

Theorems on the Computation of the Transient Response of Nonlinear Networks Containing Transistors and Diodes

By I. W. SANDBERG

(Manuscript received June 15, 1970)

We consider in detail the nonlinear equations encountered at each time step when certain implicit numerical-integration algorithms are used. In terms of only the properties of the Jacobian matrix of the pertinent set of differential equations, we present necessary and sufficient conditions for the existence and uniqueness of the solution of the nonlinear equations for all continuous forcing functions and any given step size. Since engineers often think about dynamic nonlinear transistor network problems in terms of the eigenvalues of the relevant Jacobian matrix, the results described are of immediate conceptual value. In particular, it is possible to carry out the algorithms whenever the conditions presented are satisfied.

Several other types of results are also presented. For example, for a special but significant and useful numerical-integration formula, theorems are proved concerning properties of the computed sequence such as the extent to which the sequence is relatively immune to small local errors introduced at each step as a result of the fact that it is ordinarily not possible to compute the solution of a certain equation exactly.

All of the results are concerned with network models that are often used in computer simulations. In fact, we heavily exploit some special properties possessed by the nonlinear functions associated with such models.

I. INTRODUCTION

The set P_0 of all real square matrices each with all principal minors nonnegative plays a key role in the study¹⁻³ of nonlinear equations of the form $F(x) + Ax = B$, and more generally⁴ of equations of the form $CF(x) + Ax = B$, in which $F(\cdot)$ is a "diagonal monotone-nondecreasing mapping" of real Euclidean n -space E^n into itself, A and C are real

$n \times n$ matrices and B is an element of E^n . Such equations arise in the dc analysis of transistor networks, the computation of the transient response of transistor networks, and the numerical solution of certain nonlinear partial-differential equations.

In Ref. 3 a nonuniqueness theorem is proved which focuses attention on a simple special property of transistor-type nonlinearities. It shows that for any transistor-type exponential $F(\cdot)$ the equation $F(x) + Ax = B$ has at least two solutions x for some $B \in E^n$ whenever $A \notin P_0$. The theorem shows that some earlier conditions^{1,2} for the existence of a unique solution cannot be improved by taking into account more information concerning the nonlinearities, and therefore makes more clear that the set of matrices P_0 plays a basic role in the theory of nonlinear transistor networks. Ref. 3 also contains material concerned with the convergence of algorithms for computing the solution of $F(x) + Ax = B$ as well as of more general equations, and some related problems concerning the numerical integration of the ordinary differential equations which govern the transient response of nonlinear transistor networks are considered briefly.

The primary purpose of this paper is to present the results of a continuation of the numerical integration study initiated in Ref. 3. Here we further exploit the special property of transistor-type exponential nonlinearities used in Ref. 3.

We consider in detail the nonlinear equations encountered at each time step when certain implicit numerical-integration algorithms are used, and, in terms of only the properties of the Jacobian matrix of the pertinent set of differential equations, we present necessary and sufficient conditions for the existence and uniqueness of the solution of the nonlinear equations for all continuous forcing functions and any given step size. Since engineers often think about dynamic nonlinear transistor network problems in terms of the location of the eigenvalues of the relevant Jacobian matrix, the results described in Section 2.2 are of immediate conceptual value. In particular, these results are of a very different character than those that appear in the literature, and whenever the conditions presented are satisfied, it is possible to carry out the algorithms. Under the assumption that the conditions are satisfied, we also show that there are convergent algorithms for solving the nonlinear equations, and that the Jacobian matrix of the nonlinear equations is essentially always at least weakly well-conditioned in a significant sense.

A part of Section 2.3 reports on a general result concerning conditions under which it is possible to invert nonlinear mappings in E^n . More

explicitly, we show that a proposition proved by G. H. Meyer enables us to give a short proof of a new theorem which is a considerably stronger result than that described and used in Ref. 11.

We also present a set of results concerning properties of an important class of transistor-diode networks for which certain implicit numerical-integration algorithms can be carried out for all values of the step size, and, for a special but significant and useful numerical-integration formula, theorems are proved concerning some properties of the computed sequence such as the extent to which the sequence is relatively immune to small local errors introduced at each step as a result of the fact that it is ordinarily not possible to compute the solution of a certain equation exactly.

Finally, in addition to other results, we present new theorems concerning the existence of solutions of the nonlinear dc equation under very realistic assumptions from the viewpoint of models often used in computer simulations.[†]

Section II contains a detailed discussion of the results and their significance.

II. TRANSIENT RESPONSE OF TRANSISTOR-DIODE NETWORKS AND IMPLICIT NUMERICAL-INTEGRATION FORMULAS

2.1 Introduction

We shall consider explicitly only networks containing transistors, diodes, and resistors. However, the material to be presented can be extended to take into account other types of elements as well. In addition, we shall focus attention on the use of linear multipoint integration formulas of closed (i.e., of implicit) type, since such formulas are of considerable use in connection with the typically "stiff systems" of differential equations encountered.

A very large class of networks containing resistors, transistors, and diodes modeled in a standard manner is governed by the equation^{5,‡}

$$\frac{du}{dt} + TF[C^{-1}(u)] + GC^{-1}(u) = B(t), \quad t \geq 0 \quad (1)$$

[†] Results concerning the dc equation are directly relevant to the problem of computing the transient response to the extent that in order to numerically integrate the differential equations it is ordinarily necessary to first solve a dc problem to determine the initial conditions.

[‡] As a practical matter, the models of transistors and diodes employed here are often used in computer simulations. Of course in some cases it is necessary to use more complicated models.

with $G = \hat{G}(I + R\hat{G})^{-1}$ and where, assuming that there are q diodes and p transistors,

- (i) $T = T_1 \oplus T_2 \oplus \cdots \oplus T_p \oplus I_q$, the direct sum of the identity matrix of order q and p 2×2 matrices T_k in which

$$T_k = \begin{bmatrix} 1 & -\alpha_r^{(k)} \\ -\alpha_f^{(k)} & 1 \end{bmatrix}$$

with $0 < \alpha_r^{(k)} < 1$ and $0 < \alpha_f^{(k)} < 1$ for $k = 1, 2, \dots, p$.

- (ii) $R = R_1 \oplus R_2 \oplus \cdots \oplus R_p \oplus R_0$, the direct sum of a diagonal matrix $R_0 = \text{diag}(r_1, r_2, \dots, r_q)$ with $r_k \geq 0$ for $k = 1, 2, \dots, q$ and p 2×2 matrices R_k in which for all $k = 1, 2, \dots, p$

$$R_k = \begin{bmatrix} r_e^{(k)} + r_b^{(k)} & r_b^{(k)} \\ r_b^{(k)} & r_c^{(k)} + r_b^{(k)} \end{bmatrix}$$

with $r_e^{(k)} \geq 0$, $r_b^{(k)} \geq 0$, and $r_c^{(k)} \geq 0$. (The matrix R takes into account the presence of bulk resistance in series with the diodes and the emitter, base, and collector leads of the transistors.)

- (iii) \hat{G} is the short-circuit conductance matrix associated with the resistors of the network. (It does not take into account the bulk resistances of the semiconductor devices.)

- (iv) $F(\cdot)$ is a mapping of $E^{(2p+q)}$ into $E^{(2p+q)}$ defined by the condition that

$$F(x) = [f_1(x_1), f_2(x_2), \dots, f_{2p+q}(x_{2p+q})]^{\text{tr}}$$

for all $x \in E^{(2p+q)}$ with each $f_i(\cdot)$ a continuously-differentiable mapping of E^1 into E^1 such that $f'_i(\alpha) > 0$ for all $\alpha \in E^1$.

- (v) $C^{-1}(\cdot)$ is the inverse of the mapping $C(\cdot)$, of $E^{(2p+q)}$ into itself, defined by

$$C(x) = cx + \tau F(x)$$

for all $x \in E^{(2p+q)}$ with $c = \text{diag}(c_1, c_2, \dots, c_{(2p+q)})$, $\tau = \text{diag}(\tau_1, \tau_2, \dots, \tau_{(2p+q)})$, and with each τ_i and each c_i a positive constant.

- (vi) $B(t)$ is a $(2p+q)$ -vector which takes into account the voltage and current generators present in the network, and

- (vii) u is related to v the vector of ideal-junction voltages of the semiconductor devices (v does not take in account the voltage drops across the bulk resistors) through $C(v) = u$ for all $v \in E^{(2p+q)}$.

Equation (1) is equivalent to[†]

[†] In Ref. 5 it is shown if $B(\cdot)$ is a continuous mapping of $[0, \infty)$ into $E^{(2p+q)}$, then for any initial condition $u^{(0)} \in E^{(2p+q)}$ there exists a unique continuous $(2p+q)$ -vector-valued function $u(\cdot)$ such that $u(0) = u^{(0)}$ and (1) is satisfied for all $t > 0$.

$$\dot{u} + f(u, t) = \theta_{(2p+q)}, \quad t \geq 0 \quad (2)$$

in which of course

$$f(u, t) = TF[C^{-1}(u)] + GC^{-1}(u) - B(t) \quad (3)$$

and $\theta_{(2p+q)}$ is the zero vector of order $(2p + q)$.

It is well known that certain specializations of the general multipoint formula^{6,7}

$$y_{n+1} = \sum_{k=0}^r a_k y_{n-k} + h \sum_{k=-1}^r b_k \tilde{y}_{n-k} \quad (4)$$

in which

$$\tilde{y}_{n-k} = -f(y_{n-k}, (n-k)h) \quad (5)$$

can be used as a basis for computing the solution of equation (2). Here h , a positive number, is the step size, the a_k and the b_k are real numbers, and of course y_n is the approximation to $u(nh)$ for $n \geq 1$.

In the literature dealing with formulas of the type (4) in connection with systems of equations of the type (2), information concerning the location of the eigenvalues of the Jacobian matrix J_u of $f(u, t)$ with respect to u plays an important role in determining whether or not a given formula will be (in some suitable sense) stable. In particular, an assumption often made is that all of the eigenvalues of J_u lie in the strict right-half plane for all $t \geq 0$ and all u . For $f(u, t)$ given by equation (3), we have

$$J_u = T \operatorname{diag} \left\{ \frac{f'_i[g_i(u_i)]}{c_i + \tau_i f'_i[g_i(u_i)]} \right\} + G \operatorname{diag} \left\{ \frac{1}{c_i + \tau_i f'_i[g_i(u_i)]} \right\} \quad (6)$$

in which for $j = 1, 2, \dots, (2p + q)$ $g_j(u_j)$ is the j th component of $C^{-1}(u)$. Thus here J_u is a matrix of the form

$$TD_1 + GD_2 \quad (7)$$

where D_1 and D_2 are diagonal matrices with positive diagonal elements. A simple result concerning (7), Theorem 4 of Ref. 3, asserts that if there exists a diagonal matrix D with positive diagonal elements such that†

(i) DT is strongly column-sum dominant, and

(ii) DG is weakly column-sum dominant,

then for all diagonal matrices D_1 and D_2 with positive diagonal elements,

† The terms "strongly-column sum dominant" and "weakly column-sum dominant" are reasonably standard. However, they are defined in Section III.

all eigenvalues of (7) lie in the strict right-half plane. This condition on T and G is often satisfied.[†]

The subclass of numerical integration formulas (4) defined by the condition that $b_{-1} > 0$ are of considerable use⁸⁻¹⁰ in applications involving the typically "stiff systems" of differential equations encountered in the analysis of nonlinear transistor networks. With $b_{-1} > 0$, y_{n+1} is defined *implicitly* through

$$y_{n+1} + hb_{-1}f(y_{n+1}, (n+1)h) = \sum_{k=0}^r a_k y_{n-k} + h \sum_{k=0}^r b_k \tilde{y}_{n-k}$$

in which the right side depends on y_{n-k} only for $k \in \{0, 1, 2, \dots, r\}$, and for $f(u, t)$ given by equation (3), we have

$$y_{n+1} + hb_{-1}\{TF[C^{-1}(y_{n+1})] + GC^{-1}(y_{n+1})\} = q_n \quad (8)$$

in which

$$q_n = \sum_{k=0}^r a_k y_{n-k} + h \sum_{k=0}^r b_k \tilde{y}_{n-k} + hb_{-1}B[(n+1)h].$$

Obviously, the numerical integration formula (8) makes sense only if there exists for each n a $y_{n+1} \in E^{(2p+q)}$ such that (8) is satisfied.

2.2 The Jacobian Matrix J_u and Necessary and Sufficient Conditions for the Existence of a Unique Solution y_{n+1} of (8) for All $q_n \in E^{(2p+q)}$

Here we shall make the additional assumption that the functions $f_j(\cdot)$ are such that the mapping $F(\cdot)$ belongs to the set $\mathcal{F}_0^{(2p+q)}$ defined in Section 3.1. This assumption is satisfied whenever the $f_j(\cdot)$ are the usual Ebers-Moll exponential-type nonlinearities. That is, $\mathcal{F}_0^{(2p+q)}$ contains all of the mappings $F(\cdot)$ such that for each j

$$f_j(x_j) = a_j[\exp(b_j x_j) - 1] \quad \text{or} \quad f_j(x_j) = a_j[1 - \exp(-b_j x_j)]$$

for all $x_j \in E^1$ with a_j and b_j positive constants.

Our first result, Theorem 1 of Section III, is a rather strong result concerning the relation between properties of the Jacobian matrix J_u and properties of equation (8). Let Ξ denote the set of all real numbers σ such that $\det(\sigma I + J_u) = 0$ for some $u \in E^{(2p+q)}$. In other words, let Ξ denote the set of all real numbers σ such that $-\sigma$ is an eigenvalue of J_u at some point u . According to Theorem 1, equation (8) possesses a unique solution y_{n+1} for each $q_n \in E^{(2p+q)}$ (and hence each $B[(n+1)h] \in E^{(2p+q)}$) if and only if $(hb_{-1})^{-1} \notin \Xi$, and also if $(hb_{-1})^{-1} \in \Xi$ then equation (8) possesses at least two solutions for some $q_n \in E^{(2p+q)}$ (and hence for

[†] See Ref. 5 for examples.

some $B[(n+1)h] \in E^{(2p+q)}$. Therefore, in particular, equation (8) possesses a unique solution for all $q_n \in E^{(2p+q)}$ and all $h \in (0, \bar{h}]$, in which \bar{h} is an arbitrary positive constant, if and only if the intersection of the interval $[(\bar{h}b_{-1})^{-1}, \infty)$ and Ξ is the null set, and equation (8) possesses a unique solution for all $q_n \in E^{(2p+q)}$ and all $h > 0$ if and only if Ξ contains no points of the interval $(0, \infty)$. Finally, as a somewhat peripheral matter, according to Theorem 1, the dc equation $TF(v) + Gv = B$ has at most one solution v for each $B \in E^{(2p+q)}$ if and only if $0 \notin \Xi$.

The statements made in the preceding paragraph are surprising to the extent that on the one hand they are rather definitive and on the other hand they involve only the location of the real eigenvalues of J_u .[†] Since engineers often find it helpful to think about nonlinear systems in terms of the location of the eigenvalues of a pertinent Jacobian matrix, it is also of interest to note here that equation (8) can possess more than one solution y_{n+1} for some q_n and some $h > 0$ only if the transistor-diode network is locally exponentially unstable at some operating point, that is, only if at some operating point u , $-J_u$ has a real positive eigenvalue.

