THE AUTOMATED DESIGN OF LINEAR INTEGRATED CIRCUIT AMPLIFIERS

By

WILLIAM ARTHUR BRISTOL

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1972

To

Susan, my wife

ACKNOWLEDGMENTS

The author is indebted to Dr. Stephen W. Director, the chairman of the supervisory commitee, for his encouragement, constructive criticism, and continuous assistance. The author is also grateful for the valuable contributions made by Dr. Arthur J. Brodersen. Appreciation is extended to Dr. Thomas E. Bullock and Dr. Frank D. Vickers for their time and interest.

Special gratitude is extended to his wife for her excellent typing of this dissertation. Her patience, love and hard work have significantly contributed to the successful completion of this study.

The author is grateful for the support of the National Science Foundation through Grant GK-27615 and for the computer funds made available by the University of Florida Computing Center.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy

THE AUTOMATED DESIGN OF LINEAR INTEGRATED CIRCUIT AMPLIFIERS

By

William Arthur Bristol

December, 1972

Chairman: Dr. Stephen W. Director
Major Department: Electrical Engineering

A procedure for automated design of monolithic integrated circuit amplifiers is described. This procedure yields a circuit which satisfies both ac and dc design specifications; the ac specifications include requirements on input and output impedances, gain, and bandwidth, while dc specifications include requirements on temperature stability and power dissipation. The design problem is transformed into a mathematical programming problem by embodying the specifications in a flexible least pth performance function. All ac and dc specifications are considered simultaneously because the contributions to the performance function from all specifications are summed to form a concise measure of the network's ability to exhibit the desired characteristics. Gradients, which are determined by the adjoint network method, are used to adjust the network parameters and topology so as to minimize this performance function. Topology adjustment is

achieved by growing and removing circuit elements, including transistors. Transistor growth and removal is indicated by gradient expressions that determine where ac gain is necessary. Since each transistor can be connected in one of three configurations, the most appropriate configuration must be selected through use of the gradients. A computer program, FROLIC (Fabrication Related Optimization of Linear Integrated Circuits), has been written which implements the design procedure, and several examples are presented to demonstrate its ability.

CHAPTER 1

INTRODUCTION

In recent years, the use of linear integrated circuits has increased significantly. Development of these circuits is complicated by their compact nature and the presence of numerous nonlinear elements such as diodes and transistors. In the past, the circuit designer typically would build a breadboard of his circuit to test and evaluate the circuit's performance. Often individual elements would be changed to improve the performance. With an integrated circuit, where the entire circuit is on a small chip, this procedure is not feasible because of the amount of time and money required to build each circuit. Therefore, a computerized procedure which automatically evaluates an integrated circuit's performance and optimizes it without requiring the circuit to be built furnishes a substantially less expensive design in less time.

Background

Network analysis, performance evaluation, and parameter adjustment are three basic functions of an automated network design algorithm. Network analysis is used to determine the responses of a network to a given set of excitations. A scalar performance function is then used to compare the actual

responses with the desired responses. Systematic adjustment of the network's parameters optimizes this performance function and therefore obtains the network that best satisfies the design specifications. Each of these operations must be performed accurately and efficiently.

Design is based upon optimization of the scalar performance function. Two types of optimization methods can be used to adjust methodically the network's parameters: direct methods and gradient methods. Direct methods which optimize functions without calculation of parameter gradients generally converge very slowly to an optimal set of parameter values when more than five parameters are considered. Gradient methods, potentially faster than direct methods, require the determination of the gradients of the performance function with respect to all parameters. The classical technique used to determine gradients is differencing which requires perturbation of each parameter, one at a time, and reevaluation of the network's behavior after each perturbation. Because of the large number of analyses required using perturbation, this method is not much faster than direct search methods.

Prior to 1967, computer-aided design algorithms were generally limited to network analysis[1-4] because of the large amounts of computational time that were required to adjust network parameters. A visual display which interacted with an analysis program was often used to allow a circuit designer to adjust network parameters and determine the effect of the adjustment on the network responses. Through a trial and error process, a final network was obtained.

In 1967, Rohrer[5] proposed a method for the efficient determination of partial derivatives of a scalar performance function with respect to the network parameters. This work was oriented toward lumped, linear, time-invariant network design problems in two domains: frequency domain (ac analysis) and time domain (transient analysis). Variational calculus was used to derive expressions that enabled the calculation of all parameter gradients after one analysis of the original network and one analysis of an associated network, known as the adjoint network. Thus, regardless of the size of the network under consideration, all gradients can be determined after two network analyses, an impressive advance beyond the method of perturbation. This work was later extended by Director and Rohrer[6-8] to include consideration of generalized network elements.

With an efficient gradient calculation method available, many researchers[9-16] developed automated design procedures. Existing techniques, such as network analysis, were easily adapted to network design. Some work was done on ac and dc procedures, but more effort was devoted to transient design. The dominance of transient design was due to its ability to handle nonlinear digital circuits which have been used extensively in recent high-speed computers.

Linear Amplifier Design

An automated procedure for linear amplifier design will be described which utilizes the efficient "adjoint network"

method for gradient calculation. This design procedure differs from previous work because it simultaneously considers the ac and dc characteristics of a circuit.

A basic premise upon which we base simultaneous ac and dc design is that the total operation of a linear amplifier can be adequately represented by the superposition of its ac and dc responses. When only static inputs are applied to a linear amplifier, dc analysis provides a complete description of the amplifier's operation. The dynamic behavior of a linear amplifier can be ascertained through ac analysis if two prerequisites are satisfied: the static operation of the amplifier must be known so that ac models can be specified and all dynamic excitations must consist of sums of sinusoids. These dynamic excitations must also cause responses which are small compared to existing dc conditions. Since satisfaction of these conditions is common, the total operation of a linear amplifier can be obtained through the superposition of ac and dc analyses.

Resistors, capacitors, diodes and bipolar transistors are the most common circuit elements used in integrated circuits. Resistors and capacitors are naturally linear and therefore may be represented by linear models. But diodes and transistors, especially for dc analysis, are nonlinear. An iterative procedure is used to update a linearization of the network equations so that conventional linear solution techniques can be employed.

A performance function is developed that is capable of representing all of the ac and dc specifications with a scalar measure. Flexibility is built into the function so that many different criteria may be specified. Typically, the design specifications include requirements on amplifier gain and input and output impedance over a frequency range and requirements on power dissipation and temperature sensitivity over a temperature range.

Gradient information is determined by use of the adjoint network method. All parameter gradients can be obtained with only two network analyses at each frequency and two analyses at each temperature. Element growth and removal have been shown to be feasible for resistors and capacitors with these gradients. Growth and removal of transistors are explored in order to allow the algorithm total flexibility.

Parameter adjustment is effected by a sophisticated algorithm which utilizes the gradients to optimize the performance function. In order to obtain realizable designs, the network parameters are constrained within reasonable boundaries. The presence of these constraints is considered when the optimization algorithm is selected.

The main objective of this research has been to develop an automated design algorithm that simultaneously meets ac and dc specifications. The computer program FROLIC (Fabrication Related Optimization of Linear Integrated Circuits) has been written and will be used to demonstrate the effectiveness of the algorithm. The techniques that are incorporated in the

algorithm will be described in this dissertation. Chapter 2 describes the models of circuit elements and the method of analysis. Chapter 3 discusses the performance function, the optimization strategy, the derivation of the gradient expressions, and the numerical integration procedure. Chapter 4 develops indicators that are used to suggest transistor growth and removal. Chapter 5 describes the overall design algorithm and some methods that are necessary to link together the ac and dc segments. Chapter 6 contains some concluding remarks and makes some suggestions for further research.

CHAPTER 2

MODELS AND ANALYSIS TECHNIQUES

A basic part of an automated network design package is the analysis subroutine which is used repetitively to assess network performance. Since analysis is performed often, it must be accomplished efficiently. We start with a discussion of the physical elements which can be fabricated in a monolithic integrated circuit and the ideal elements used to model them. Then we proceed to a discussion of equation formulation and solution techniques.

Modeling of Physical Elements

It is important to distinguish between the physical elements present in an integrated circuit and the ideal elements which are used to model them. We consider networks which are made up of integrated resistors, capacitors, diodes, and bipolar transistors. Each of these physical elements is modeled by some of the following ideal elements: resistors, capacitors, inductors, voltage-controlled current sources, diodes, and independent voltage and current sources. The ideal element values are specified by parameters which are related to the characteristics of the physical elements, i.e., the planar dimensions of the masks, the fabrication procedure, and possibly the dc operating point. A linear temperature

dependence is associated with each model to allow for the effect of temperature on the performance of each physical element.

The ac steady-state, or frequency domain, behavior of diodes and transistors depends upon the dc, or large signal, operating conditions of the network. Therefore, before the frequency response of an integrated circuit can be determined, a dc analysis must be performed to ascertain the operating points. The models used to represent each element may be different under dc and ac conditions. This fact is discussed below in relation to each element type.

Integrated resistors are typically realized as base diffusions. The v-i relationship for a resistor is approximated by

$$v_R(T) = \rho_S[\ell/w]\{1+\alpha_R[T-T_0]\}i_R(T)$$

where T is temperature in degrees Kelvin,

$v_R(T)$ is the branch voltage,

ρ_S is the sheet resistivity,

ℓ/w is the length to width ratio of the base diffusion,

α_R is a temperature coefficient,

T_0 is 298° Kelvin, and

$i_R(T)$ is the branch current.

For a given temperature, an integrated resistor can be modeled by an ideal resistor with resistance

$$R = \rho_S[\ell/w]\{1+\alpha_R[T-T_0]\}.$$

Since this expression is independent of frequency and operating point, it is used for both dc and ac analysis.

Integrated capacitors are realized as reverse-biased base-collector or base-emitter junctions or metal-oxide-silicon structures whose area A is proportional to the capacitance of the element. The i-v relationship of an integrated capacitor is approximated by

$$i_C(t,T) = \theta A\{1+\alpha_C[T-T_0]\}\partial v_C(t,T)/\partial t$$

where t is time and θ is a processing parameter. For a given temperature, an integrated capacitor can be modeled by an ideal capacitor with capacitance

$$C = \theta A\{1+\alpha_C[T-T_0]\}.$$

Under dc conditions,

$$\partial v_C(t,T)/\partial t = 0.$$

Therefore, for dc analysis

$$i_C(t,T) = 0$$

and the capacitor may be considered to be a zero-valued current source or an open circuit. For sinusoidal steady-state analysis, we employ a complex-valued branch relationship

$$I_C(\omega) = j\omega C V_C(\omega)$$

where ω is the angular frequency of the excitation and $I_C(\omega)$

and $V_C(\omega)$ are complex-valued branch current and voltage phasors, respectively.

Integrated diodes are described by the nonlinear V-I relationship

$$V_D(T) = [nkT/q]\ell n\{I_D(T)/[J_S A(1+\alpha_D(T-T_0))]+1\}+I_D(T)R_S$$

where n is the emission coefficient,

k is Boltzmann's constant,

q is the magnitude of an electronic charge,

J_S is a processing dependent current density, and

R_S is a series resistance.

The dynamic effects that influence the operation of a diode will be considered later. For a given temperature, an integrated diode is represented by an ideal diode whose relationship is

$$V_D = [nkT/q]\ell n[I_D/I_S+1],$$

or its inverse

$$I_D = I_S\{exp[qV_D/(nkT)]-1\},$$

in series with resistance R_S. The saturation current I_S is defined as

$$I_S = J_S A[1+\alpha_D(T-T_0)].$$

During ac analysis, a diode is replaced by a linear RC network whose dominant low-frequency element is the small signal resistance of the diode about the dc operating point:

$$R = \partial V_D / \partial I_D \Big|_{V_D} = nkT/(qI_S) exp[-qV_D/(nkT)]$$

where V_D is the dc junction voltage. Two capacitances are added to the ac diode model to represent the dominant dynamic effects in an integrated diode: a diffusion capacitance to represent the changes in minority carriers stored in the quasi-neutral regions and a transition capacitance to represent the changes in charge stored in the space-charge region. The diffusion capacitance is defined by the relationship

$$C_D = qI_S/(nkT\omega_X) exp[qV_D/(nkT)]$$

where ω_X is an angular frequency processing constant. The transition capacitance is defined by the relationship

$$C_T = C_0/(1-V_D/V_0)**(1/m)$$

where C_0 is the transition capacitance at $V_D=0$, V_0 is the contact potential, and m is the grading coefficient. The complete ac diode model consists of the small signal resistance in parallel with two capacitances, all of which are in series with the resistance R_S, as shown in Fig. 2-1.

The integrated circuits under consideration employ both NPN and PNP three-layer bipolar transistors. The following discussion describes only the NPN transistor models, but the PNP models are developed similarly.

Separate consideration of the ac and dc performance of a transistor with distinct models for each mode of operation allows simpler models to be used without sacrificing accuracy.

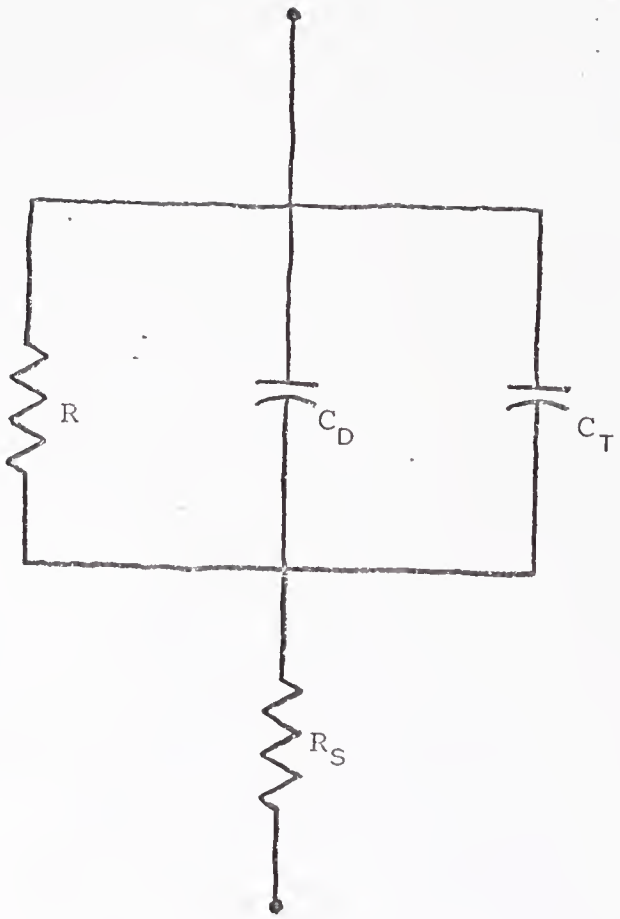


Fig. 2-1

Model used to represent a diode for ac analysis

Since it is desirable to use the same network configuration for each mode of operation, all transistor models are reformulated as two-port networks represented by y-parameters, as shown in Fig. 2-2. This formulation allows all transistor models to be represented by two branches, two voltage-controlled current sources, and for the dc case, two independent sources. (Note that no internal nodes are present in this model.)

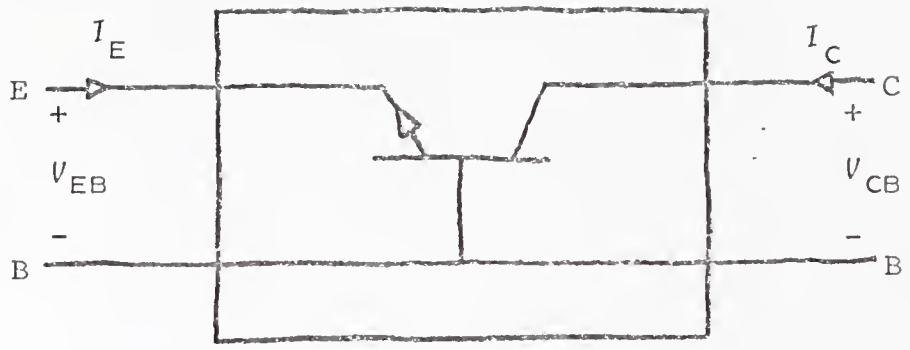
The transistor model can be separated into intrinsic and extrinsic parts. Extrinsic effects are due to the bulk material between the intrinsic transistor and the emitter, base, and collector contacts and are modeled by the three resistors shown in Fig. 2-3. The remainder of this discussion is concerned only with the intrinsic transistor.

The dc or large signal, performance of a transistor may be adequately represented by the Ebers-Moll model,[17,18] as shown in Fig. 2-4, when exposed to typical operating conditions. Each diode in this model is assumed to obey the standard diode relation

$$I = I_X\{\exp[qV_{DX}/(n_X kT)]-1\} \qquad (2.1)$$

where I_X, V_{DX}, and n_X are I_{ES}, V_{D1}, and n_1 for the emitter-base diode and I_{CS}, V_{D2}, and n_2 for the collector-base diode. Temperature dependence is incorporated in this model through the following relationships:

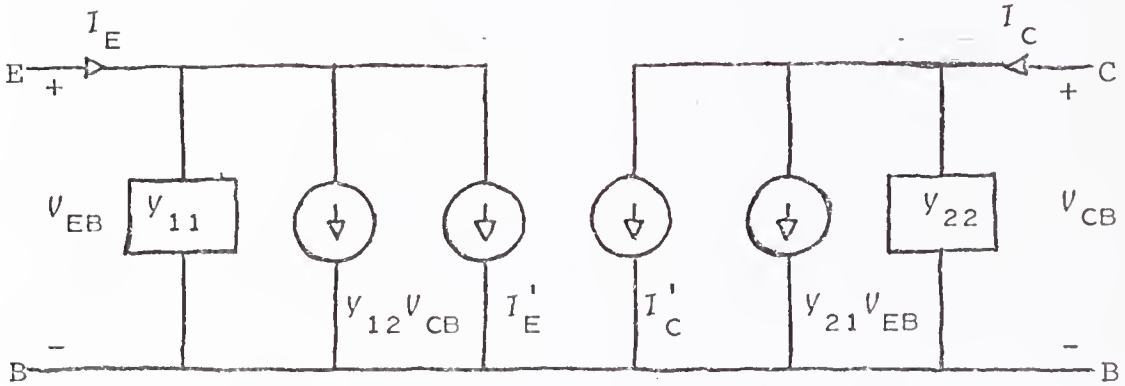
$$I_{ES} = I_{ES0}[1+\alpha_1(T-T_0)],$$



(a)

$$\begin{bmatrix} I_E \\ I_C \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_{EB} \\ V_{CB} \end{bmatrix} + \begin{bmatrix} I_E' \\ I_C' \end{bmatrix}$$

(b)



(c)

Fig. 2-2

Each transistor is described as a two-port
(a) Terminal voltages and currents
(b) Y-parameter description
(c) Circuit description

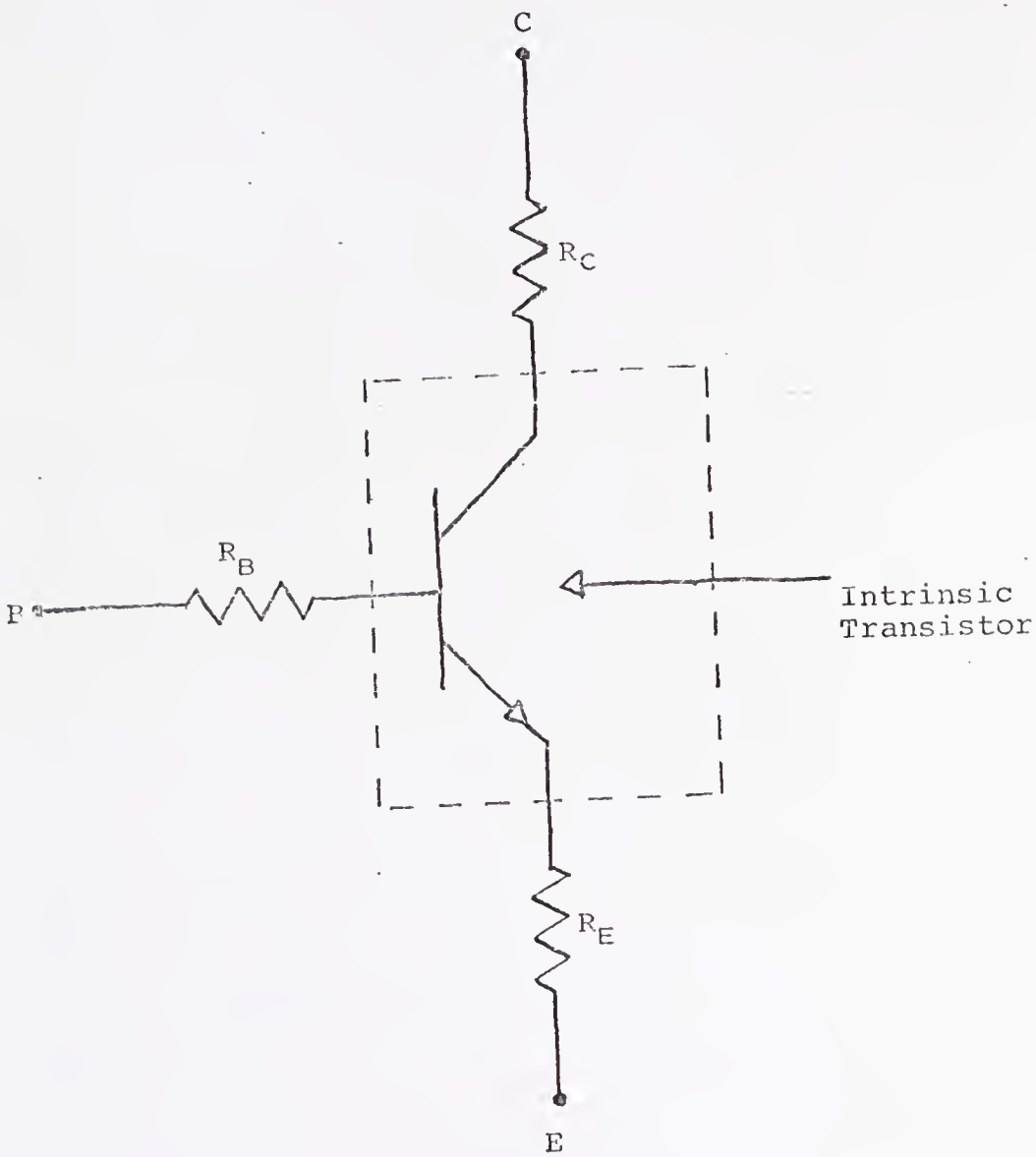


Fig. 2-3

Three resistors represent the extrinsic effects of a transistor

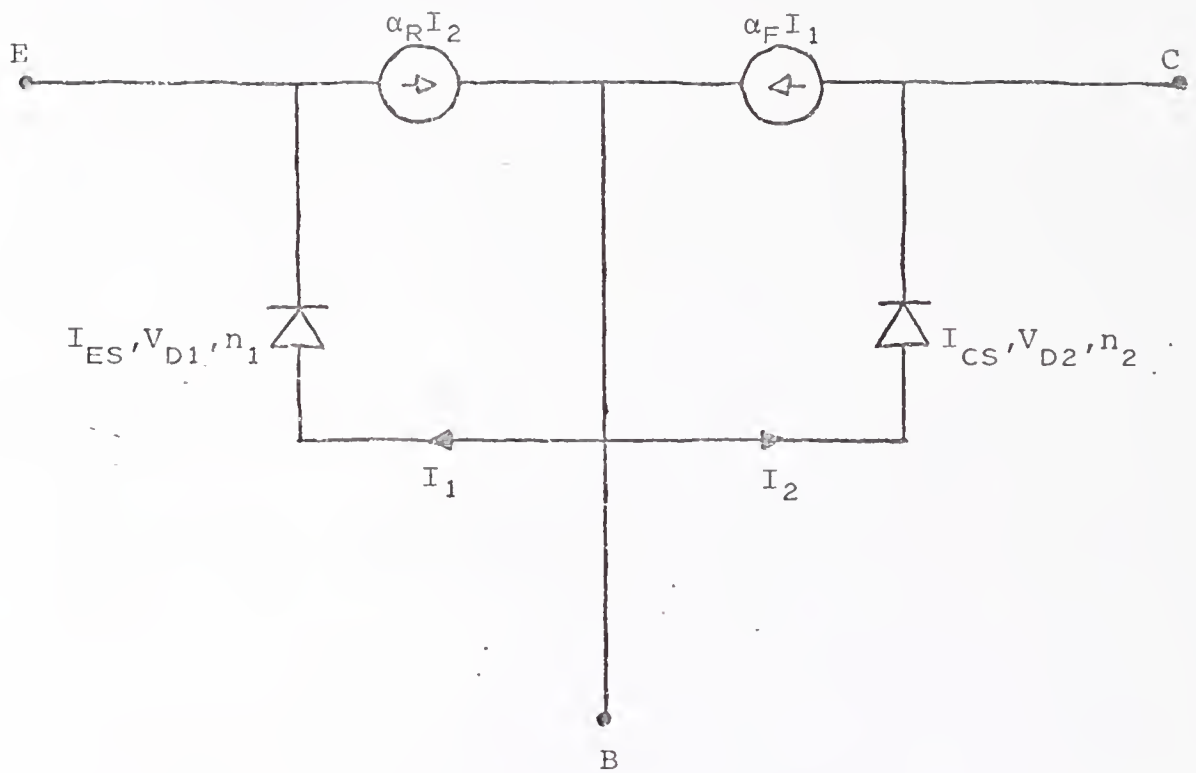


Fig. 2-4

Ebers-Moll model used to model
the dc behavior of a NPN transistor

$$I_{CS} = I_{CSO}[1+\alpha_2(T-T_0)],$$

and

$$\beta_F = \beta_{FO}[1+\alpha_3(T-T_0)]$$

where I_{ESO}, I_{CSO}, and β_{FO} are parameter values at T_0 and α_1, α_2 and α_3 are temperature coefficients. Since β_F is not directly used in the Ebers-Moll model, the following relationship

$$\alpha_F = \beta_F/(\beta_F+1)$$

defines α_F which is a controlled-source gain coefficient. It should be noted that dynamic effects have been omitted in the above characterization.

