ and hence $x = (f + D)^{-1}z$. Thus, x^* is the solution of equation (1) if $x^* = (f + D)^{-1}z^*$, where z^* is the solution of $z = -\Delta(f + D)^{-1}z$. The theorem is thus proved if it is proved that the operator $T \equiv -\Delta(f + D)^{-1}$ is a contraction.

Let P denote the operator $(f + D)^{-1}$, and let x and y be arbitrary points in E^n. Then, proceeding as in the proof of Theorem A, we obtain

$$\mid (Px)_j - (Py)_j \mid \leqq \frac{1}{a_{jj} + \epsilon} \mid x_j - y_j \mid, \qquad \text{for} \quad j = 1, \cdots, n.$$

Thus, if $g = Tx$ and $h = Ty$, then for $i = 1, \cdots, n$,

$$g_i = -\sum_{j \neq i} a_{ij}(Px)_j \qquad \text{and} \qquad h_i = -\sum_{j \neq i} a_{ij}(Py)_j\ .$$

Hence

$$\begin{aligned} \mid g_i - h_i \mid &= \mid \sum_{j \neq i} a_{ij}((Px)_j - (Py)_j) \mid \\ &\leqq \sum_{j \neq i} (\mid a_{ij} \mid \cdot \mid (Px)_j - (Py)_j \mid) \\ &\leqq \sum_{j \neq i} \left(\mid a_{ij} \mid \cdot \frac{1}{a_{jj} + \epsilon} \cdot \mid x_j - y_j \mid \right). \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{i=1}^{n} \mid g_i - h_i \mid &\leqq \sum_{i=1}^{n} \sum_{j \neq i} \frac{\mid a_{ij} \mid}{a_{jj} + \epsilon} \mid x_j - y_j \mid \\ &= \sum_{j=1}^{n} \left(\sum_{i \neq j} \frac{\mid a_{ij} \mid}{a_{jj} + \epsilon} \right) \mid x_j - y_j \mid. \end{aligned}$$

But, there exists K, $0 \leqq K < 1$, such that, for $j = 1, \cdots, n$,

$$\sum_{i \neq j} \frac{\mid a_{ij} \mid}{a_{jj} + \epsilon} \leqq K,$$

and hence

$$\sum_{i=1}^{n} | g_i - h_i | \leqq K \sum_{j=1}^{n} | x_j - y_j |.$$

Defining the metric ρ on E^n by

$$\rho(x, y) = \sum_{i=1}^{n} | x_i - y_i |,$$

we therefore have

$$\rho(Tx, Ty) = \sum_{i=1}^{n} | g_i - h_i | \leqq K \cdot \rho(x, y),$$

and hence T is a contraction. □

APPENDIX B

Determination of Bounds on the Solution of Equation (1)

In this appendix we present a method for determining bounds on the solution of equation (1) when $F \in \mathfrak{F}^n$, A is weakly row-sum dominant, and (for given $\alpha \equiv (\alpha_1, \cdots, \alpha_n)^t$, $\beta \equiv (\beta_1, \cdots, \beta_n)^t \in E^n$) $B \equiv (b_1, \cdots, b_n)^t$ satisfies $\alpha_i \leqq b_i \leqq \beta_i$ for $i = 1, \cdots, n$. The solution bounds are, in general, easier to compute than those of Theorem 5. The method presented here is a generalization of an idea presented in Ref. 3.

The computation of the solution bounds proceeds in two steps. First, one solves each of the equations

$$F(x) = \alpha \tag{14a}$$

and

$$F(x) = \beta. \tag{14b}$$

Denoting the solutions of equations (14a) and (14b) by $\mu \equiv (\mu_1, \cdots, \mu_n)^t$ and $\nu \equiv (\nu_1, \cdots, \nu_n)^t$, respectively, and defining

$$\lambda = \max\{| \mu_1 |, \cdots, | \mu_n |, | \nu_1 |, \cdots, | \nu_n |\},$$

and

$$B' = (\sum_{j \neq 1} | a_{1j} |, \cdots, \sum_{j \neq n} | a_{nj} |)^t,$$

one then solves each of the equations

$$F(x) + \text{diag}\,[a_{11}, \cdots, a_{nn}]x = \alpha - \lambda B', \tag{15a}$$

$$F(x) + \text{diag}\,[a_{11}\,,\ \cdots\,,\ a_{nn}]x = \beta + \lambda B'. \tag{15b}$$

Denoting the solutions of equations (15a) and (15b) by $\gamma \equiv (\gamma_1\,,\ \cdots\,,\ \gamma_n)^t$ and $\delta \equiv (\delta_1\,,\ \cdots\,,\ \delta_n)^t$, respectively, one has $\gamma_i \leqq x_i^0 \leqq \delta_i$ for $i = 1, \cdots, n$, where x^0 is the solution of equation (1) that corresponds to any B satisfying $\alpha_i \leqq b_i \leqq \beta_i$ for $i = 1, \cdots, n$.

To prove that the components of the vectors γ and δ, determined by the above procedure, are indeed bounds for the corresponding components of the solution x^0 involves no more than a word-for-word repetition of the proof of Theorem 2 of Ref. 3, with several quite obvious modifications. We omit the details.

REFERENCES

1. Sandberg, I. W., "On the Properties of Some Systems that Distort Signals-I," B.S.T.J., *42,* No. 5 (September 1963), pp. 2033–2046.
2. Minty, G. J., "Two Theorems on Nonlinear Functional Equations in Hilbert Space," Bull. Amer. Math. Soc., *69,* No. 5 (September 1963), pp. 691–692.
3. Willson, Jr., A. N., "On the Solutions of Equations for Nonlinear Resistive Networks," B.S.T.J., *47,* No. 8 (October 1968), pp. 1755–1773.
4. Calahan, D. A., *Computer-Aided Network Design* (Preliminary Ed.), New York: McGraw-Hill, 1968.
5. Fiedler, M. and Pták, V., "On Matrices with Non-Positive Off-Diagonal Elements and Positive Principal Minors," Czech. Math. J., *12,* No. 3 (1962), pp. 382–400.
6. Fiedler, M. and Pták, V., "Some Generalizations of Positive Definiteness and Monotonicity," Numer. Math., *9,* No. 2 (1966), pp. 163–172.
7. Palais, R. S., "Natural Operations on Differential Forms," Trans. Amer. Math. Soc., *92,* No. 1 (1959), pp. 125–141.
8. Holzmann, C. A. and Liu, R., "On the Dynamical Equations of Nonlinear Networks with n-Coupled Elements," Proc. Third Ann. Allerton Conf. on Circuit and System Theory., (U. of Illinois, 1965), pp. 536–545.
9. Marcus, M., *Basic Theorems in Matrix Theory,* Washington, D. C., National Bureau of Standards Applied Mathematics Series, *57* (1960).