2.3 Existence of Convergent Algorithms for Computing the Solution of (8)

Throughout this section we assume that the $f_i(\cdot)$ are such that the additional condition that $F(\cdot) \in \mathfrak{F}_0^{(2p+q)}$ is satisfied.

Whenever $(hb_{-1})^{-1}$ is not contained in the set Ξ of Section 2.2, equation (8), which we shall write as $Q(y_{n+1}) = q_n$, possesses a unique solution y_{n+1} for any $q_n \in E^{(2p+q)}$. We show here that when $(hb_{-1})^{-1} \notin \Xi$ and each $f_i(\cdot)$ is twice continuously differentiable on E^1 ,[‡] there exist steepest descent as well as Newton-type algorithms each of which generates a sequence in $E^{(2p+q)}$ which converges to y_{n+1} .

Assume that $(hb_{-1}) \notin \Xi$. The Jacobian matrix $(I + hb_{-1}J_{y_{n+1}})$ of $Q(\cdot)$ satisfies

$$\det(I + hb_{-1}J_{y_{n+1}}) \neq 0 \quad \text{for all } y_{n+1} \in E^{(2p+q)}. \quad (9)$$

Hence $Q(\cdot)$ is a local homeomorphism on $E^{(2p+q)}$ and since there exists a unique $y_{n+1} \in E^{(2p+q)}$ such that $Q(y_{n+1}) = q_n$ for each $q_n \in E^{(2p+q)}$, $Q(\cdot)$

[†] Indeed, while we can write (8) as $Q(y_{n+1}) = q_n$ with $Q(\cdot)$ a continuously-differentiable mapping of $E^{(2p+q)}$ into itself with Jacobian matrix $(I + hb_{-1}J_{y_{n+1}})$ recall that for $R(\cdot)$ a general continuously-differentiable mapping of E^n into itself with Jacobian matrix J , $\det J \neq 0$ throughout E^n does not imply that (and is not implied by the statement that) for each $x \in E^n$ there exists a unique $y \in E^n$ such that $R(y) = x$, even for $n = 1$.

[‡] This differentiability condition is obviously satisfied if the $f_i(\cdot)$ are the usual exponential functions.

is a homeomorphism of $E^{(2p+q)}$ onto itself. Thus, with $\|\cdot\|$ any norm on $E^{(2p+q)}$,

$$\|Q(y)\| \rightarrow \infty \quad \text{as} \quad \|y\| \rightarrow \infty.^\dagger$$

Let $R(\cdot)$ be defined by the condition that $R(y) = Q(y) - q_n$ for all $y \in E^{(2p+q)}$. Then $R(\cdot)$ satisfies $\|R(y)\| \rightarrow \infty$ as $\|y\| \rightarrow \infty$ and the determinant of the Jacobian matrix of $R(\cdot)$ does not vanish throughout $E^{(2p+q)}$. Therefore, assuming that $R(\cdot)$ is twice continuously differentiable on $E^{(2p+q)}$, it follows (see the Appendix) that the solution y_{n+1} of $R(y_{n+1}) = \theta_{(2p+q)}$ can be computed by using certain steepest descent or Newton-type algorithms.

2.4 The Jacobian Matrix ($I + hb_{-1}J_{y_{n+1}}$), and Inversion of Nonlinear Operators on E^m and Jacobian Matrices

As in Section 2.3, let the additional condition that $F(\cdot) \in \mathcal{F}_0^{(2p+q)}$ be satisfied and let $Q(\cdot)$ be the mapping of $E^{(2p+q)}$ into itself with the property that equation (8) can be written as $Q(y_{n+1}) = q_n$. According to Theorem 2 of Section III the Jacobian matrix ($I + hb_{-1}J_{y_{n+1}}$) possesses the property that there exists a constant $\epsilon > 0$ such that

$$\det(I + hb_{-1}J_{y_{n+1}}) \geq \epsilon \quad \text{for all} \quad y_{n+1} \in E^{(2p+q)} \quad (10)$$

if and only if the matrix

$$[(hb_{-1})^{-1}\tau + T]^{-1}[(hb_{-1})^{-1}c + G],$$

which we shall call S , belongs to the set P of all real square matrices each with all principal minors positive. Thus when $S \in P$ the matrix ($I + hb_{-1}J_{y_{n+1}}$) is well conditioned in at least the weak sense of (10). This fact is of some interest for two reasons. First, certain standard algorithms require that the matrix ($I + hb_{-1}J_{y_{n+1}}$) be inverted along a sequence of points $\{y_{n+1}^{(k)}\}$ in order to compute the solution y_{n+1} of equation (8), and, secondly, Theorem 3 of Section III shows that if $\det[(hb_{-1})^{-1}I + J_u] \neq 0$ for all $u \in E^{(2p+q)}$ and all $(hb_{-1})^{-1} \in g'$ in which g' denotes either $(0, \infty)$ or any interval contained in $(0, \infty)$, then $S \in P$ for all but at most a finite number of points $(hb_{-1})^{-1}$ contained in g' . Therefore, referring to the material of Section 2.2, if $Q(y_{n+1}) = q_n$ possesses a unique solution y_{n+1} for all $q_n \in E^{(2p+q)}$ and all $(hb_{-1})^{-1} \in g'$, then ($I + hb_{-1}J_{y_{n+1}}$) is at least weakly well conditioned at all but at most a finite number of points contained in g' .

[†] Since $Q(\cdot)$ is a homeomorphism of $E^{(2p+q)}$ onto itself, $Q(\cdot)^{-1}$ exists and is continuous. Therefore, the image of any closed ball in $E^{(2p+q)}$ under $Q(\cdot)^{-1}$ is contained in some closed ball in $E^{(2p+q)}$, and hence $\|Q(y)\| \rightarrow \infty$ as $\|y\| \rightarrow \infty$.

Since the elements of $(I + hb_{-1}J_{v_{n+1}})$ are bounded on $y_{n+1} \in E^{(2p+q)}$, it follows from a theorem described by M. Vohovec¹¹ that for each $q_n \in E^{(2p+q)}$ there exists a unique $y_{n+1} \in E^{(2p+q)}$ such that $Q(y_{n+1}) = q_n$ if $S \in P$. More explicitly, the theorem described[†] by Vohovec asserts that if $R(\cdot)$ is a continuously-differentiable mapping of E^n into E^n with $J(R)_q$ the Jacobian matrix of $R(\cdot)$ at an arbitrary point $q \in E^n$, if the elements of $J(R)_q$ are bounded on E^n , and if there exists a positive constant ϵ such that $\det J(R)_q \geq \epsilon$ for all $q \in E^n$, then $R(\cdot)$ is a homeomorphism. Thus, using the theorem of Ref. 11 and Theorems 2 and 3 of Section III, we are able to show that if $\det [(hb_{-1})^{-1}I + J_u] \neq 0$ for all $u \in E^{(2p+q)}$ and all $(hb_{-1})^{-1} \in \mathcal{S}'$, then for all but at most a finite number of points $(hb_{-1})^{-1} \in \mathcal{S}'$, (8) possesses a unique solution y_{n+1} for each $q_n \in E^{(2p+q)}$. Although this result is obviously much weaker than the existence proposition presented in Section 2.2, it shows that the theorem of Ref. 11 can be exploited to provide some insight in connection with the specific problem considered here.

The theorem of Ref. 11 is of interest primarily because the key hypothesis concerns only the determinant $\det J(R)_q$ (as opposed to the condition of Palais[†] that $\|R(q)\| \rightarrow \infty$ as $\|q\| \rightarrow \infty$). Theorem 4 of Section III is a general result which is considerably stronger than the theorem of Ref. 11. It shows that the condition of the theorem of Ref. 11 that there exist a positive constant ϵ such that $\det J(R)_q \geq \epsilon$ for all q can be replaced with the condition that there exist real constants $a > 0$ and $b \geq 0$ such that

$$\det J(R)_q \geq \frac{1}{a + b \|q\|} \quad \text{for all } q \in E^n.$$

2.5 A Class of Networks for Which (8) Possesses a Unique Solution for All Values of the Step Size

There is an interesting class of transistor-diode-resistor networks with the property that for each network in the class, equation (8) possesses a unique solution for all $h > 0$ (i.e., for all $h > 0$, all $q_n \in E^{(2p+q)}$, and all diagonal matrices c and τ with positive diagonal elements). In order to define and discuss that class, consider the dc equation $TF(v) + Gv = B$ in which v is the $(2p + q)$ -vector of semiconductor ideal-junction voltages and $B \in E^{(2p+q)}$. If $p > 0$ and the matrix R of Section 2.1 is the zero matrix, v_1 is the emitter-to-base voltage of transistor one, v_2 is the collector-to-base voltage of transistor one, and so forth. By port

[†] According to Vohovec, the theorem was recently proved by I. Vidar, and the proof is expected to appear in the journal *Glasnik Matematički*.

[‡] See Ref. 12 and the appendix of Ref. 13. Here $\|\cdot\|$ denotes any norm on E^n .

j of the transistor-diode-resistor network we mean the terminal pair between which the voltage v_j appears. Again we shall make the assumption that $F(\cdot) \in \mathcal{F}_0^{(2p+q)}$.

In Ref. 3 it is proved that $TF(v) + Gv = B$ possesses at most one solution v for each $B \in E^{(2p+q)}$ if and only if $T^{-1}G \in P_0$. It is also proved in Ref. 3 that equation (8) possesses a unique solution y_{n+1} for each $q_n \in E^{(2p+q)}$ and each $h > 0$ if $M^{-1}G \in P_0$ for all $M \in \mathcal{J}(T)$ in which here $\mathcal{J}(T)$ denotes the set of all real matrices having the same form as T and with the " α 's" of M not larger than those of T .[†] In other words, it was also proved in Ref. 3 that equation (8) possesses a unique solution y_{n+1} for each $q_n \in E^{(2p+q)}$ and each $h > 0$ if the dc equation possesses at most one solution for each $B \in E^{(2p+q)}$ for "the original set of α 's as well as for an arbitrary set of not-larger α 's." Before proceeding, and for the sake of completeness, we mention here that the same result can be obtained by way of the approach of Section 2.2; a direct corollary of Theorem 5 of Section III, Corollary 1, shows that if $M^{-1}G \in P_0$ for all $M \in \mathcal{J}(T)$, then $\det(\sigma I + J_w) \neq 0$ for all real $\sigma \geq 0$ and all $u \in E^{(2p+q)}$.

Theorem 5 of Section III provides considerable information concerning the nature of the class of networks for which $M^{-1}G \in P_0$ for all $M \in \mathcal{J}(T)$. In particular, the theorem shows that $M^{-1}G \in P_0$ for all $M \in \mathcal{J}(T)$ if and only if $M^{-1}G \in P_0$ for all $M \in \mathcal{J}_0(T)$ in which $\mathcal{J}_0(T)$ is the set of all 2^{2p} real square matrices M having the same form as T and with each " α " of M either zero or the corresponding " α " of T .[‡] The theorem also shows that " $M^{-1}G \in P_0$ for all $M \in \mathcal{J}(T)$ " is equivalent to each of six other statements involving T and G . For example, according to Theorem 5, we have $M^{-1}G \in P_0$ for all $M \in \mathcal{J}(T)$ if and only if either $T^{-1}(G + D) \in P_0$ for all diagonal matrices D with positive diagonal elements, which has an obvious network interpretation in terms of the addition of resistors to the network characterized by G , or $T^{-1}G \in P_0$ and $(T_w)^{-1}G_w \in P_0$ for all pairs of matrices T_w and G_w obtained from T and G , respectively, by deleting an arbitrary set w of rows, and the same set of columns, of both T and G .

When the matrix R of Section 2.1 is the zero matrix, the last condition on T and G of the preceding paragraph also has a simple network interpretation: Given T and G , we have $T^{-1}G \in P_0$, and any network obtained from the network characterized by T and G by short-circuiting an arbitrary set w of at most all but one of the $(2p + q)$ semiconductor junctions possesses the following property. With respect to the voltage vector v_w associated with the junctions not short-circuited, and with

[†] See Definition 4 of Section III for a precise definition of $\mathcal{J}(T)$.

[‡] See Definition 5 of Section III for a precise definition of $\mathcal{J}_0(T)$.

the components of v_w taken in the same order as those of v , the "new T and G " matrices[†] T_w and G_w satisfy $(T_w)^{-1}G_w \in P_0$. As reasonable as this condition or any of the other seven equivalent conditions of Theorem 5 might seem, and even though, as Theorem 6 of Section III shows, $T^{-1}G \in P_0$ implies that $(T_w)^{-1}G_w \in P_0$ whenever w has the property that if the port number associated with one junction of a given transistor is contained in w , then the port number associated with the other junction of that transistor is also contained in w , it is the case that there are transistor-diode-resistor networks for which $T^{-1}G \in P_0$ and $M^{-1}G \notin P_0$ for some $M \in \mathcal{J}(T)$. In fact, Ref. 14 presents an example in which $p = 3$, $q = 0$, $T^{-1}G \in P_0$, and $T^{-1}(G + D) \notin P_0$ for some diagonal matrix D with positive diagonal elements. However, the class of networks for which $T^{-1}G \in P_0$ implies that $M^{-1}G \in P_0$ for all $M \in \mathcal{J}(T)$ is clearly quite large; it obviously includes all networks in which $p = 0$, it includes all networks in which the base terminals of all transistors are connected to a common point, and as Theorem 7 of Section III shows, the class includes all networks in which $T^{-1}G \in P_0$ and $p = 1$ or $p = 2$.^{††}

2.6 Results Concerning the Numerical-Integration Formula $y_{n+1} = y_n + h\tilde{y}_{n+1}$

The general multipoint formula (4) reduces to the well-known implicit numerical-integration formula $y_{n+1} = y_n + h\tilde{y}_{n+1}$ when $a_0 = b_{-1} = 1$, $b_0 = 0$, and $a_k = b_k = 0$ for $k = 1, 2, \dots, r$. For that important special case, and with \tilde{y}_{n+1} given by equations (3) and (5), $\{y_{n+1}\}$ is defined implicitly through

$$y_{n+1} + h\{TF[C^{-1}(y_{n+1})] + GC^{-1}(y_{n+1})\} = y_n + hB_n \quad (11)$$

for all $n \geq 0$, in which $B_n = B[(n+1)h]$. Here we describe some detailed results concerning the relation between the sequences $\{y_{n+1}\}$ and $\{B_n\}$. We assume throughout this section that G is such that there exists a diagonal matrix D with positive diagonal elements with the property that both DT and DG are strongly column-sum dominant. This condition, which is often satisfied,[‡] guarantees that there exists a unique solution[†] y_{n+1} of equation (11) for each $(y_n + hB_n) \in E^{(2p+q)}$.

[†] It is a simple matter to show that the "new T and G " matrices are T_w and G_w .

^{††} It is proved in Ref. 14 that if $q = 0$ and if $p = 1$ or $p = 2$, then $T^{-1}G \in P_0$ implies that $T^{-1}(G + D) \in P_0$ for all diagonal matrices with positive diagonal elements. Thus, by the equivalence of statements (i) and (v) of Theorem 5 of Section III, it follows at once that if $T^{-1}G \in P_0$ then $M^{-1}G \in P_0$ for all $M \in \mathcal{J}(T)$ if $q = 0$ and $p = 1$ or $p = 2$. The proof of essentially the same end result given here is of a very different nature and is quite short.