A model that represents the ac, or small signal, performance of a transistor can be obtained directly from the Ebers-Moll model by linearizing the diodes about their dc operating points and adding components to represent the dynamic characteristics of each diode. Although this model would be suitable, a more convenient model can be developed through a simple transformation.[19] The first step of the transformation is accomplished by substituting two equal current sources for each of the two sources already in the Ebers-Moll model, as shown in Fig. 2-5(a), so that the currents flowing in the model are unchanged. The second step is to replace each parallel combination of a diode and current source with a diode that has the I_X of (2.1) defined by

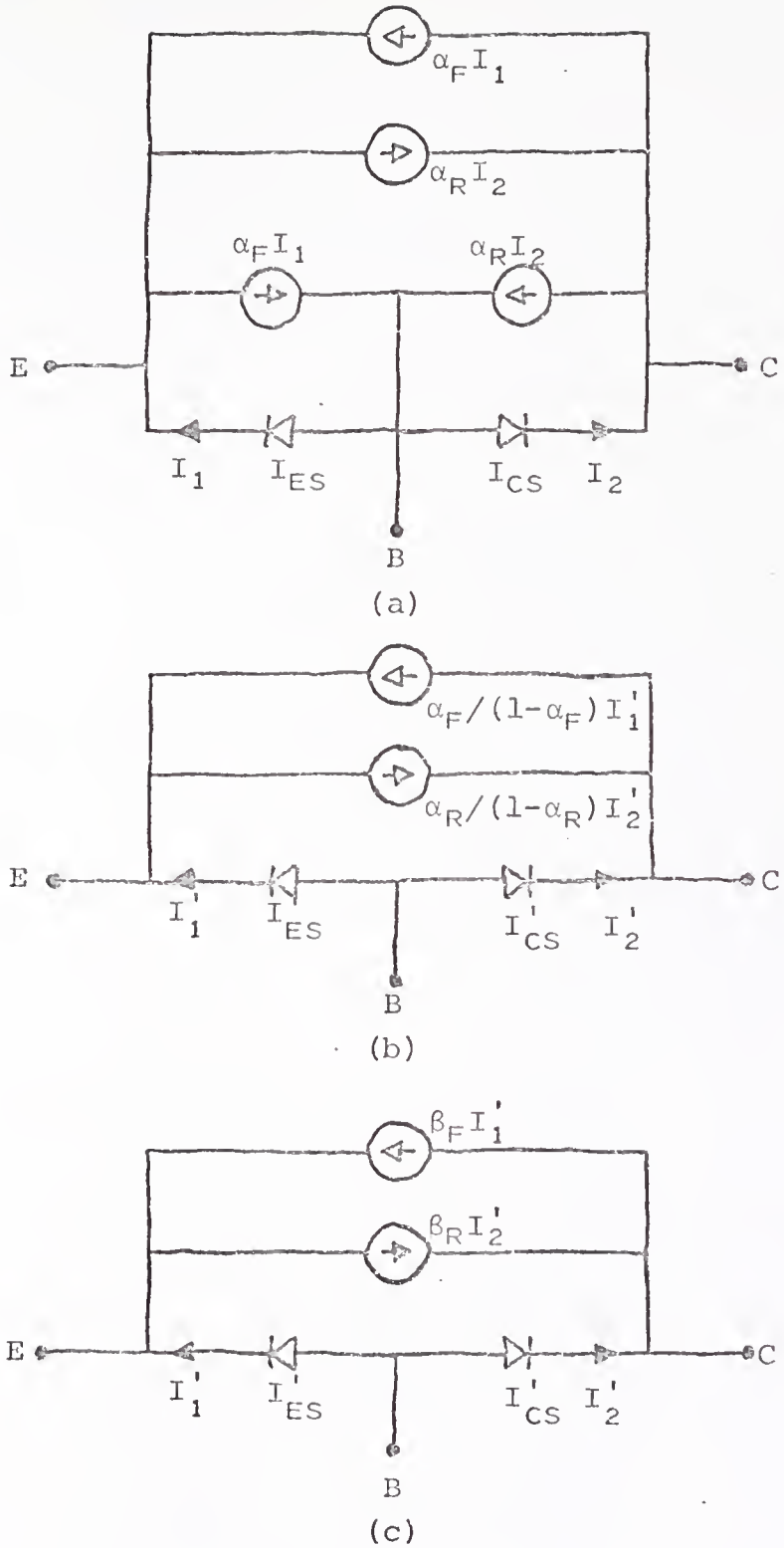


Fig. 2-5

Transformation of Ebers-Moll model
(a) First step
(b) Second step
(c) Final version

$$I'_{ES} = I_{ES}(1-\alpha_F)$$

for the emitter-base combination and

$$I'_{CS} = I_{CS}(1-\alpha_R)$$

for the collector-base combination. If we define

$$I'_1 = I_1(1-\alpha_F)$$

and

$$I'_2 = I_2(1-a_R),$$

then we have the network of Fig. 2-5(b). By defining

$$\beta_F = \alpha_F/(1-\alpha_F)$$

and

$$\beta_R = \alpha_R/(1-\alpha_R),$$

we have the final version of the transformed Ebers-Moll model shown in Fig. 2-5(c).

To obtain a small signal model from this transformed model, each diode must be replaced by a small signal resistance and two capacitors, as shown in Fig. 2-6. The diodes' small signal resistances are

$$R_1 = [n_1 kT\beta_F/(\alpha_F q I_{ES})]exp[-qV_{D1}/(n_1 kT)]$$

and

$$R_2 = [n_2 kT\beta_R/(\alpha_R q I_{CS})]exp[-qV_{D2}/(n_2 kT)].$$

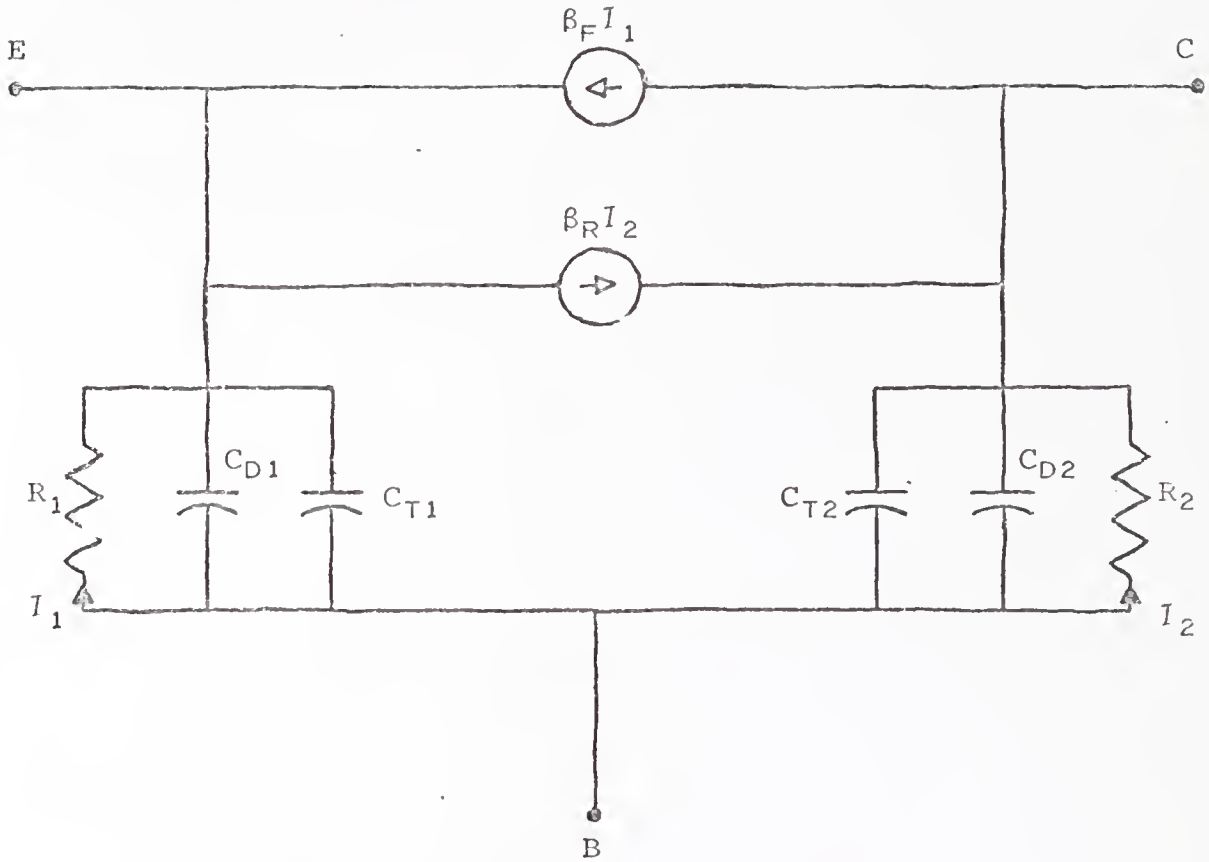


Fig. 2-6

Linearized transformed Ebers-Moll model
used to model the small-signal behavior of a transistor

Two capacitors for each diode represent the change in minority carriers stored in the quasi-neutral regions and the change in charge stored in the space-charge region. The diffusion capacitances which model the quasi-neutral regions are defined by the relations

$$C_{D1} = \beta_F/(R_1\omega_{T1})$$

and

$$C_{D2} = \beta_R/(R_2\omega_{T2})$$

where ω_{T1} and ω_{T2} are angular frequency parameters which are dependent upon the processing and the planar dimensions of the masks. The transition capacitances which model the space-charge layer are defined by the relations

$$C_{T1} = C_{01}/(1-V_{D1}/V_{01})**(1/m_1)$$

and

$$C_{T2} = C_{02}/(1-V_{D2}/V_{02})**(1/m_2)$$

where C_{01} and C_{02} are the transition capacitances measured at $V_{D1}=0$ and $V_{D2}=0$, respectively, V_{01} and V_{02} are contact potentials, and m_1 and m_2 are the grading coefficients.

The linearized transformed Ebers-Moll model is a suitable small-signal transistor model for use in all four regions of transistor operation. Two problems with this model are the large number of parameters that need to be specified and the absence of consideration of the Early effect.[20]

For transistors known to operate in the forward-active region, a simpler small-signal model, the hybrid-pi model[21] shown in Fig. 2-7, is available. This model requires fewer parameters and includes consideration of the Early effect. This model should not be used for any transistor which may operate in a region other than forward-active, as inaccurate results will be obtained.

Nonlinear DC Analysis Procedure

The presence of diodes in a network requires the solution of nonlinear equations in order to obtain large signal responses. Since efficient methods for the solution of linear equations are available, we are motivated to employ an analysis scheme which is based upon repeated linearizations of the nonlinear branch relationships about estimates of their operating points. In effect, what results is a technique which replaces each diode with a series combination of a conductor and an independent current source.

The nonlinear equations are of the form

$$\underline{f}(\underline{V}) = 0 \tag{2.2}$$

where each equation is the sum of currents leaving a node of a network and $\underline{V}$ is the vector of node voltages. An equation and a voltage are associated with each node of a network except one node which is specified as a reference or ground. A suitable solution technique is the iterative Newton-Raphson algorithm which can be derived by expanding $\underline{f}(\underline{V})$ in

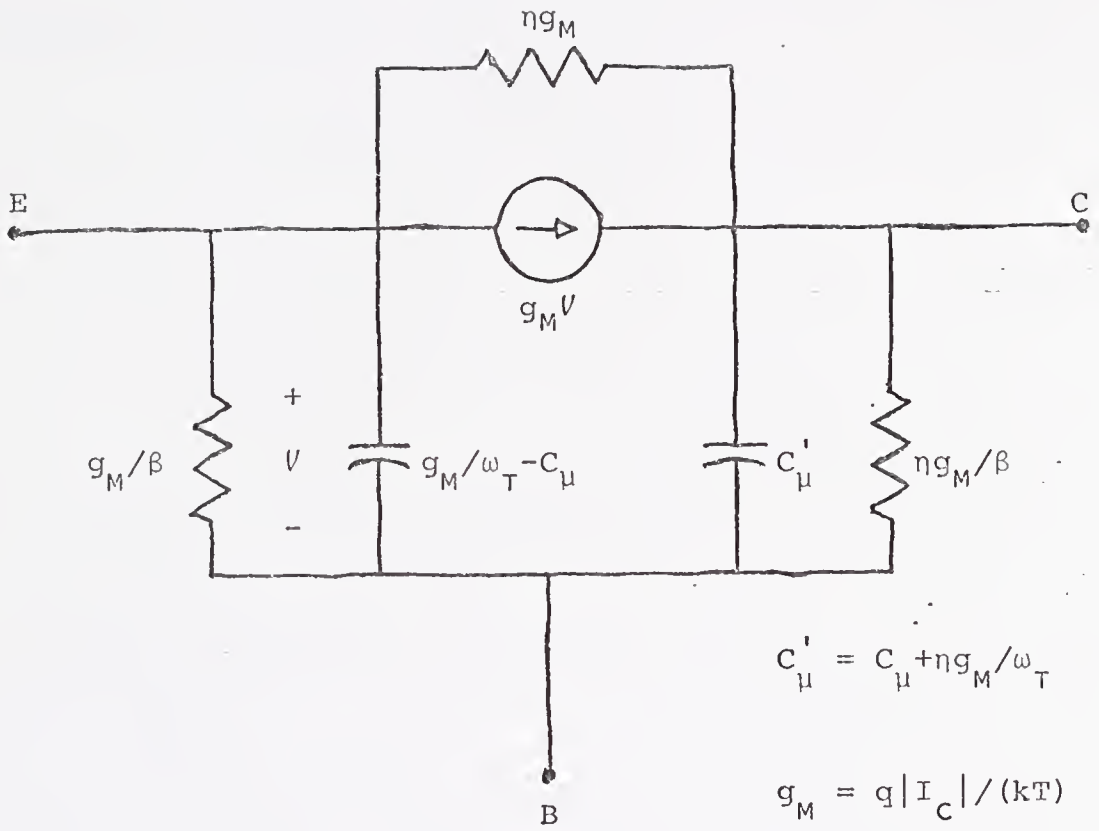


Fig. 2-7

The hybrid-pi model
used to model the small-signal behavior of transistors

a Taylor's series:

$$\underline{f}(\underline{V}^{M+1}) = \underline{f}(\underline{V}^{M})+\underline{J}(\underline{V}^{M})[\underline{V}^{M+1}-\underline{V}^{M}]+\ldots. \qquad (2.3)$$

where $\underline{V}^{M+1}$ is the node voltage vector at iteration M+1, $\underline{V}^{M}$ is the node voltage vector at iteration M, $\underline{J}(\underline{V})$ is the Jacobian of $\underline{f}(\underline{V})$, and higher order terms are neglected. If $\underline{V}^{M+1}$ is assumed to satisfy (2.2), then (2.3) can be rewritten in the form

$$\underline{V}^{M+1} = \underline{V}^{M}-[\underline{J}(\underline{V}^{M})]^{-1}\underline{f}(\underline{V}^{M}).$$

This expression allows an initial estimate of the node voltages to be improved iteratively until the actual node voltages are obtained.

An equivalent procedure will now be described which allows the nonlinear equations to be solved by repetitive analysis of a linearized network in which each nonlinearity is replaced by its Taylor's series expansion. The diode relationship is

$$I_D = I_S[exp(V_D/V_X)-1]$$

and its Taylor's series expansion is

$$I_D^{M+1} = I_S[exp(V_D^M/V_X)-1]+I_S/V_X\ exp(V_D^M/V_X)(V_D^{M+1}-V_D^M)+\ldots$$

where

$$V_X = nkT/q.$$

If we define

$$G_D^M = I_S/V_X\ exp(V_D^M/V_X),$$

then

$$I_D^{M+1} = I_S[exp(V_D^M/V_X)-1]+G_D^M(V_D^{M+1}-V_D^M).$$

This expression can be put in the form

$$I_D^{M+1} = G_D^M V_D^{M+1}+\{I_S[exp(V_D^M/V_X)-1]-G_D^M V_D^M\}$$

which is the I-V relationship of a current source of value

$$\hat{I}_D^M = I_S[exp(V_D^M/V_X)-1]-G_D^M V_D^M$$

in parallel with a conductance G_D^M. This configuration, shown in Fig. 2-8, replaces each diode in the network and, upon repeated linear analyses, yields the actual operating points of the network.

An example will now show the equivalence of these two procedures. For the network in Fig. 2-9(a), we write equations of the form (2.2) which are

$$f_1(\underline{V}) = 0 = -I+G_A V_1+I_{S1}[exp(V_1/V_X)-1] -I_{S2}[exp((V_2-V_1)/V_X)-1] \quad (2.4)$$

and

$$f_2(\underline{V}) = 0 = I_{S2}[exp((V_2-V_1)/V_X)-1]+G_B V_2. \quad (2.5)$$

The Jacobian in (2.3) is defined as

$$J = \begin{bmatrix} \partial f_1/\partial V_1 & \partial f_1/\partial V_2 \\ \partial f_2/\partial V_1 & \partial f_2/\partial V_2 \end{bmatrix} = \begin{bmatrix} G_A+G_{D1}^M+G_{D2}^M & -G_{D1}^M \\ -G_{D1}^M & G_{D2}^M+G_B \end{bmatrix}$$

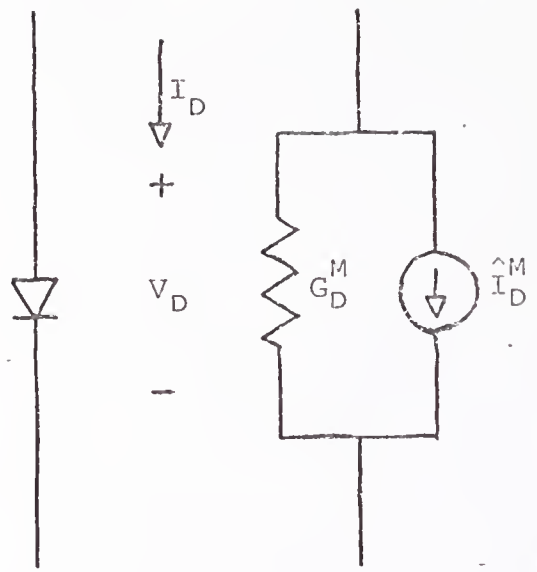
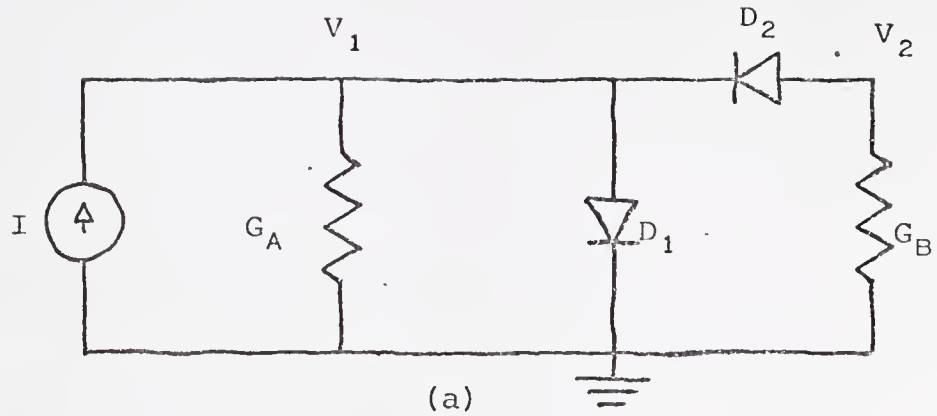
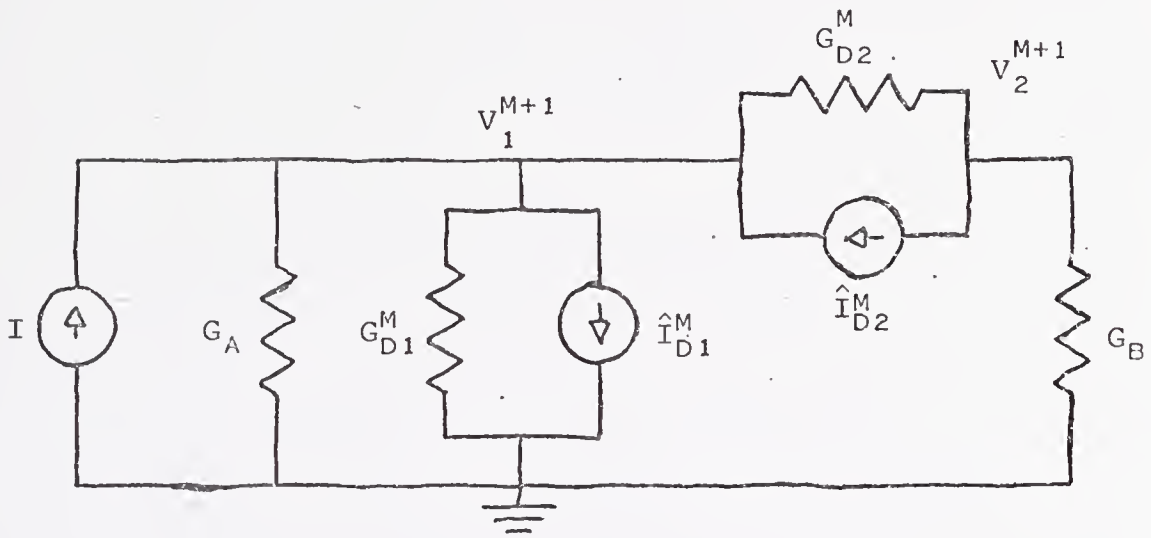


Fig. 2-8

Linearized diode model on right replaces all diodes in networks to be analyzed



(a)



(b)

Fig. 2-9

(a) Simple diode network used in example
(b) Linearized version of the same network

where

$$G_{D1}^{M} = I_{S1}/V_X exp(V_1^M/V_X) \quad .$$

and

$$G_{D2}^{M} = I_{S2}/V_X exp((V_2^M - V_1^M)/V_X) \, .$$

We now put (2.4) and (2.5) into the linearized form of (2.3):

$$0 = -I+G_A V_1^M + I_{S1}[exp(V_1^M/V_X)-1] - I_{S2}[exp((V_2^M-V_1^M)/V_X)-1]$$

$$+ \begin{bmatrix} G_A+G_{D1}^M+G_{D2}^M \\ \\ -G_{D2}^M \end{bmatrix}^T \begin{bmatrix} V_1^{M+1}-V_1^M \\ \\ V_2^{M+1}-V_2^M \end{bmatrix} .$$

and

$$0 = I_{S2}[exp((V_2^M-V_1^M)/V_X)-1]+G_B V_2^M$$

$$+ \begin{bmatrix} -G_{D2}^M \\ \\ G_{D2}^M+G_B \end{bmatrix}^T \begin{bmatrix} V_1^{M+1}-V_1^M \\ \\ V_2^{M+1}-V_2^M \end{bmatrix} .$$

where the superscript T indicates transposition. The vectors are multiplied and terms are grouped to form the following expressions:

$$0 = -I+\hat{I}_{D1}^M - \hat{I}_{D2}^M$$

$$+(G_A+G_{D1}^M+G_{D2}^M)V_1^{M+1} - G_{D2}^M V_2^{M+1} \qquad (2.6)$$

and

$$0 = \hat{I}^M_{D2} - G^M_{D2} V^{M+1}_1 + (G^M_{D2} + G_B) V^{M+1}_2 \qquad (2.7)$$

where

$$\hat{I}^M_{D1} = I_{S1}[exp(V^M_1/V_X) - 1] - G^M_{D1} V^M_1$$

and

$$\hat{I}^M_{D2} = I_{S2}[exp((V^M_2 - V^M_1)/V_X) - 1] - G^M_{D2}(V^M_2 - V^M_1).$$

A matrix form of (2.6) and (2.7) is

$$\begin{bmatrix} G_A + G^M_{D1} + G^M_{D2} & -G^M_{D2} \\ -G^M_{D2} & G^M_{D2} + G_B \end{bmatrix} \begin{bmatrix} V^{M+1}_1 \\ V^{M+1}_2 \end{bmatrix} = \begin{bmatrix} I - \hat{I}^M_{D1} + \hat{I}^M_{D2} \\ -\hat{I}^M_{D2} \end{bmatrix}.$$

This expression could have been obtained directly from the network in Fig. 2-9(b) where each diode is replaced by its linearized model of Fig. 2-8. The node voltages obtained from the solution of these equations allow determination of new linearized models and the process is repeated until the actual operating points are obtained.

The two Newton-Raphson procedures that have been discussed will produce the same result in the same number of iterations. The linearized model scheme is used in this algorithm.

A modification to the Newton-Raphson algorithm is implemented to avoid numerical problems and speed convergence. The linearized model requires the evaluation of an exponential function of the diode voltage. If the diode voltage is large compared to V_X, the exponential function could be

prohibitively large and cause numerical problems or possible divergence by the iterative procedure. Therefore, at all iterations where the diode voltage increases with respect to the previous iteration, an alternate model is used which is based on the current that was flowing in the linearized diode model. This alternate model is defined by the relations

$$G_D^M = (I_D^M + I_S)/V_X$$

and

$$\hat{I}_D^M = I_D^M - G_D^M[V_X \ell n(I_D^M/I_S + 1)].$$

At iterations where the diode voltage decreases, the standard Newton-Raphson method is used. The two estimated operating points about which the diode model is linearized are shown in Fig. 2-10. This modified Newton-Raphson algorithm has been found to converge reliably.

Another numerical problem that is encountered during use of the linearized model is caused by the wide range of values that can be assigned to the resistance. This problem is avoided by linearizing the diode curve to the right of the point where the resistance is equal to a defined short circuit and to the left of the origin, as shown in Fig. 2-11.

The number of linear analyses needed to obtain an accurate analysis of a nonlinear network can be reduced by the use of passive initial models for the nonlinear elements. A satisfactory initial model for non-transistor-model diodes and emitter-base diodes is a one-kilohm resistor. This

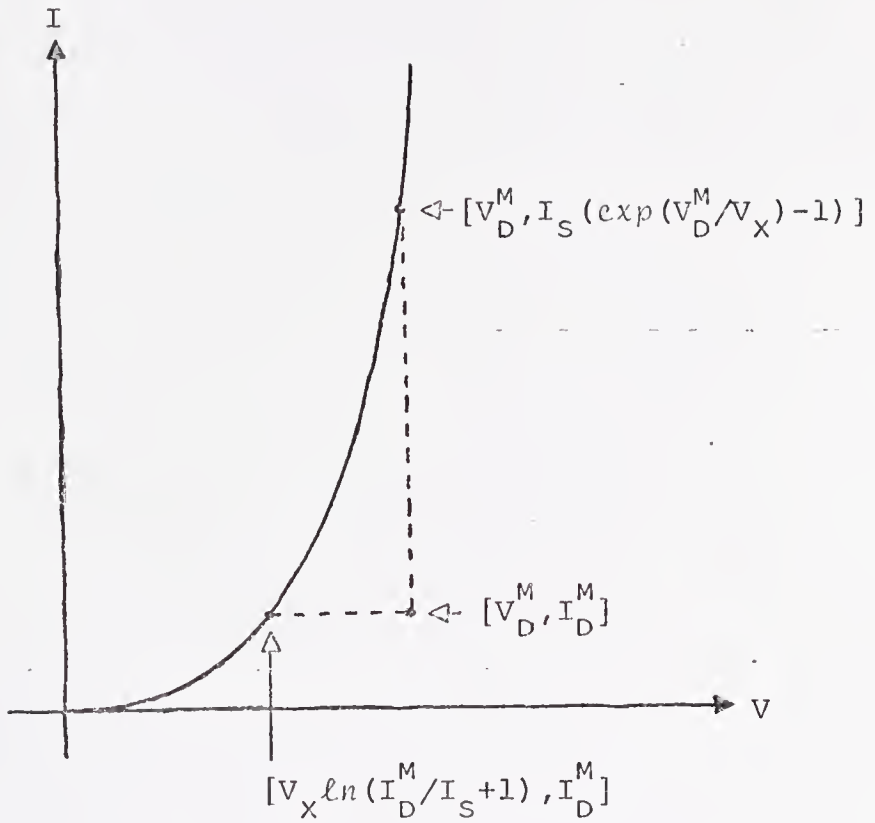


Fig. 2-10

Two possible estimated operating points
about which to linearize the diode relationship

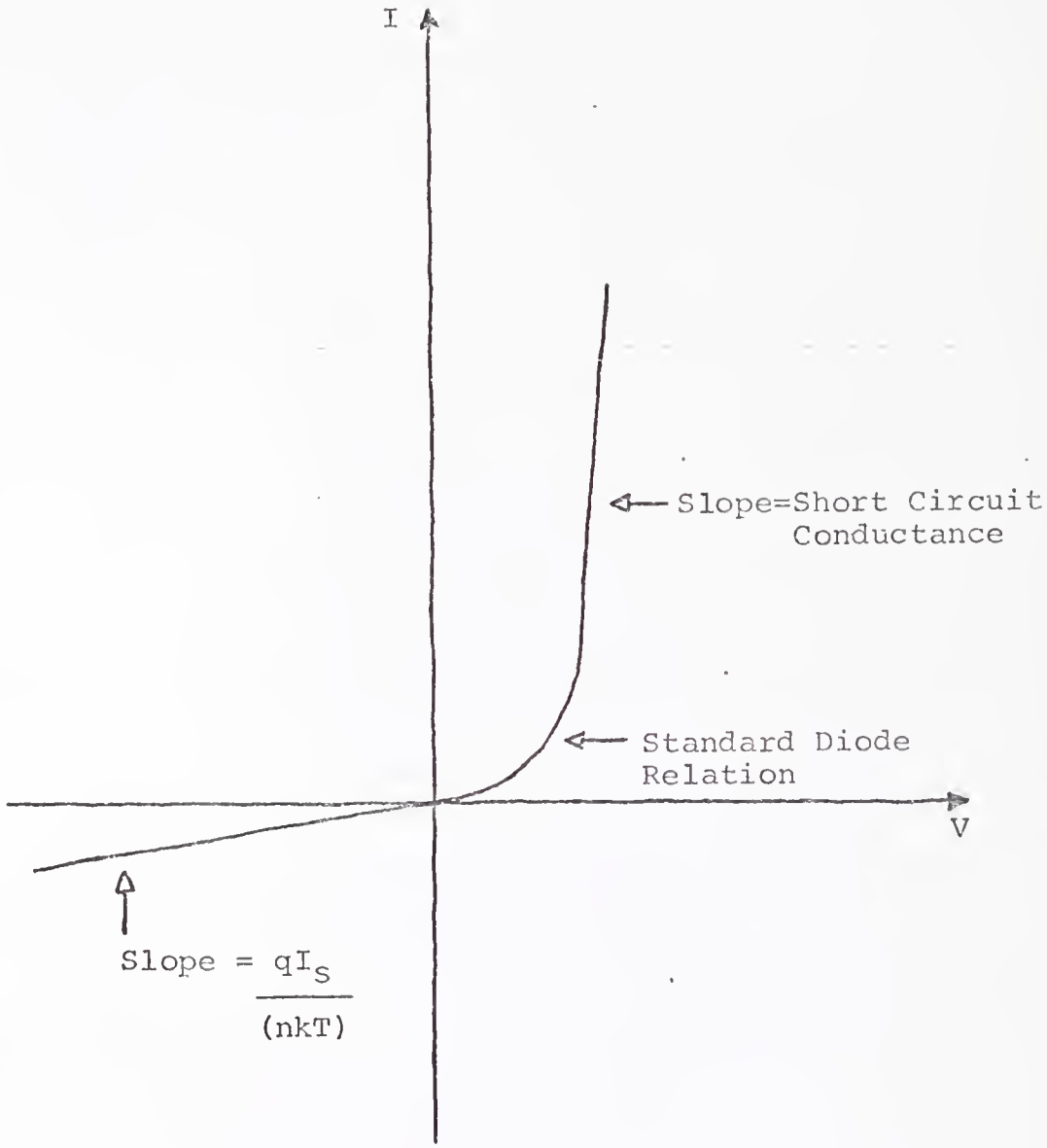


Fig. 2-11

Current-voltage relationship
used to avoid numerical inaccuracy

initial model assumes that the diode is forward-biased, but if the diode is reverse-biased, the model will be corrected after the first linear analysis. Collector-base diodes are assumed to be reverse-biased and their initial model is a resistor with the value defined above for reverse-biased diodes. These passive initial models are superior to initial models with current sources because of the possibility of initial divergence with the active model when a wrong assumption is made.