[‡] See Ref. 5 for examples.

[†] A result mentioned in Section 2.1 implies that if DT and DG are both strongly column-sum dominant, then $\det [(h)^{-1}I + J_u] \neq 0$ for all $u \in E^{(2p+q)}$ and all $h > 0$.

Let $\|\cdot\|_1$ be defined by the condition that $\|v\|_1 = \sum_{j=1}^{(2p+q)} |v_j|$ for all $v \in E^{(2p+q)}$. According to Theorem 8 of Section III, there exists a positive constant δ depending only on the c_i , the τ_i , T , G , and D such that

$$\|Dy_n\|_1 \leq (1 + \delta h)^{-n} \|Dy_0\|_1 + h \sum_{k=1}^n (1 + \delta h)^{-k} \|DB_{(n-k)}\|_1$$

for all $n \geq 1$. Therefore, it follows that for all $h > 0$, the sequence y_1, y_2, \dots is bounded whenever the sequence B_1, B_2, \dots is bounded, and y_1, y_2, \dots approaches $\theta_{(2p+q)}$ the zero vector of $E^{(2p+q)}$ whenever B_1, B_2, \dots approaches $\theta_{(2p+q)}$.

Typically at each step an iterative algorithm is employed to compute the solution y_{n+1} of equation (11). Since it is ordinarily not possible to compute y_{n+1} with infinite precision, it is important to consider the effects of the errors which are introduced. While, ideally, we would like to determine the sequence $\{y_{n+1}\}$ defined by equation (11) and some initial-condition vector y_0 , suppose that we determine instead a sequence $\{\hat{y}_{n+1}\}$ such that, with ϵ an arbitrary positive constant, $\|D(\hat{y}_n - y_n^*)\|_1 \leq \epsilon$ for all $n \geq 1$ and

$$y_{n+1}^* + h\{TF[C^{-1}(y_{n+1}^*)] + GC^{-1}(y_{n+1}^*)\} = \hat{y}_n + hB_n \quad (12)$$

for all $n \geq 0$. That is, suppose that at each step the local error $\|D(\hat{y}_n - y_n^*)\|_1$ in solving for " y_{n+1} " is at most ϵ . Then, according to Theorem 8, and with δ the positive constant referred to above,

$$\begin{aligned} \|D(y_n - \hat{y}_n)\|_1 &\leq (1 + \delta h)^{-n} \|D(y_0 - \hat{y}_0)\|_1 \\ &\quad + \epsilon \sum_{k=0}^n (1 + \delta h)^{-k} \quad \text{for all } n \geq 1 \end{aligned}$$

in which \hat{y}_0 is the approximation to y_0 . Therefore, given an arbitrarily small positive constant ρ , for any $h > 0$ it is possible to choose \hat{y}_0 and $\epsilon > 0$ such that the accumulated-error vector $(y_n - \hat{y}_n)$ satisfies $\|y_n - \hat{y}_n\|_1 \leq \rho$ for all $n \geq 1$.

Finally, Theorem 9 of Section III provides us with a conceptually interesting uniform bound on the norm of the difference between corresponding elements of the sequences $\{y_n\}$ and $\{u_n\}$ in which $u_n = u(nh)$ for all $n \geq 0$ and $u(\cdot)$ satisfies the differential equation (1). According to Theorem 9, there exist positive constants δ and ρ , both independent of h , such that

$$\|D(u_n - y_n)\|_1 \leq (1 + \delta h)^{-n} \|D(u_0 - y_0)\|_1 + \rho h$$

for all $n \geq 1$, assuming that the elements of $B(\cdot)$ and $(d/dt)B(\cdot)$ are

bounded and continuous on $[0, \infty)$. In particular, if $y_0 = u_0$ we see that there exists a positive constant ρ' , independent of h , such that $\|u_n - y_n\|_1 \leq \rho'h$ for all $n \geq 1$, provided only that the assumptions of this section are satisfied and that $B(\cdot)$ and $(d/dt)B(\cdot)$ are bounded and continuous on $[0, \infty)$.

2.7 Conditions Which Imply That $T^{-1}\hat{G}(I + R\hat{G})^{-1} \in P_0$

In this section and in Section 2.8 we present some results concerning properties of the dc equation $TF(v) + Gv = B$. These results are directly relevant to the problem of computing the transient response of transistor-diode networks to the extent that in order to numerically integrate the differential equation (1) it is ordinarily necessary to first solve a dc problem to determine the initial conditions.

As indicated in Section 2.1, $G = \hat{G}(I + R\hat{G})^{-1}$ in which R takes into account the bulk resistances associated with the semiconductor devices. Here we present some material concerning conditions which imply that $T^{-1}\hat{G}(I + R\hat{G})^{-1}$ belongs to P_0 .

Let $p > 0$. Theorem 10 of Section III asserts that $T^{-1}\hat{G}(I + R\hat{G})^{-1} \in P_0$ whenever $T^{-1}\hat{G} \in P_0$ and R satisfies

$$\begin{aligned}\alpha_r^{(k)}(1 - \alpha_r^{(k)})^{-1}r_e^{(k)} &= r_b^{(k)} \\ \alpha_f^{(k)}(1 - \alpha_f^{(k)})^{-1}r_c^{(k)} &= r_b^{(k)}\end{aligned}$$

for $k = 1, 2, \dots, p$. This rather special result shows that if $F(\cdot)$ satisfies the additional condition that $F(\cdot)$ belongs to the set $\mathcal{F}_0^{(2p+q)}$ defined in Section 3.1, and if the network associated with T and \hat{G} possesses the property that there is at most one solution v of the dc equation $TF(v) + \hat{G}v = B$ for each $B \in E^{(2p+q)}$, then it is always possible to add certain resistors of positive value in series with each transistor lead such that the dc equation of the resulting network possesses at most one solution.

Theorem 11 of Section III directs attention to the fact that there is a nontrivial class of transistor networks for which $T^{-1}\hat{G}(I + R\hat{G})^{-1} \in P_0$ for all R . According to Theorem 11, if $p > 0$ and \hat{G} is such that $T^{-1}\hat{G} \in P_0$ for all " α 's" (i.e., for all $\alpha_r^{(k)}$ and $\alpha_f^{(k)}$ belonging to $(0, 1)$), then for any particular set of " α 's" $T^{-1}\hat{G}(I + R\hat{G})^{-1} \in P_0$ for all R .[†]

Given T , an interesting characterization of the class of short-circuit-conductance matrices \hat{G} such that $M^{-1}\hat{G} \in P_0$ for all $M \in \mathcal{J}(T)$ is provided by Theorem 12 of Section III.[‡] According to Theorem 12, $M^{-1}\hat{G} \in P_0$ for all $M \in \mathcal{J}(T)$ if and only if $T^{-1}\hat{G}(I + R\hat{G})^{-1} \in P_0$ for all R satisfying certain inequality-type conditions. In particular, if the base-lead

[†] A similar result is proved in Ref. 2 under the assumption that \hat{G} is not singular.

[‡] The set $\mathcal{J}(T)$ is described in Section 2.5.

resistance of each transistor is taken to be zero, then $M^{-1}\hat{G} \in P_0$ for all $M \in \mathfrak{J}(T)$ implies that $T^{-1}\hat{G}(I + R\hat{G})^{-1} \in P_0$ for all nonnegative values of each emitter-lead resistor and each collector-lead resistor.

2.8 Ebers-Moll Models and the Existence of a Solution of $TF(v) + Gv = B$

In Section III, a set \mathfrak{F}_3 of mappings $F(\cdot)$ is defined such that each element of \mathfrak{F}_3 possesses certain important properties possessed by an arbitrary $F(\cdot)$ of the type that arises when an Ebers-Moll exponential-nonlinear-function model is used for each transistor and diode. In contrast with the set of all $F(\cdot)$ such that each $f_j(\cdot)$ is a strictly-monotone-increasing mapping of E^1 onto E^1 , an arbitrary element $F(\cdot)$ of \mathfrak{F}_3 possesses the properties that for each j , $f_j(\cdot)$ is bounded on either $[0, \infty)$ or $(-\infty, 0]$, and the two nonlinear functions associated with the same transistor are both bounded on either $[0, \infty)$ or $(-\infty, 0]$. The set \mathfrak{F}_3 is contained in $\mathfrak{F}_0^{(2p+q)}$ and contains every Ebers-Moll exponential-nonlinear-function-type $F(\cdot)$.

The first part of Theorem 13 of Section III asserts that the equation $TF(v) + Gv = B$ possesses a unique solution v for each $F(\cdot) \in \mathfrak{F}_3$ and each $B \in E^{(2p+q)}$ if and only if $T^{-1}G \in P_0$ and $\det G \neq 0$. It is the "only if" part of this proposition which is the new result presented here. The proof exploits some special properties of transformerless resistor networks; it shows that if $T^{-1}G \in P_0$ but $\det G = 0$, then there are functions $t(\cdot)$ and $d(\cdot)$, both functions taking on only the values 1 or -1 , such that there is no solution v of $TF(v) + Gv = B$ for some $B \in E^{(2p+q)}$ for any set of Ebers-Moll-modeled transistors and diodes with the property that for all k transistor k is a pnp device (as opposed to a npn device) if and only if $t(k) = 1$, and for all j diode j is a p-n junction if and only if $d(j) = 1$.[†]

The discussion of the preceding paragraph concerning the proof of Theorem 13 shows that it is not possible to make stronger assertions concerning the existence of a unique solution of $TF(v) + Gv = B$ for all $B \in E^{(2p+q)}$ for Ebers-Moll-modeled transistors and diodes unless we take into account more information about the nature of the semiconductor junctions. A good deal of progress in this direction has recently been made, and we state here without proof the following complete result dealing with diode-resistor networks.

Theorem 14:[‡] Let $p = 0$ and $q > 0$. Let $F(\cdot) \in \mathfrak{F}_3$ (see Definition 12 of

[†] In contrast, the proof of the "only if" part of Theorem 3 of Ref. 1 shows that if $A \notin P_0$ then there is a mapping $F(\cdot)$ with each $f_j(\cdot)$ a linear function such that $F(x) + Ax = B$ does not possess a unique solution for all $B \in E^n$.

[‡] The proof of Theorem 14 will be presented in a subsequent paper.

Section 3.31), and for $j = 1, 2, \dots, q$ let s_j equal either 1 or -1 depending on whether $f_j(\cdot)$ is bounded on $[0, \infty)$ or $(-\infty, 0]$, respectively. Then, with A any real symmetric nonnegative-definite matrix of order q , there exists a unique solution v of $F(v) + Av = B$ for all $B \in E^q$ if and only if there is no real q -vector η such that $\eta \neq \theta_q$, $A\eta = \theta_q$, and $\eta \in S$, in which

$$S = \{y: y \in E^q \text{ and } y_j s_j \geq 0 \text{ for } j = 1, 2, \dots, q\}^\dagger$$

III. THEOREMS AND PROOFS

3.1 Notation and Definitions

Throughout Section III,

- (i) unless stated otherwise, p and q denote nonnegative integers such that $(p + q) > 0$, and n denotes an arbitrary positive integer;
- (ii) the set of all real n -vectors is denoted by E^n , θ is the zero element of E^n , and if $v \in E^n$ and j is an integer such that $1 \leq j \leq n$, then v_j denotes the j th component of v ;
- (iii) $\|v\| = (\sum_{i=1}^n v_i^2)^{1/2}$ and $\|v\|_1 = \sum_{i=1}^n |v_i|$ for all $v \in E^n$; for any real $n \times n$ matrix M , $\|M\|$ denotes $\sup \{m: \|Mx\| \leq m \|x\|, x \in E^n\}$;
- (iv) the transpose of an arbitrary (not necessarily square) matrix M is denoted by M^{tr} ;
- (v) I_n denotes the identity matrix of order n , and I denotes the identity matrix of order determined by the context in which the symbol is used; if Q_1, Q_2, \dots, Q_n are square matrices, then $Q_1 \oplus Q_2 \oplus \dots \oplus Q_n$ denotes the direct sum of Q_1, Q_2, \dots, Q_n , in the order indicated;
- (vi) if D is a real diagonal matrix, then $D > 0 (D \geq 0)$ means that the diagonal elements of D are positive (nonnegative); and
- (vii) we say that a real $n \times n$ matrix M is strongly (weakly) column-sum dominant if and only if for $j = 1, 2, \dots, n$

$$m_{jj} > (\geq) \sum_{i \neq j} |m_{ij}|.$$

Definition 1: The set of all real square matrices M such that every principal minor of M is nonnegative (positive) is denoted by $P_0(P)$.

Definition 2: Let $\mathfrak{F}_0^{(2p+q)}$ denote that collection of mappings of $E^{(2p+q)}$ into itself defined by: $F(\cdot) \in \mathfrak{F}_0^{(2p+q)}$ if and only if there exist for $j =$

[†] In the network case, $A = G$, and it is often possible to determine by inspection whether or not there exists an $\eta \neq \theta_q$ such that $G\eta = \theta_q$ and $\eta \in S$.

1, 2, \dots , $(2p + q)$ continuous functions $f_j(\cdot)$ mapping E^1 into E^1 such that for each $x \in E^{(2p+q)}$, $F(x) = [f_1(x_1), f_2(x_2), \dots, f_{(2p+q)}(x_{(2p+q)})]^T$, and

$$(i) \quad \inf_{\alpha \in (-\infty, \infty)} [f_j(\alpha + \beta) - f_j(\alpha)] = 0,$$

$$(ii) \quad \sup_{\alpha \in (-\infty, \infty)} [f_j(\alpha + \beta) - f_j(\alpha)] = +\infty$$

for all $\beta > 0$ and all $j = 1, 2, \dots, (2p + q)$.

Definition 3: Let \mathfrak{J} denote the set of all real matrices M such that $M = M_1 \oplus M_2 \oplus \dots \oplus M_p \oplus I_q$ with

$$M_k = \begin{bmatrix} 1 & -\alpha_r^{(k)} \\ -\alpha_f^{(k)} & 1 \end{bmatrix},$$

$0 \leq \alpha_r^{(k)} < 1$, and $0 \leq \alpha_f^{(k)} < 1$ for all $k = 1, 2, \dots, p$. As suggested, if $q = 0$, then $M = M_1 \oplus M_2 \oplus \dots \oplus M_p$, while if $p = 0$, then $M = I_q$.

Assumption 1: Throughout Section III, G denotes a real nonnegative-definite matrix of order $(2p + q)$.

A tool that we shall use often is:

Lemma 1: A real square matrix M is an element of P_0 if and only if $\det(D + M) \neq 0$ for all real diagonal matrices $D > 0$.

Lemma 1 is proved in Ref. 2.

3.2 Theorem 1: Let $F(\cdot) \in \mathfrak{F}_0^{(2p+q)}$ with each $f_j(\cdot)$ continuously differentiable on $(-\infty, \infty)$ and $f'_j(\alpha) > 0$ for all $\alpha \in (-\infty, \infty)$. Let $T \in \mathfrak{J}$, let $C(\cdot)$ [that is, $c + \tau F(\cdot)$], G , and J_u be as defined in Section 2.1, and let σ be a real nonnegative constant. Then

$$\sigma y + TF[C^{-1}(y)] + GC^{-1}(y) = r \quad (13)$$

possesses at most one solution y for each $r \in E^{(2p+q)}$ if and only if

$$\det(\sigma I + J_u) \neq 0 \quad \text{for all } u \in E^{(2p+q)}, \quad (14)$$

and if $\sigma > 0$ and condition (14) is satisfied then for each $r \in E^{(2p+q)}$ there exists a solution y of (13).