Many criteria could be used to determine when the linearized network has accurately reproduced the voltages and currents of the nonlinear network. A possible choice, to specify a maximum change allowable in node voltages at two consecutive iterations, would be satisfactory for some networks. Problems with this criterion are the requirement that the user know the range of approximate magnitudes of the node voltages in advance and each node voltage will have a different required accuracy that is dependent on the magnitude of the voltage. An improved method is to determine the relative change in node voltages between iterations by the relation

$$\left| (V_N^{L+1} - V_N^L) / V_N^{L+1} \right| < \varepsilon \qquad (2.8)$$

where V_N^{L+1} is the voltage of node N at iteration L+1, V_N^L is the voltage of node N at iteration L, and ε is a parameter specified by the user. If any node voltage fails to satisfy this criterion, another iteration should be started unless a maximum number of iterations has been reached. This criterion

is sufficient for almost all networks. When $V_N^{L+1}=0$, an alternate criteria, $|V_N^L|<\varepsilon$, is used which avoids division by zero.

Some networks contain transistors or diodes whose linearized models, during the iterative process, are biased so that their approach to the true operating point is extremely slow. Unless the value of ε is very small, the node voltages change so slowly that termination criterion (2.8) could be satisfied before the true operating conditions are determined. Such a situation is demonstrated by the network of Fig. 2-12. The node voltages for each iteration of the analysis scheme, as obtained by the computer program FROLIC whose flow charts are shown in the appendix, are listed in Table 2-1. The value of ε for this analysis was 1.E-4, a value suitable for most problems. The node voltages at iterations five and six, identical for five digits of accuracy, would have satisfied the criteria in (2.8) at iteration six and yielded an incorrect voltage for node twelve. This problem could be alleviated by reducing the value of ε which would have forced the analysis beyond iteration twelve. A more satisfactory solution is to require all diode relationships to be satisfied within ten percent before node voltages are checked, forcing the network to the neighborhood of the final solution before it is possible to terminate analysis and yet not overriding the flexibility of the check on the node voltages. This combination approach is used in FROLIC and was in effect during the determination of the node voltages in Table 2-1. When the test

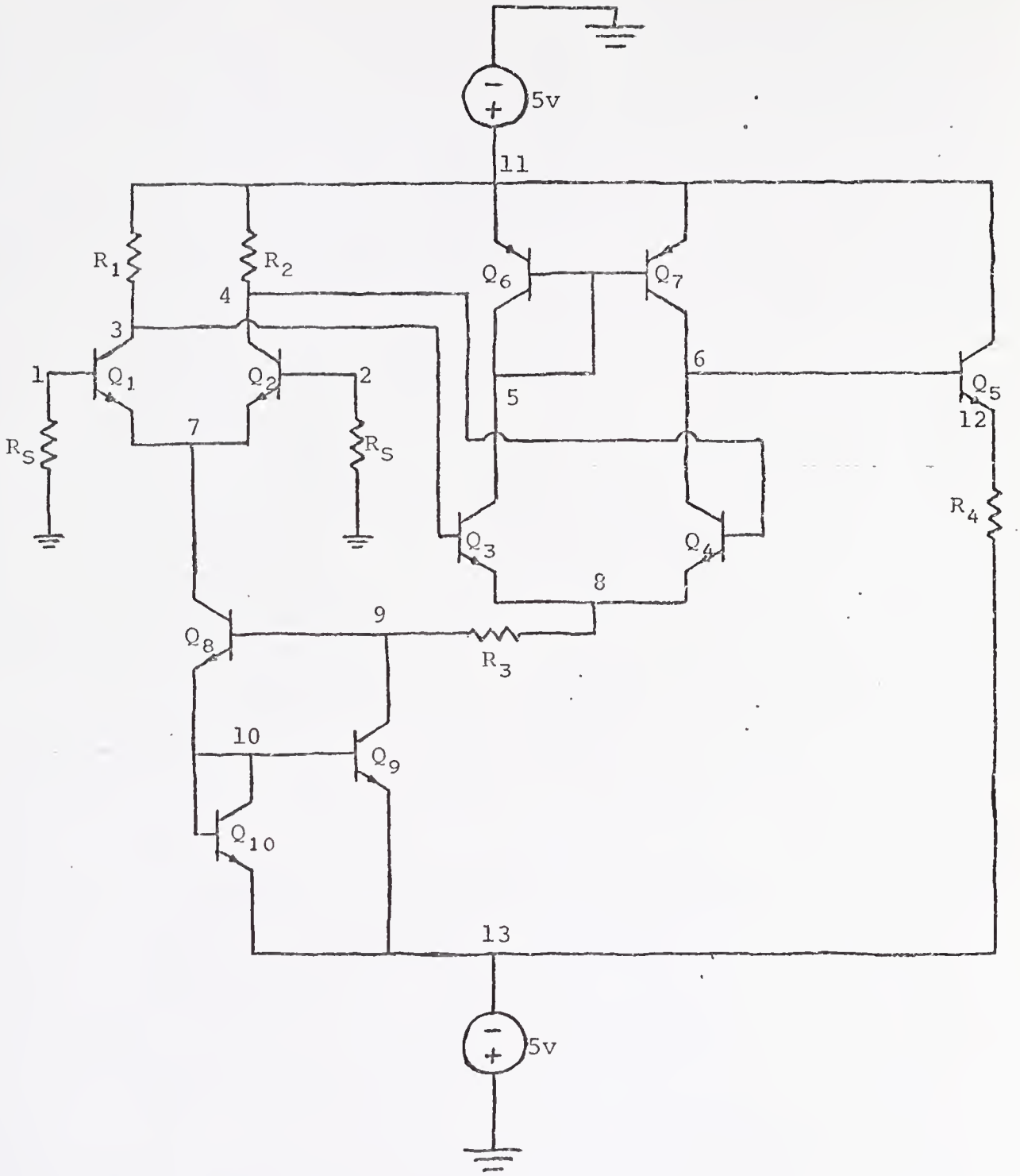


Fig. 2-12

Ten-transistor differential amplifier

Table 2-1

Intermediate Node Voltages For Network Shown In Fig. 2-12

ITERATION NUMBER	NODES 1,8	2,9	3,10	4,11	5,12	6,13	7
1	-1.3593D-5 -1.8754	-1.3593D-5 -4.9544	-1.8617 -4.9773	-1.8665 5.0000	4.9855 9.5886	9.9453 -5.0000	-1.1348D-2
2	1.8435D-5 1.9413	1.8467D-5 -3.9323	2.4970 -4.4665	2.5011 5.0000	4.4454 4.1246	4.7374 -5.0000	-0.51533
3	-4.9822D-6 1.7949	-4.9822D-6 -3.9469	2.3455 -4.4740	2.3432 5.0000	4.4507 -5.6277	-5.0427 -5.0000	-0.50877
4	-5.1595D-6 1.7069	-5.1595D-6 -3.9476	2.2571 -4.4744	2.2547 5.0000	4.4510 -5.0000	-4.2302 -5.0000	-0.50864
5	-5.1595D-6 1.7069	-5.1595D-6 -3.9476	2.2571 -4.4744	2.2547 5.0000	4.4510 -5.0000	1.9589 -5.0000	-0.50864
6	-5.1595D-6 1.7069	-5.1595D-6 -3.9476	2.2571 -4.4744	2.2547 5.0000	4.4510 -5.0000	1.9589 -5.0000	-0.50864
7	-5.1595D-6 1.7069	-5.1595D-6 -3.9476	2.2571 -4.4744	2.2547 5.0000	4.4510 -5.0000	1.9589 -5.0000	-0.50864
8	-5.1595D-6 1.7069	-5.1595D-6 -3.9476	2.2571 -4.4744	2.2547 5.0000	4.4510 -4.9905	1.9580 -5.0000	-0.50864
9	-5.1586D-6 1.7059	-5.1586D-6 -3.9476	2.2563 -4.4744	2.2534 5.0000	4.4508 -3.2829	1.7995 -5.0000	-0.50864
10	-5.1562D-6 1.7032	-5.1562D-6 -3.9476	2.2543 -4.4744	2.2500 5.0000	4.4501 1.2093	1.8444 -5.0000	-0.50863
11	-5.1561D-6 1.7032	-5.1561D-6 -3.9476	2.2543 -4.4744	2.2500 5.0000	4.4501 1.2710	1.8719 -5.0000	-0.50863
12	-5.1561D-6 1.7032	-5.1561D-6 -3.9476	2.2543 -4.4744	2.2500 5.0000	4.4501 1.2710	1.8719 -5.0000	-0.50863

for satisfaction of the diode relationship is the only termination criterion, the agreement necessary in some diodes is too severe for others, causing prolonged analysis in many networks, while the node voltages do not change significantly.

Solution of Linear Network Equations

It has been shown that nonlinear dc network equations can be solved by repeated analysis of a sequence of linearized networks. In addition, the ac models which were described in a previous section have linear I-V relationships and therefore ac networks are linear. Nodal analysis is employed to generate linear equations which are solved by sparse matrix techniques.

The ac and dc linearized networks could be analyzed with the same set of computer instructions because the general procedure is identical for both cases. However, equations describing the small signal ac behavior have complex coefficients whereas the equations describing dc behavior have real coefficients. Since complex operations consume more time, separate solution subroutines are desirable for ac and dc analysis. Although separate subroutines are used, the techniques employed by them are the same and will not be distinguished in the discussion that follows.

To avoid certain manipulative difficulties, we assume each branch to be of the form shown in Fig. 2-13 and described by the relationship

$$I_K = Y_K (V_K - V_{SK}) + I_{SK} + Y_{CK} (V_J - V_{SJ})$$

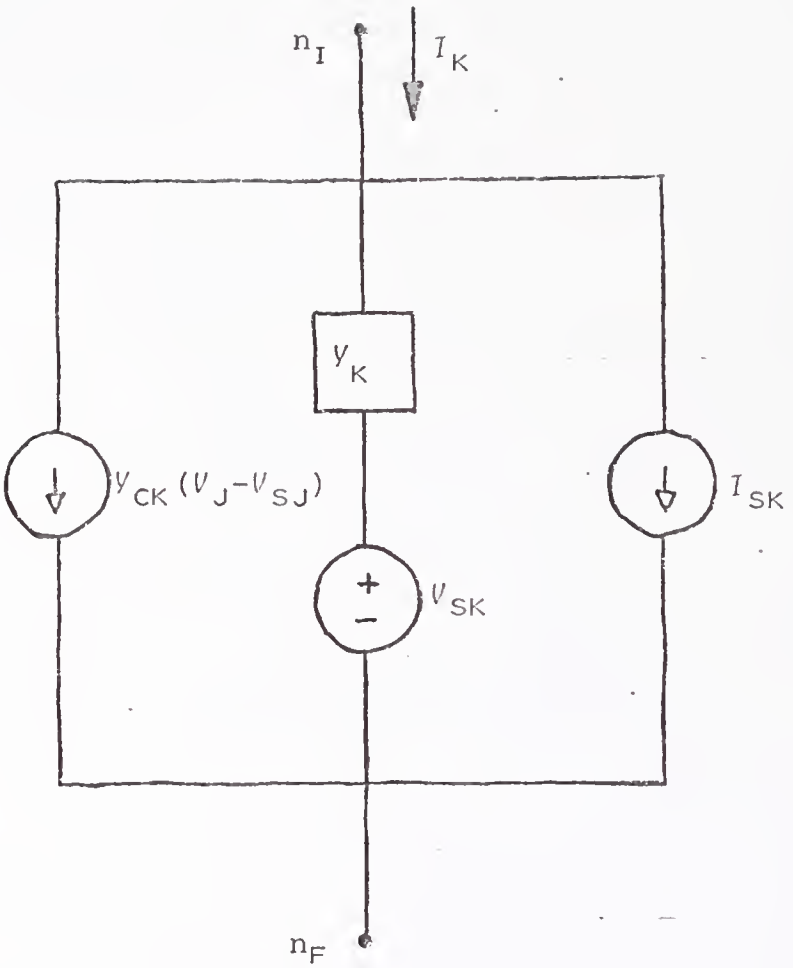


Fig. 2-13

General branch used in the design algorithm where (V_J-V_{SJ}) are voltages from another branch

where I_K is the branch current,

Y_K is the branch self-admittance,

V_K is the branch voltage,

V_{SK} is the source voltage,

I_{SK} is the source current,

Y_{CK} is a transadmittance,

V_J is the branch voltage of branch J, $J \neq K$, and

V_{SJ} is the source voltage of branch J.

All of the individual branch relations may be combined into the matrix relation

$$\underline{I}_B = \underline{Y}_B (\underline{V}_B - \underline{V}_{SB}) + \underline{I}_{SB} + \underline{Y}_C (\underline{V}_B - \underline{V}_{SB}).$$

We may rewrite this expression as

$$\underline{I}_B = (\underline{Y}_B + \underline{Y}_C)\underline{V}_B + \underline{I}_{SB} - (\underline{Y}_B + \underline{Y}_C)\underline{V}_{SB} \quad (2.9)$$

Kirchhoff's current law can be written in the matrix form

$$\underline{A}\underline{I}_B = \underline{0} \quad (2.10)$$

where $\underline{A}$ is the nodal incidence matrix. Multiplication of (2.9) by $\underline{A}$ and substitution of (2.10) yields the relationship

$$\underline{A}(\underline{Y}_B + \underline{Y}_C)\underline{V}_B = \underline{A}(\underline{Y}_B + \underline{Y}_C)\underline{V}_{SB} - \underline{A}\underline{I}_{SB}. \quad (2.11)$$

Kirchhoff's voltage law can be written in the matrix form

$$\underline{A}^T \underline{V}_N = \underline{V}_B \quad (2.12)$$

where $\underline{A}^T$ is the transpose of the nodal incidence matrix and $\underline{V}_N$ is the vector of node voltages. Substitution of (2.12)

into (2.11) yields the relationship

$$\underline{A}(\underline{y}_B + \underline{y}_C)\underline{A}^T\underline{V}_N = \underline{A}(\underline{y}_B + \underline{y}_C)\underline{V}_{SB} - \underline{A}\underline{I}_{SB}. \tag{2.13}$$

If we define

$$\underline{Y} = \underline{A}(\underline{y}_B + \underline{y}_C)\underline{A}^T$$

and

$$\underline{I} = \underline{A}[(\underline{y}_B + \underline{y}_C)\underline{V}_{SB} - \underline{I}_{SB}],$$

then (2.13) can be written

$$\underline{Y}\underline{V}_N = \underline{I}. \tag{2.14}$$

(Notice that $\underline{I}$ is not used to represent the identity matrix anywhere in this work.)

Significant importance can be attributed to the ease and efficiency with which the nodal equations are generated. The node numbers and general branch's component values are sufficient information to form the coefficients that constitute the nodal admittance matrix $\underline{Y}$ and current vector $\underline{I}$. No effort is required to consider tree branches, links, or any other topological designation, except nodes, in the selection of appropriate unknowns.

The accuracy of nodal analysis is limited by the accuracy of the models used to represent the physical elements and by the accuracy of the digital computer that is used to solve the nodal equations. The accuracy of the models can be improved to some extent if desired, but the accuracy of the computer

is fixed unless special programming techniques are employed which may severly reduce the efficiency of analysis. An example of the numerical limitations of a digital computer can be shown for a computer that is able to retain three digits of accuracy for each number it stores. If the network in Fig. 2-14 were to be analyzed, then the following equations must be solved:

$$\begin{bmatrix} 1000.01 & -1000 \\ -1000 & 1000.01 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix}.$$

But the digital computer for this example will retain only three digits of accuracy, so the equations will be stored as

$$\begin{bmatrix} 1000 & -1000 \\ -1000 & 1000 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix}.$$

The equations are no longer independent; the nodal admittance matrix is singular and the equations can not be solved. This problem can be avoided by applying limits to the network parameters such as defined values for a short circuit resistance and an open circuit resistance. The short circuit resistance must be greater than zero to avoid division by zero when using the general branch relationship. The range of values, specified in powers of ten, between a short and an open circuit must be limited to the number of digits of accuracy the computer is capable of retaining. These parameter boundaries do not reduce the efficiency or applicability of nodal analysis, but they do eliminate numerical inaccuracy of the type described above. It is possible that

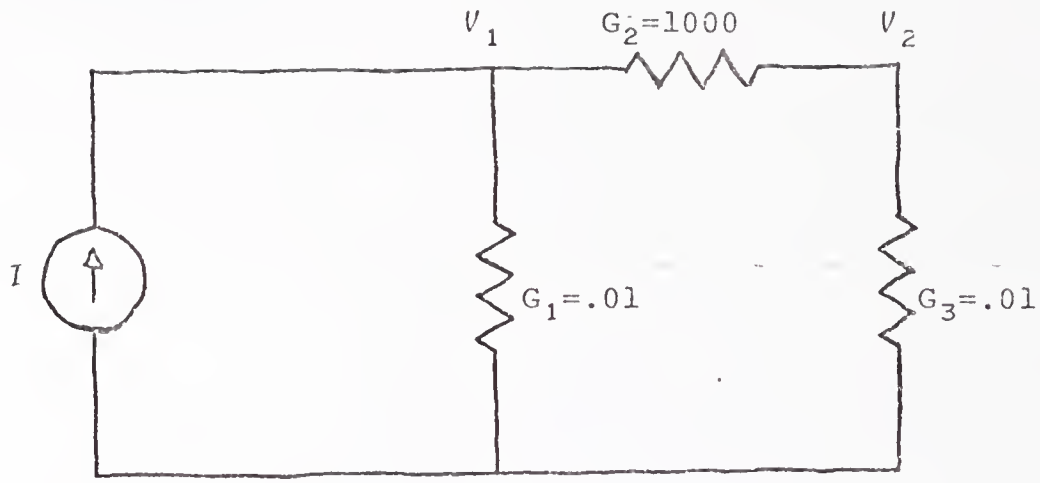


Fig. 2-14

Sample network to be analyzed

the boundaries will alter the value of a parameter, but the effect is usually negligible.

A possible method for the solution of (2.14) is to determine the inverse of the nodal admittance matrix such that

$$\underline{V}_N = \underline{Y}^{-1}\underline{I}. \tag{2.15}$$

The inversion of an n by n matrix requires on the order of n^3 operations where an operation is defined as a multiplication of a division. Additions and subtractions are not counted because the amount of time required to perform these operations is insignificant compared to the time required for multiplications and divisions. The matrix-vector multiplication in (2.15) requires n^2 operations. The total number of operations to solve (2.14) by matrix inversion is approximately

$$n^3+n^2.$$

LU factorization, a method that produces a lower triangular matrix, $\underline{L}$, and an upper triangular matrix, $\underline{U}$, which satisfy the relation

$$\underline{Y} = \underline{LU},$$

allows the solution of nodal equations in fewer operations than matrix inversion. This factorization is accomplished by a Crout reduction.[22] The $\underline{L}$ matrix consists of zeros in the upper right triangle and the $\underline{U}$ matrix has ones on the main diagonal and zeros in the lower left triangle. If only the

nonzero and nonunity elements of $\underline{L}$ and $\underline{U}$ are saved, the $\underline{U}$ matrix can be superimposed on the $\underline{L}$ matrix allowing the two matrices to be stored in the same area that was occupied by the original nodal admittance matrix. Once the LU factorization has been effected, the determination of $\underline{V}_N$ is reduced to a forward substitution and a back substitution.[23] The forward substitution,

$$\underline{LX} = \underline{I},$$

which determines an intermediate vector $\underline{X}$ is followed by a back substitution,

$$\underline{UV}_N = \underline{X},$$

to yield the desired node voltage vector. The LU factorization which eliminates the need for the matrix inversion requires approximately $n^3/3$ operations. The forward and back substitutions are performed in n^2 operations. Consequently, the total operation count is about

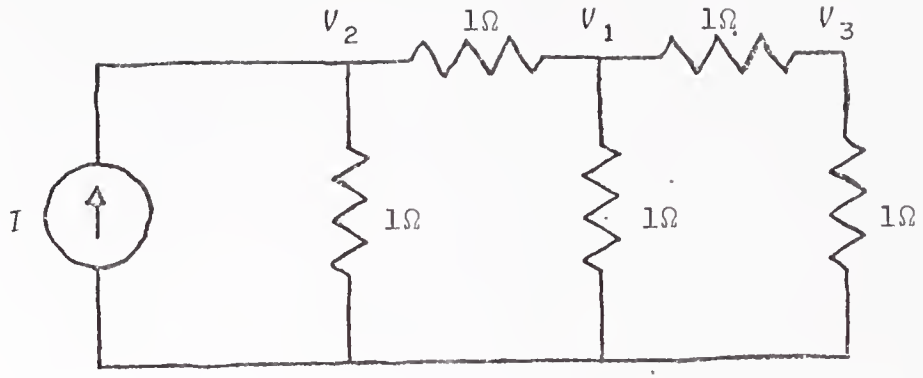
$$n^3/3+n^2$$

and the related execution time for the solution of (2.14) is significantly reduced by the use of LU factorization, especially for large n. If additional sets of network responses are to be determined for different excitations, only n^2 more operations need be performed.

The nodal admittance matrix of a typical electrical network consists of many zero coefficients. If an optimum order is determined for the nodal equations that causes this

sparsity to be maintained in the $\underline{L}$ and $\underline{U}$ matrices, many operations can be saved if operations involving multiplication by zero are avoided. An algorithm[24,25] is used that orders the nodal equations with the only objective being to retain as much sparsity as possible in the $\underline{L}$ and $\underline{U}$ matrices. This algorithm orders the nodal admittance matrix coefficients one row at a time while checking ahead in a simulated LU factorization to determine which order would introduce the fewest new nonzero terms. A simple example demonstrates the advantage of this ordering scheme for the network in Fig. 2-15(a). The nodal admittance matrix of this network, shown in Fig. 2-15(b), has two zeros but its LU factorization, shown in Fig. 2-15(c), has no zeros other than those which are fixed by the triangular form of the matrices because of the poor order of the equations. The reordered nodal admittance matrix, shown in Fig. 2-15(d), still has the two zeros and so does its LU factorization, shown in Fig. 2-15(e).

After the equations have been ordered, a program SPCRED[26] is used to generate three subroutines for matrices with complex (ac) coefficients and three subroutines for matrices with real (dc) coefficients. These subroutines are generated in machine code that is immediately executable. Each set of three subroutines has one that does LU factorization, a second for the forward and back substitutions to determine the solution to the original nodal equations, and a third does the forward and back substitutions to determine the solution to the equations generated by transposing the nodal



(a)

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

(b)

$$\begin{bmatrix} 3 & 0 & 0 \\ -1 & 5/3 & 0 \\ -1 & -1/3 & 24/15 \end{bmatrix} \quad \begin{bmatrix} 1 & -1/3 & -1/3 \\ 0 & 1 & -1/5 \\ 0 & 0 & 1 \end{bmatrix}$$

(c)

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

(d)

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ -1/2 & -1/2 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{bmatrix}$$

(e)

Fig. 2-15

(a) Simple network
(b) Nodal admittance matrix for network in (a)
(c) LU factorization of (b)
(d) Reordered nodal admittance matrix of network in (a)
(e) LU factorization of (d)

admittance matrix. This third set of equations will be used for sensitivity analysis, to be discussed in the next chapter. The second and third subroutines may be executed multiple times for different source vectors.

The benefits of these sparse matrix programs are a reduction in execution time and roundoff error because of fewer operations and reduced storage space because of fewer nonzero coefficients. Execution time is also decreased because general expressions with variable subscripts are not necessary. The operands associated with each operation are specified by exact addresses. Therefore, the time used to determine addresses of operands through arithmetic operations on subscripts is saved. Retaining only nonzero coefficients saves a significant amount of storage for large values of n. Another feature of the sparse matrix technique, the variable length of the generated subroutines, is dependent on the sparsity of the nodal admittance matrix. For a sparse matrix, the subroutines will have a length comparable to that of a subroutine written to solve general sets of equations; if the matrix is not sparse, the three subroutines could get prohibitively large. Few networks are described by equations with a dense nodal admittance matrix so this problem is seldom significant.

Since the analysis subroutine is to be used frequently during design, we require that the techniques which are employed be efficient, accurate, and suited to the needs of the design algorithm. Nodal analysis, as will be seen later,

meets these requirements when used in conjunction with the adjoint network[6] method for sensitivity calculation.

CHAPTER 3

PERFORMANCE FUNCTIONS AND MINIMIZATION METHODS

Automated network design is based upon the minimization of a performance function that suitably reflects the design specifications. In this chapter we discuss the least pth performance function, the conjugate gradients minimization scheme, and the efficient and accurate determination of the parameter gradients by the adjoint network method.

Evaluation of Network Performance

A performance function assigns a scalar measure to the ability of a network to meet design specifications. In particular, the performance function represents a comparison of a network's actual responses with the desired responses. For the case of linear integrated circuit design, this functional relationship must account for both ac and dc performance. Flexibility can be built into the function through weights which can emphasize the more important parts of the responses.

Since ac and dc responses are determined separately, the ac and dc portions of the performance function are determined in conjunction with the appropriate analyses and the resulting error values are summed:

$$E = E_{AC} + E_{DC}. \tag{3.1}$$

Because of the form of the performance function, a compromise can be reached between the ac and dc design specifications. A desirable performance function is based on the difference, or response error, between the actual and the desired responses. All components of the performance function should be nonnegative so that some components do not cancel others. Raising each response error to an even integer power guarantees that no negative components exist. A suitable ac performance function is

$$E_{AC} = \sum_{J=1}^{n_{AC}} \int_{\omega_I}^{\omega_F} 1/p[W_J(\omega)|X_J(\omega)/\hat{X}_J(\omega)-1|]^p \, d\omega \tag{3.2}$$

where n_{AC} is the number of ac ports,

ω_I is the initial angular frequency,

ω_F is the final angular frequency,

p is a positive even integer,

W_J is the nonnegative weight at port J,

$X_J(\omega)$ is the actual voltage or current response, and

$\hat{X}_J(\omega)$ is the desired voltage or current response.

The magnitude of the response error is raised to the pth power to avoid a complex performance function as the responses are complex phasors. A suitable dc performance function is

$$E_{DC} = \sum_{J=n_{AC}+1}^{n_{AC}+n_{DC}} \int_{T_I}^{T_F} 1/p\{W_J(T)[X_J(T)/\hat{X}_J(T)-1]\}^p \, dT \tag{3.3}$$

where n_{DC} is the number of dc ports,

T_I is the initial temperature,

T_F is the final temperature,

$X_J(T)$ is the actual voltage or current response, and

$\hat{X}_J(T)$ is the desired voltage or current response.

The dependence of the performance function on the network parameters is implicit in that the actual responses are dependent on the network parameters. To show this dependence, (3.1) can be written as

$$E(\underline{r},\omega,T) = E_{AC}(\underline{r},\omega)+E_{DC}(\underline{r},T)$$

where $\underline{r}$ is the vector of network parameters. The dependence of the ac performance function on temperature is ignored because it is assumed that the dc specifications will include temperature desensitization requirements, thus minimizing the effect of temperature on the network.

The flexibility of this performance function is not limited to the use of weights. The power, p, to which the normalized response difference is raised, may be assigned any positive even integer value. For p=2, the function represents a least-squares criteria. As p is increased, the larger response errors experience increased emphasis, and as p becomes very large, the performance function approaches a minimax criterion[27] of the form

$$E = \max_J [W_J |X_J/\hat{X}_J - 1|].$$

The normalization included in the performance function equalizes the effect of each response on the total error. For example, the contribution from a response of one milliamp

is comparable to the contribution from a response of one volt. Therefore, responses with smaller magnitudes will not be neglected. To avoid division by zero, desired responses of zero magnitude are not normalized and the quantity that is raised to the pth power is of the form

$$W_I|X_I|. \tag{3.4}$$

This alternate form has been found to be adequate for the few occasions where it was necessary.

Frequently, only the magnitude of an ac response is specified and the phase is ignored. For these cases, an alternate form of the ac portion of the performance function is used:

$$E_{AC} = \sum_{J=1}^{n_{AC}} \int_{\omega_I}^{\omega_F} 1/p\{W_J(\omega)[|X_J(\omega)|/|\hat{X}_J(\omega)|-1]\}^p \, d\omega. \tag{3.5}$$

As before, if a magnitude of zero is desired, the alternate form of (3.4) is used.

The frequency response of a network is often specified over several decades of frequency. So that errors at the lower frequencies are weighted the same as those at the higher frequencies, a logarithmic scaling of frequency can be employed. The performance functions (3.2) and (3.5) become

$$E_{AC} = \sum_{J=1}^{n_{AC}} \int_{u_I}^{u_F} 1/p\{W_J(u)|X_J(u)/\hat{X}_J(u)-1|\}^p \, du$$

and

$$E_{AC} = \sum_{J=1}^{n_{AC}} \int_{u_I}^{u_F} 1/p\{W_J(u)[|X_J(u)|/|\hat{X}_J(u)|-1]\}^p \, du, \qquad (3.6)$$

respectively, where

$$u = \ell og\ \omega.$$

The use of large values of p, $p>6$, has caused numerical overflows and underflows for problems that have responses that are not close to the desired responses at the determination of the initial error. This problem can be avoided by using a small value of p to obtain a partial design and then performing the remainder of the design with a larger p.

Minimization of the Performance Function

Systematic adjustment of the network parameters by a sophisticated multidimensional search algorithm minimizes the performance function and yields a network whose responses approximate the desired responses. A desirable feature of such a minimization algorithm is that it require as few function evaluations as possible.

Many algorithms are available that effectively minimize general functions. Such algorithms are separable into two types: gradient methods, i.e., those that use partial derivations to locate a minimum,[28-30] and direct methods.[31-33] A comparison of the number of function evaluations needed to minimize some test functions gives an indication as to whether or not it is desirable to use a minimization algorithm that requires partial derivatives. A comparison has been

made[34] that shows the two types of methods are comparable if the determination of each parameter gradient requires a function evaluation. Therefore, n+1 function evaluations are assumed necessary to evaluate a function of n variables and determine the associated first partial derivatives. If a method was available that required fewer function evaluations to obtain the parameter gradients, the gradient methods would be superior.

An efficient method of determining partial derivatives has been developed[6] and through use of this method, all first partial derivatives of the performance function with respect to the network parameters can be determined with one function evaluation and the simultaneous analysis of a related network. This method will be discussed in the next section.

Because partial derivatives can be computed so efficiently, gradient type minimization methods are employed. In particular, two such methods are implemented: conjugate gradients[30] and steepest descent.[35] Other methods, such as the methods of Fletcher and Powell,[29] Jacobson and Oksman,[36] and Fletcher,[28] are available and have been found to be more efficient than either conjugate gradients or steepest descent when used to locate an unconstrained minimum. These superior methods are not used here because of the presence of constraints on the network parameters. Although all of the methods which have been mentioned are designed for minimization of functions of unconstrained parameters, conjugate gradients and steepest descent methods adjust each parameter along a direction that is mostly independent of the adjustment of the other parameters.

This independence prevents an accumulation of erroneous data which is due to parameter changes that were suggested at previous iterations and had been ignored in order to keep parameters within boundaries. For instance, the method of Fletcher and Powell updates the direction vector with a vector that is dependent upon all parameter gradients and all parameter changes. Once a parameter encounters a boundary, all components of the direction vector become distorted by the inconsistancy caused by the constraint. The method of steepest descent adjusts each parameter in the negative gradient direction at all times and thus allows no interaction between parameters. The method of conjugate gradients adjusts parameters along a direction which is the negative of the sum of the present gradient component and a portion of recent gradient components; the interaction between parameters is limited to the determination of a number that is used to scale the past gradient information before it is combined with the current gradient. The scale factor is dependent on the magnitudes of the gradients and not on the actual parameter changes, thus avoiding introduction of significant errors when a parameter encounters a constraint.

The network parameter values must be forced to remain within specified boundaries in order to consistantly produce network designs that can actually be realized with a predefined fabrication procedure. In particular, if the beta of a transistor was an unconstrained parameter and its value was specified to be a million, the design would have to be discarded because the large beta was not obtainable.

The method of steepest descent is based on the adjustment of parameters in the direction of the negative gradient:

$$\underline{r}' = \underline{r} - \alpha \underline{g}$$

where $\underline{r}'$ is the vector of new parameter values,

$\underline{r}$ is the vector of old parameter values,

α is a variable of the minimization algorithm,

$\underline{g}$ is the parameter gradient vector.

This method is very effective when the initial parameter values are poorly chosen, but is often ineffective when searching through a narrow valley of the performance function. Information concerning previous parameter adjustments or previous gradients is not retained and the method has a tendency to oscillate instead of following the contours of the performance function.

The method of conjugate gradients is different from the method of steepest descent in that the direction of search is a combination of past and present gradient information. The direction vector along which the parameters are varied is initially the negative of the gradient vector and is modified at each following iteration by the formula

$$\underline{d}_{I+1} = -\underline{g}_{I+1} + \beta_I \underline{d}_I$$

where

$$\beta_I = \underline{g}_{I+1}^T \underline{g}_{I+1} / (\underline{g}_I^T \underline{g}_I),$$

$\underline{d}_{I+1}$ is the direction vector at iteration I+1 and $\underline{d}_I$ is the

direction vector at iteration I. The vectors $\underline{d}_1$, $\underline{d}_2$, ..., $\underline{d}_n$ are $\underline{A}$-conjugate where the performance function is assumed to be of the form

$$E(\underline{r}) = E(\hat{\underline{r}}) + \tfrac{1}{2}(\underline{r}-\hat{\underline{r}})^T \underline{A}(\underline{r}-\hat{\underline{r}}) + \ldots$$

where $\hat{\underline{r}}$ is the parameter vector at the minimum of the performance function and $\underline{A}$ is a symmetric positive definite matrix. The minimum of such a quadratic function of n variables will be obtained in n+1 iterations if no constraints are present. Typically, the performance function we use is not quadratic and requires more than n+1 iterations. It has been found that after n+1 iterations, if the direction vector is again set equal to the negative of the gradient, faster convergence is achieved. There are three major reasons for this procedure: the function is generally not quadratic and thus the direction vectors are not $\underline{A}$-conjugate, erroneous data accumulate due to constraints and the accumulation of rounding errors. The direction vector is also reset to the negative gradient if the value of the performance function is not reduced a specified percentage.

The minimization strategy used in this design algorithm is based on the conversion of a multidimensional search into a one-dimensional search. The parameters are adjusted along a direction vector, obtained through one of the above methods, until the value of the performance function no longer decreases, thus bounding the location of the minimum along this direction. Cubic interpolation is then used to obtain the minimum from

whence a new direction is generated. This process is repeated until one of four criteria is satisfied: the value of the performance function is less than desired, the value of the performance function is changing less than desired, an excessive number of iterations have been executed, or an excessive amount of time has been consumed.

A change of variable is made to avoid the elongated contours of the performance function that arise when the designable parameters have values that differ by many orders of magnitude. In particular we use

$$Y_I = \ell n(r_I), \quad I=1,2,3,\ldots,n$$

where Y_I is the new parameter and r_I is the old parameter. The new gradients can be determined directly from the old gradients by the relationship

$$\partial E/\partial Y_I = \partial E/\partial r_I [\partial Y_I/\partial r_I]^{-1} = [1/r_I]^{-1}\partial E/\partial r_I = r_I \partial E/\partial r_I .$$

This transformation has improved convergence characteristics and allows all parameters to be varied similiarly during the minimization of the performance function.

Efficient Determination of Parameter Gradients

As indicated earlier, we require knowledge of the first-order partial derivatives of the performance function with respect to the network parameters. A method proposed by Director and Rohrer[6] for evaluation of partial derivatives through the use of an auxiliary network related to the

original network, has been shown to be very efficient and accurate when used for network optimization.

All the elements (resistors, capacitors, diodes, and transistors) of a monolithic integrated circuit are ultimately specified in terms of planar dimensions (the masks) and a series of diffusions, oxidations, etc. (the processing steps). Typically for economic reasons the processing, or impurity profiles, are fixed and the mask geometries varied to realize the designable elements. The only constraint on the geometries is dictated by the resolution capability of the processing. In particular, a resistor with resistance R is realized as a base diffusion whose length to width ratio is such that

$$R = (\ell/w)\rho_S[1+\alpha_R(T-T_0)]$$

where the sheet resistivity and the temperature coefficient are determined by the processing. Thus we can consider R to be the designable electrical parameter of a resistor and length to width ratio the corresponding designable mask parameter. A capacitor is realized as a reverse biased base-collector or emitter-base junction or a metal-oxide-silicon structure whose area is such that

$$C = \Theta A[1+\alpha_C(T-T_0)]$$

where Θ and the temperature coefficient are processing dependent constants. The designable electrical parameter of a capacitor is C while the corresponding designable mask parameter is area. A junction diode has area such that

$$I_S = J_S A[1+\alpha_D(T-T_0)]$$

where the current density and the temperature coefficient are determined by the processing. Thus we can consider I_S the designable electrical parameter of a diode and area the corresponding designable mask parameter. DC bias parameters will be of importance in the sequel; for the diode, such a parameter is the junction voltage.

It is not so apparent what the designable electrical parameters for a transistor are, but a suitable choice for designable mask parameters is the emitter area A_E and perimeter P_E. Actually any two independent specifications of length and width of the emitter suffice; these two prove most convenient. Once the emitter geometry is given, the base and collector geometries are determined by the masking tolerances. Both β_F and ω_T are completely specified for a given A_E, P_E, I_C, and the processing parameters a_I and b_J:[37]

$$1/\beta_F = a_1+a_2(P_E/A_E)+a_3(I_C/A_E)+a_4(1/I_C)$$

$$1/\omega_T = b_1+b_2(A_E/I_C)+b_3A_E.$$

It is apparent that the designable electrical parameters for a transistor are $1/\beta_F$, $1/\omega_T$, and at least one bias parameter to specify I_C. For the linearized Ebers-Moll ac model, we choose the two diode voltages at T_0, V_{D1} and V_{D2}. For the hybrid-pi ac model, we choose the magnitude of the collector current at T_0, $|I_C|$, which is a component of g_M.

An auxiliary network will be defined that is topologically identical to the original network. It will be shown that this auxiliary network, called the adjoint network, consists of elements that are related to the elements of the original network and is excited by error signals determined from responses of the original network. The nodal admittance matrix of the adjoint network has been shown to be the transpose of the nodal admittance matrix of the original network.[23] Since the nodal admittance matrix is generated for the analysis of the original network, the analysis of the adjoint network reduces to a forward and back substitution using the subroutine described in the previous chapter to solve the equations for the transposed nodal admittance matrix. This feature of the adjoint network in combination with the use of nodal analysis makes this method of gradient evaluation very efficient.

Tellegen's theorem[38] which is used to develop the relations needed to calculate the first-order partial derivatives can be written

$$\sum_{I=1}^{n_B} V_I \Phi_I = 0 \tag{3.7}$$

and

$$\sum_{I=1}^{n_B} I_I \Psi_I = 0 \tag{3.8}$$

where n_B is the number of branches in the original network,

V_I is the voltage of branch I in the original network,

Φ_I is the current of branch I in the adjoint network,

I_I is the current of branch I in the original network,

Ψ_I is the voltage of branch I in the adjoint network.

If the parameters of the original network are perturbed, Tellegen's theorem is still valid:

$$\sum_{I=1}^{n_B} (V_I+\Delta V_I)\Phi_I = 0 \qquad (3.9)$$

and

$$\sum_{I=1}^{n_B} (I_I+\Delta I_I)\Psi_I = 0. \qquad (3.10)$$

If the difference between (3.7) and (3.8) is subtracted from the difference between (3.9) and (3.10), the following relation is obtained:

$$\sum_{I=1}^{n_B} (\Delta V_I \Phi_I - \Delta I_I \Psi_I) = 0. \qquad (3.11)$$

Substitution of the branch relationships for each element type with perturbed parameters into (3.11) will suggest gradient and adjoint element relationships.

The branch relations that are considered in the ac and dc cases for a resistor are

$$V_R = RI_R$$

and

$$I_G = GV_G$$

where G=1/R. The purpose of the two branch representations for a resistor will become clear in the next chapter. Perturbing the parameters and substituting the perturbed branch

relations into (3.11) yields[7]

$$\Delta V_R \Phi_R - \Delta I_R \Psi_R = (R\Phi_R - \Psi_R)\Delta I_R + I_R \Phi_R \Delta R \tag{3.12}$$

for resistances and

$$\Delta V_G \Phi_G - \Delta I_G \Psi_G = (\Phi_G - G\Psi_G)\Delta V_G - V_G \Psi_G \Delta G. \tag{3.13}$$

for conductances. Two branch relations are also used for both capacitances and inductances in the ac case to yield

$$\Delta V_C \Phi_C - \Delta I_C \Psi_C = (\Phi_C - j\omega C\Psi_C)\Delta V_C - j\omega V_C \Psi_C \Delta C \tag{3.14}$$

for capacitances,

$$\Delta V_S \Phi_S - \Delta I_S \Psi_S = [1/(j\omega)S\Phi_S - \Psi_S]\Delta I_S + 1/(j\omega) I_S \Phi_S \Delta S \tag{3.15}$$

for elastances where S=1/C,

$$\Delta V_L \Phi_L - \Delta I_L \Psi_L = (j\omega L\Phi_L - \Psi_L)\Delta I_L + j\omega I_L \Phi_L \Delta L \tag{3.16}$$

for conductances, and

$$\Delta V_\Gamma \Phi_\Gamma - \Delta I_\Gamma \Psi_\Gamma = [\Phi_\Gamma - 1/(j\omega)\Gamma\Psi_\Gamma]\Delta V_\Gamma - 1/(j\omega) V_\Gamma \Psi_\Gamma \Delta\Gamma \tag{3.17}$$

for reciprocal inductances where Γ=1/L. In each of the above relations, dependence on ΔV and ΔI can be eliminated by choosing the adjoint network element to be identical to the element in the original network.

A voltage-controlled current source is represented by the two-branch model shown in Fig. 3-1 and the branch relationships

$$I_{VDI} = g_M V_{VCI}$$

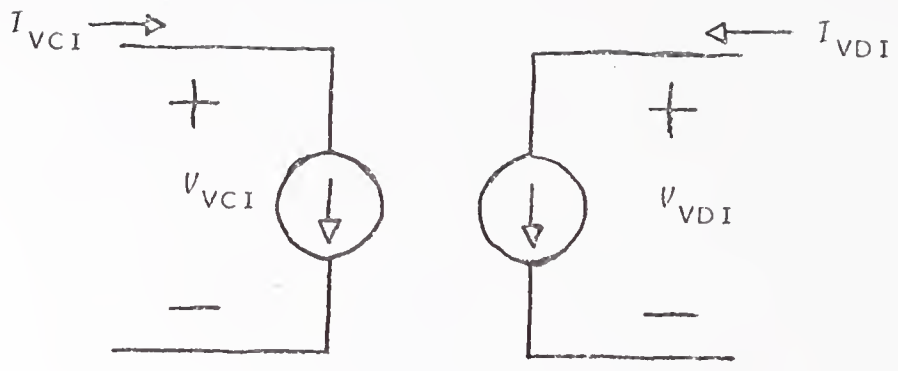


Fig. 3-1

Two-branch model for a voltage-controlled current source where the subscripts VCI and VDI indicate the controlling and dependent branches in the original network, respectively

and

$$I_{VCI} = 0$$

for both ac and dc cases. Substituting these expressions into (3.11) yields

$$\Delta V_{VCI}\Phi_{VCI} - \Delta I_{VCI}\Psi_{VCI} + \Delta V_{VDI}\Phi_{VDI} - \Delta I_{VDI}\Psi_{VDI}$$

$$= (\Phi_{VCI} - g_M\Psi_{VDI})\Delta V_{VCI} + \Phi_{VDI}\Delta V_{VDI} - V_{VCI}\Psi_{VDI}\Delta g_M. \qquad (3.18)$$

The adjoint network element is chosen to be a voltage-controlled current source with the controlling and dependent branches reversing rolls as indicated in the branch relations

$$\Phi_{VCI} = g_M\Psi_{VDI}$$

and

$$\Phi_{VDI} = 0.$$

The dc diode relation with the value of the saturation current perturbed can be written

$$V_D + \Delta V_D = V_X \ell n[(I_D + \Delta I_D)/(I_S + \Delta I_S) + 1] + (I_D + \Delta I_D)R_S.$$

Substituting into this expression the first-order terms of a Taylor's series for the logarithm yields

$$V_D + \Delta V_D = V_X\{\ell n(I_D/I_S + 1) + \Delta I_D/(I_D + I_S)$$

$$- I_D\Delta I_S/[I_S(I_D + I_S)]\} + (I_D + \Delta I_D)R_S.$$

Therefore,

$$\Delta V_D = V_X \Delta I_D/(I_D+I_S) - V_X I_D \Delta I_S/[I_S(I_D+I_S)] + R_S \Delta I_D.$$

Substituting this expression into (3.11) yields

$$\Delta V_D \Phi_D - \Delta I_D \Psi_D = \{V_X \Delta I_D/(I_D+I_S) - V_X I_D \Delta I_S/[I_S(I_D+I_S)] + R_S \Delta I_D\} \Phi_D - \Delta I_D \Psi_D$$

which can be rearranged into the form

$$\{[V_X/(I_D+I_S)+R_S]\Phi_D - \Psi_D\}\Delta I_D - V_X I_D \Phi_D/[I_S(I_D+I_S)]\Delta I_S. \qquad (3.19)$$

At this point, the goal is to eliminate dependence of this expression on ΔI_D by choosing the adjoint branch relationship to be

$$\Psi_D = [V_X/(I_D+I_S)+R_S]\Phi_D.$$

Thus the adjoint network element of the dc diode relationship is a linear resistor whose value is the resistance of the linearized diode model at its operating point plus the series resistance. The remainder of the expression will be used to determine the gradient later in the section.

For ac analysis a diode is represented by the combination of two resistors and two capacitors shown in Fig. 3-2. We will omit consideration of the series resistance for this derivation as it can be represented as an identical resistance in the adjoint network and it is independent of the saturation current and the junction voltage which are the parameters of interest in the diode. The branch relationship of the

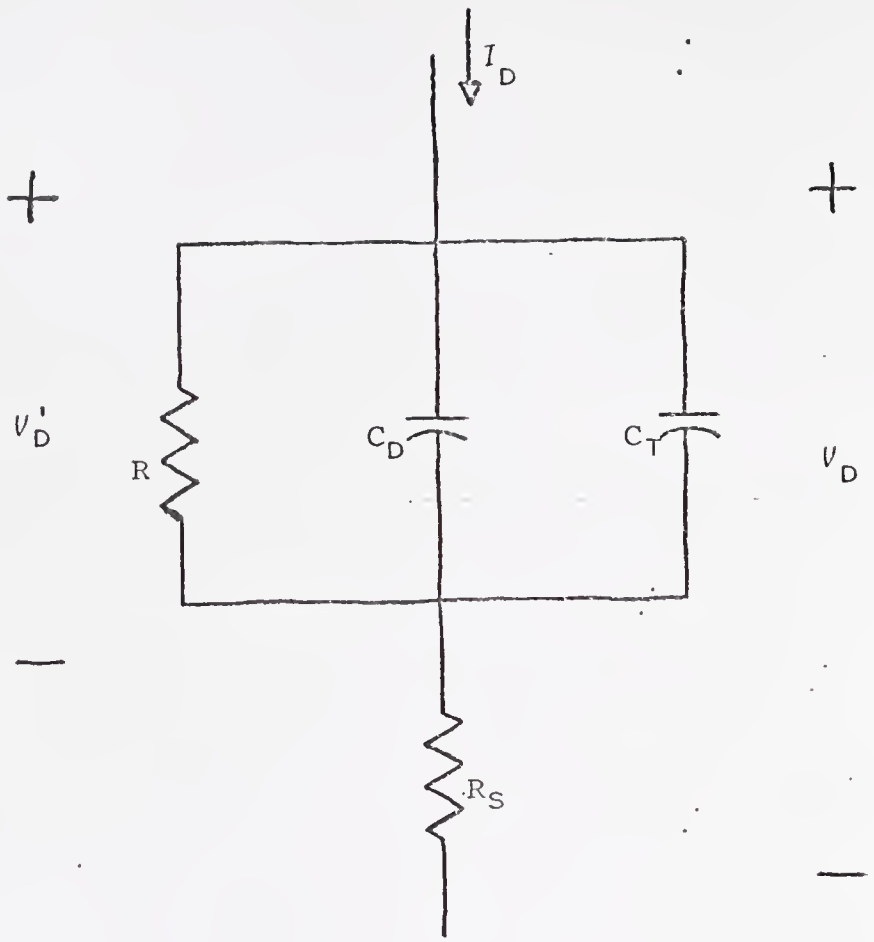


Fig. 3-2

Model used to represent a diode for ac analysis

remainder of the ac diode model with the value of the saturation current and the junction voltage perturbed is

$$I_D+\Delta I_D = \{(I_S+\Delta I_S)/V_X exp[(V_D+\Delta V_D)/V_X][1+j\omega/\omega_X]$$

$$+j\omega C_0[1-(V_D+\Delta V_D)/V_{D0}]**(1/m)\}(V_D'+\Delta V_D')$$

where V_D' is the ac diode voltage less the ac voltage dropped across the series resistance. If we substitute the linear terms of a Taylor's series where necessary, subtract the unperturbed expression from the perturbed form, and neglect second-order terms, we obtain the following expression:

$$\Delta I_D = \{I_S/V_X^2 exp(V_D/V_X)[1+j\omega/\omega_X]$$

$$-j\omega C_0/[mV_{D0}(1-V_D/V_{D0})**(1+1/m)]\ V_D'\Delta V_D$$

$$+1/V_X exp(V_D/V_X)[1+j\omega/\omega_X]V_D'\Delta I_S$$

$$+\{I_S/V_X exp(V_D/V_X)[1+j\omega/\omega_X]$$

$$+j\omega C_0/(1-V_D/V_{D0})**(1/m)\}\Delta V_D'.$$

Substitution of this expression into (3.11) yields the following:

$$\Delta V_D'\Phi_D - \Delta I_D\Psi_D' = -\{I_S/V_X^2 exp(V_D/V_X)[1+j\omega/\omega_X]$$

$$-j\omega C_0/[mV_{D0}(1-V_D/V_{D0})^{**}(1+1/m)]\}V_D'\Psi_D'\Delta V_D$$

$$-1/V_X exp(V_D/V_X)[1+j\omega/\omega_X]V_D'\Psi_D'\Delta I_S$$

$$+\{\Phi_D - I_S/V_X exp(V_D/V_X)[1+j\omega/\omega_X]\Psi_D'$$

$$-j\omega C_0/(1-V_D/V_{D0})^{**}(1/m)\Psi_D'\}\Delta V_D' \qquad (3.20)$$

where Ψ_D' is the ac diode voltage in the adjoint network less the voltage dropped across the series resistance. The adjoint network ac branch relationship of a diode is therefore

$$\Phi_D = \{I_S/V_X exp(V_D/V_X)[1+j\omega/\omega_X] \cdot$$

$$+j\omega C_0/(1-V_D/V_{D0})^{**}(1/m)\}\Psi_D'$$

which is identical to the ac diode branch relationship in the original network. (Remember that the series resistance was set aside as an individual element and is still present.)

For the determination of the adjoint network elements and gradient expressions of the transistor models, we will consider only the intrinsic portions of the models. The three bulk resistors that have been designated the extrinsic elements have corresponding resistors as adjoint elements. Since these resistors are not dependent upon the parameters of the transistor, their omission from gradient consideration is justified.