3.3 Proof of Theorem 1

We have

$$\begin{aligned} \det(\sigma I + J_u) &= \det(\sigma I + TF'[g(u)]\{c + \tau F'[g(u)]\}^{-1} + G\{c + \tau F'[g(u)]\}^{-1}) \\ &= \det\{c + \tau F'[g(u)]\}^{-1} \cdot \det\{\sigma c + \sigma \tau F'[g(u)] + TF'[g(u)] + G\}, \end{aligned}$$

in which $g(\cdot)$ is the mapping of $E^{(2p+q)}$ onto itself defined by $g(u) = C^{-1}(u)$ for all $u \in E^{(2p+q)}$, and $F'[g(u)] = \text{diag}\{f'_i[g_i(u_i)]\}$. Since $\det\{c + \tau F'[g(u)]\} > 0$ for all u , $\det(\sigma I + J_u) \neq 0$ for all u if and only if

$$\det\{(\sigma\tau + T)F'[g(u)] + (\sigma c + G)\} \neq 0 \quad \text{for all } u.$$

For each j $g_j(\cdot)$ maps E^1 onto E^1 , and since $F(\cdot) \in \mathcal{F}_0^{(2p+q)}$ with each $f_i(\cdot)$ continuously differentiable on $(-\infty, \infty)$ and $f'_i(\alpha) > 0$ for all $\alpha \in (-\infty, \infty)$, the image of E^1 under the mapping $f'_i[g_i(\cdot)]$ is $(0, \infty)^+$ for all j . Thus, by Lemma 1 (since $\det(\sigma\tau + T) \neq 0$) $(\sigma\tau + T)^{-1}(\sigma c + G) \in P_0$ if and only if

$$\det(\sigma I + J_u) \neq 0 \quad \text{for all } u. \quad (15)$$

The equation

$$\sigma y + TF[C^{-1}(y)] + GC^{-1}(y) = r$$

possesses a solution y if and only if $x = C^{-1}(y)$ satisfies

$$\sigma C(x) + TF(x) + Gx = r,$$

that is, if and only if

$$(\sigma\tau + T)F(x) + (\sigma c + G)x = r. \quad (16)$$

But equation (16) possesses at most one solution for each $r \in E^{(2p+q)}$ if and only if $(\sigma\tau + T)^{-1}(\sigma c + G) \in P_0$ (see pp. 105–107 of Ref. 3) and hence if and only if condition (15) is met.

Suppose now that $\sigma > 0$. Since G is nonnegative definite, $\det(\sigma c + G) \neq 0$. If condition (15) is satisfied then $(\sigma\tau + T)^{-1}(\sigma c + G) \in P_0$ and hence for each $r \in E^{(2p+q)}$, equation (16) possesses a solution x (see p. 99 of Ref. 3). \square

3.4 Theorem 2: Let $T \in \mathcal{S}$, and let $F(\cdot) \in \mathcal{F}_0^{(2p+q)}$ with each $f_i(\cdot)$ continuously differentiable on $(-\infty, \infty)$ and $f'_i(\alpha) > 0$ for all $\alpha \in (-\infty, \infty)$. Then for each $\sigma \geq 0$ there exists a positive constant ϵ such that $\det(\sigma I + J_u) \geq \epsilon$ for all $u \in E^{(2p+q)}$ if and only if $(\sigma\tau + T)^{-1}(\sigma c + G) \in P$.

[†] For any $\beta > 0$ and any $\alpha \in (-\infty, \infty)$, $f_j(\alpha + \beta) - f_j(\alpha) = \beta f'_j(\delta)$ for some $\delta \in [\alpha, \alpha + \beta]$.

3.5 Proof of Theorem 2

We have

$$\begin{aligned} \det(\sigma I + J_u) &= \det(\sigma I + TF'[g(u)]\{c + \tau F'[g(u)]\}^{-1} + G\{c + \tau F'[g(u)]\}^{-1}) \\ &= \det\{c + \tau F'[g(u)]\}^{-1} \cdot \det\{(\sigma\tau + T)F'[g(u)] + (\sigma c + G)\} \\ &= \det(\sigma\tau + T) \frac{\det(F'[g(u)] + A)}{\prod_{j=1}^{(2p+q)} (c_j + \tau_j f'_j[g_j(u_j)])} \end{aligned} \quad (17)$$

in which $A = (\sigma\tau + T)^{-1}(\sigma c + G)$.

For each sequence $e_1, e_2, \dots, e_{(2p+q)}$ with each e_j either zero or unity and $e_1, e_2, \dots, e_{(2p+q)}$ not the sequence $1, 1, \dots, 1$: let $m_{e_1, e_2, \dots, e_{(2p+q)}}$ denote the determinant obtained from A by deleting rows $\rho_1, \rho_2, \dots, \rho_i$ and columns $\rho_1, \rho_2, \dots, \rho_i$ in which $\{\rho_1, \rho_2, \dots, \rho_i\} = \{j: e_j = 1\}$. Thus for each sequence $e_1, e_2, \dots, e_{(2p+q)}$ other than the sequence $1, 1, \dots, 1$ $m_{e_1, e_2, \dots, e_{(2p+q)}}$ is a principal minor of A . Let $m_{1,1,\dots,1} = 1$, and let $d_j = f'_j[g_j(u_j)]$ for all j . Then by a standard expression¹⁵ for the determinant of the sum of two matrices

$$\det(F'[g(u)] + A) = \sum' d_1^{e_1} d_2^{e_2} \dots d_{(2p+q)}^{e_{(2p+q)}} m_{e_1, e_2, \dots, e_{(2p+q)}}$$

in which \sum' denotes a summation over all $2^{(2p+q)}$ sequences $e_1, e_2, \dots, e_{(2p+q)}$ and $d_j^0 = 1$ for all j . It is clear that

$$\prod_{j=1}^{(2p+q)} (c_j + \tau_j f'_j[g_j(u_j)]) = \sum' d_1^{e_1} d_2^{e_2} \dots d_{(2p+q)}^{e_{(2p+q)}} c_{e_1, e_2, \dots, e_{(2p+q)}}$$

in which each $c_{e_1, e_2, \dots, e_{(2p+q)}}$ is a positive constant. Thus with $\eta = \det(\sigma\tau + T)$,

$$\eta^{-1} \det(\sigma I + J_u) = \frac{\sum' d_1^{e_1} d_2^{e_2} \dots d_{(2p+q)}^{e_{(2p+q)}} m_{e_1, e_2, \dots, e_{(2p+q)}}}{\sum' d_1^{e_1} d_2^{e_2} \dots d_{(2p+q)}^{e_{(2p+q)}} c_{e_1, e_2, \dots, e_{(2p+q)}}} \quad (18)$$

Suppose that all principal minors of A are positive. Then there is a positive constant δ such that

$$m_{e_1, e_2, \dots, e_{(2p+q)}} \geq \delta c_{e_1, e_2, \dots, e_{(2p+q)}}$$

for all $e_1, e_2, \dots, e_{(2p+q)}$ and hence (since $d_j > 0$ for all j) $\det(\sigma I + J_u) \geq \eta\delta$ for all $u \in E^{(2p+q)}$.

As in the proof of Theorem 1, the range of each $d_j = f'_j[g_j(u_j)]$ is $(0, \infty)$, and for any positive constants $p_1, p_2, \dots, p_{(2p+q)}$ there exists a $u \in E^{(2p+q)}$ such that $d_j = p_j$ for all j . If $A \notin P$ then at least one principal

minor of A is not positive. If $A \notin P_0$, then $\det(F'[g(u) + A]) = 0$ for some u . Therefore to complete the proof it is sufficient to show that if $A \in P_0$ but $A \notin P$ then there is no constant $\epsilon > 0$ such that $\det(\sigma I + J_u) \geq \epsilon$ for all u .

With $A \in P_0$ and $A \notin P$, for at least one sequence $e'_1, e'_2, \dots, e'_{(2p+q)}$

$$m_{e'_1, e'_2, \dots, e'_{(2p+q)}} = 0.$$

If $\det A = m_{0,0,\dots,0} = 0$ we have

$$\inf_{u \in E^{(2p+q)}} \det(\sigma I + J_u) = 0$$

since $\det(\sigma I + J_u) \rightarrow 0$ as $d_j \rightarrow 0$ for all j . Suppose now that $\det A > 0$ and that $m_{e'_1, e'_2, \dots, e'_{(2p+q)}} = 0$ for some sequence $e'_1, e'_2, \dots, e'_{(2p+q)}$. Then with $d_j = d$ for all j for which $e'_j = 1$ and $d_j = d^{-1}$ for all j for which $e'_j = 0$, we have [see equation (18)] $\det(\sigma I + J_u) \rightarrow 0$ as $d \rightarrow \infty$. \square

3.6 *Theorem 3: Let $T \in \mathfrak{S}$, let $F(\cdot) \in \mathfrak{F}_0^{(2p+q)}$ with each $f_j(\cdot)$ continuously differentiable on $(-\infty, \infty)$ and $f'_j(\alpha) > 0$ for all $\alpha \in (-\infty, \infty)$, and let \mathcal{S} denote $[0, \infty)$ or an interval contained in $[0, \infty)$. Then for all but at most a finite number of points σ contained in \mathcal{S} , there is a real constant $\epsilon_\sigma > 0$ such that $\det(\sigma I + J_u) \geq \epsilon_\sigma$ for all $u \in E^{(2p+q)}$ if and only if $\det(\sigma I + J_u) \neq 0$ for all $\sigma \in \mathcal{S}$ and all $u \in E^{(2p+q)}$.*

3.7 *Proof of Theorem 3*

As in the proof of Theorem 1, $(\sigma\tau + T)^{-1}(\sigma c + G) \in P_0$ for all $\sigma \in \mathcal{S}$ if and only if $\det(\sigma I + J_u) \neq 0$ for all $\sigma \in \mathcal{S}$ and all u . We shall also use the fact that since $\det(\sigma\tau + T) > 0$ for all $\sigma \geq 0$, each principal minor of $(\sigma\tau + T)^{-1}(\sigma c + G)$ is a finite-valued rational function of σ for all $\sigma \geq 0$.

(if) If $\det(\sigma I + J_u) \neq 0$ for all u and all $\sigma \in \mathcal{S}$, then $(\sigma\tau + T)^{-1}(\sigma c + G) \in P_0$ for all $\sigma \in \mathcal{S}$. It is clear that $(\sigma\tau + T)^{-1}(\sigma c + G) \in P$ for all sufficiently large $\sigma > 0$. Thus each principal minor of $(\sigma\tau + T)^{-1}(\sigma c + G)$ is nonnegative for all $\sigma \in \mathcal{S}$ and is positive for all sufficiently large $\sigma > 0$. They are therefore positive for all but at most a finite number of values of $\sigma \in \mathcal{S}$. Thus, by Theorem 2, if $\det(\sigma I + J_u) \neq 0$ for all $\sigma \in \mathcal{S}$ and all u there exist for all but at most a finite number of points $\sigma \in \mathcal{S}$ a positive constant ϵ_σ such that $\det(\sigma I + J_u) \geq \epsilon_\sigma$ for all u .

(only if) If $\det(\sigma I + J_u) = 0$ for some $\sigma \in \mathcal{S}$ and some u , then, for that σ , $(\sigma\tau + T)^{-1}(\sigma c + G) \notin P_0$. That is, for that σ at least one principal minor of $(\sigma\tau + T)^{-1}(\sigma c + G)$ is negative. This means that $(\sigma\tau + T)^{-1}(\sigma c + G) \notin P_0$ for all σ contained in some interval $\mathcal{S}' \subset \mathcal{S}$, and by Theorem 2, for all $\sigma \in \mathcal{S}'$ there is no $\epsilon_\sigma > 0$ such that $\det(I + J_u) \geq \epsilon_\sigma$ for all u . \square

3.8 *Theorem 4:* Let $R(\cdot)$ be a continuously differentiable mapping of E^n into E^n , and let $J(R)_q$ denote the Jacobian matrix of $R(\cdot)$ at an arbitrary point $q \in E^n$. If the elements of $J(R)_q$ are bounded on E^n , and if there exist real constants $a > 0$ and $b \geq 0$ such that $\det J(R)_q \geq (a + b \|q\|)^{-1}$ for all $q \in E^n$, then $R(\cdot)$ is a homeomorphism of E^n onto E^n .

3.9 Proof of Theorem 4

If Ref. 16 Meyer proves[†] that $R(\cdot)$ is a homeomorphism of E^n onto E^n if $J(R)_q^{-1}$ exists for all $q \in E^n$ and there exist real constants $\alpha > 0$ and $\beta \geq 0$ such that $\|J(R)_q^{-1}\| \leq \alpha + \beta \|q\|$ for all $q \in E^n$.

With q an arbitrary element of E^n , let $\lambda_1, \lambda_2, \dots, \lambda_n$ denote the eigenvalues of $J(R)_q^{\text{tr}} J(R)_q$, and let $\lambda_1 = \min_j \{\lambda_j\}$. Then $\lambda_1 \lambda_2 \dots \lambda_n = [\det J(R)_q]^2 \geq (a + b \|q\|)^{-2}$, and since the elements of $J(R)_q$ are bounded on E^n , there is a constant $\lambda > 0$ such that $\lambda_j \leq \lambda$ for all j and all $q \in E^n$. Thus

$$(\lambda_1)^{1/2} \geq \lambda^{-(1/2)(n-1)} (a + b \|q\|)^{-1} \quad (19)$$

for all q . For any $x \in E^n$ and any $q \in E^n$, $x^{\text{tr}} J(R)_q^{\text{tr}} J(R)_q x \geq \lambda_1 x^{\text{tr}} x$; that is,

$$\|J(R)_q x\| \geq (\lambda_1)^{1/2} \|x\| \geq \lambda^{-(1/2)(n-1)} (a + b \|q\|)^{-1} \|x\|.$$

With $x = J(R)_q^{-1} y$ in which y is an arbitrary element of E^n , we have

$$\|J(R)_q^{-1} y\| \leq \lambda^{(1/2)(n-1)} (a + b \|q\|) \|y\|,$$

which shows that our hypothesis concerning $\det J(R)_q$ ensures that Meyer's condition on $\|J(R)_q^{-1}\|$ is satisfied. \square

3.10 Some Further Definitions

Definition 4: For each $T \in \mathfrak{J}$, let $\mathfrak{J}(T)$ denote the set of all matrices M such that $M = M_1 \oplus M_2 \oplus \dots \oplus M_p \oplus I_q$ with

$$M_k = \begin{bmatrix} 1 & -\delta_r^{(k)} \\ -\delta_f^{(k)} & 1 \end{bmatrix}$$

and

$$0 < \delta_r^{(k)} \leq \alpha_r^{(k)} \quad \text{if } \alpha_r^{(k)} > 0 \quad \text{and} \quad \delta_r^{(k)} = 0 \quad \text{if } \alpha_r^{(k)} = 0,$$

$$0 < \delta_f^{(k)} \leq \alpha_f^{(k)} \quad \text{if } \alpha_f^{(k)} > 0 \quad \text{and} \quad \delta_f^{(k)} = 0 \quad \text{if } \alpha_f^{(k)} = 0,$$

for all $k = 1, 2, \dots, p$. As suggested, if $q = 0$, then $M = M_1 \oplus M_2 \oplus \dots \oplus M_p$, while if $p = 0$, then $M = I_q$.