First we consider the dc Ebers-Moll model for an NPN transistor which can be expressed by the following equations:

$$I_E = -I_{ES}[exp(-V_{EB}/V_{X1})-1]+\alpha_R I_{CS}[exp(-V_{CB}/V_{X2})-1] \quad (3.21)$$

$$I_C = \alpha_F I_{ES}[exp(-V_{EB}/V_{X1})-1]-I_{CS}[exp(-V_{CB}/V_{X2})-1] \quad (3.22)$$

where

$$V_{X1} = n_1 kT/q$$

and

$$V_{X2} = n_2 kT/q.$$

For this model, we consider the effect of the perturbation of only $1/\beta_F$ which can be determined by perturbing the forward alpha and transforming the result. The perturbation of the forward alpha and use of the first-order terms of the Taylor's series for the exponential yields

$$I_E+\Delta I_E = -I_{ES}[exp(-V_{EB}/V_{X1})-1-exp(-V_{EB}/V_{X1})\Delta V_{EB}/V_{X1}]$$

$$+\alpha_R I_{CS}[exp(-V_{CB}/V_{X2})-1-exp(-V_{CB}/V_{X2})\Delta V_{CB}/V_{X2}] \quad (3.23)$$

and

$$I_C+\Delta I_C = (\alpha_F+\Delta\alpha_F)I_{ES}[exp(-V_{EB}/V_{X1})-1-exp(-V_{EB}/V_{X1})\Delta V_{EB}/V_{X1}]$$

$$-I_{CS}[exp(-V_{CB}/V_{X2})-1-exp(-V_{CB}/V_{X2})\Delta V_{CB}/V_{X2}]. \quad (3.24)$$

If we subtract (3.21) from (3.23), subtract (3.22) from (3.24),

and neglect second-order terms, we arrive at the following expressions:

$$\Delta I_E = I_{ES}/V_{X1}[exp(-V_{EB}/V_{X1})]\Delta V_{EB} - \alpha_R (I_{CS}/V_{X2})[exp(-V_{CB}/V_{X2})]\Delta V_{CB}$$

and

$$\Delta I_C = -\alpha_F (I_{ES}/V_{X1})[exp(-V_{EB}/V_{X1})]\Delta V_{EB} + I_{ES}[exp(-V_{EB}/V_{X1})-1]\Delta\alpha_F$$

$$+I_{CS}/V_{X2}[exp(-V_{CB}/V_{X2})]\Delta V_{CB}.$$

Substituting these relations into (3.11) for the two branches yields

$$\Delta V_{EB}\Phi_E - \Delta I_E \Psi_{EB} + \Delta V_{CB}\Phi_C - \Delta I_C \Psi_{CB}$$

$$= \Delta V_{EB}\Phi_E - \{I_{ES}/V_{X1}[exp(-V_{EB}/V_{X1})]\Delta V_{EB}$$

$$-\alpha_R I_{CS}/V_{X2}[exp(-V_{CB}/V_{X2})]\Delta V_{CB}\}\Psi_{EB}$$

$$+\Delta V_{CB}\Phi_C - \{-\alpha_F I_{ES}/V_{X1}[exp(-V_{EB}/V_{X1})]\Delta V_{EB}$$

$$+I_{ES}[exp(-V_{EB}/V_{X1})-1]\Delta\alpha_F$$

$$+I_{CS}/V_{X2}[exp(-V_{CB}/V_{X2})]\Delta V_{CB}\}\Psi_{CB}.$$

Combining like terms yields

$$\{\Phi_E - I_{ES}/V_{X1}[exp(-V_{EB}/V_{X1})]\Psi_{EB}$$

$$+\alpha_F I_{ES}/V_{X1}[exp(-V_{EB}/V_{X1})]\Psi_{CB}\}\Delta V_{EB}$$

$$+\{\Phi_C + \alpha_R I_{CS}/V_{X2}[exp(-V_{CB}/V_{X2})]\Psi_{EB}$$

$$-I_{CS}/V_{X2}[exp(-V_{CB}/V_{X2})]\Psi_{CB}\}\Delta V_{CB}$$

$$-I_{ES}[exp(-V_{EB}/V_{X1})-1]\Psi_{CB}\Delta\alpha_F. \qquad (3.25)$$

We eliminate the terms containing ΔV_{EB} and ΔV_{CB} by defining the adjoint network relationship to be

$$\begin{bmatrix}\Phi_E\\ \\ \Phi_C\end{bmatrix} = \underline{Y} \begin{bmatrix}\Psi_{EB}\\ \\ \Psi_{CB}\end{bmatrix}$$

where

$$\underline{Y} = \begin{bmatrix} I_{ES}/V_{X1}exp(-V_{EB}/V_{X1}) & -\alpha_F I_{ES}/V_{X1}exp(-V_{EB}/V_{X1}) \\ \\ -\alpha_R I_{CS}/V_{X2}exp(-V_{CB}/V_{X2}) & I_{CS}/V_{X2}exp(-V_{CB}/V_{X2}) \end{bmatrix}$$

which are the expressions describing a linear two-port. The y-parameter matrix of this two-port is the transpose of the y-parameter matrix of the final linearization of the transistor model in the original network. We may transform the remaining term of (3.25) to be dependent on $\Delta(1/\beta_F)$ instead of $\Delta\alpha_F$ by realizing that

$$\partial\alpha_F/\partial(1/\beta_F) = \partial(1+1/\beta_F)^{-1}/\partial(1/\beta_F) = -\alpha_F^2.$$

Therefore, (3.25) becomes

$$\alpha_F^2 I_{ES}[exp(-V_{EB}/V_{X1})-1]\Psi_{CB}\Delta(1/\beta_F). \qquad (3.26)$$

The ac transistor models are linear and may be represented by the two-port expression

$$\begin{bmatrix} I_E \\ I_C \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_{EB} \\ V_{CB} \end{bmatrix} . \qquad (3.27)$$

We will now perturb the y-parameters and determine an expression that can be used to specify the adjoint network representation and gradient information for each transistor model individually. Perturbing the y-parameters, subtracting (3.27), and neglecting second-order terms yields

$$\Delta I_E = y_{11}\Delta V_{EB} + V_{EB}\Delta y_{11} + y_{12}\Delta V_{CB} + V_{CB}\Delta y_{12}$$

and

$$\Delta I_C = y_{21}\Delta V_{EB} + V_{EB}\Delta y_{21} + y_{22}\Delta V_{CB} + V_{CB}\Delta y_{22}.$$

Substitution of these expressions into (3.11) yields

$$\Delta V_{EB}\Phi_E - \Delta I_E \Psi_{EB} + \Delta V_{CB}\Phi_C - \Delta I_C \Psi_{CB}$$

$$= \Delta V_{EB}\Phi_E - [Y_{11}\Delta V_{EB} + V_{EB}\Delta Y_{11}$$

$$+ Y_{12}\Delta V_{CB} + V_{CB}\Delta Y_{12}]\Psi_{EB} + \Delta V_{CB}\Phi_C$$

$$- [Y_{21}\Delta V_{EB} + V_{EB}\Delta Y_{21} + Y_{22}\Delta V_{CB} + V_{CB}\Delta Y_{22}]\Psi_{CB}$$

which can be rearranged into the form

$$[\Phi_E - Y_{11}\Psi_{EB} - Y_{21}\Psi_{CB}]\Delta V_{EB}$$

$$+ [\Phi_C - Y_{12}\Psi_{EB} - Y_{22}\Psi_{CB}]\Delta V_{CB}$$

$$- [V_{EB}\Psi_{EB}\Delta Y_{11} + V_{CB}\Psi_{EB}\Delta Y_{12}$$

$$+ V_{EB}\Psi_{CB}\Delta Y_{21} + V_{CB}\Psi_{CB}\Delta Y_{22}]. \qquad (3.28)$$

The adjoint network representation for a network represented by Y-parameters is a network with the transpose of the original Y-parameter matrix:

$$\begin{bmatrix} \Phi_E \\ \Phi_C \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{21} \\ Y_{12} & Y_{22} \end{bmatrix} \begin{bmatrix} \Psi_{EB} \\ \Psi_{CB} \end{bmatrix}.$$

The linearized Ebers-Moll model can be described by the following Y-parameters:

$$y_{11} = (\beta_F+1)[(1/\beta_F+j\omega/\omega_{T1})\alpha_F I_{ES}/V_{X1}\exp(V_{D1}/V_{X1})$$

$$+j\omega C_{01}/(1-V_{D1}/V_{01})^{**}(1/m_1)],$$

$$y_{12} = -\alpha_R I_{CS}/V_{X2}\exp(V_{D2}/V_{X2}),$$

$$y_{21} = -\alpha_F I_{ES}/V_{X1}\exp(V_{D1}/V_{X1}),$$

and

$$y_{22} = (\beta_R+1)[(1/\beta_R+j\omega/\omega_{T2})\alpha_R I_{CS}/V_{X2}\exp(V_{D2}/V_{X2})$$

$$+j\omega C_{02}/(1-V_{D2}/V_{02})^{**}(1/m_2)].$$

The four parameters for this transistor model are $1/\beta_F$, $1/\omega_{T1}$, V_{D1}, and V_{D2}. An expression in terms of perturbations of these parameters can be obtained from (3.28) by the following conversion:

$$\Delta y_{11} = [\partial y_{11}/\partial(1/\beta_F)]\Delta(1/\beta_F)+[\partial y_{11}/\partial(1/\omega_{T1})]\Delta(1/\omega_{T1})$$

$$+[\partial y_{11}/\partial V_{D1}]\Delta V_{D1}+[\partial y_{11}/\partial V_{D2}]\Delta V_{D2}.$$

Expressions for Δy_{12}, Δy_{21}, and Δy_{22} are similar and yield

$$-\alpha_F^2 I_{ES}/V_{X1}\, exp(V_{D1}/V_{X1})\, V_{EB}[-j\omega/\omega_{T1}\Psi_{EB}+\Psi_{CB}]\Delta(1/\beta_F)$$

$$-j\omega\alpha_F I_{ES}/V_{X1}\, exp(V_{D1}/V_{X1})\, V_{EB}\Psi_{EB}\Delta(1/\omega_{T1})$$

$$-V_{EB}\{\Psi_{EB}[(1/\alpha_F+j\omega/\omega_{T1})\alpha_F I_{ES}/V_{X1}^2\, exp(V_{D1}/V_{X1})$$

$$+j\omega C_{01}/[m_1 V_{D01}(1-V_{D1}/V_{D01})^{**}(1+1/m_1)]]$$

$$-\Psi_{CB}\alpha_F I_{ES}/V_{X1}^2\, exp(V_{D1}/V_{X1})\}\Delta V_{D1}$$

$$-V_{CB}\{-\Psi_{EB}\alpha_R I_{CS}/V_{X2}^2\, exp(V_{D2}/V_{X2})$$

$$+\Psi_{CB}[(1/\alpha_R+j\omega/\omega_{T2})\alpha_R I_{CS}/V_{X2}^2\, exp(V_{D2}/V_{X2})$$

$$+j\omega C_{02}/[m_2 V_{D02}(1-V_{D2}/V_{D02})^{**}(1+1/m_2)]\}\Delta V_{D2}. \qquad (3.29)$$

The hybrid-pi model can be described by the following Y-parameters:

$$Y_{11} = q|I_C|/(kT)[1/\beta_F+j\omega/\omega_{T1}+1+\eta]-j\omega C_\mu,$$

$$Y_{12} = -\eta q|I_C|/(kT),$$

$$Y_{21} = -q|I_C|(1+\eta)/(kT),$$

and

$$Y_{22} = \eta q|I_C|/(kT)(1/\beta_F+j\omega/\omega_{T1}+1)+j\omega C_\mu.$$

The three parameters for this transistor model are $1/\beta_F$, $1/\omega_{T1}$, and $|I_C|$. An expression in terms of perturbations of these parameters can be obtained from (3.28) by the following conversion:

$$\Delta Y_{11} = [\partial Y_{11}/\partial(1/\beta_F)]\Delta(1/\beta_F)+[\partial Y_{11}/\partial(1/\omega_{T1})]\Delta(1/\omega_{T1})$$

$$+[\partial Y_{11}/\partial|I_C|]\Delta|I_C|.$$

Expressions for ΔY_{12}, ΔY_{21}, and ΔY_{22} are similar and yield

$$-q|I_C|/(kT)(V_{EB}\Psi_{EB}+\eta V_{CB}\Psi_{CB})\Delta(1/\beta_F)$$

$$-j\omega q|I_C|/(kT)(V_{EB}\Psi_{EB}+\eta V_{CB}\Psi_{CB})\Delta(1/\omega_{T1})$$

$$-q/(kT)[(1/\beta_F+j\omega/\omega_{T1}+1+\eta)V_{EB}\Psi_{EB}-\eta V_{CB}\Psi_{EB}$$

$$-(\eta+1)V_{EB}\Psi_{CB}+\eta(1/\beta_F+j\omega/\omega_{T1}+1)V_{CB}\Psi_{CB}]\Delta|I_C|. \qquad (3.30)$$

The only other elements that are present in the original network are independent voltage and current sources, some of which are zero-valued and introduced solely for the purpose of measuring a response. An independent current source is required for every desired voltage response and an independent voltage source is required for every desired current response because they will be needed as sources in the adjoint network even if they are zero-valued in the original network. We assume that the values of the independent sources remain constant so that

$$\Delta V_V = \Delta I_I = 0.$$

Expression (3.11) for ac considerations reduces to

$$-\sum_V \Delta I_V \Psi_V + \sum_I \Delta V_I \Phi_I = -\sum_{J=1}^{n} G_J \Delta r_J \tag{3.31}$$

where n is the number of designable parameters, r_J is the Jth designable parameter, and G_J is the corresponding coefficient which can be found in (3.12) through (3.18), (3.20), (3.29), and (3.30). For dc considerations, (3.11) reduces to

$$-\sum_V \Delta I_V \psi_V + \sum_I \Delta V_I \phi_I = -\sum_{J=1}^{n} G_J \Delta r_J \tag{3.32}$$

where G_J can be found in (3.12), (3.13), (3.18), (3.19), and (3.26). We can associate independent sources in the original network with independent sources in the adjoint network of the same type. The values of these sources are still to be found.

We now consider the ac performance function (3.6). Consideration of other ac performance functions yields similar results. The change in error due to the change in designable elements is, to first order,

$$\begin{aligned}\Delta E_{AC} = \sum_I \int_{u_I}^{u_F} & \mathrm{Re}\{[W_I(u)(|V_I(u)|/|\hat{V}_I(u)|-1)]^{p-1} \\ & \cdot W_I(u) V_I^*(u)/[|\hat{V}_I(u)||V_I(u)|]\Delta V_I(u)\}du \\ +\sum_V \int_{u_I}^{u_F} & \mathrm{Re}\{[W_V(u)(|I_V(u)|/|\hat{I}_V(u)|-1)]^{p-1} \\ & \cdot W_V(u) I_V^*(u)/[|\hat{I}_V(u)||I_V(u)|]\Delta I_V(u)\}du\end{aligned} \tag{3.33}$$

where * indicates complex conjugate. Comparison of (3.31) and (3.33) reveals that if all current sources in the adjoint network corresponding to desired responses are set to

$$\Phi_I(u) = [W_I(u)(|V_I(u)|/|\hat{V}_I(u)|-1)]^{p-1}$$

$$\cdot[W_I(u)V_I^*(u)/(|\hat{V}_I(u)||V_I(u)|)],$$

all voltage sources in the adjoint network corresponding to desired responses are set to

$$\Psi_V(u) = -[W_V(u)(|I_V(u)|/|\hat{I}_V(u)|-1)]^{p-1}$$

$$\cdot[W_V(u)I_V^*(u)/(|\hat{I}_V(u)||I_V(u)|)],$$

and all remaining sources in the adjoint network are set to zero, then

$$\Delta E_{AC} = \sum_{J=1}^{n} \{\int_{u_I}^{u_F} \mathrm{Re}(-G_J)\,du\}\Delta r_J.$$

Since

$$\Delta E_{AC} = \sum_{J=1}^{n} \{\partial E_{AC}/\partial r_J\}\Delta r_J,$$

the desired gradient components result:

$$\partial E_{AC}/\partial r_J = \int_{u_I}^{u_F} \mathrm{Re}(-G_J)\,du.$$

A similar process specifies the adjoint network excitations for dc considerations to be

$$\phi_I(T) = [W_I(T)(V_I(T)/\hat{V}_I(T)-1)]^{p-1} W_I(T)/\hat{V}_I(T)$$

and

$$\psi_V(T) = -[W_V(T)(I_V(T)/\hat{I}_V(T)-1)]^{p-1} W_V(T)/\hat{I}_V(T)$$

when using performance function (3.3). The dc gradient components are

$$\partial E_{DC}/\partial r_J = \int_{T_I}^{T_F} -G_J dT.$$

Since the total error is the sum of the ac error and the dc error, the total gradients are the sum of the ac gradients and dc gradients:

$$\partial E/\partial r_J = [\partial E_{AC}/\partial r_J]+[\partial E_{DC}/\partial r_J].$$

We have determined first-order partial derivatives of the performance function with respect to each network parameter. Since each of these parameters has a temperature dependence of the form

$$r_J = r_J'[1+\alpha_J(T-T_0)]$$

where r_J' is the component of the parameter that must be varied, we determine gradient information with respect to r_J and transform it:

$$\partial E/\partial r_J' = [\partial E/\partial r_J][\partial r_J/\partial r_J']$$

$$= [1+\alpha_J(T-T_0)][\partial E/\partial r_J].$$

It has been shown that the gradients of all network parameters of interest are appropriately weighted products of voltages and currents of the original and adjoint networks. All gradient information can be determined upon performing two ac analyses per frequency and two dc analyses per temperature. (It should be noted that the adjoint network is linear and requires only one linear dc analysis.) The gradients obtained by this method have been shown to be exact.[8]

Accurate Integration of the Performance Function and Gradients

The evaluation of the performance functions and gradients discussed above requires integration of functions of network responses. Because of the numerical nature of the network functions, the ac performance function and the ac gradient components must be numerically integrated over the frequency range and the dc performance function and dc gradient components must be numerically integrated over the temperature range. A set of frequencies and a set of temperatures must be selected by the designer or the design algorithm that accurately represent the entire domain of operation. The quality of the final network depends on the accuracy of the numerical integration algorithms.

All of the functions to be integrated are explicit functions of frequency or temperature and it is not possible for the numerical integral to be unstable unless the function itself is unstable.

The dc performance function and dc gradient components are generally smooth functions of temperature that can be integrated accurately by selecting three or more fixed temperatures and using trapezoidal rule. Desired dc responses are usually constant voltage and current specifications. Actual dc responses are usually linear within the range of temperatures over which the element models are valid.

The ac performance function and ac gradient components may be rapidly varying functions of frequency and in such cases, a fixed set of frequencies will not be adequate unless a large number of frequencies are specified. An algorithm that automatically selects frequencies in order to maintain a specified accuracy enables the integration to be performed with a smaller number of frequencies. The network response shown in Fig. 3-3 is typical of the type of response that best utilizes the flexibility of an automatic frequency selection algorithm. The response from ω_A to ω_B is unchanged and few frequencies are necessary to numerically integrate this portion of the response accurately. On the other hand, the response from ω_B to ω_C varies considerably and many frequencies must be used to accurately integrate this portion of the response.

Integration of the dc functions is accomplished by use of trapezoidal rule at a set of temperatures specified by the user. The formula used is

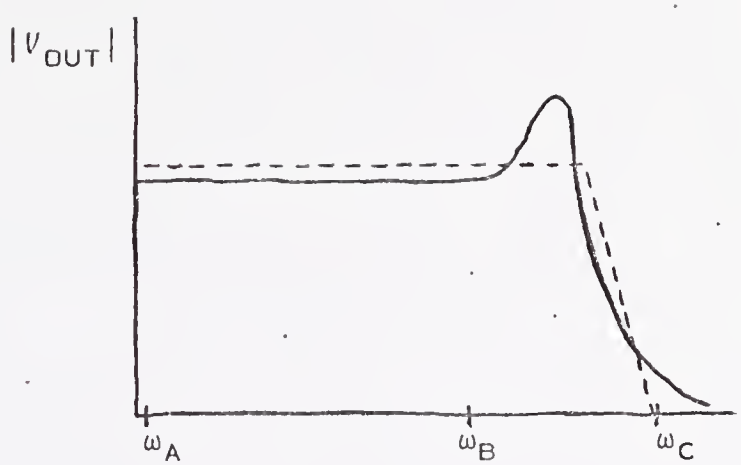


Fig. 3-3

Network response over a frequency range
with peaking where solid line is actual response
and dashed line is desired response

$$\int_{T_1}^{T_N} f(T)dT = f(T_1)(T_2-T_1)/2+f(T_2)(T_3-T_1)/2+\ldots$$

$$+f(T_{N-1})(T_N-T_{N-2})/2+f(T_N)(T_N-T_{N-1})/2 \qquad (3.34)$$

where N is the number of temperatures and the function f(T) can be either the performance function or a gradient component. Accuracy may be increased by specifying more temperatures, but there is seldom need for more than ten.

Three different numerical integration algorithms were tested for evaluation of the ac performance function and gradients. Two of the algorithms were found to be more efficient for the functions of interest here.

The most straightforward algorithm uses trapezoidal rule at a set of frequencies specified by the user. As with the dc case, (3.34) is used with angular frequency as the variable of integration instead of temperature. It was found that, except for some simple networks, a large number of frequencies had to be specified to achieve acceptable results.

The second integration algorithm also employed trapezoidal rule, but included an automatic frequency selection feature. The frequency selection process is based on a halving or doubling of the step size depending on the degree to which a specified error criterion is satisfied. The error criterion compares the partial integral evaluated over a double step size to the sum of two partial integrals, each evaluated over a single step size, and indicates whether the step size is suitable.

To begin the integration, three analyses are performed. The first analysis is performed at the initial frequency of the range and all function values for that frequency are saved. The frequency is updated by adding to it one-tenth of the entire frequency range. (When the frequency is scaled logarithmically, the step size is logarithmic also.) One-tenth of the range is specified to be the initial step size and the maximum step size because it is small enough to prevent most aberrations from being overlooked and yet it is large enough to prevent the use of too many frequencies when they are unnecessary. After analysis at the second frequency, all function values are stored and a third frequency is obtained by adding the step size to the previous frequency. Analysis at this frequency allows two values of the integral to be calculated and compared:

$$Q = |h/2[f(\omega_1)+2f(\omega_2)+f(\omega_3)]-h[f(\omega_1)+f(\omega_3)]|/\Lambda \quad (3.35)$$

where h is the step size,

ω_1 is the first frequency

ω_2 is the second frequency,

ω_3 is the third frequency,

Λ is the smallest maximum value previously assigned to the integral of $f(\omega)$, and

Q is the measure of the acceptability of the current step size.

If Λ is zero, as it would be on the first comparison of the first analysis, the normalization of the difference is not

performed. Since many functions are being integrated simultaneously, the values of all Q's are averaged to obtain Q_A which is used to judge the acceptability of the step size. If $Q_A > 2\varepsilon$ where ε is a specified error criterion, the step size is unacceptably large and is halved unless it is already at the minimum step size. If $\varepsilon \leq Q_A \leq 2\varepsilon$, the step size is acceptable and its value is retained. If $Q_A < \varepsilon$, the step size is unacceptably small and is doubled unless it is already at the maximum step size. When the step size is retained or increased, integration proceeds from the third frequency; if the step size is decreased, integration resumes at the first frequency.

The ac performance function and gradients are explicit functions of frequency. By explicit functions, we mean that the value of such a function at a specific frequency is not dependent on the value of the function at another frequency. Thus, if a slight inaccuracy is allowed early in an integration, this inaccuracy will not affect the remainder of the integration. Since only explicit functions are considered, error accumulation is not examined in the halving and doubling integration scheme and a straightforward algorithm results.

The third integration algorithm is a direct descendant of the variable step size, variable order algorithm developed by Gear.[39] This algorithm is a predictor-corrector multistep numerical integration scheme based on two variable-order equations. The first equation is used to predict the partial

integral for a specified step size and order and the second equation is used iteratively to correct the step size and order to reduce the error in integration below a specified value. Gear's algorithm is intended to integrate implicit functions and has an elaborate error control mechanism built into it. Because of the complexity of this algorithm, it is the most inefficient of the three algorithms although it generally performs acceptably. An error criterion, ε', is specified for this algorithm which is similar to the error criterion of the other variable step size algorithm but not identical.

A problem common to both variable step size algorithms is the possibility of obtaining a slightly different value of the performance function on two passes even though the network parameters are the same. This inconsistency occurs because the value of Λ in (3.35) is constantly being updated. Since the optimization algorithm becomes confused at this fluctuation, a reasonably tight error criteria should be maintained to reduce the significance of this possible variation.

The three algorithms were compared through some typical design statistics obtained through use of the computer program FROLIC. The first network to be considered is a seven-pole, LC, passive, low-pass filter with initial parameter values as shown in Fig. 3-4(a). The desired response and the initial actual response to a constant one ampere excitation are shown in Fig. 3-4(b). The variable network para-

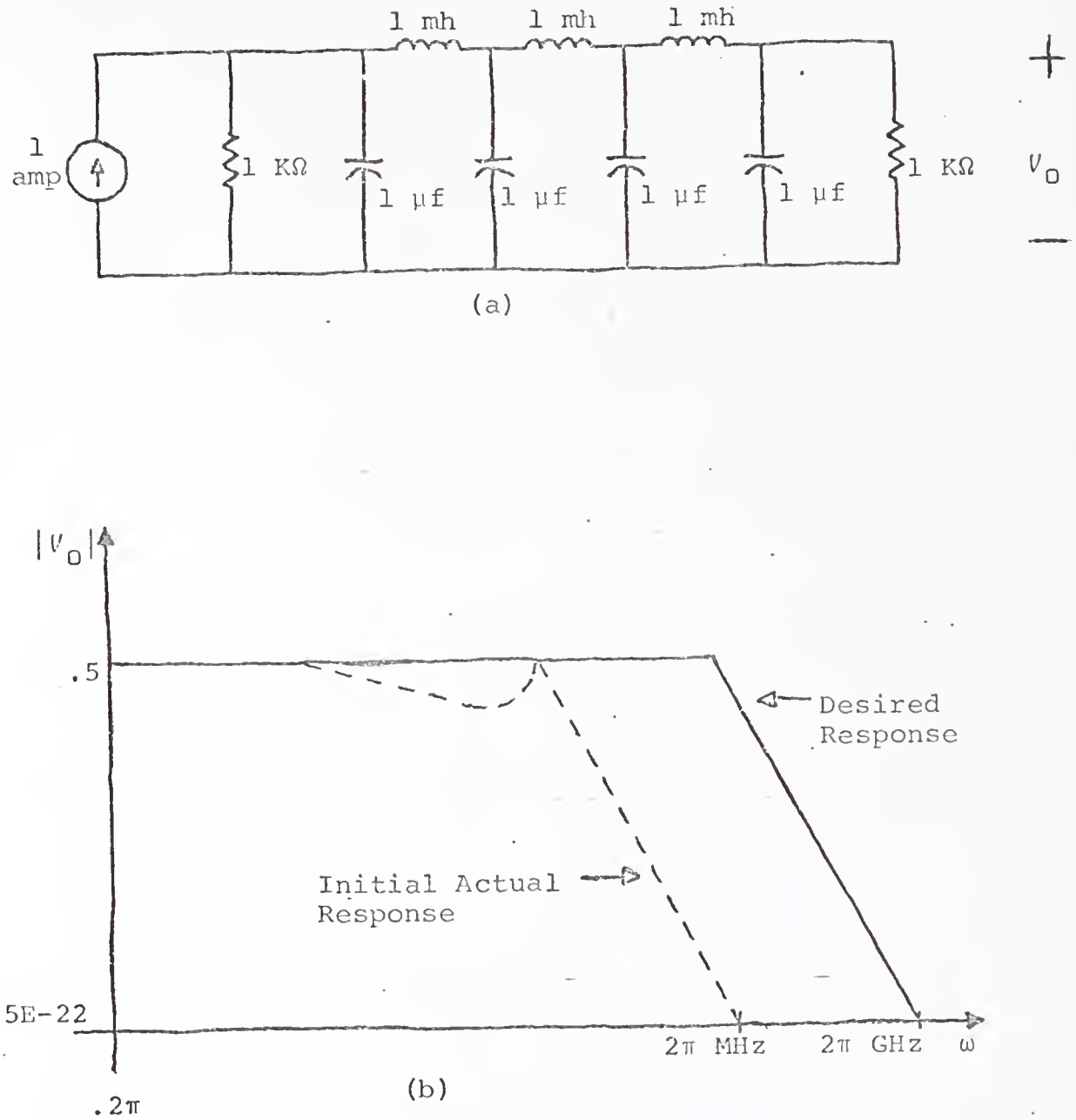


Fig. 3-4

(a) Seven-pole passive low-pass filter
(b) Desired and initial actual responses

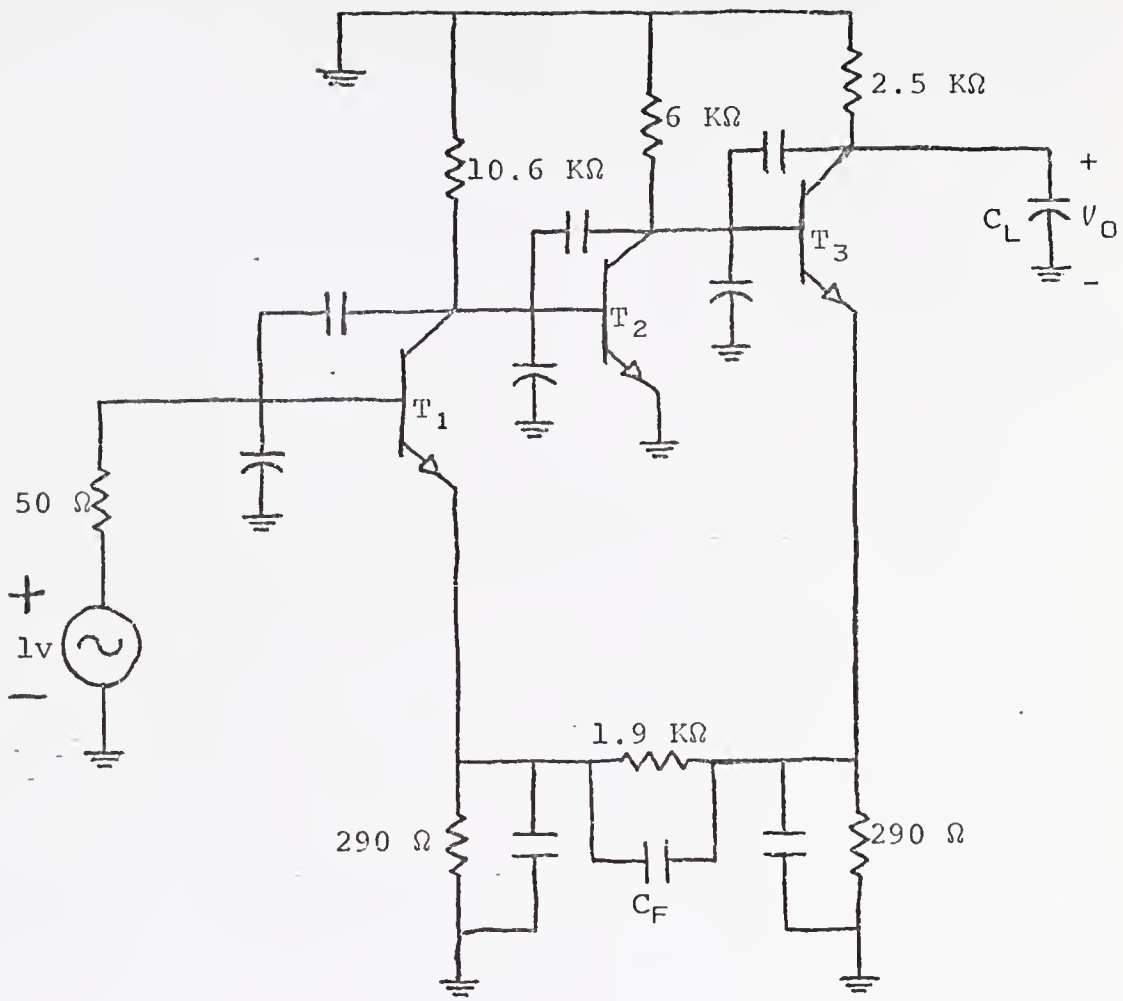
meters are the four capacitances and the three inductances. For all of the designs, performance function (3.6) was used with p=2, optimization was by means of the conjugate gradients algorithm, the parameters were logarithmically scaled, and the parameter constraints were set so that they did not affect the design process. The values of ε and ε' are varied for the variable step size algorithms and the number of frequencies is varied for the fixed frequencies algorithm. Comparisons are made in Table 3-1 of average time per analysis, average time per frequency, average number of frequencies per analysis, and whether the desired result was obtained. The fixed frequencies were chosen so that they were more dense in the vicinity of the break frequency. It is shown that more frequencies may be used in the fixed frequency algorithm because it has no overhead for frequency selection; although fewer frequencies may be just as acceptable because of their appropriate placement with the variable step size algorithm. It is also shown that the trapezoidal rule variable step size algorithm is more efficient than the Gear algorithm for this type of problem.