[†] Meyer's result is a generalization of a well-known result of Hadamard.¹⁷ Hadamard proved that $R(\cdot)$ is a homeomorphism if $J(R)_q^{-1}$ exists for all $q \in E^n$ and satisfies $\|J(R)_q^{-1}\| \leq \alpha$ for all $q \in E^n$ for some positive constant α .¹⁷

Definition 5: For each $T \in \mathfrak{S}$, let $\mathfrak{S}_0(T)$ denote the set of all 2^{2p} matrices M such that $M = M_1 \oplus M_2 \oplus \cdots \oplus M_p \oplus I_q$ with

$$M_k = \begin{bmatrix} 1 & -\delta_r^{(k)} \\ -\delta_f^{(k)} & 1 \end{bmatrix}$$

and

$$\delta_r^{(k)} = \alpha_r^{(k)} \quad \text{or} \quad \delta_r^{(k)} = 0,$$

$$\delta_f^{(k)} = \alpha_f^{(k)} \quad \text{or} \quad \delta_f^{(k)} = 0,$$

for all $k = 1, 2, \dots, p$. As suggested, if $q = 0$, then $M = M_1 \oplus M_2 \oplus \cdots \oplus M_p$, while if $p = 0$, then $M = I_q$.

Definition 6: Let $Q_{(2p+q)}$ denote the family of all $2^{(2p+q)} - 1$ sets $w = \{i_1, i_2, \dots, i_r\}$, including the null set, such that $r < (2p + q)$ and $w \subset \{1, 2, \dots, (2p + q)\}$.

Definition 7: For M an arbitrary square matrix of order $(2p + q)$, and for each $w \in Q_{(2p+q)}$, let M_w denote the principal submatrix obtained from M by deleting rows i_1, i_2, \dots, i_r and columns i_1, i_2, \dots, i_r . (If w is the null set, then $M_w = M$.)

Definition 8: For each $j \in \{1, 2, \dots, (2p + q)\}$, let U_j denote the $(2p + q)$ -column-vector with unity in the j th position and zeros in all other positions.

Definition 9: For each $T \in \mathfrak{S}$ and each $w \in Q_{(2p+q)}$, let T^w denote the matrix obtained from T by replacing the j th column of T with U_j for all $j \in w$.

3.11 *Theorem 5:* Let $T \in \mathfrak{S}$. Then the following statements are equivalent.

- (i) $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}(T)$.
- (ii) $(D_a + T)^{-1}(D_b + G) \in P_0$ for all diagonal $D_a \geq 0$ and all diagonal $D_b \geq 0$.
- (iii) $T^{-1}(G + D) \in P_0$ for all diagonal $D \geq 0$.
- (iv) $(D_a + T)^{-1}(D_b + G) \in P_0$ for all diagonal $D_a > 0$ and all diagonal $D_b > 0$.
- (v) $T^{-1}(G + D) \in P_0$ for all diagonal $D > 0$.
- (vi) $(T_w)^{-1}G_w \in P_0$ for all $w \in Q_{(2p+q)}$.
- (vii) $[(T^w)^{-1}G]_w \in P_0$ for all $w \in Q_{(2p+q)}$.
- (viii) $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}_0(T)$.

3.12 Proof of Theorem 5

[(i) and (ii) are equivalent]

By Lemma 1, $(D_a + T)^{-1}(D_b + G) \in P_0$ if and only if $\det [(D_a + T)^{-1}(D_b + G) + D] \neq 0$ for all diagonal $D > 0$. Thus $(D_a + T)^{-1}(D_b + G) \in P_0$ for all $D_a \geq 0$ and all $D_b \geq 0$ if and only if

$$\det [(D_b D^{-1} + D_a + T)D + G] \neq 0$$

for all $D_a \geq 0$, all $D_b \geq 0$, and all $D > 0$, and hence if and only if

$$\det [(\Lambda + T)D + G] \neq 0$$

for all diagonal $\Lambda \geq 0$ and $D > 0$. Let $T_\Lambda = (\Lambda + T)(I + \Lambda)^{-1}$. Then $(D_a + T)^{-1}(D_b + G) \in P_0$ for all $D_a \geq 0$ and all $D_b \geq 0$ if and only if

$$\det [T_\Lambda(I + \Lambda)D + G] \neq 0$$

for all $\Lambda \geq 0$ and all $D > 0$, and hence if and only if $\det (T_\Lambda \hat{D} + G) \neq 0$ for all diagonal $\hat{D} > 0$ and all $\Lambda \geq 0$. By Lemma 1, this means that $T_\Lambda^{-1}G \in P_0$ for all $\Lambda \geq 0$ if and only if $(D_a + T)^{-1}(D_b + G) \in P_0$ for all $D_a \geq 0$ and all $D_b \geq 0$. We observe that $T_\Lambda = (T_\Lambda)_1 \oplus (T_\Lambda)_2 \oplus \cdots \oplus (T_\Lambda)_p \oplus I_q$ in which, with $\Lambda = \text{diag} (\lambda_1, \lambda_2, \dots, \lambda_{(2p+q)})$,

$$(T_\Lambda)_k = \begin{bmatrix} 1 & \frac{-\alpha_r^{(k)}}{1 + \lambda_{2k}} \\ \frac{-\alpha_f^{(k)}}{1 + \lambda_{2k-1}} & 1 \end{bmatrix}$$

for $k = 1, 2, \dots, p$. Thus for each $\Lambda \geq 0$, $T_\Lambda \in \mathfrak{S}(T)$; and if M is an arbitrary element of $\mathfrak{S}(T)$, there is a $\Lambda \geq 0$ such that $M = T_\Lambda$. Therefore $(D_a + T)^{-1}(D_b + G) \in P_0$ for all $D_a \geq 0$ and all $D_b \geq 0$ if and only if $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}(T)$.

[(i) and (iii) are equivalent]

Repeat the proof of "(i) is equivalent to (ii)" with each statement that $D_a \geq 0$ replaced with $D_a = \text{diag} (0, 0, \dots, 0)$.

[(ii) and (iv) are equivalent and (iii) and (v) are equivalent]

Suppose that (ii) and (iv) are not equivalent. Then $(D_a + T)^{-1}(D_b + G) \in P_0$ for all $D_a > 0$ and all $D_b > 0$, and for some $D_a^* \geq 0$ and some $D_b^* \geq 0$, with $D_a^* \succ 0$ or $D_b^* \succ 0$ or $D_a^* \succ 0$ and $D_b^* \succ 0$, $(D_a^* + T)^{-1}(D_b^* + G) \notin P_0$. Thus some principal minor of $(D_a^* + T)^{-1}(D_b^* + G)$, and hence of $(D_a^* + T)^{-1}(D_b^* + G) \det (D_a^* + T)$, is negative. Let

$m(D_a^*, D_b^*)$ be some negative principal minor of $(D_a^* + T)^{-1}(D_b^* + G) \det (D_a^* + T)$, and let $m(D_a^* + \epsilon I, D_b^* + \epsilon I)$ be the corresponding principal minor of $(D_a^* + \epsilon I + T)^{-1}(D_b^* + \epsilon I + G) \det (D_a^* + \epsilon I + T)$ for all real $\epsilon \geq 0$. Thus $m(D_a^* + \epsilon I, D_b^* + \epsilon I)$ is a polynomial $p(\epsilon)$ in ϵ for $\epsilon \geq 0$, and $p(\epsilon) \geq 0$ for all $\epsilon > 0$. Therefore $p(0) \geq 0$, which contradicts $m(D_a^*, D_b^*) < 0$.

A proof that (iii) and (v) are equivalent can be obtained by modifying the previous paragraph in an obvious manner.

[(vi) is equivalent to (v)]

By Lemma 1, $T^{-1}(G + D) \in P_0$ for all diagonal $D > 0$ if and only if $\det [T^{-1}(G + D) + D^*] \neq 0$ for all diagonal $D^* > 0$ and $D > 0$, and hence if and only if $\det (G + TD^* + D) \neq 0$ for all $D^* > 0$ and all $D > 0$. Therefore, by Lemma 1, $T^{-1}(G + D) \in P_0$ for all $D > 0$ if and only if $(G + TD^*) \in P_0$ for all $D^* > 0$, that is, if and only if $\det [G_w + (TD^*)_w] \geq 0$ for all $w \in Q_{(2p+q)}$ and all $D^* > 0$. Since $(TD^*)_w = T_w D_w^*$, we see that $T^{-1}(G + D) \in P_0$ for all $D > 0$ if and only if

$$\det [(T_w)^{-1}G_w + D_w^*] \geq 0 \quad \text{for all } w \in Q_{(2p+q)} \quad \text{and all } D^* > 0. \quad (20)$$

But, by Lemma 2 (which follows) condition (20) is equivalent to the condition that $\det [(T_w)^{-1}G_w + D_w^*] > 0$, and hence that $\det [(T_w)^{-1}G_w + D_w^*] \neq 0$, for all $w \in Q_{(2p+q)}$ and all $D^* > 0$. Thus by Lemma 1, $T^{-1}(G + D) \in P_0$ for all $D > 0$ if and only if $(T_w)^{-1}G_w \in P_0$ for all $w \in Q_{(2p+q)}$.

Lemma 2: If A is a real square matrix of order n such that $\det (D + A) = 0$ for some diagonal $D > 0$, then $\det (D + A) < 0$ for some diagonal $D > 0$.

Proof: Using the notation of the proof of Theorem 2,

$$\det (D + A) = \sum' d_1^{e_1} d_2^{e_2} \cdots d_n^{e_n} m_{e_1, e_2, \dots, e_n} \quad (21)$$

for all $D > 0$. Since $m_{1, 1, \dots, 1} = 1$, if $\det (D + A) = 0$ for some $D > 0$, then for at least one sequence e'_1, e'_2, \dots, e'_n we have $m_{e'_1, e'_2, \dots, e'_n} < 0$. If $m_{0, 0, \dots, 0} = \det A < 0$, then there exists a positive constant σ_1 such that $\det (D + A) < 0$ whenever $0 < d_j < \sigma_1$ for all j . If $\det A \geq 0$, then, with $d_j = d$ for all j such that $e'_j = 1$ and $d_j = d^{-1}$ for all j such that $e'_j = 0$, there exists a positive constant σ_2 such that $\det (D + A) < 0$ for all $d > \sigma_2$ [see (21)]. \square

[(vi) and (vii) are equivalent]

We shall prove that

$$[(T^w)^{-1}G]_w = (T_w)^{-1}G_w \quad \text{for all } w \in Q_{(2p+q)}. \quad (22)$$

Obviously the equality of (22) is satisfied if w is the null set.

It is convenient to introduce the following notation. Let u denote the 1×1 matrix containing the entry 1. Let φ denote what might be called the empty matrix, a matrix with no rows or columns; by this we mean that φ is to be interpreted in the following manner: $\varphi \oplus \varphi = \varphi$, $I_s = \varphi$ when $s = 0$, $\varphi^{-1} = \varphi$, and if M_1 and M_2 are any two (ordinary) matrices, then $\varphi \oplus M_1 = M_1$, $M_1 \oplus \varphi = M_1$, and $M_1 \oplus \varphi \oplus M_2 = M_1 \oplus M_2$.

Let $w \in Q_{(2p+q)}$ and let w not be the null set. The matrix T can be written as the direct sum $T_1 \oplus T_2 \oplus \cdots \oplus T_p \oplus I_q$. In terms of u and φ , $T_w = t_1 \oplus t_2 \oplus \cdots \oplus t_p \oplus I_s$, in which $s = q - \bar{q}$ where \bar{q} is the number of elements contained in the intersection of the sets w and $\{2p+1, 2p+2, \dots, 2p+q\}$, and for $k = 1, 2, \dots, p$: $t_k = T_k$ if both $(2k-1)$ and $2k$ are not elements of w , $t_k = \varphi$ if both $(2k-1)$ and $2k$ are elements of w , and $t_k = u$ if either $(2k-1) \in w$ and $2k \notin w$ or $(2k-1) \notin w$ and $2k \in w$. Thus $(T_w)^{-1} = t_1^{-1} \oplus t_2^{-1} \oplus \cdots \oplus t_p^{-1} \oplus I_s$. But $(T^w)^{-1} = \hat{T}_1^{-1} \oplus \hat{T}_2^{-1} \oplus \cdots \oplus \hat{T}_p^{-1} \oplus I_q$, in which for $k = 1, 2, \dots, p$: $\hat{T}_k = T_k$ if both $(2k-1)$ and $2k$ are not elements of w ,

$$\hat{T}_k^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

if both $(2k-1)$ and $2k$ are contained in w ,

$$\hat{T}_k^{-1} = \begin{bmatrix} 1 & \alpha_r^{(k)} \\ 0 & 1 \end{bmatrix}$$

if $(2k-1) \in w$ and $2k \notin w$, and

$$\hat{T}_k^{-1} = \begin{bmatrix} 1 & 0 \\ \alpha_f^{(k)} & 1 \end{bmatrix}$$

if $(2k-1) \notin w$ and $2k \in w$. Thus we see that $[(T^w)^{-1}]_w = (T_w)^{-1}$. Let ${}_{(w)}(T^w)^{-1}$ denote the $(2p+q-r) \times (2p+q)$ matrix obtained from $(T^w)^{-1}$ by deleting rows i_1, i_2, \dots, i_r . But all elements of columns i_1, i_2, \dots, i_r of ${}_{(w)}(T^w)^{-1}$ are zeros, and hence, with $G_{(w)}$ the matrix obtained from G by deleting columns i_1, i_2, \dots, i_r ,

$$\begin{aligned} [(T^w)^{-1}G]_w &= {}_{(w)}(T^w)^{-1}G_{(w)} \\ &= [(T^w)^{-1}]_w G_w = (T_w)^{-1}G_w. \end{aligned}$$

[(viii) and (i) are equivalent]

If $M^{-1}G \in P_0$ for all $M \in \mathfrak{J}_0(T)$, then $[(T^w)^{-1}G]_w \in P_0$ for all $w \in Q_{(2p+q)}$. Thus, statement (viii) implies statement (vii). Since we have proved that (vii) is equivalent to (i), it suffices to prove that (i) implies (viii).

Suppose that $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}(T)$. Let \hat{M} be an arbitrary element of $\mathfrak{S}_0(T)$. Then $[\hat{M} + \delta(T - \hat{M})] \in \mathfrak{S}(T)$ for all $\delta \in (0, 1]$, and therefore $[\hat{M} + \delta(T - \hat{M})]^{-1}G \in P_0$ for all $\delta \in (0, 1]$. At this point a continuity-type argument similar to that used in the proof of [(ii) and (iv) are equivalent] shows that $\hat{M}^{-1}G \in P_0$. \square

3.13 Corollary 1 (Corollary to Theorem 5):

If $T \in \mathfrak{S}$ and $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}(T)$, then $\det(\sigma I + J_u) \neq 0$ for all $\sigma \geq 0$ and all $u \in E^{(2p+q)}$ provided that for all j $f_j(\cdot)$ is continuously differentiable on $(-\infty, \infty)$ and $f'_j(\alpha) > 0$ for all $\alpha \in (-\infty, \infty)$.

3.14 Proof of Corollary 1.

If $T \in \mathfrak{S}$ and $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}(T)$, then, by the equivalence of (i) and (ii) of Theorem 5, $(\sigma T + T)^{-1}(\sigma c + G) \in P_0$ for all $\sigma \geq 0$. The first portion of the proof of Theorem 1 shows that if $(\sigma T + T)^{-1}(\sigma c + G) \in P_0$ for all $\sigma \geq 0$ and if for all j $f_j(\cdot)$ is continuously differentiable on $(-\infty, \infty)$ and $f'_j(\alpha) > 0$ for all $\alpha \in (-\infty, \infty)$, then $\det(\sigma I + J_u) \neq 0$ for all $\sigma \geq 0$ and all $u \in E^{(2p+q)}$.

3.15 Definition 10: For $p > 0$ let $Q'_{(2p+q)}$ denote the subset of $Q_{(2p+q)}$ containing all sets w belonging to $Q_{(2p+q)}$ such that w is not the null set and $2k \in w$ if and only if $(2k - 1) \in w$ for $k = 1, 2, \dots, p$. For $p = 0$, let $Q'_{(2p+q)}$ denote the family of all sets contained in $Q_{(2p+q)}$ with the exception of the null set.

3.16 Theorem 6: If $T \in \mathfrak{S}$ and $T^{-1}G \in P_0$, then $(T_w)^{-1}G_w \in P_0$ for all $w \in Q'_{(2p+q)}$.

3.17 Proof of Theorem 6

Let $T \in \mathfrak{S}$, and let $T^{-1}G \in P_0$. By Lemma 1, $\det(TD + G) \neq 0$ (and hence $\det(TD + G) > 0$) for all diagonal $D > 0$. Let $w = \{i_1, i_2, \dots, i_r\} \in Q'_{(2p+q)}$, and let $d_{i_k} = d$ for $k = 1, 2, \dots, r$.

It may be the case that $(TD + G)$ is a block matrix of the form

$$\begin{bmatrix} (TD + G)_w & H_{12} \\ H_{21} & (d\hat{T} + H_{22}) \end{bmatrix} \tag{23}$$

in which \hat{T} is a direct sum of all 2×2 and 1×1 block matrices on the diagonal of T which do not appear in T_w , and H_{12} , H_{21} , and H_{22} are independent of D . Clearly $\det \hat{T} > 0$. If $(TD + G)$ is not of the form (23), then by a sequence of interchanges of rows and corresponding columns of $(TD + G)$ we obtain a matrix of that form.

Thus, for some \hat{T} of the form indicated above and for the corresponding constant matrices H_{12} , H_{21} , and H_{22} whose elements are elements of G ,

$$\det(TD + G) = \det \begin{bmatrix} (TD+G)_w & H_{12} \\ H_{12} & (d\hat{T} + H_{22}) \end{bmatrix}$$

for all $d_i > 0$ for $j \neq w$. For all sufficiently large $d > 0$, $\det(d\hat{T} + H_{22}) > 0$, and then

$$0 < \det(TD + G) = \det(d\hat{T} + H_{22}) \cdot \det[(TD + G)_w - H_{12}(d\hat{T} + H_{22})^{-1}H_{21}]$$

for all $d_i > 0$ for $j \neq w$. Since $H_{12}(d\hat{T} + H_{22})^{-1}H_{21}$ approaches the zero matrix of order $(2p + q - r)$ as $d \rightarrow \infty$, we must have $\det(TD + G)_w \geq 0$ for all $d_i > 0$ for $j \neq w$. Therefore, since $(TD)_w = T_w D_w$, we must have $\det(T_w D_w + G_w) \geq 0$ for all $D_w > 0$. But this means (see Lemma 2) that $\det(T_w D_w + G_w) \neq 0$ for all $D_w > 0$. Thus, by Lemma 1, $(T_w)^{-1}G_w \in P_0$. \square

3.18 Theorem 7: *If $T \in \mathfrak{S}$ with $p = 1$ or $p = 2$, and if $T^{-1}G \in P_0$ with G the short-circuit conductance matrix of a transformerless positive-element resistance network, then $(T_w)^{-1}G_w \in P_0$ for all $w \in Q_{(2p+q)}$.*

3.19 Proof of Theorem 7

Suppose that $T^{-1}G \in P_0$ with $p = 2$. Theorem 6 asserts that $(T_w)^{-1}G_w \in P_0$ for all $w \in Q'_{(2p+q)}$. But, aside from the null set, the sets $w = \{i_1, i_2, \dots, i_r\}$ that are contained in $Q_{(2p+q)}$ but not in $Q'_{(2p+q)}$ possess the property that $T_w = T_1 \oplus I_{(2+q-r)}$, or $T_w = u \oplus T_2 \oplus I_{(1+q-r)}$ where u is the 1×1 matrix containing the element 1, or $T_w = I_{(4+q-r)}$.

If $T_w = I_{(4+q-r)}$, then obviously $(T_w)^{-1}G_w \in P_0$. If $T_w = T_1 \oplus I_{(2+q-r)}$, then for any $D_w = \text{diag}[D_2 \oplus D_{(2+q-r)}]$ with $D_2 > 0$ and $D_{(2+q-r)} > 0$ diagonal matrices of order 2 and $(2 + q - r)$ respectively,

$$\det(T_w D_w + G_w) = \begin{bmatrix} T_1 D_2 + G_{11} & G_{12} \\ G_{21} & D_{(2+q-r)} + G_{22} \end{bmatrix} \quad (24)$$

in which G_{11} , G_{12} , G_{21} , and G_{22} are the appropriate block matrices of G_w . Since $\det[D_{(2+q-r)} + G_{22}] > 0$, we have

$$\det(T_w D_w + G_w) = \det[D_{(2+q-r)} + G_{22}] \cdot \det\{T_1 D_2 + G_{11} - G_{12}[D_{(2+q-r)} + G_{22}]^{-1}G_{21}\}.$$

But $G_{11} - G_{12}[D_{(2+q-r)} + G_{22}]^{-1}G_{21}$ is the short-circuit conductance matrix of a transformerless common-ground 2-port network; it is of the form

$$\begin{bmatrix} g_{11} & -g_{12} \\ -g_{12} & g_{22} \end{bmatrix}$$

with $g_{11} \geq 0, g_{22} \geq 0, g_{12} \geq 0, g_{11} \geq g_{12},$ and $g_{22} \geq g_{12}.$ Therefore¹

$$\det \{T_1 D_2 + G_{11} - G_{12}[D_{(2+q-r)} + G_{22}]^{-1}G_{21}\} > 0$$

for all $D_2 > 0$ and all $D_{(2+q-r)} > 0, \det (T_w D_w + G_w) \neq 0$ for all $D_w > 0,$ and hence, by Lemma 1, $(T_w)^{-1}G_w \in P_0.$ Finally, the case in which $T_w = u \oplus T_2 \oplus I_{(1+q-r)}$ can be treated in a manner similar to that used to show that $(T_w)^{-1}G_w \in P_0$ when $T_w = T_1 \oplus I_{(2+q-r)},$ since, with w such that $T_w = u \oplus T_2 \oplus I_{(1+q-r)},$ and with D an arbitrary diagonal matrix of order $(4 + q - r),$ a sequence of interchanges of rows and corresponding columns of $(T_w D + G_w)$ can be performed to obtain a matrix of the type that appears on the right side of equation (24). Therefore $(T_w)^{-1}G_w \in P_0$ for all $w \in Q_{(2p+q)}.$

When $p = 1,$ aside from the null set, the sets $w = \{i_1, i_2, \dots, i_r\}$ that are contained in $Q_{(2p+q)}$ but not in $Q'_{(2p+q)}$ possess the property that $T_w = I_{(2+q-r)}$ and obviously when $T_w = I_{(2+q-r)}, (T_w)^{-1}G_w \in P_0. \square$

3.20 Theorem 8: Let $T \in \mathfrak{E}$ and let G possess the property that for some diagonal matrix $D > 0,$ both DT and DG are strongly-column-sum dominant. For each $j = 1, 2, \dots, (2p + q)$ let $f_j(\cdot)$ be a continuous monotone-nondecreasing mapping of E^1 into itself such that $f_j(0) = 0,$ let $h \in (0, \infty),$ and, with $F(\cdot)$ and $C(\cdot)$ defined relative to the $f_j(\cdot)$ as in Section 2.1, suppose that the sequences $\{y_n\}$ and $\{w_n\}$ in $E^{(2p+q)}$ satisfy

$$y_{n+1} + h\{TF[C^{-1}(y_{n+1})] + GC^{-1}(y_{n+1})\} = y_n + w_n$$

for all $n \geq 0.$ Then there exists a positive constant δ depending only on the $c_i,$ the $\tau_i, T, G,$ and D such that

$$(i) \quad \|Dy_n\|_1 \leq (1 + \delta h)^{-n} \|Dy_0\|_1 + \sum_{k=1}^n (1 + \delta h)^{-k} \|Dw_{(n-k)}\|_1$$

for all $n \geq 1,$ and

$$(ii) \quad \|D(y_n - \tilde{y}_n)\|_1 \leq (1 + \delta h)^{-n} \|D(y_0 - \tilde{y}_0)\|_1 + \epsilon \sum_{k=0}^n (1 + \delta h)^{-k}$$

for all $n \geq 1,$ in which $\{\tilde{y}_n\}$ is any sequence in $E^{(2p+q)}$ with the property that $\|D(\tilde{y}_n - y_n^*)\|_1 \leq \epsilon$ for all $n \geq 1$ with ϵ a positive constant and the

sequence $\{y_n^*\}$ such that

$$y_{n+1}^* + h\{TF[C^{-1}(y_{n+1}^*)] + GC^{-1}(y_{n+1}^*)\} = \tilde{y}_n + w_n$$

for all $n \geq 0$.

3.21 Proof of Theorem 8

We shall first prove part (ii). With D such that DT and DG are strongly-column-sum dominant, we have for all $n \geq 0$

$$Dy_{n+1} + h\{DTF[C^{-1}(y_{n+1})] + DGC^{-1}(y_{n+1})\} = Dy_n + Dw_n$$

and

$$Dy_{n+1}^* + h\{DTF[C^{-1}(y_{n+1}^*)] + DGC^{-1}(y_{n+1}^*)\} = Dy_n^* + D(\tilde{y}_n - y_n^*) + Dw_n$$

in which we shall take y_0^* to be \tilde{y}_0 . As in the proof of Theorem 2 of Ref. 3, we write

$$F[C^{-1}(y_{n+1})] - F[C^{-1}(y_{n+1}^*)] = \text{diag} \left(\frac{r(n)_j}{c_j + \tau_j r(n)_j} \right) (y_{n+1} - y_{n+1}^*) \quad (25)$$

and

$$C^{-1}(y_{n+1}) - C^{-1}(y_{n+1}^*) = \text{diag} \left(\frac{1}{c_j + \tau_j r(n)_j} \right) (y_{n+1} - y_{n+1}^*) \quad (26)$$

in which $r(n)_j$ depends on the j th components of y_{n+1} and y_{n+1}^* , and $r(n)_j \geq 0$ for all $n \geq 0$ and all j .

Thus, with $Q = DTD^{-1}$ and $R = DGD^{-1}$,

$$\begin{aligned} \left\{ I + hQ \text{diag} \left(\frac{r(n)_j}{c_j + \tau_j r(n)_j} \right) + hR \text{diag} \left(\frac{1}{c_j + \tau_j r(n)_j} \right) \right\} D(y_{n+1} - y_{n+1}^*) \\ = D(y_n - y_n^*) - D(\tilde{y}_n - y_n^*) \end{aligned}$$

for all $n \geq 0$. At this point we shall use the proposition that if M is any real matrix of order $(2p + q)$ with the property that there exists a positive constant η such that $m_{jj} - \sum_{i \neq j} |m_{ji}| \geq \eta$ for all j , then $\|Mx\|_1 \geq \eta \|x\|_1$ for all $x \in E^{(2p+q)}$. Now let

$$M = \left\{ I + hQ \text{diag} \left(\frac{r(n)_j}{c_j + \tau_j r(n)_j} \right) + hR \text{diag} \left(\frac{1}{c_j + \tau_j r(n)_j} \right) \right\}$$

for arbitrary $n \geq 0$. Then for arbitrary j

$$\begin{aligned} m_{jj} - \sum_{i \neq j} |m_{ji}| &= 1 + hq_{ji} \left(\frac{r(n)_j}{c_j + \tau_j r(n)_j} \right) + hr_{ji} \left(\frac{1}{c_j + \tau_j r(n)_j} \right) \\ &\quad - h \sum_{i \neq j} \left| q_{ii} \frac{r(n)_j}{c_j + \tau_j r(n)_j} + r_{ii} \frac{1}{c_j + \tau_j r(n)_j} \right| \end{aligned}$$

$$\begin{aligned} &\geq 1 + h \left(q_{ii} - \sum_{i \neq j} |q_{ij}| \right) \frac{r^{(n)}_i}{c_i + \tau_i r^{(n)}_i} \\ &\quad + h \left(r_{ii} - \sum_{i \neq j} |r_{ij}| \right) \frac{1}{c_i + \tau_i r^{(n)}_i} \\ &\geq 1 + \delta h, \end{aligned}$$

in which

$$\delta = \min \left\{ \min_i c_i^{-1} \left(r_{ii} - \sum_{i \neq j} |r_{ij}| \right), \min_i \tau_i^{-1} \left(q_{ii} - \sum_{i \neq j} |q_{ij}| \right) \right\}.$$

Therefore

$$\begin{aligned} &\| D(y_{n+1} - y_{n+1}^*) \|_1 \\ &\leq (1 + \delta h)^{-1} \| D(y_n - y_n^*) - D(\tilde{y}_n - y_n^*) \|_1 \\ &\leq (1 + \delta h)^{-1} \| D(y_n - y_n^*) \|_1 + (1 + \delta h)^{-1} \| D(\tilde{y}_n - y_n^*) \|_1 \\ &\leq (1 + \delta h)^{-1} \| D(y_n - y_n^*) \|_1 + \epsilon (1 + \delta h)^{-1} \end{aligned}$$

for all $n \geq 0$, and hence

$$\| D(y_n - y_n^*) \|_1 \leq (1 + \delta h)^{-n} \| D(y_0 - y_0^*) \|_1 + \epsilon \sum_{k=1}^n (1 + \delta h)^{-k}$$

for all $n \geq 1$. Finally, since $\| D(y_n - \tilde{y}_n) \|_1 \leq \| D(y_n - y_n^*) \|_1 + \| D(y_n^* - \tilde{y}_n) \|_1 \leq \| D(y_n - y_n^*) \|_1 + \epsilon$, and since $y_0^* = \tilde{y}_0$,

$$\| D(y_n - \tilde{y}_n) \|_1 \leq (1 + \delta h)^{-n} \| D(y_0 - \tilde{y}_0) \|_1 + \epsilon \sum_{k=0}^n (1 + \delta h)^{-k}$$

for all $n \geq 1$, which completes the proof of part (ii) of the theorem.