For a second example, we design the response of a series-series triple feedback amplifier with parameters and initial responses shown in Fig. 3-5. Three designs are run with each of the three integration algorithms. For all of the designs, performance function (3.6) was used with p=2, optimization was supervised by the conjugate gradients algorithm and the parameters were logarithmically scaled. The values of ε and

Table 3-1

Comparison of Numerical Integration Statistics
for Three Available Algorithms When Used on the Design
of the Network of Fig. 3-4

	Trapezoidal Rule Fixed Frequencies			Trapezoidal Rule Variable Step Size					Variable Order Variable Step Size				
	11	101	595	ε=1.	ε=.5	ε=.1	ε=.01	$\varepsilon=10^{-4}$	ε'=1.	ε'=.1	ε'=.07	ε'=.03	ε'=.01
Average Time Per Analysis (sec)	.10	.91	5.36	.25	.22	.56	.93	2.19	.38	1.13	1.28	1.71	1.97
Average Time Per Frequency (msec)	9	9	9	10	10	10	10	10	14	14	14	14	14
Average Number of Frequencies Per Analysis	11	101	595	25	22	56	93	219	27	81	91	122	141
Successful Design	NO	YES	YES	NO	NO	YES	YES	YES	NO	YES	YES	YES	YES



	T_1	T_2	T_3
β	120	120	120
f_T	470 MHz	560 MHz	550 MHz
$\|I_C\|$	.5 ma	.77 ma	.73 ma
R_B	258 Ω	293 Ω	238 Ω
R_C	29 Ω	30 Ω	29 Ω
η	.00065	.00065	.00065
C_μ	1 pf	1 pf	1 pf

Fig. 3-5

Series-series triple feedback amplifier

ε' are set to .01 except in one run where Gear's algorithm was being used and unacceptable results were obtained. In this case, a value of ε'=.005 was used and satisfactory results were obtained. The value .01 has been found to be acceptable for most problems. Comparisons are made in Table 3-2 of average time per analysis, average time per frequency, average number of frequencies per analysis, and whether the desired result was obtained. The first desired response was a flat gain of fifty over a bandwidth of one hertz to a hundred megahertz. All capacitors and resistors in the network except the source resistor were allowed to vary. The results showed all algorithms to be capable of simultaneously integrating the seventeen functions, but the variable step size, trapezoidal rule algorithm is superior. If eight of the capacitors are removed from the network and only C_L, C_F, and the resistors remain as network parameters, similar results are obtained. If all capacitors are removed from the network and the desired bandwidth reduced slightly, the two trapezoidal rule algorithms are successful. Gear's algorithm had to have a reduced value of ε' to achieve the desired result.

The fixed frequency approach is a desirable algorithm for some problems and will work for all problems if enough frequencies are specified. It has the advantage of always yielding the same result for identical functions. The variable step size trapezoidal rule algorithm has performed well and frees the user of the chore of predicting areas of

Table 3-2

Comparison of Numerical Integration Statistics
for Three Available Algorithms
When Used on the Design of the Network of Fig. 3-5

	All Capacitors $1\text{-}10^8$ Hz (17 functions)			C_L & C_E $1\text{-}10^8$ Hz (9 functions)			No Capacitors $1\text{-}5X10^7$ Hz (7 functions)		
	Fixed 71 Frequencies	Variable Step Size ε=.01	Variable Step Size and Order ε'=.01	Fixed 71 Frequencies	Variable Step Size ε=.01	Variable Step Size and Order ε'=.01	Fixed 41 Frequencies	Variable Step Size ε=.01	Variable Step Size and Order ε'=.005
Average Time Per Analysis (sec)	1.27	.72	1.19	1.57	.96	1.74	.72	.50	.92
Average Time Per Frequency (msec)	18	21	23	22	26	31	18	19	21
Average Number of Frequencies Per Analysis	71	35	52	71	37	56	40	26	44
Successful Design	YES	YES	YES	YES	YES	YES	YES	YES	YES

the frequency range where problems might occur. This algorithm is superior to Gear's algorithm because it uses fewer frequencies and is more efficient when integrating the explicit functions considered here.

An additional feature of a variable step size algorithm is that if aberrations in the frequency response move around within the frequency range, the algorithm can adjust to their movement. If these aberrations disappear during the course of the design, the algorithm can adjust for this also.

CHAPTER 4

TOPOLOGY DESIGN THROUGH ELEMENT GROWTH AND REMOVAL

Parameter adjustment to obtain desired network responses solves only part of the network design problem. An initial network configuration must be determined that has the possibility of yielding the desired responses. In this chapter, a practical approach to the design of network topology is discussed with emphasis on the growth and removal of transistors to satisfy ac specifications.

Resistor and Capacitor Growth and Removal

A fixed network configuration limits the flexibility of an automated design procedure. The ability to add and remove elements allows new configurations to be developed and old configurations to be improved. The growth and removal of linear elements has been discussed by Director and Rohrer[7] in conjunction with the adjoint network method of gradient calculation.

The gradient for a resistor,

$$\partial E/\partial R = \int_{\omega_I}^{\omega_F} \mathrm{Re}(-I_R \Phi_R)[1+\alpha_R(T-T_0)]d\omega + \int_{T_I}^{T_F} -I_R \phi_R [1+\alpha_R(T-T_0)]dT,$$

does not explicitly depend on the value of resistance. If

$R=0$, I_R, Φ_R, I_R, and ϕ_R can be calculated and the desirability of "growing" a resistor in a short circuit is determined. The gradient for a conductor, actually a resistor with $G=1/R$, is

$$\partial E/\partial G = \int_{\omega_I}^{\omega_F} \mathrm{Re}(V_G \Psi_G)[1+\alpha_G(T-T_0)]d\omega$$

$$+\int_{T_J}^{T_F} V_G \psi_G [1+\alpha_G(T-T_0)]dT$$

which is not dependent on the value of conductance. If $G=0$, V_G, Ψ_G, V_G, and ψ_G can be calculated and the desirability of growing a conductor in an open circuit is determined. Although a resistor and a conductor are simple inverses of each other at any non-zero values of resistance and conductance, it is the zero values that are of interest when considering growth. Thus two representations of a resistive branch are considered, one to grow from a short circuit and one to grow from an open circuit. Similarly, the gradient for a capacitor will indicate possible growth of a capacitor in an open circuit and the gradient for an elastance indicates possible growth of a capacitor in a short circuit.

With indicators available to determine if a resistor or a capacitor should be added in any lead or between any pair of nodes, it initially appears to be possible to check all locations for resistor and capacitor growth. Actual application of this procedure would require more time than is feasible and an alternate approach is used: possible growth locations are specified by entering components with minimum

parameter values. It is not possible to have components with zero-valued parameters because of the general branch model assumed earlier.

Since zero-valued parameters are not used, only resistances and capacitances need to be considered. A minimum-valued conductance is equivalent to a maximum-valued resistance, and similarly, a minimum-valued elastance is equivalent to a maximum-valued capacitance.

Resistor and capacitor removal is presently simulated by the adjustment of their parameters to their lower or upper limits. When these parameters encounter a boundary, the elements are not removed from consideration as they may again become necessary to achieve the desired responses and "grow" by their parameters moving away from the boundaries. This procedure has proven to be satisfactory.

Diode Growth and Removal

The gradient for a diode (the contribution due to the junction voltage is omitted from growth consideration) is

$$\partial E/\partial I_S = \int_{\omega_I}^{\omega_F} \mathrm{Re}[1/V_X \exp(V_D/V_X)(1+j\omega/\omega_X) V_D \Psi_D][1+\alpha_D(T-T_0)]d\omega$$

$$+\int_{T_I}^{T_F} [\exp(V_D/V_X)-1]\psi_D[1+\alpha_D(T-T_0)]dT$$

where V_D, Ψ_D, V_D, and ψ_D are voltages across the junction and the expression in the second integrand is a simplified form of (3.19). Since the expression contains no dependence on the parameter I_S, it is possible to grow a diode between any

two nodes. The problem with growing a diode is that the gradient is actually indicating conductor growth. Thus a diode is grown as a variable conductance and its voltage setting characteristics which are a consequence of the diode's nonlinearity are ignored. An alternate indicator to measure the desirability of a fixed voltage has not been determined so diode growth and removal are not considered.

Transistor Growth and Removal

Bipolar junction transistors have become the most commonly used components in integrated circuits because transistors are small, dissipate little power, and perform many functions. Transistor growth and removal is therefore a relevant subject to consider. In this work, dc considerations are not used to determine transistor growth or removal. The hybrid-pi transistor model is used because it is convenient and yields results comparable to those obtained through use of other models.

Initially, we approach transistor growth in a manner similar to that applied to other components: we investigate the gradient expressions for a zero-valued transistor. A zero-valued transistor may be envisioned as one with

$$1/\beta_F = 1/\omega_{T1} = |I_C| = 0.$$

In this case, the only ac transistor gradient component not identically equal to zero is

$$\partial E_{AC}/\partial |I_C| = q/(kT) \int_{\omega_I}^{\omega_F} \mathrm{Re}[V_{EB}\Psi_{EB} - V_{EB}\Psi_{CB}]\, d\omega$$

which can be rewritten as

$$\partial E_{AC}/\partial|I_C| = q/(kT)\int_{\omega_I}^{\omega_F} \mathrm{Re}[V_{EB}\Psi_{EC}]d\omega \qquad (4.1)$$

where $\eta=0$ for a nonexistant transistor. The gradient of a voltage-controlled current source shown in Fig. 4-1 is known to be

$$\partial E_{AC}/\partial g_M = \int_{\omega_I}^{\omega_F} \mathrm{Re}[V_{EB}\Psi_{EC}]d\omega.$$

Since $g_M=[q/(kT)]|I_C|$, the gradient computation associated with a zero-valued transistor is the same as that for a voltage-controlled current source.

At this juncture it is important to assess the practicality of what has been indicated in the previous paragraph. The desirability of growing a voltage-controlled current source between three distinct nodes is easily determined. Since the voltages V_{EB} and Ψ_{EC} can always be ascertained, even for the case of $g_M=0$, the gradient expression (4.1) may be evaluated. But for many situations it proves more convenient to consider the possibility of growing transistors simultaneously in a short circuit and across an open circuit. For instance, consider the single common-emitter stage of Fig. 4-2(a). An additional common-emitter stage might be required if more gain is needed. If so, another common-emitter stage should then be grown as indicated in Fig. 4-2(b).

In general, there are three possibilities for transistor growth as illustrated in Fig. 4-3. Observe that for the common-emitter configuration, the short circuit is between the

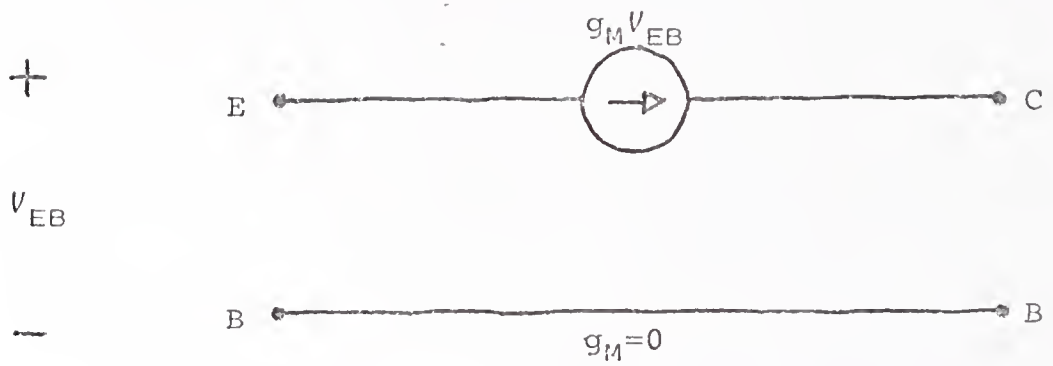
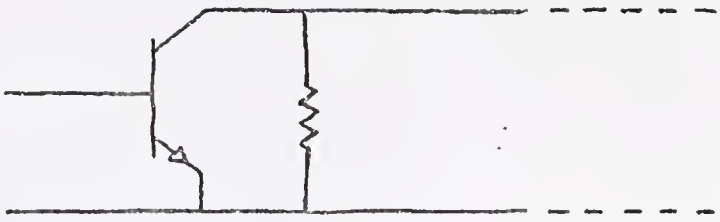
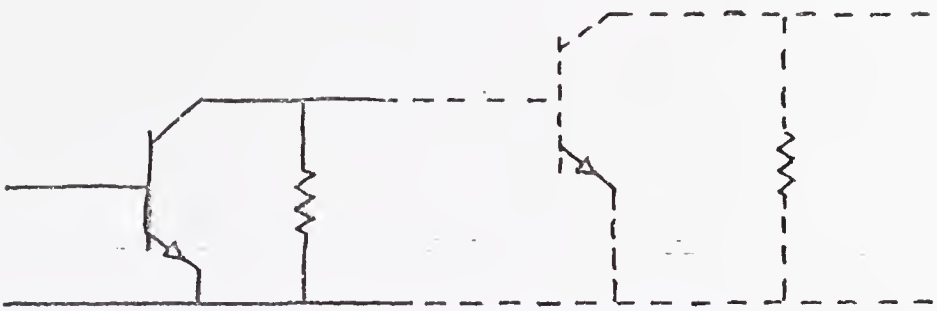


Fig. 4-1

Voltage-controlled current source configuration that is equivalent to a zero-valued transistor



(a)



(b)

Fig. 4-2

(a) Single-stage transistor amplifier with growth site
(b) The growth of an additional common-emitter stage

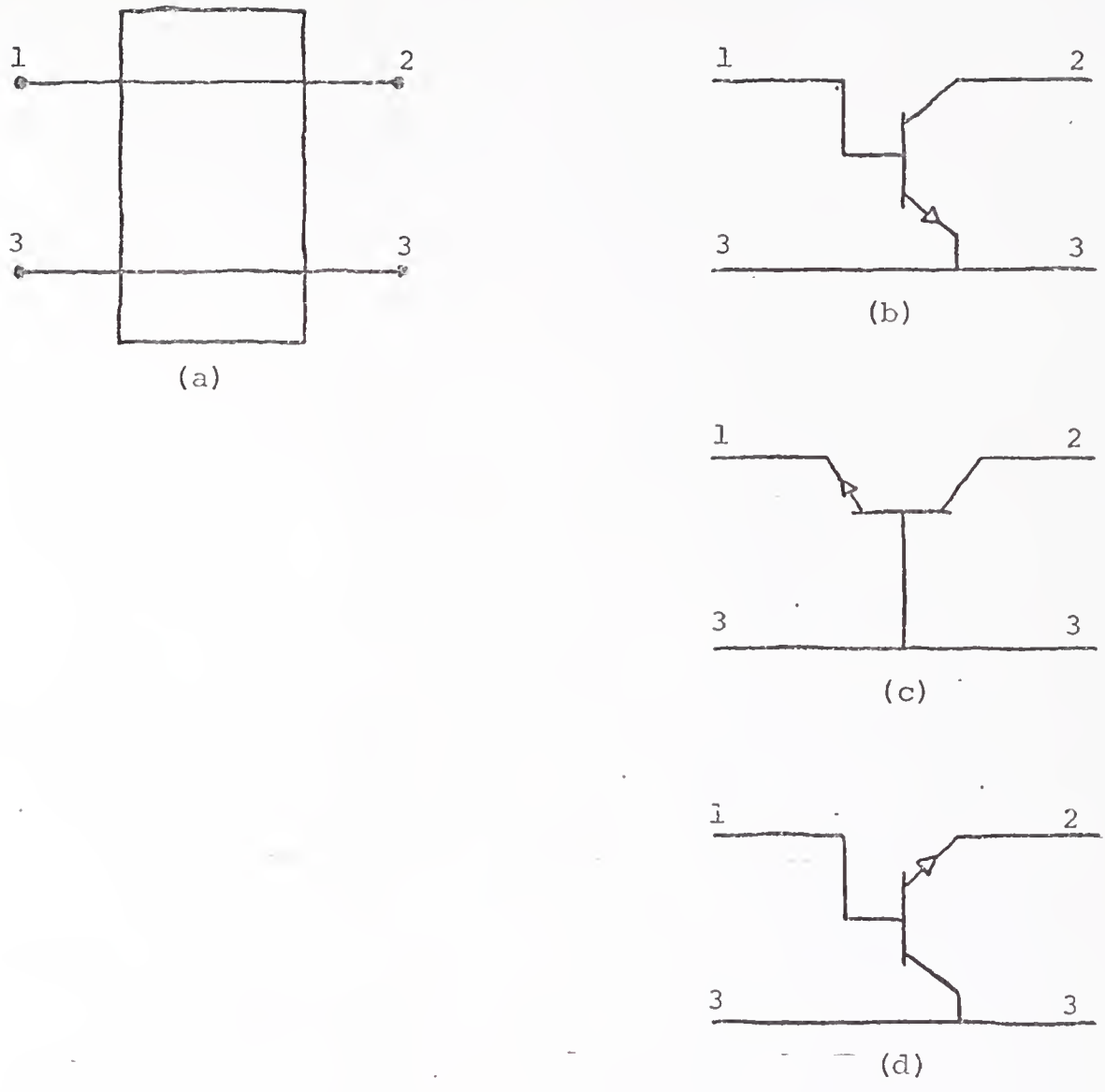


Fig. 4-3

(a) Transistors grown simultaneously in a short circuit (between terminals 1 and 2) and across an open circuit (between terminals 1 or 2 and 3). The three possible ways in which a transistor may be grown are shown.
(b) Common-emitter stage.
(c) Common-base stage.
(d) Common-collector stage.

base and collector nodes and the gradient component (4.1) reduces to that of a conductance since $V_{EC}=V_{EB}$ and $\Psi_{EB}=\Psi_{EC}$. Moreover, the gradients associated with the common-base and common-collector are identically zero since $V_{EC}=\Psi_{EC}=0$ and $V_{EB}=\Psi_{EB}=0$, respectively. Thus the choice of a voltage-controlled current source to indicate transistor growth is a poor one. We are therefore forced to alter our approach.

It is reasonable to ask what causes a need for transistors, the answer being the desirability for voltage gain and/or current gain. It can be shown that in an ideal situation, each transistor configuration yields gain. The gain for each stage can be determined by use of an ideal mid-frequency transistor model in each of the three possible configurations as shown in Fig. 4-4. The voltage and current gain expressions for each configuration are shown in Table 4-1. The common-emitter configuration yields a large voltage gain for typical values of transconductance and load resistance and a large current gain for typical values of beta. Both the voltage and current gains of the common-emitter include phase inversion that can be ignored for a magnitude fit (3.5) but must be considered for a complete fit (3.2). The common-base configuration also yields a large voltage gain but it yields no current gain. The common-collector configuration yields no voltage gain and large current gain. Therefore, transistor growth should be dependent on the need for voltage and/or current gain.

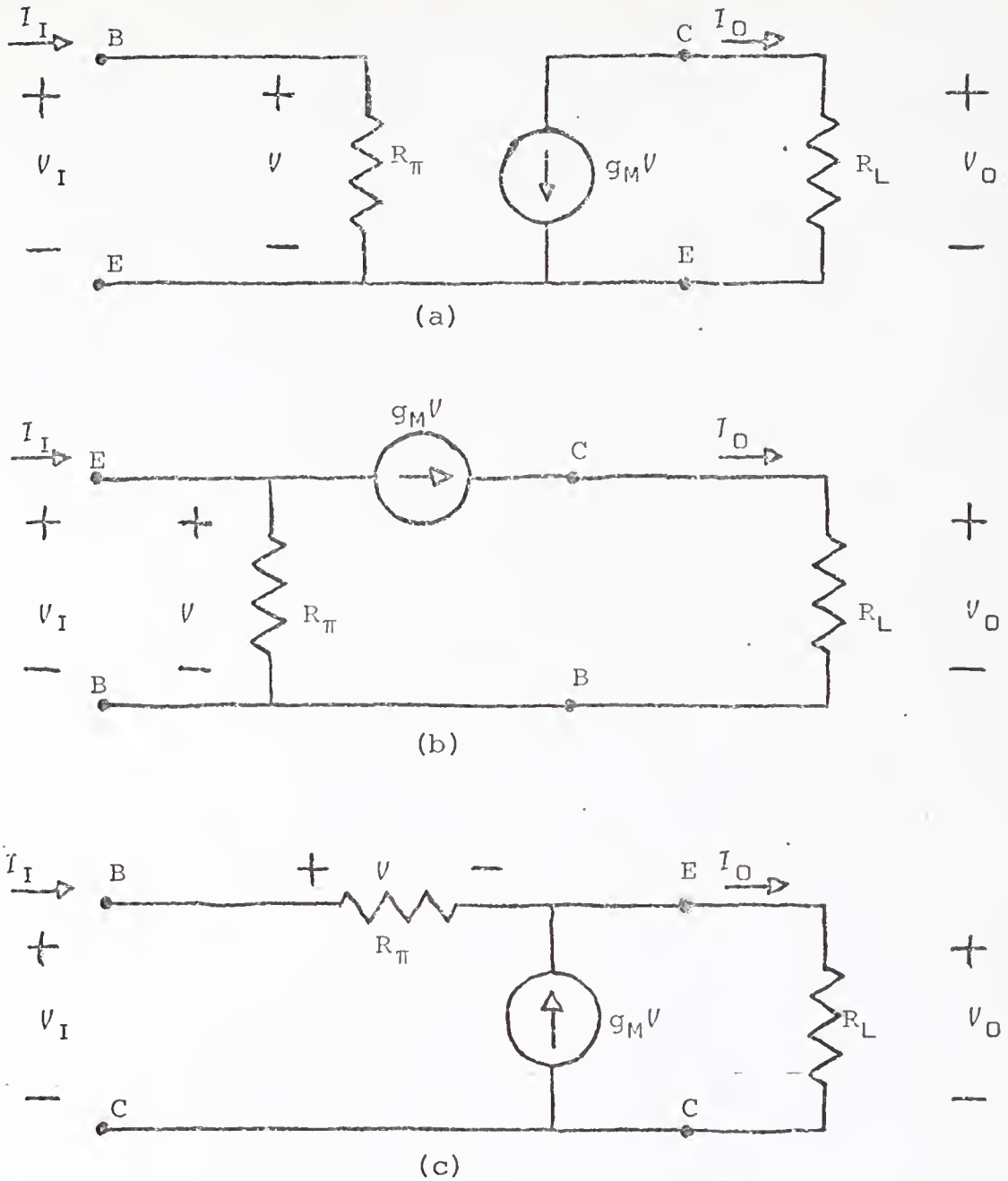


Fig. 4-4

Networks used to determine voltage and current gain for each transistor configuration
(a) Common-emitter
(b) Common-base
(c) Common-collector

Table 4-1

Voltage and Current Gain for Ideal Low-Frequency Transistor Model in Common-Emitter, Common-Base, and Common-Collector Configurations

	Common-Emitter	Common-Base	Common-Collector
Voltage Gain V_O/V_I	$-g_M R_L$	$g_M R_L$	$\dfrac{\beta_F+1}{\beta_F+1+\beta_F/(g_M R_L)}$
Current Gain I_O/I_I	$-\beta_F$	α_F	β_F+1

Gradients that indicate growth of controlled sources also indicate need for voltage and current gain. A voltage-controlled voltage source connected as shown in Fig. 4-5(a) simulates the voltage gain of the common-base configuration while a current-controlled current source connected as shown in Fig. 4-5(b) simulates the current gain of the common-collector configuration. Both a voltage-controlled voltage source and a current-controlled current source connected as shown in Fig. 4-5(c)-(d) simulate the voltage and current gains of the common-emitter configuration for magnitude fit as in (3.5) and complete fit as in (3.2), respectively. The gradient components associated with each of these configurations are[7]

$$\partial E_{AC}/\partial \mu = \int_{\omega_I}^{\omega_F} \mathrm{Re}[V_1 \Phi_1]\, d\omega$$

for the voltage gain,

$$\partial E_{AC}/\partial \beta = \int_{\omega_I}^{\omega_F} \mathrm{Re}[-I_1 \Psi_1]\, d\omega$$

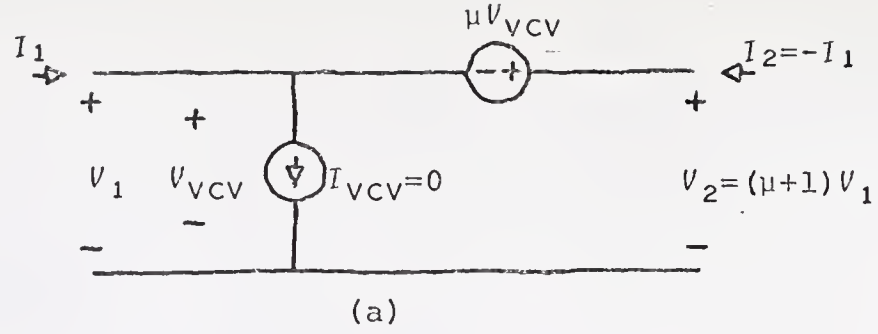
for the current gain,

$$\partial E_{AC}/\partial (\mu\beta) \equiv \partial E_{AC}/\partial \mu + \partial E_{AC}/\partial \beta$$

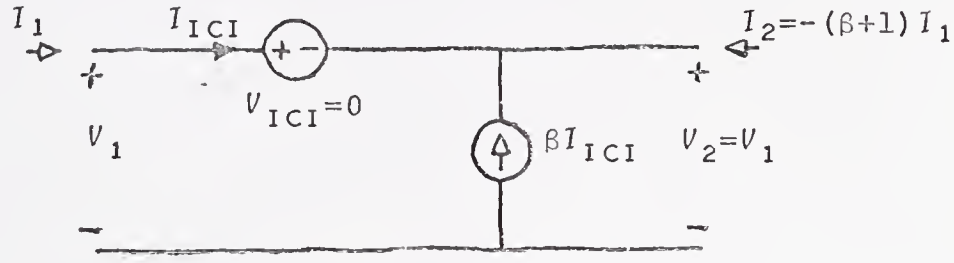
$$= \int_{\omega_I}^{\omega_F} \{\mathrm{Re}[V_1 \Phi_1] + \mathrm{Re}[-I_1 \Psi_1]\}\, d\omega$$

for voltage and current gain for a magnitude fit, and

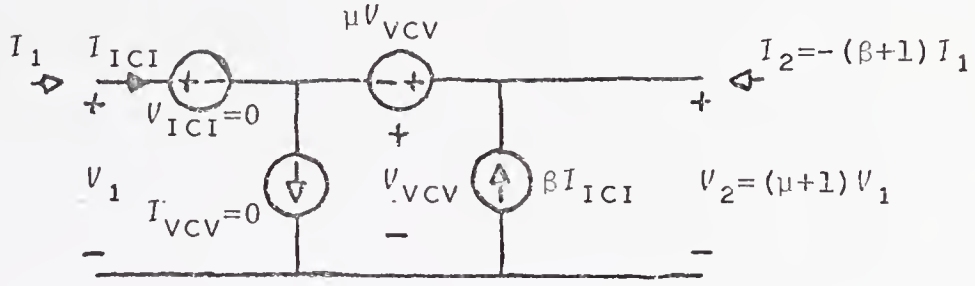
$$\partial E_{AC}/\partial (\mu\beta) = \int_{\omega_I}^{\omega_F} \{\mathrm{Re}[-V_1 \Phi_1] + \mathrm{Re}[I_1 \Psi_1]\}\, d\omega$$



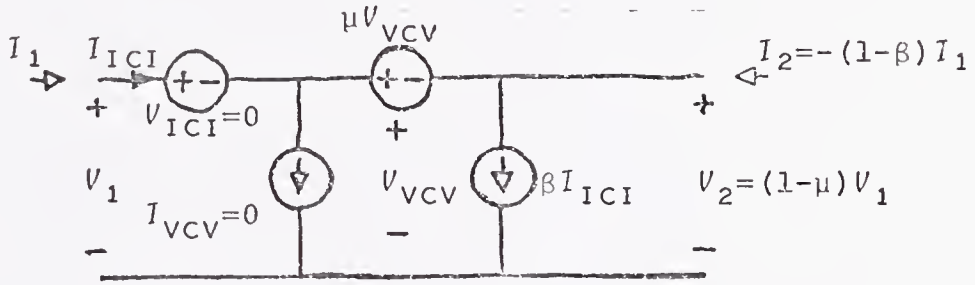
(a)



(b)



(c)



(d)

Fig. 4-5

Controlled source configurations
(a) Voltage gain
(b) Current gain
(c) Both voltage and current gain with inversion for magnitude fit
(d) Both voltage and current gain with inversion for complete fit

for voltage and current gain for a complete fit. The gradients for common-base and common-collector are not changed by use of magnitude or complete fit because the desired response is positive in both cases. For a common-emitter, the desired response must change signs to change from one to the other so the gradient expressions must also differ in sign. Observe that for the zero-valued case ($\mu=\beta=0$), each of these gradients exist and, moreover, each of the models of Fig. 4-5 reduce to the appropriate initial model of Fig. 4-3(a). Furthermore, after two circuit analyses, one on the original network and the other on the adjoint network, it is not only possible to determine if a transistor should be grown, but the appropriate configuration is also indicated as shown in Tables 4-2 and 4-3.