The proof of part (i) is similar to that of part (ii). Using

$$Dy_{n+1} + h \{ DTF[C^{-1}(y_{n+1})] + DGC^{-1}(y_{n+1}) \} = Dy_n + Dw_n$$

for all $n \geq 0$, and equations (25) and (26) with $y_{n+1}^* = \theta$ for all n , we find that

$$\| Dy_{n+1} \|_1 \leq (1 + h\delta)^{-1} \| Dy_n \|_1 + (1 + h\delta)^{-1} \| Dw_n \|_1$$

for all $n \geq 0$. Therefore

$$\| Dy_n \|_1 \leq (1 + h\delta)^{-n} \| Dy_0 \|_1 + \sum_{k=1}^n (1 + h\delta)^{-k} \| Dw_{(n-k)} \|_1$$

for all $n \geq 1$. \square

3.22 *Theorem 9:* Let $T \in \mathfrak{S}$ and let G possess the property that for some diagonal matrix $D > 0$, both DT and DG are strongly-column-sum dominant. Let $B(\cdot)$ denote a real continuously-differentiable $(2p + q)$ -vector-valued function of t for $t \in [0, \infty)$ such that both $B(\cdot)$ and $(d/dt)B(\cdot)$ are bounded on $[0, \infty)$. With $F(\cdot)$ such that each $f_i(0) = 0$, and with $C(\cdot)$ defined relative to $F(\cdot)$ as in Section 2.1, let $u(\cdot)$ satisfy

$$\frac{du}{dt} + TF[C^{-1}(u)] + GC^{-1}(u) = B(t), \quad t \geq 0$$

and, with h an arbitrary positive constant, let u_n denote $u(nh)$ for all $n \geq 0$. Let $\{y_n\}$ be a sequence in $E^{(2p+q)}$ such that

$$y_{n+1} + h\{TF[C^{-1}(y_{n+1})] + GC^{-1}(y_{n+1})\} = y_n + hB[(n+1)h], \quad n \geq 0.$$

Then there exist positive constants δ and ρ , both independent of h , such that

$$\|D(u_n - y_n)\|_1 \leq (1 + \delta h)^{-n} \|D(u_0 - y_0)\|_1 + \rho h$$

for all $n \geq 1$.

3.23 Proof of Theorem 9

The sequence $\{u_n\}$ satisfies

$$\begin{aligned} u_{n+1} + h\{TF[C^{-1}(u_{n+1})] + GC^{-1}(u_{n+1})\} \\ = u_n + B[(n+1)h] + \xi_n, \quad n \geq 0 \end{aligned}$$

in which ξ_n is often referred to as "the local-truncation error at step n ." We shall first bound ξ_n .

Since $B(\cdot)$ is bounded on $[0, \infty)$, and since for some $D > 0$, both DT and DG are strongly-column-sum dominant, a direct modification of the proof of Theorem 1 of Ref. 5 shows that $u(\cdot)$ is bounded on $[0, \infty)$; and hence since

$$\frac{d^2u}{dt^2} = J_u\{TF[C^{-1}(u)] + GC^{-1}(u)\} - J_u B(t) + \frac{d}{dt} B(t), \quad t \geq 0 \quad (27)$$

with $(d/dt)B(\cdot)$ and the elements of the Jacobian matrix J_u bounded, it is clear that (d^2u/dt^2) is bounded on $[0, \infty)$. By the usual Taylor-series-type argument we can show that for arbitrary $n \geq 0$, $\xi_n = \frac{1}{2}h^2 U_n$ in which for each j the j th component of U_n is the j th component of (d^2u/dt^2) evaluated at some point contained in the interval $[nh, (n+1)h]$. Thus there exists a positive constant ρ_1 such that

$$\|D\xi_n\|_1 \leq \frac{1}{2}h^2 \rho_1 \quad \text{for all } n \geq 0. \quad (28)$$

Therefore, using (28) and the equations

$$u_{n+1} + h\{TF[C^{-1}(u_{n+1})] + GC^{-1}(u_{n+1})\} = u_n + B[(n + 1)h] + \xi_n, \quad n \geq 0$$

$$y_{n+1} + h\{TF[C^{-1}(y_{n+1})] + GC^{-1}(y_{n+1})\} = y_n + B[(n + 1)h], \quad n \geq 0$$

by an argument similar to that used in the proof of part (ii) of Theorem 8, and with δ as defined there, we find that

$$\|D(u_{n+1} - y_{n+1})\|_1 \leq (1 + \delta h)^{-1} \|D(u_n - y_n)\|_1 + (1 + \delta h)^{-1} \frac{1}{2} h^2 \rho_1$$

for all $n \geq 0$, and hence that

$$\begin{aligned} \|D(u_n - y_n)\|_1 &\leq (1 + \delta h)^{-n} \|D(u_0 - y_0)\|_1 + \frac{1}{2} h^2 \rho_1 \sum_{k=1}^n (1 + \delta h)^{-k} \\ &\leq (1 + \delta h)^{-n} \|D(u_0 - y_0)\|_1 + \frac{1}{2} h^2 \rho_1 \sum_{k=1}^{\infty} (1 + \delta h)^{-k} \\ &\leq (1 + \delta h)^{-n} \|D(u_0 - y_0)\|_1 + \frac{1}{2} h \delta^{-1} \rho_1 \end{aligned}$$

for all $n \geq 1$. \square

3.24 *Definition 11:* Let $R = R_1 \oplus R_2 \oplus \dots \oplus R_p \oplus R_0$ in which $R_0 = \text{diag}(r_1, r_2, \dots, r_q)$ with $r_j \geq 0$ for $j = 1, 2, \dots, q$ and

$$R_k = \begin{pmatrix} r_a^{(k)} + r_b^{(k)} & r_b^{(k)} \\ r_b^{(k)} & r_c^{(k)} + r_b^{(k)} \end{pmatrix}$$

with $r_a^{(k)} \geq 0, r_b^{(k)} \geq 0$, and $r_c^{(k)} \geq 0$ for all $k = 1, 2, \dots, p$. As suggested, if $q = 0$, then $R = R_1 \oplus R_2 \oplus \dots \oplus R_p$, while if $p = 0$, then $R = R_0$.

3.25 *Theorem 10:* Let $T \in \mathfrak{S}$. If $p > 0$ and if R satisfies

$$\begin{aligned} \alpha_r^{(k)}(1 - \alpha_r^{(k)})^{-1} r_a^{(k)} &= r_b^{(k)} \\ \alpha_f^{(k)}(1 - \alpha_f^{(k)})^{-1} r_c^{(k)} &= r_b^{(k)} \end{aligned}$$

for $k = 1, 2, \dots, p$, then $T^{-1}G(I + RG)^{-1} \in P_0$ whenever $T^{-1}G \in P_0$.

3.26 *Proof of Theorem 10*

By Lemma 1, $T^{-1}G(I + RG)^{-1} \in P_0$ if and only if

$$\det [T^{-1}G(I + RG)^{-1} + D^*] \neq 0 \tag{29}$$

for all diagonal $D^* > 0$. But (29) is satisfied if and only if

$$\det (T^{-1}G + D^*RG + D^*) \neq 0.$$

Here, since

$$\alpha_r^{(k)}(1 - \alpha_r^{(k)})^{-1}r_e^{(k)} = r_b^{(k)}$$

$$\alpha_f^{(k)}(1 - \alpha_f^{(k)})^{-1}r_c^{(k)} = r_b^{(k)}$$

for $k = 1, 2, \dots, p$ we have $R = DT^{-1}$ for some diagonal matrix $D \geq 0$. Thus (29) is satisfied if and only if

$$\det [(I + DD^*)T^{-1}G + D^*] \neq 0.$$

When $T^{-1}G \in P_0$ we have

$$\det (T^{-1}G + \tilde{D}) \neq 0$$

for all diagonal $\tilde{D} > 0$. Thus (29) is satisfied for all $D^* > 0$ whenever $T^{-1}G \in P_0$. \square

3.27 Theorem 11: *If $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}$, then for any $T \in \mathfrak{S}$, $T^{-1}G(I + RG)^{-1} \in P_0$ for all R .*

3.28 Proof of Theorem 11

Let $T \in \mathfrak{S}$. As in the proof of Theorem 10, $T^{-1}G(I + RG)^{-1} \in P_0$ if and only if

$$\det [(T^{-1} + D^*R)G + D^*] \neq 0$$

for all diagonal $D^* > 0$. It is a simple matter to verify that for each $D^* > 0$ and each R there exists an $\tilde{M} \in \mathfrak{S}$ and a diagonal matrix $D > 0$ such that $(T^{-1} + D^*R)G = D\tilde{M}^{-1}$. Since $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}$, we have (by Lemma 1)

$$\det (D\tilde{M}^{-1}G + D^*) \neq 0$$

for all $D^* > 0$. \square

3.29 Theorem 12: *Let $T \in \mathfrak{S}$ with $p > 0$ and $q \geq 0$. Then $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}(T)$ if and only if $T^{-1}G(I + RG)^{-1} \in P_0$ for all R such that*

$$\alpha_r^{(k)}(1 - \alpha_r^{(k)})^{-1}r_e^{(k)} \geq r_b^{(k)}$$

$$\alpha_f^{(k)}(1 - \alpha_f^{(k)})^{-1}r_c^{(k)} \geq r_b^{(k)}$$

for $k = 1, 2, \dots, p$ and $r_i \geq 0$ for all j such that $1 \leq j \leq q$.

3.30 Proof of Theorem 12

As in the proof of Theorem 10, $T^{-1}G(I + RG)^{-1} \in P_0$ if and only if

$$\det (T^{-1}G + D^*RG + D^*) \neq 0 \quad (30)$$

for all diagonal $D^* > 0$. The inequalities $r_j \geq 0$ for all j such that $1 \leq j \leq q$ and

$$\begin{aligned}\alpha_r^{(k)}(1 - \alpha_r^{(k)})^{-1} r_e^{(k)} &\geq r_b^{(k)} \\ \alpha_f^{(k)}(1 - \alpha_f^{(k)})^{-1} r_c^{(k)} &\geq r_b^{(k)}\end{aligned}$$

for $k = 1, 2, \dots, p$ are equivalent to the condition that $R = D_1 T^{-1} + D_2$ for some diagonal matrix $D_2 \geq 0$ and some diagonal matrix $D_1 \in S$, in which S is the set of all diagonal matrices $D \geq 0$ such that DT^{-1} is symmetric. Hence $T^{-1}G(I + RG)^{-1} \in P_0$ for all such R if and only if

$$\det \{[(I + D_1 D^*)T^{-1} + D^* D_2]G + D^*\} \neq 0 \quad (31)$$

for all diagonal $D^* > 0$, $D_2 \geq 0$, and $D_1 \in S$.

Let $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{(2p+q)})$ be such that

$$D_2 = D^{*-1} \Lambda^{-1} \Lambda (I + D_1 D^*)$$

in which

$$\Delta = \text{diag}(\delta_1, \delta_1, \delta_2, \delta_2, \dots, \delta_p, \delta_p) \oplus I_q$$

if $q > 0$, $\Delta = \text{diag}(\delta_1, \delta_1, \delta_2, \delta_2, \dots, \delta_p, \delta_p)$ if $q = 0$, and

$$\delta_k = 1 - \alpha_f^{(k)} \alpha_r^{(k)} \quad \text{for } k = 1, 2, \dots, p.$$

The left side of (31) is

$$\det [(I + D_1 D^*)(T^{-1} + \Delta^{-1} \Lambda)G + D^*]$$

which can be written as

$$\det [(I + D_1 D^*) \Delta^{-1} (I + \Lambda) \Delta_\Lambda T_\Lambda^{-1} G + D^*] \quad (32)$$

with

$$T_\Lambda^{-1} = \Delta_\Lambda^{-1} \Delta (I + \Lambda)^{-1} (T^{-1} + \Delta^{-1} \Lambda)$$

and

$$\Delta_\Lambda = \text{diag}(\delta'_1, \delta'_1, \delta'_2, \delta'_2, \dots, \delta'_p, \delta'_p) \oplus I_q$$

if $q > 0$ and $\Delta_\Lambda = \text{diag}(\delta'_1, \delta'_1, \delta'_2, \delta'_2, \dots, \delta'_p, \delta'_p)$ if $q = 0$, in which for $k = 1, 2, \dots, p$

$$\delta'_k = 1 - \alpha_f^{(k)} \alpha_r^{(k)} (1 + \lambda_{(2k-1)})^{-1} (1 + \lambda_{2k})^{-1}.$$

But (32) vanishes if and only if $\det (T_\Lambda^{-1} G + \tilde{D})$ vanishes, in which $\tilde{D} = \Delta_\Lambda^{-1} (I + \Lambda)^{-1} \Delta (I + D_1 D^*)^{-1} D^*$. We observe that \tilde{D} is a positive diagonal matrix and that given any diagonal $\tilde{D}' > 0$ and given any

$\Lambda \geq 0$ we can choose $D^* > 0$ and $D_1 \in S$ so that $\tilde{D} = \tilde{D}'$. Thus $T^{-1}G(I + RG)^{-1} \in P_0$ for all $R = (D_1 T^{-1} + D_2)$ with $D_1 \in S$ and $D_2 \geq 0$ if and only if

$$\det (T_{\Lambda}^{-1}G + \tilde{D}) \neq 0$$

for all $\Lambda \geq 0$ and $\tilde{D} > 0$, that is, if and only if $T_{\Lambda}^{-1}G \in P_0$ for all $\Lambda \geq 0$ (see Lemma 1 of Section 3.1). But

$$T_{\Lambda} = T_1 \oplus T_2 \oplus \cdots \oplus T_p \oplus I_q \quad \text{if } q > 0$$

and

$$T_{\Lambda} = T_1 \oplus T_2 \oplus \cdots \oplus T_p \quad \text{if } q = 0$$

with

$$T_k = \begin{bmatrix} 1 & \frac{-\alpha_r^{(k)}}{1 + \lambda_{2k-1}} \\ \frac{-\alpha_f^{(k)}}{1 + \lambda_{2k}} & 1 \end{bmatrix}$$

for all $k = 1, 2, \dots, p$. Therefore $T^{-1}G(I + RG)^{-1} \in P_0$ for all $R = (D_1 T^{-1} + D_2)$ with $D_2 \geq 0$ and $D_1 \in S$ if and only if $M^{-1}G \in P_0$ for all $M \in \mathfrak{S}(T)$. \square

3.31 *Definition 12*: Let \mathfrak{F}_3 denote the set of all $F(\cdot)$ such that

- (i) $F(\cdot) \in \mathfrak{F}_0^{(2p+q)}$, and
- (ii) for each $j = 1, 2, \dots, (2p + q)$ there exists a real constant β_j such that $f_j(\cdot)$ is a strictly-monotone-increasing mapping of E^1 onto either (β_j, ∞) or $(-\infty, \beta_j)$, and
- (iii) whenever $p > 0$, $f_{(2k-1)}(\cdot)$ and $f_{2k}(\cdot)$ are both bounded on either $[0, \infty)$ or $(-\infty, 0]$ for $k = 1, 2, \dots, p$.

3.32 *Theorem 13*: Let $T \in \mathfrak{S}$, and, referring to the network of Fig. 1 in which it is assumed that R (see Section 2.1) is the zero matrix, let G denote the short-circuit conductance matrix of the linear portion of the network. (The linear portion is assumed to contain only sources and linear resistors of nonnegative resistance.) Then the equation $F(x) + T^{-1}Gx = B$ possesses a unique solution x for each $F(\cdot) \in \mathfrak{F}_3$ and each $B \in E^{(2p+q)}$ if and only if $T^{-1}G \in P_0$ and $\det G \neq 0$. If $T^{-1}G \in P_0$ and $\det G = 0$, then there exists a real $(2p + q)$ -vector η such that (i) $\eta \neq \theta$, and for some $F(\cdot) \in \mathfrak{F}_3$ all of the components of $F(\alpha\eta)$ are bounded on $\alpha \in [0, \infty)$, and (ii) for any $F(\cdot) \in \mathfrak{F}_3$ with the property that all of the components of $F(\alpha\eta)$ are bounded on $\alpha \in [0, \infty)$ the equation $F(x) + T^{-1}Gx = B$ does not possess a solution for some $B \in E^{(2p+q)}$.