If transistor growth is indicated, appropriate initial parameter values must be defined to simulate as closely as possible the short and open relationship previously in effect. The dc bias parameters are directly related to the elements connecting each pair of nodes. Therefore, appropriate values of these parameters may not be at boundaries so that a condition of all shorts or all opens is avoided; the initial values for these parameters are chosen to be the values specified as input by the user for all three configurations. The initial values of $1/\beta_F$ and $1/\omega_{T1}$ for the common-collector configuration are chosen to be the upper limits for these parameters in order to simulate the short from base-to-emitter. The value of η is typically very small so the base-to-collector

Table 4-2

Transistor Growth Indicators for Magnitude Fit

$\partial E_{AC}/\partial \mu$	$\partial E_{AC}/\partial \beta$	Transistor Configuration
≥ 0	≥ 0	No Transistor Needed
< 0	≥ 0	Grow Common-Base Stage
≥ 0	< 0	Grow Common-Collector Stage
< 0	< 0	Grow Common-Emitter Stage

Table 4-3

Transistor Growth Indicators for Complete Fit

$\partial E_{AC}/\partial \mu$	$\partial E_{AC}/\partial \beta$	Transistor Configuration
≥ 0	≥ 0	Grow Common-Emitter Stage
< 0	≥ 0	Grow Common-Base Stage
≥ 0	< 0	Grow Common-Collector Stage
< 0	< 0	No Transistor Needed

connection does not simultaneously become a short circuit. For the other two configurations, initial values of $1/\beta_F$ and $1/\omega_{T1}$ are chosen to be the lower limits in order to simulate the necessary open circuits; the short circuit is difficult to simulate in these cases.

The ability to add transistors to a network must be complemented by the ability to remove transistors that are no longer desirable. It is not sufficient to allow transistor parameters to remain at a boundary as such a transistor still has significant influence on the network response. A simple approach to transistor removal is now suggested.

The transistor is no longer useful for satisfaction of ac specifications when the voltage and/or current gain it provides is not desired. An indication of this condition can be obtained by the determination of the desirability of conductor growth between the terminals of a transistor. Gradients for the zero-valued conductors shown in Fig. 4-6 can be calculated and used as removal indicators.

When a transistor has been removed, it must be replaced by the short-open combination from which transistors grow. The location of the short must be determined from the removal indicators. This can be done as follows. First removal requires that all three of the gradients for the conductors in Fig. 4-6 be negative (indicate conductor growth), and that the conductor with the most negative gradient be replaced by the short circuit. This method is simple and yet general. If the removal of a transistor causes the value of

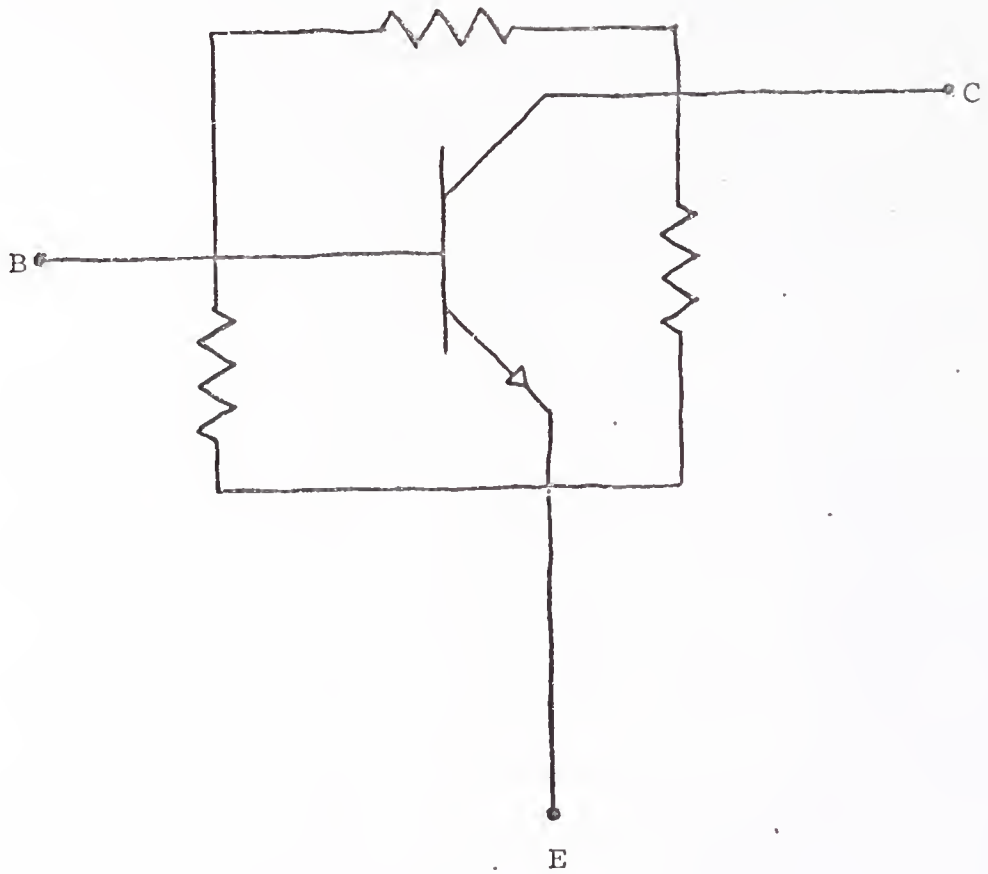


Fig. 4-6

Positions of zero-valued conductances
whose gradients are used to indicate transistor removal

the performance function to increase, the transistor is restored and other transistors are checked for removal.

Transistor growth indicators may be determined for all possible combinations of a short and an open. This approach would consume large amounts of time and is therefore discarded in favor of requiring possible growth sites to be specified a priori. Upon transistor removal, the location of the removed transistor becomes a growth site. Similarly, all transistors that are grown automatically become removable.

After each growth or removal, a design is performed and a minimum is obtained. If the value of the ac portion of the performance function at the end of the design is larger than it was before the action was taken on the transistor, the action is reversed and the old parameter values are restored.

The desirability of transistor growth is checked before transistor removal is considered because it is more often necessary to supply missing gain than it is to remove undesirable gain. It is not possible to compare growth indicators with removal indicators in order to decide which process should be done first because the two sets of indicators are unrelated.

Before a design is performed, growth indicators are determined for all possible growth sites and the largest indicator is selected. If no transistor grows, removal indicators are determined for all removable transistors and the largest indicator is selected. In most cases, only one transistor is grown or removed at a time because any action

that is indicated elsewhere in the network may be obviated by the action on the transistor with the largest indicator. An exception to this procedure is when a differential pair of transistors must be grown together to retain dc balance.

Our first approach to transistor growth-removal[40] consisted of repeated growth-removal where each such action was followed by a design without intermediate supervision by the user. The performance function remained the same throughout the process and the weights for each desired response had to be chosen very carefully in order to avoid overemphasis of a small portion of the desired responses. Many runs had to be made with different weights chosen by reducing the weight on portions of the desired responses that were receiving the most emphasis. The growth sites also had to be selected for the entire run and their best locations had to be determined over a series of runs.

It is difficult to specify appropriate sets of weights and growth sites for an entire run so a new approach is used that allows the user to specify new sets of weights and growth sites after each intermediate design. This approach simulates the use of a visual display and an interactive interface which could be added to the design algorithm. The flexibility of this approach allows the user to emphasize a portion of the desired responses for one or more growth-removal-design steps and then emphasize another portion of the desired responses at a later step. Thus the user's experience may reduce the effort expended by the computer.

The examples that follow were run with intermediate user interaction on the computer program FROLIC.

A wideband voltage amplifier is the basic design objective of the first example. The initial configuration shown in Fig. 4-7(a) is chosen and specifications are placed on the input admittance, output inpedance and voltage gain. It is convenient to express these specifications in terms of the hybrid g-parameters so that the port variables of the amplifier are given by

$$I_1(\omega) = g_{11}(\omega)V_1(\omega)$$

and

$$V_2(\omega) = g_{21}(\omega)V_1(\omega)+g_{22}(\omega)I_2(\omega).$$

The letter A in Fig. 4-7(a) indicates a possible transistor growth site. Gradients for voltage gain, A_V, and current gain, A_I, are calculated at this location. If a gradient component is negative, growth is indicated. In the later stages of the example, several growth sites will be available. So as to prevent unlimited growth or removal, only one transistor is considered at a time. The desired responses are shown in Fig. 4-7(b)-(d). Performance function (3.6) was used with p=2. The weights were specified constant over frequency at 1 for input admittance and output impedance and 1000 for voltage gain. This put a strong emphasis on the voltage gain which characterizes a voltage amplifier. Growth indicators were determined and a common-base stage was inserted. After

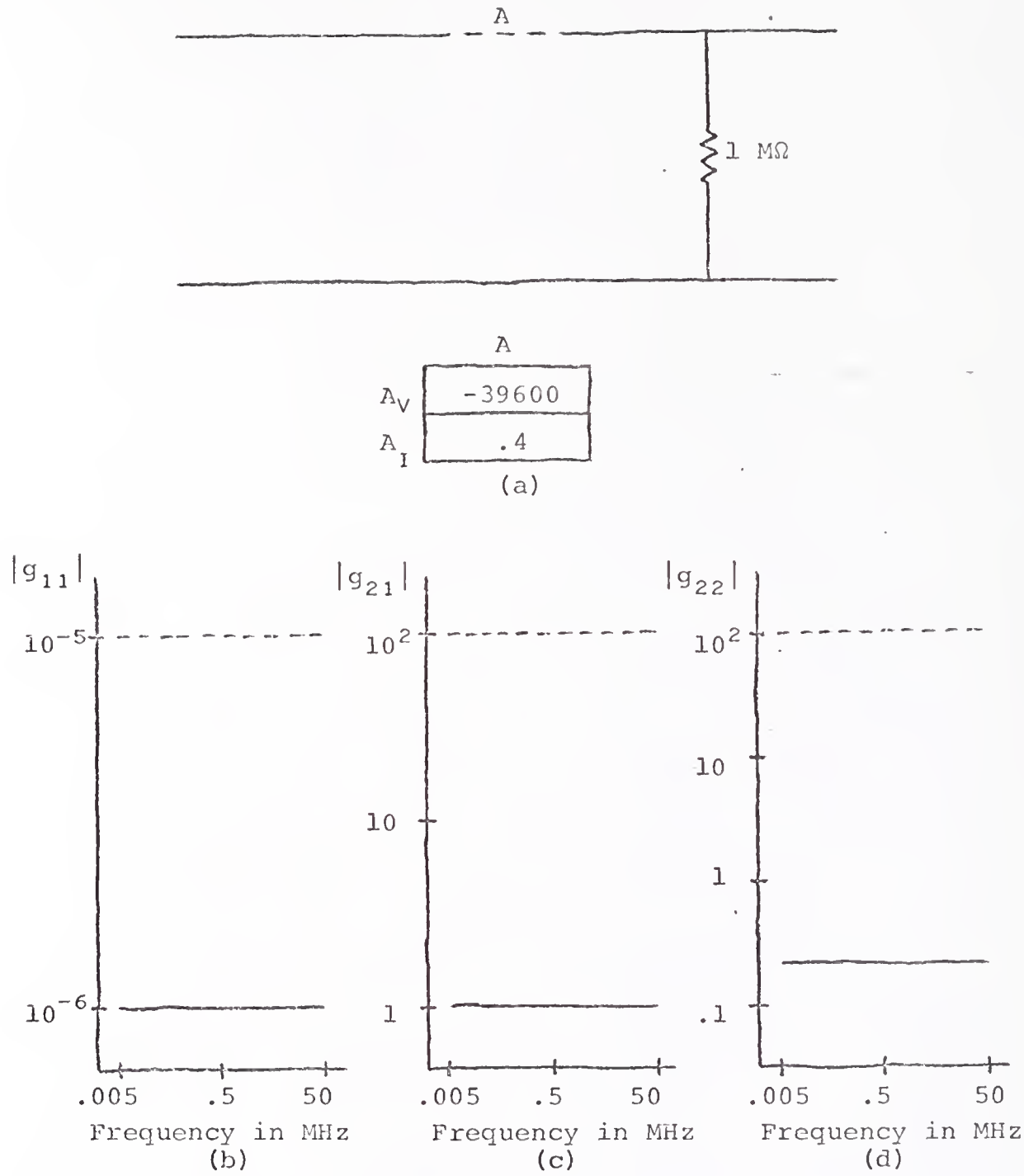
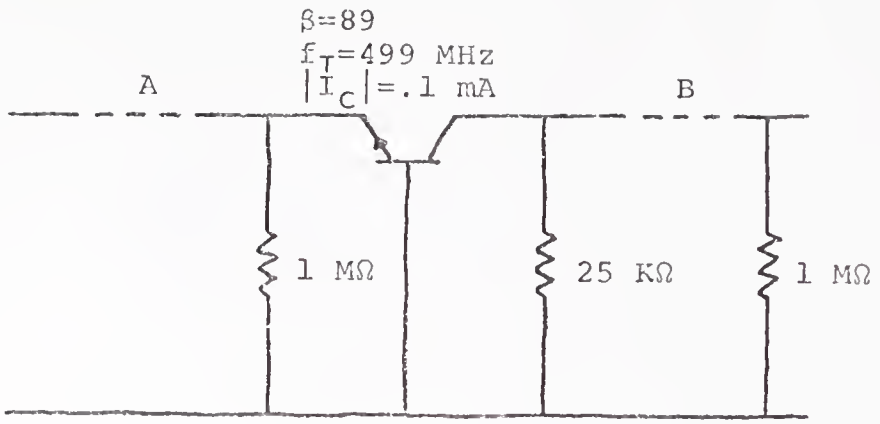


Fig. 4-7

(a) Initial configuration for voltage amplifier example
(b)-(d) Desired (dashed) and actual (solid) responses

a design, the resulting network and responses are shown in Fig. 4-8. The weights were then adjusted to emphasize the input admittance by setting the weights for $\hat{g}_{11}$ to 1000 and those for $\hat{g}_{21}$ and $\hat{g}_{22}$ to 1. Two growth sites were added and growth indicators were calculated. A common-collector stage was inserted in front of the common-base stage and a design was performed. The resulting network and responses are shown in Fig. 4-9. To emphasize the output impedance, the weights were set to 1000 for $\hat{g}_{22}$ and 1 for $\hat{g}_{11}$ and $\hat{g}_{21}$. Two new growth sites were added and growth indicators were calculated. A common-collector stage was grown and a design yielded the network and responses shown in Fig. 4-10. The weights were adjusted to again emphasize the voltage gain and growth indicators were calculated. A common-emitter stage was grown and a design yielded the network and responses shown in Fig. 4-11. In order to obtain a final network that met all specifications more equally, all weights were set to 1 and a design was performed. The final network and its responses are shown in Fig. 4-12. The voltage gain and output impedance specifications are met satisfactorily and the input admittance is lower than desired which is quite acceptable. The total design required 35 seconds of IBM 360/65 computer time.

The same example was run with an alternate order of transistor growths. After the common-base stage was grown, emphasis was placed on output impedance instead of input admittance as before. A common-collector stage grew behind the common-emitter stage and a design yielded the network and responses



	A	B
A_V	6E11	1E8
A_I	-6E11	-1E8

(a)

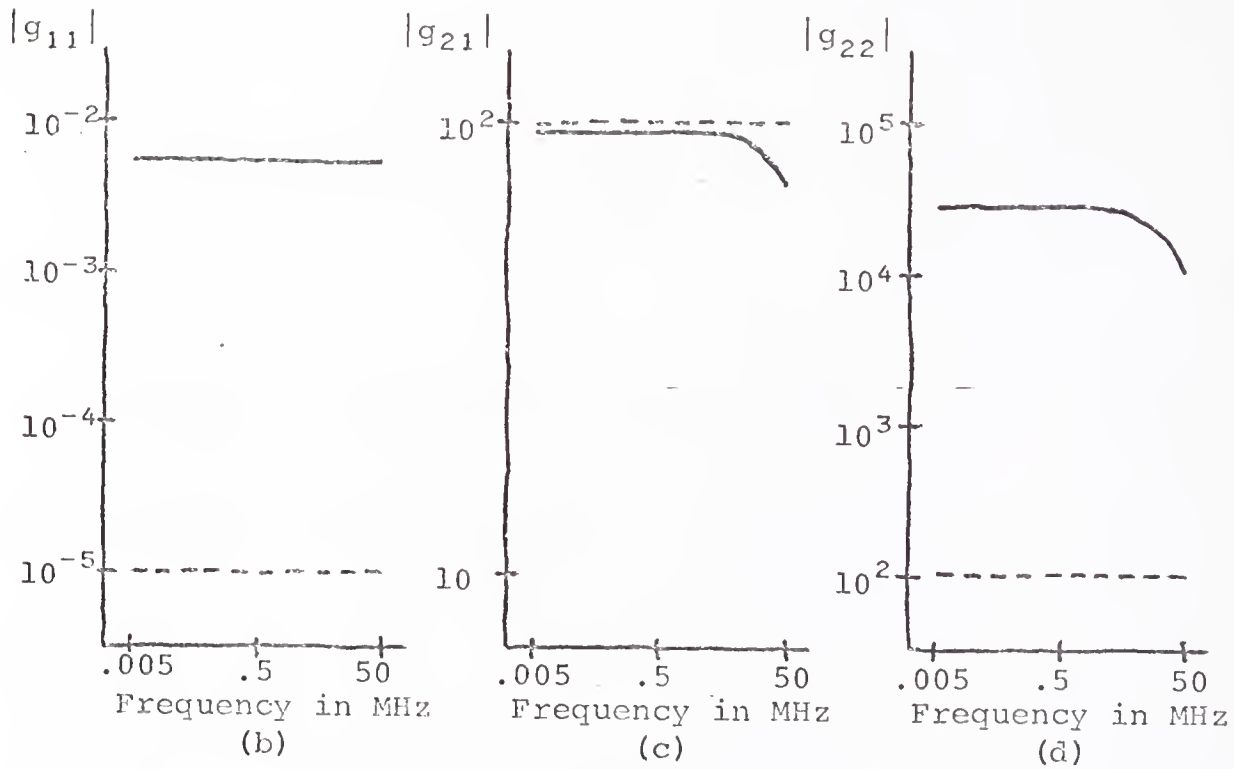
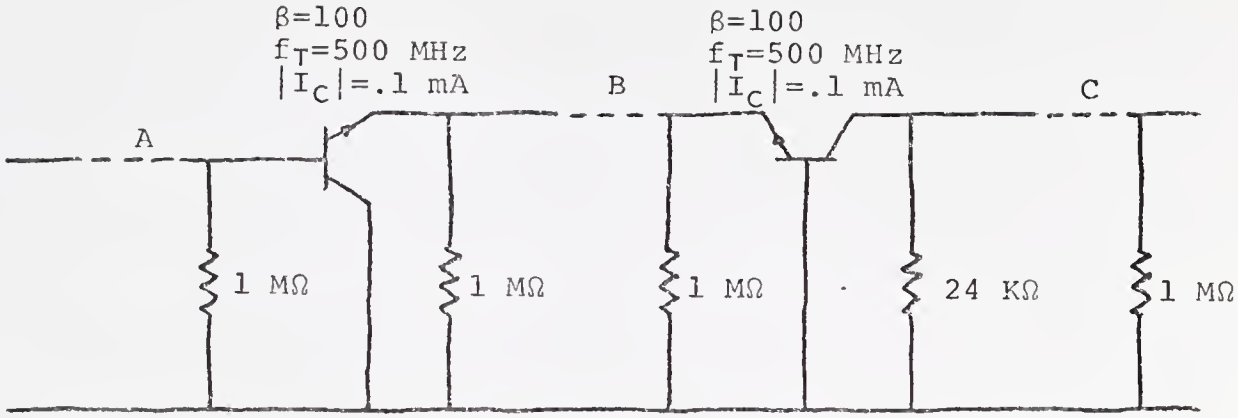


Fig. 4-8

(a) Common-base stage grew to satisfy gain requirement
(b)-(d) Desired (dashed) and actual (solid) responses



	A	B	C
A_V	3633	4E8	2E11
A_I	-6837	-4E8	-2E11

(a)

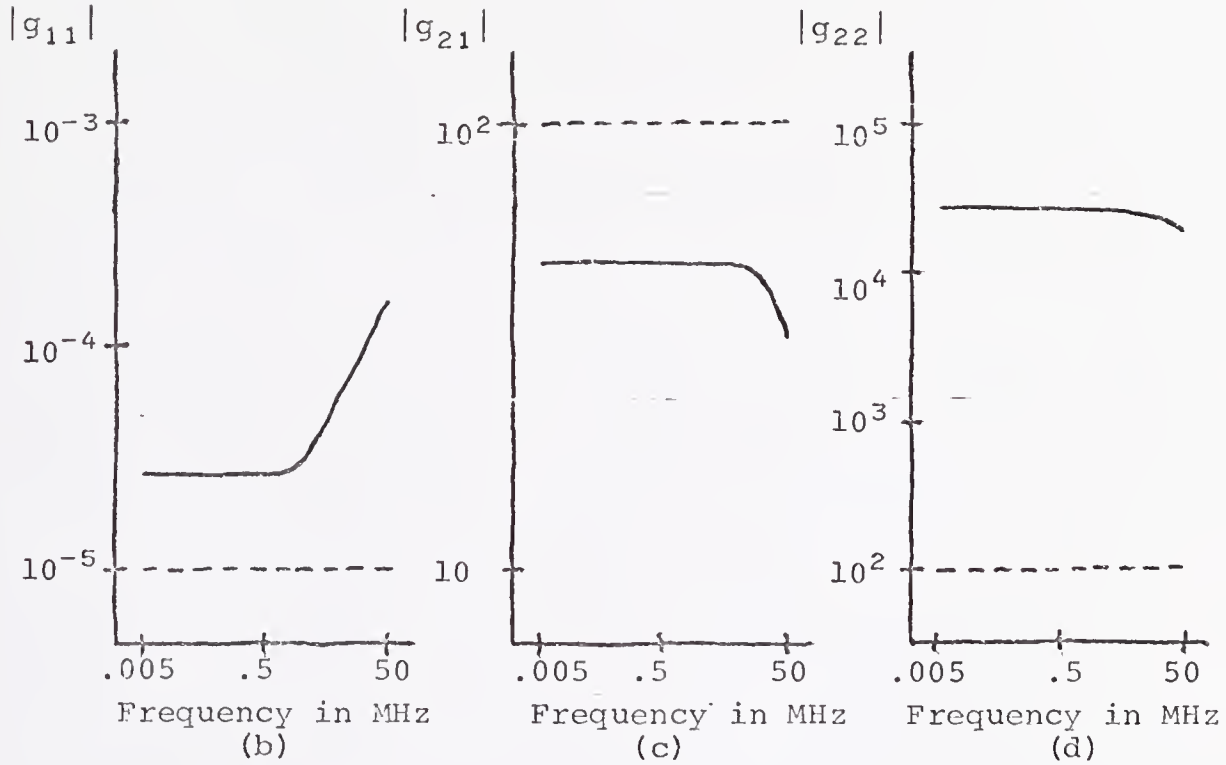
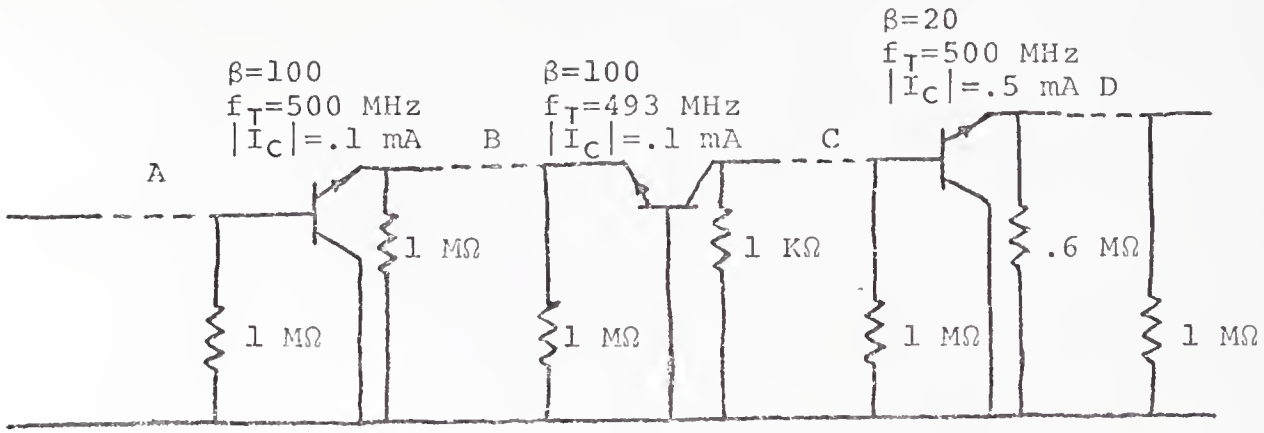