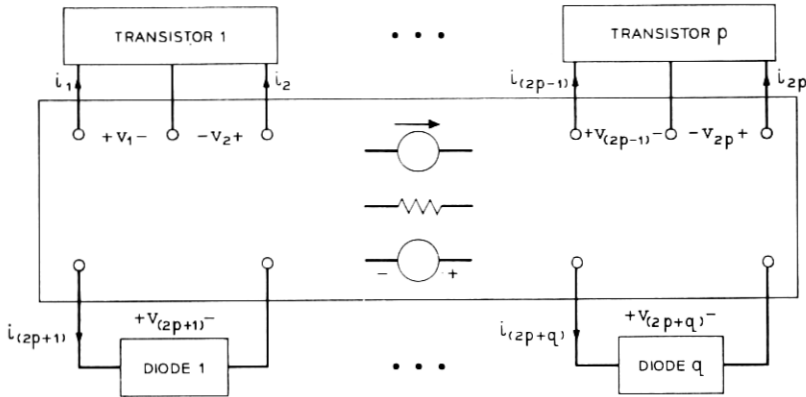


Fig. 1—General network containing transistors, diodes, resistors, and sources.

3.33 Proof of Theorem 13

(if) If $T^{-1}G \in P_0$ with $\det T^{-1}G \neq 0$, and if $F(\cdot) \in \mathcal{F}_3$, then, since each $f_i(\cdot)$ is a strictly-monotone-increasing mapping of E^1 onto (β_i, ∞) or $(-\infty, \beta_i)$ for some real constant β_i , by Theorem 4 of Ref. 2, the equation $F(x) + T^{-1}Gx = B$ possesses a unique solution x for each $B \in E^{(2p+q)}$.

(only if) Assume that $T^{-1}G \notin P_0$. Then since \mathcal{F}_3 is contained in $\mathcal{F}_0^{(2p+q)}$, by Theorem 1 of Ref. 3, for each $F(\cdot) \in \mathcal{F}_3$ there exists a $B \in E^{(2p+q)}$ such that there are at least two solutions x of $F(x) + T^{-1}Gx = B$.

Assume now that $T^{-1}G \in P_0$ and that $\det G = 0$. We shall use the proposition that if $R(\cdot)$ is any continuous mapping of $E^{(2p+q)}$ into itself, then $R(\cdot)$ is a homeomorphism of $E^{(2p+q)}$ onto itself if and only if $R(\cdot)$ is a local homeomorphism on $E^{(2p+q)}$ and $\|R(x)\| \rightarrow \infty$ as $\|x\| \rightarrow \infty$.[†]

Let $R(\cdot)$ be defined by the condition that $R(x) = F(x) + T^{-1}Gx$ for all $x \in E^{(2p+q)}$. For any $F(\cdot) \in \mathcal{F}_3$ the operator $R(\cdot)$ is a local homeomorphism on $E^{(2p+q)}$, since with $F(\cdot)$ such that each $f_i(\cdot)$ is a strictly-monotone-increasing mapping of E^1 onto E^1 the mapping $[F(\cdot) + T^{-1}G]$ is a homeomorphism of $E^{(2p+q)}$ onto itself.¹ In addition, for any $F(\cdot) \in \mathcal{F}_3$ and any $B \in E^{(2p+q)}$, there is at most one $x \in E^{(2p+q)}$ such that $R(x) = B$.¹

Let us suppose that for each $B \in E^{(2p+q)}$ and each $F(\cdot) \in \mathcal{F}_3$ there exists a solution x of $R(x) = B$. Then for all $F(\cdot) \in \mathcal{F}_3$, $R(\cdot)$ is a homeomorphism of $E^{(2p+q)}$ onto itself, and hence for all $F(\cdot) \in \mathcal{F}_3$ $\|R(x)\| \rightarrow \infty$ as $\|x\| \rightarrow \infty$. But, by Lemma 3 (which appears below) $E^{(2p+q)}$ contains a vector η such that $\eta \neq \theta$, $\eta_j \in \{0, +1, -1\}$ for all j , and $G\eta = \theta$; and if

[†] See Ref. 12 and the appendix of Ref. 13.

$p > 0$, η satisfies $\eta_{(2k-1)}\eta_{2k} \geq 0$ for all $k = 1, 2, \dots, p$. Let $\mathcal{F}_3(\eta)$ denote the subset of \mathcal{F}_3 containing all elements $F(\cdot)$ with the property that $f_j(\alpha\eta_j)$ is bounded on $\alpha \in [0, \infty)$ for all $j = 1, 2, \dots, (2p + q)$. Since $\eta_{(2k-1)}\eta_{2k} \geq 0$ for all $k = 1, 2, \dots, p$ when $p > 0$, it is clear that $\mathcal{F}_3(\eta)$ is not empty. However, for any $F(\cdot) \in \mathcal{F}_3(\eta)$ we have $\|R(\alpha\eta)\| = \|F(\alpha\eta)\|$ with $\|F(\alpha\eta)\|$ bounded on $\alpha \in [0, \infty)$, which contradicts the assumption that there exists a solution x of $R(x) = B$ for each $F(\cdot) \in \mathcal{F}_3$ and each $B \in E^{(2p+q)}$.

Lemma 3: Let G be the short-circuit conductance matrix of the linear portion of the network of Fig. 1. If $\det G = 0$, then there exists a vector $\eta \in E^{(2p+q)}$ such that $G\eta = \theta$, $\eta \neq \theta$, and $\eta_j \in \{0, +1, -1\}$ for all $j = 1, 2, \dots, (2p + q)$; and if $p > 0$ η also satisfies $\eta_{(2k-1)}\eta_{2k} \geq 0$ for $k = 1, 2, \dots, p$.

Proof of Lemma 3:

Let N denote the $(2p + q)$ -port resistor network obtained from the network of Fig. 1 by removing all transistors and diodes and by setting the value of each source to zero. The short-circuit conductance matrix G possesses the property that if $v \in E^{(2p+q)}$ denotes the vector of port voltages of N and $i \in E^{(2p+q)}$ denotes the corresponding vector of port currents (with polarities as indicated in Fig. 1), then $i = -Gv$.

Let $\det G = 0$. Then the open-circuit resistance matrix of N does not exist. Therefore there exists a port ℓ of N such that there is no path through resistors of N that connects the two terminals of port ℓ when all other ports are open-circuited. Let a one-volt source be placed at port ℓ so that $v_\ell = 1$. Then when all ports j of N with $j \neq \ell$ are open-circuited, $i_\ell = 0$ and there is zero current in every resistor of N . Let S denote a set of port numbers of N with the following properties. The number ℓ is not contained in S and when all ports j with $j \in S$ are short-circuited and all ports j with $j \notin S \cup \{\ell\}$ are open-circuited then zero current flows through the one-volt source; when any port $j_1 \notin S \cup \{\ell\}$ and all ports j with $j \in S$ are short-circuited and all ports j with $j \notin S \cup \{\ell, j_1\}$ are open-circuited then nonzero current flows through the one-volt source. It is clear that such a set S exists (with the understanding that S might be the null set). In general S contains r port numbers where $0 \leq r \leq (2p + q - 1)$.

If $r = (2p + q - 1)$, then with $v_\ell = 1$ and with all remaining components of v equal to zero, we have $Gv = \theta$. Obviously in this case we can take the vector η of the statement of Lemma 3 to be v .

If $r \neq (2p + q - 1)$, then, with $v_\ell = 1$, with $v_j = 0$ for all $j \in S$,

and with all ports $j \notin S \cup \{\ell\}$ open-circuited, there exists for each $j \in S \cup \{\ell\}$ some path through the one-volt source and the resistors of N that connects the two terminals of port j . Therefore when $r \neq (2p + q - 1)$, when all ports $j \in S \cup \{\ell\}$ are open circuited, when $v_\ell = 1$, and when $v_j = 0$ for all $j \in S$, the open-circuit voltage v_j at each port j with $j \in S \cup \{\ell\}$ is well defined and nonzero. Since no current flows in any resistor of N when $v_\ell = 1$, $v_j = 0$ for all $j \in S$, and all ports $j \in S \cup \{\ell\}$ are open-circuited, it follows that $v_j \in \{-1, +1\}$ for all $j \in S$. With $v_\ell = 1$, with $v_j = 0$ for all $j \in S$, and with v_j the corresponding open-circuit voltage for each $j \in S \cup \{\ell\}$, we have $Gv = \theta$. When $p > 0$, the vector v also satisfies the condition that $v_{(2k-1)}v_{2k} \geq 0$ for all $k = 1, 2, \dots, p$ since if $v_{(2k-1)}v_{2k}$ were negative for some k , then for that k $v_{(2k-1)} = 1$ and $v_{2k} = -1$ or $v_{(2k-1)} = -1$ and $v_{2k} = 1$; in either case $|v_{(2k-1)} - v_{2k}| = 2$ which contradicts the proposition that a network of nonnegative resistors can have no voltage gain. \square

APPENDIX*

A theorem due to R. S. Palais[†] asserts that if $R(\cdot)$ is a continuously-differentiable mapping of E^n into itself with values $R(q)$ for $q \in E^n$, then $R(\cdot)$ is a diffeomorphism[‡] of E^n onto itself if and only if

- (i) $\det J_q \neq 0$ for all $q \in E^n$, in which J_q is the Jacobian matrix of $R(\cdot)$ with respect to q , and
- (ii) $\|R(q)\| \rightarrow \infty$ as $\|q\| \rightarrow \infty$.

If $R(\cdot)$ is any twice-continuously-differentiable mapping of E^n into itself such that conditions (i) and (ii) of Palais' theorem are satisfied, then E^n contains a unique element x such that $R(x) = \theta$ in which θ is the zero element of E^n , and there are steepest decent as well as Newton-type algorithms each of which generates a sequence in E^n that converges to x . To show this, let ¹⁸ $f(y) = \|R(y)\|^2$ for all $y \in E^n$ in which $\|\cdot\|$ denotes the usual Euclidean norm (i.e., the square-root of the sum of squares). Since condition (i) of Palais' theorem is satisfied, the gradient ∇f of $f(\cdot)$ satisfies $(\nabla f)(y) \neq \theta$ unless $f(y) = 0$,[§] and since condition (ii) of Palais' theorem is satisfied, the set $S = \{y \in E^n : f(y) \leq f(x^{(0)})\}$ is bounded for any $x^{(0)} \in E^n$. Therefore we may appeal to, for example, the theorem of page 43 of Ref. 18 according to which for any $x^{(0)} \in E^n$, for any member of a certain class of mappings $\varphi(\cdot)$ of S

* The material of this appendix together with some misprints appears in Ref. 3.

† See Ref. 12 and the appendix of Ref. 13.

‡ A diffeomorphism of E_n onto itself is a continuously differentiable mapping of E_n into E_n which possesses a continuously differentiable inverse.

§ Here we have used the fact that $(\nabla f)(y) = 2J_y{}^t R(y)$ for all $y \in E_n$.¹⁸

into E^n , and for suitably chosen constants $\gamma_0, \gamma_1, \dots$, the sequence $x^{(0)}, x^{(1)}, \dots$ defined by

$$x^{(k+1)} = x^{(k)} + \gamma_k \varphi(x^{(k)}) \quad \text{for all } k \geq 0$$

belongs to S and is such that $\|R(x^{(k)})\| \rightarrow 0$ as $k \rightarrow \infty$. However, since $R^{-1}(\cdot)$ exists and is continuous, it follows from

$$x^{(k)} = R^{-1}[R(x^{(k)})] \quad \text{for all } k \geq 0$$

and the fact that $R(x^{(k)}) \rightarrow \theta$ as $k \rightarrow \infty$, that $\lim_{k \rightarrow \infty} x^{(k)}$ exists and

$$\lim_{k \rightarrow \infty} x^{(k)} = R^{-1}(\theta),$$

which means that $\lim_{k \rightarrow \infty} x^{(k)}$ is the unique solution x of $R(y) = \theta$.

REFERENCES

1. Sandberg, I. W., and Willson, A. N., Jr., "Some Theorems on Properties of DC Equations of Nonlinear Networks," B.S.T.J., 48, No. 1 (January 1969), pp. 1-34.
2. Sandberg, I. W., and Willson, A. N., Jr., "Some Network-Theoretic Properties of Nonlinear DC Transistor Networks," B.S.T.J., 48, No. 5 (May-June 1969), pp. 1293-1312.
3. Sandberg, I. W., "Theorems on the Analysis of Nonlinear Transistor Networks," B.S.T.J., 49, No. 1 (January 1970), pp. 95-114.
4. Willson, A. N., Jr., "New Theorems on the Equations of Nonlinear DC Transistor Networks," B.S.T.J., this issue, pp. 1713-1738.
5. Sandberg, I. W., "Some Theorems on the Dynamic Response of Nonlinear Transistor Networks," B.S.T.J., 48, No. 1 (January 1969), pp. 35-54.
6. Hamming, R. W., *Numerical Methods for Scientists and Engineers*, New York: McGraw-Hill Book Co., (1962).
7. Ralston, A. A., *A First Course in Numerical Analysis*, New York: McGraw-Hill Book Co., (1965).
8. Hachtel, G. D., and Rohrer, R. A., "Techniques for the Optimal Design and Synthesis of Switching Circuits," Proc. of the IEEE, 55, No. 11 (November 1967), pp. 1864-1876.
9. Sandberg, I. W., and Shichman, H., "Numerical Integration of Systems of Stiff Nonlinear Differential Equations," B.S.T.J., 47, No. 4 (April 1968), pp. 511-527.
10. Calahan, D. A., "Efficient Numerical Analysis of Non-Linear Circuits," Proc. Sixth Ann. Allerton Conf. on Circuit and System Theory, University of Illinois, 1968, pp. 321-331.
11. Vehovec, M., "Simple Criterion for the Global Regularity of Vector-Valued Functions," Elec. Letters, 5, No. 26 (December 1969), pp. 680-681.
12. Palais, R. S., "Natural Operations on Differential Forms," Trans. Amer. Math. Soc., 92, No. 1 (1959), pp. 125-141.
13. 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, University of Illinois, 1965, pp. 536-545.
14. Mitra, D., Sandberg, I. W., and Gopinath, B., "A Note on a Curious Property of the Equations of Nonlinear Networks Containing Transistors," to be published.
15. Muir, T., *A Treatise on the Theory of Determinants*, New York: Dover Publications, Inc., (1960), pp. 31-33.
16. Meyer, G. H., "On Solving Nonlinear Equations with a One-Parameter Operator Imbedding," Tech. Rep. 67-50, Comp. Science Center, University of Maryland, College Park, 1967.
17. Hadamard, J., "Sur Les Transformations Ponctuelles," Bull. Soc. Math. France, 48, (1920), pp. 13-27.
18. Goldstein, A. A., *Constructive Real Analysis*, New York: Harper and Row (1967), pp. 41-45.