Fig. 4-9

(a) Common-collector stage grew to satisfy input admittance requirement
(b)-(d) Desired (dashed) and actual (solid) responses



	A	B	C	D
A_V	-78646	-38876	-78965	-79103
A_I	-465	-40251	-146	-8

(a)

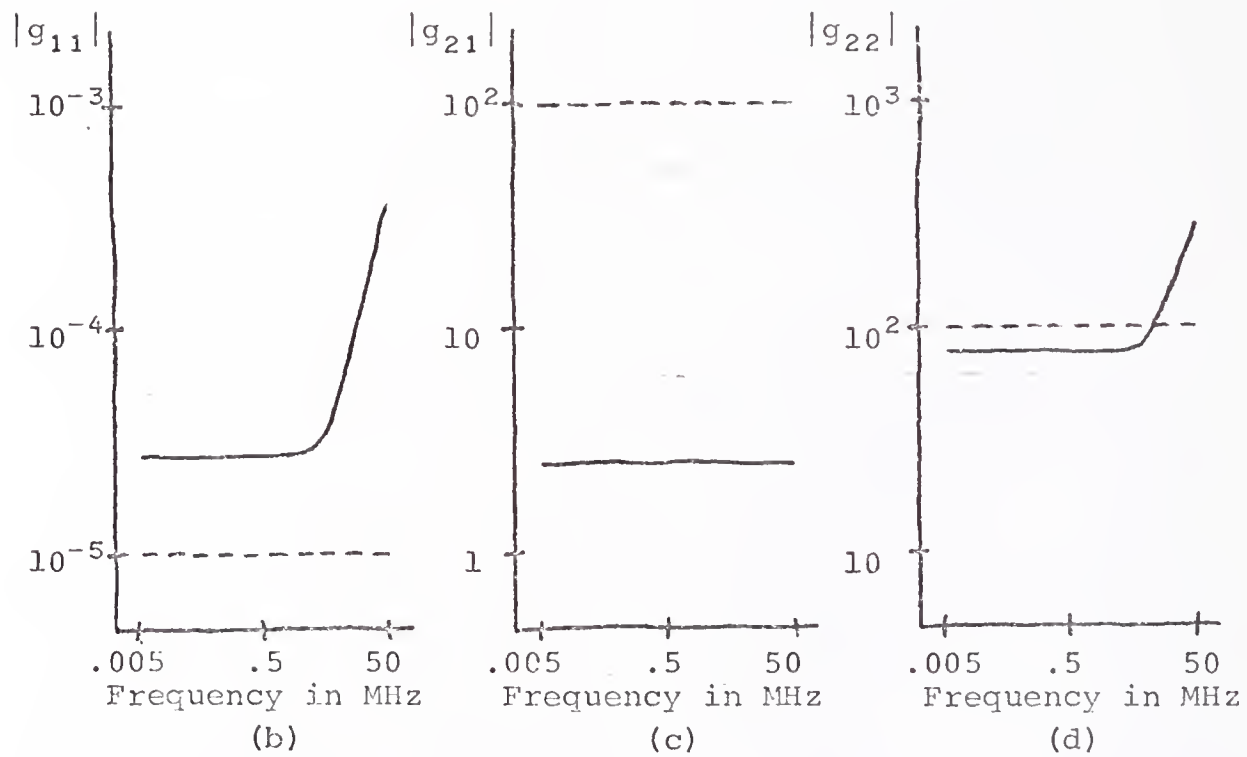


Fig. 4-10

(a) Common-collector stage grown to satisfy output impedance requirement
(b)-(d) Desired (dashed) and actual (solid) responses

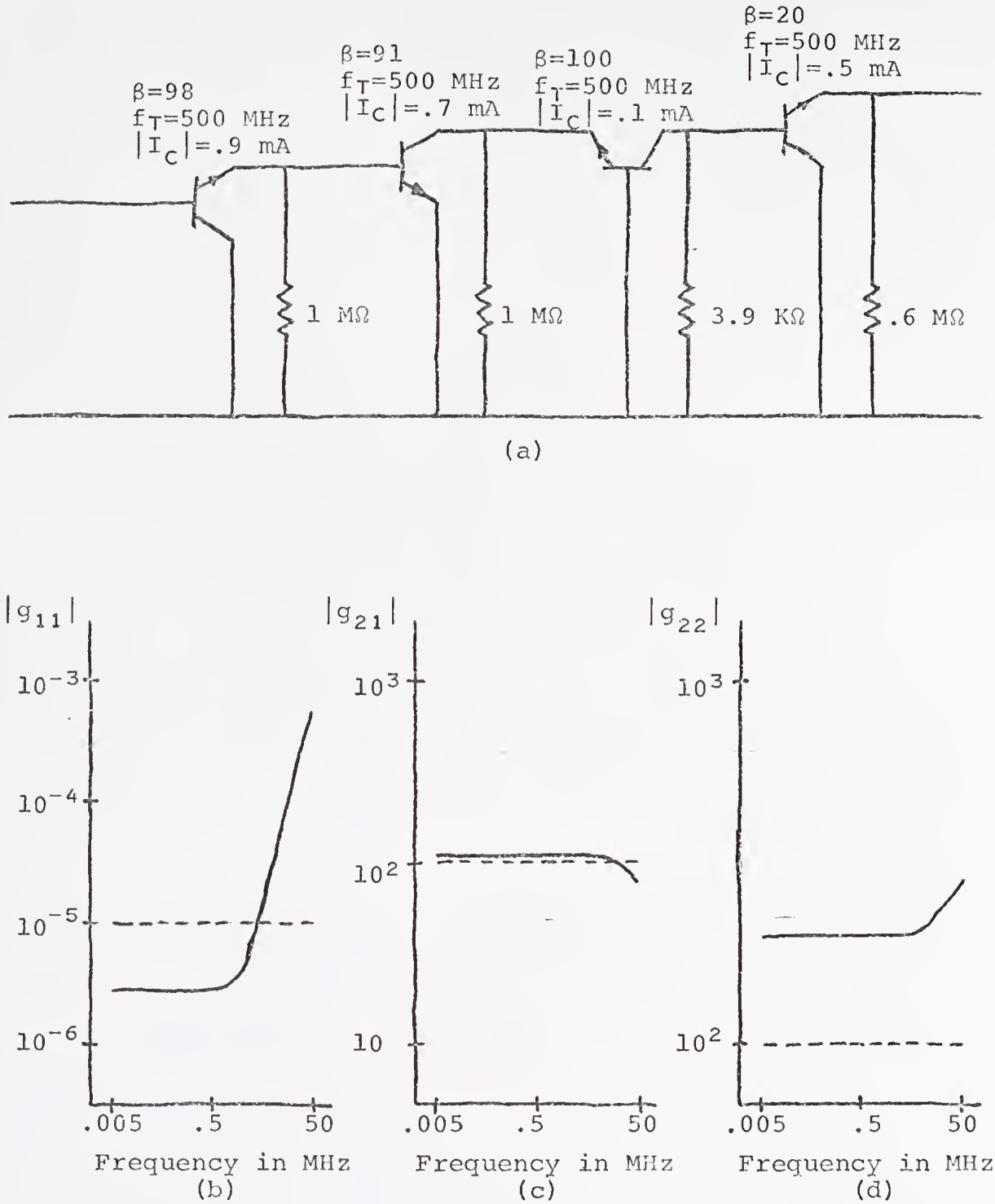
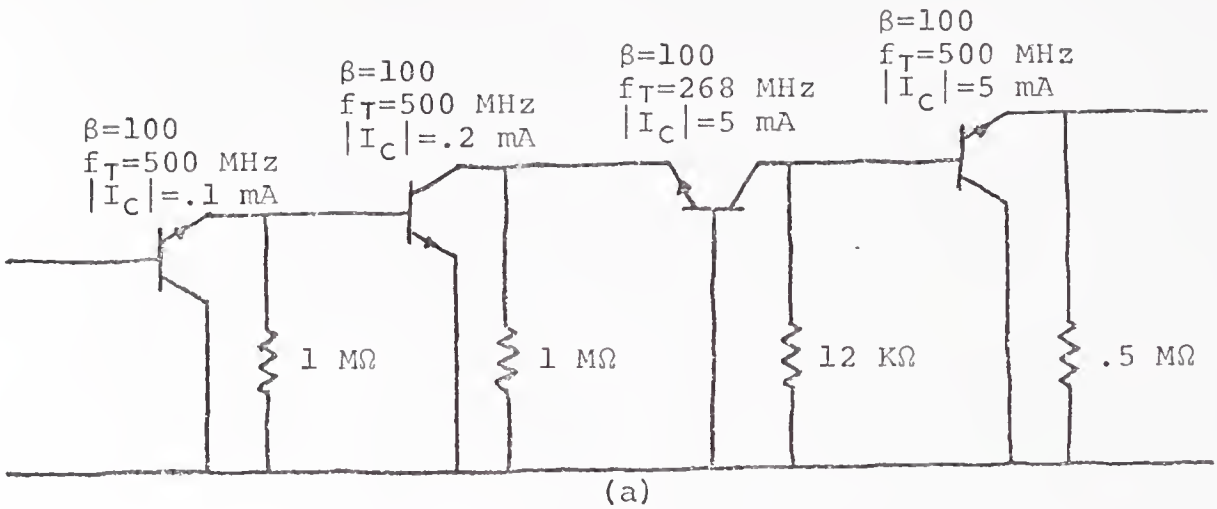


Fig. 4-11

(a) Common-emitter stage grew to satisfy gain requirement
(b)-(d) Desired (dashed) and actual (solid) responses



(a)

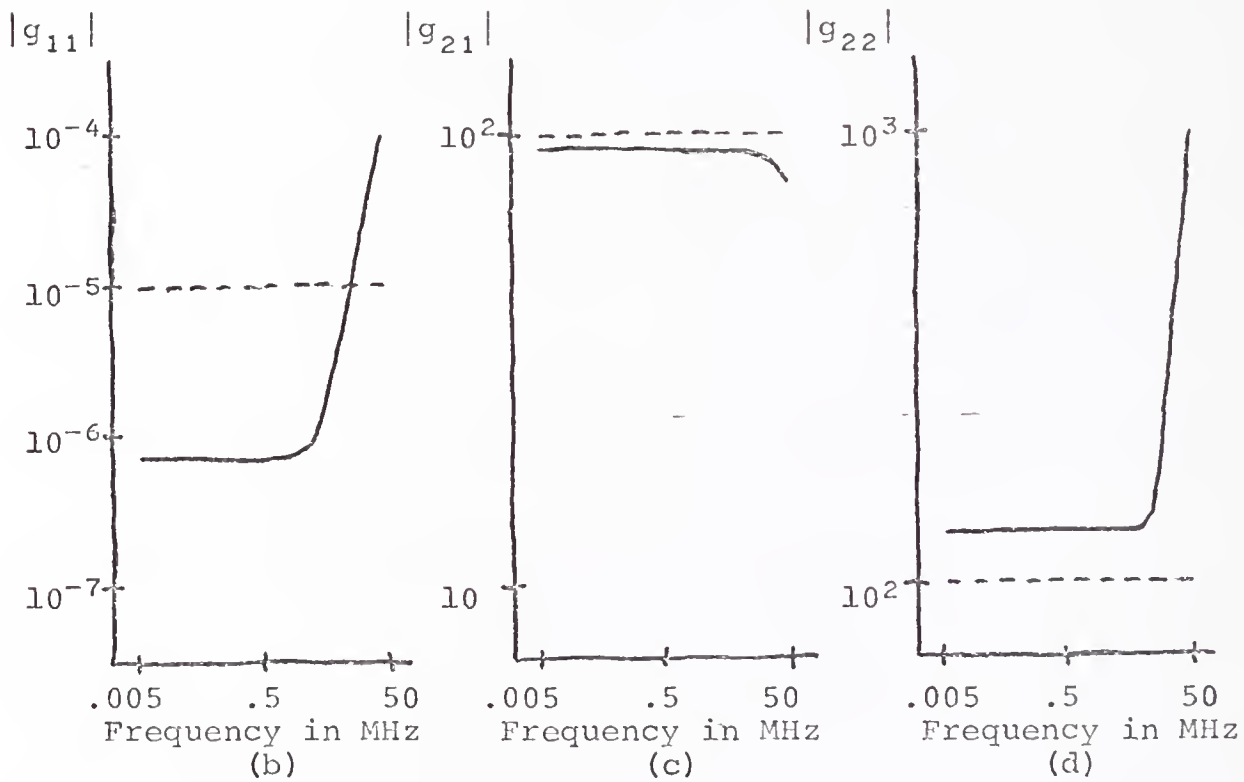


Fig. 4-12

(a) Final configuration of voltage amplifier example
(b)-(d) Desired (dashed) and actual (solid) responses

shown in Fig. 4-13. More growth sites were added and input admittance was then emphasized. A common-collector was grown and a design yielded the network and responses shown in Fig. 4-14. As with the previous example, voltage gain was again emphasized in a run and then equal weights were applied. The final network and responses were nearly identical to those shown in Fig. 4-12 for the specifications emphasized in the alternate order.

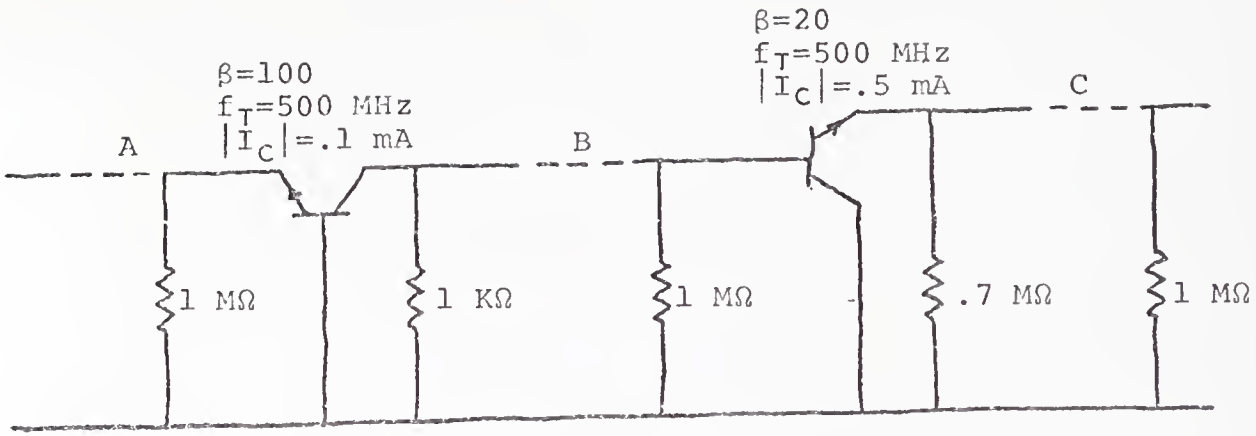
The dual of the voltage amplifier, a current gain amplifier, is now obtained using the same interactive procedure. Specifications are placed on input impedance, current gain, and output admittance. It is convenient to express these specifications in terms of the hybrid h-parameters so that the port variables of the amplifier are given by

$$V_1(\omega) = h_{11}(\omega) I_1(\omega)$$

and

$$I_2(\omega) = h_{21}(\omega) I_1(\omega) + h_{22}(\omega) V_2(\omega).$$

The desired responses, $\hat{h}_{11}(\omega)$, $\hat{h}_{21}(\omega)$, and $\hat{h}_{22}(\omega)$, along with the actual responses and initial network configuration are shown in Fig. 4-15. The weights of input impedance and output admittance were set to 1 and the weights of current gain were set to 1000 initially. A common-collector stage grew and a design yielded the network and responses in Fig. 4-16. Input impedance was emphasized to yield the network and responses shown in Fig. 4-17. Emphasis on output admittance



	A	B	C
A_V	6E11	7E5	12820
A_I	-6E11	-7E5	-12821

(a)

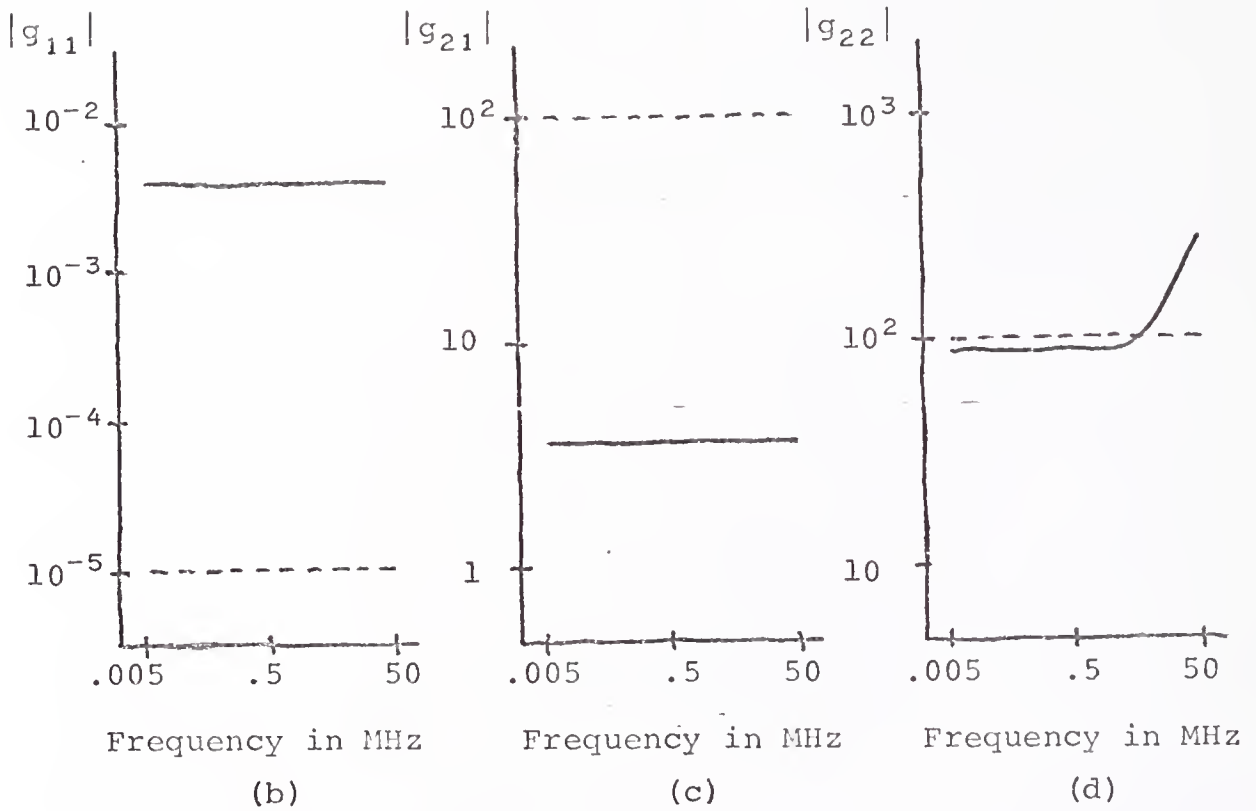
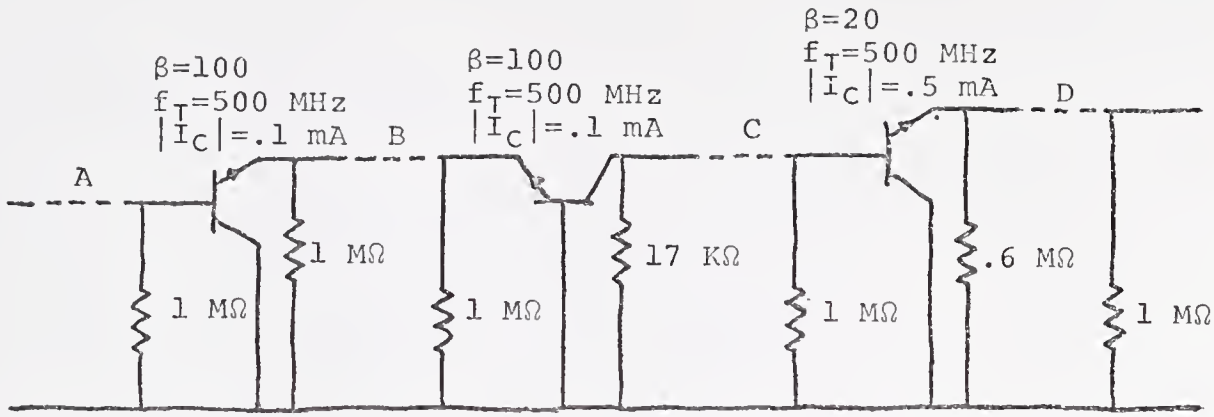


Fig. 4-13

(a) Common-collector stage grew to satisfy output impedance requirement
(b)-(d) Desired (dashed) and actual (solid) responses



	A	B	C	D
A_V	-820524	-411508	-760100	-819904
A_I	-435	-409610	-60857	-1053

(a)

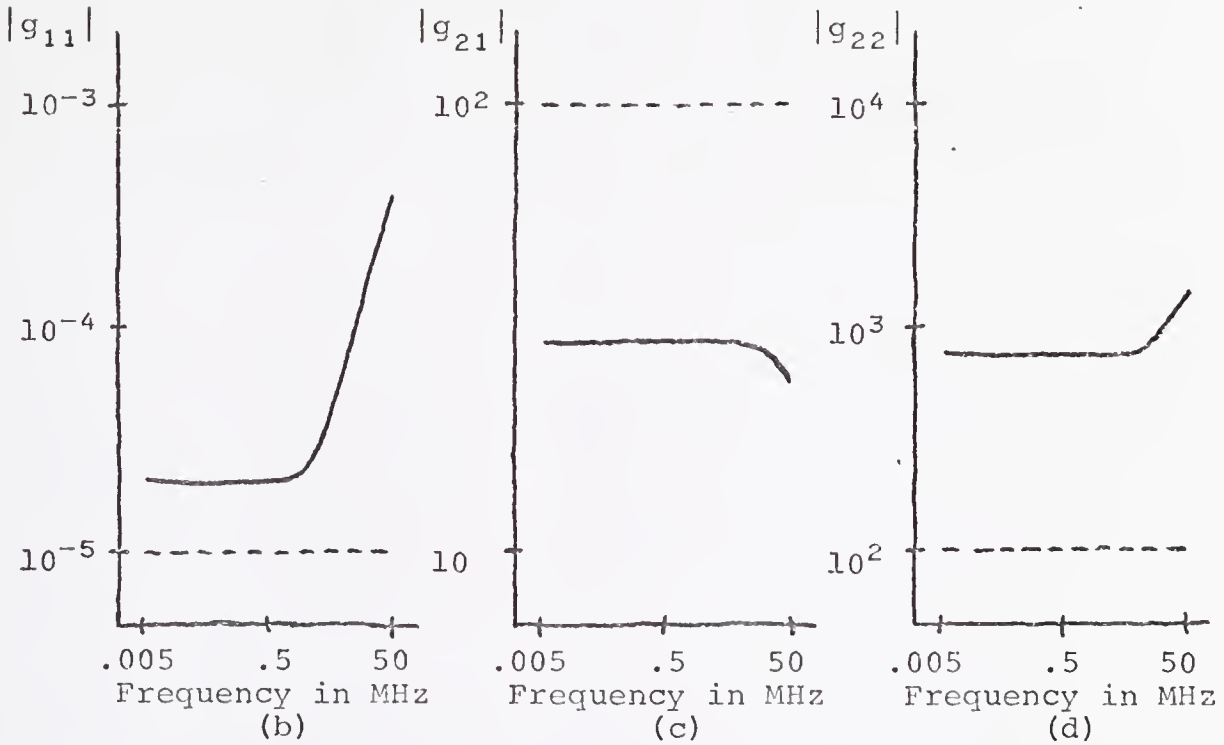


Fig. 4-14

(a) Common-collector stage grown to satisfy input admittance requirement
(b)-(d) Desired (dashed) and actual (solid) responses

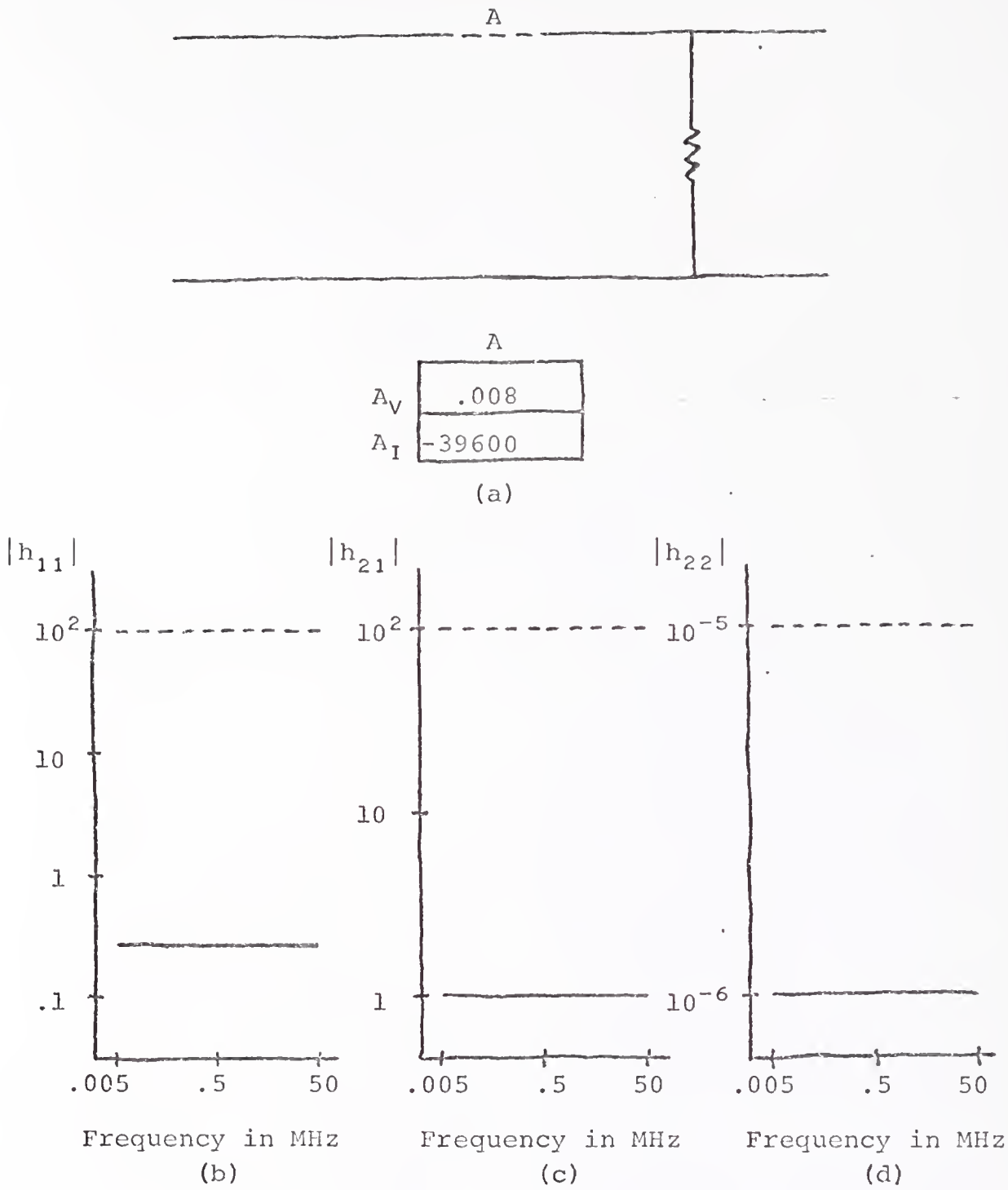


Fig. 4-15

(a) Initial configuration for current amplifier example
(b)-(d) Desired (dashed) and actual (solid) responses

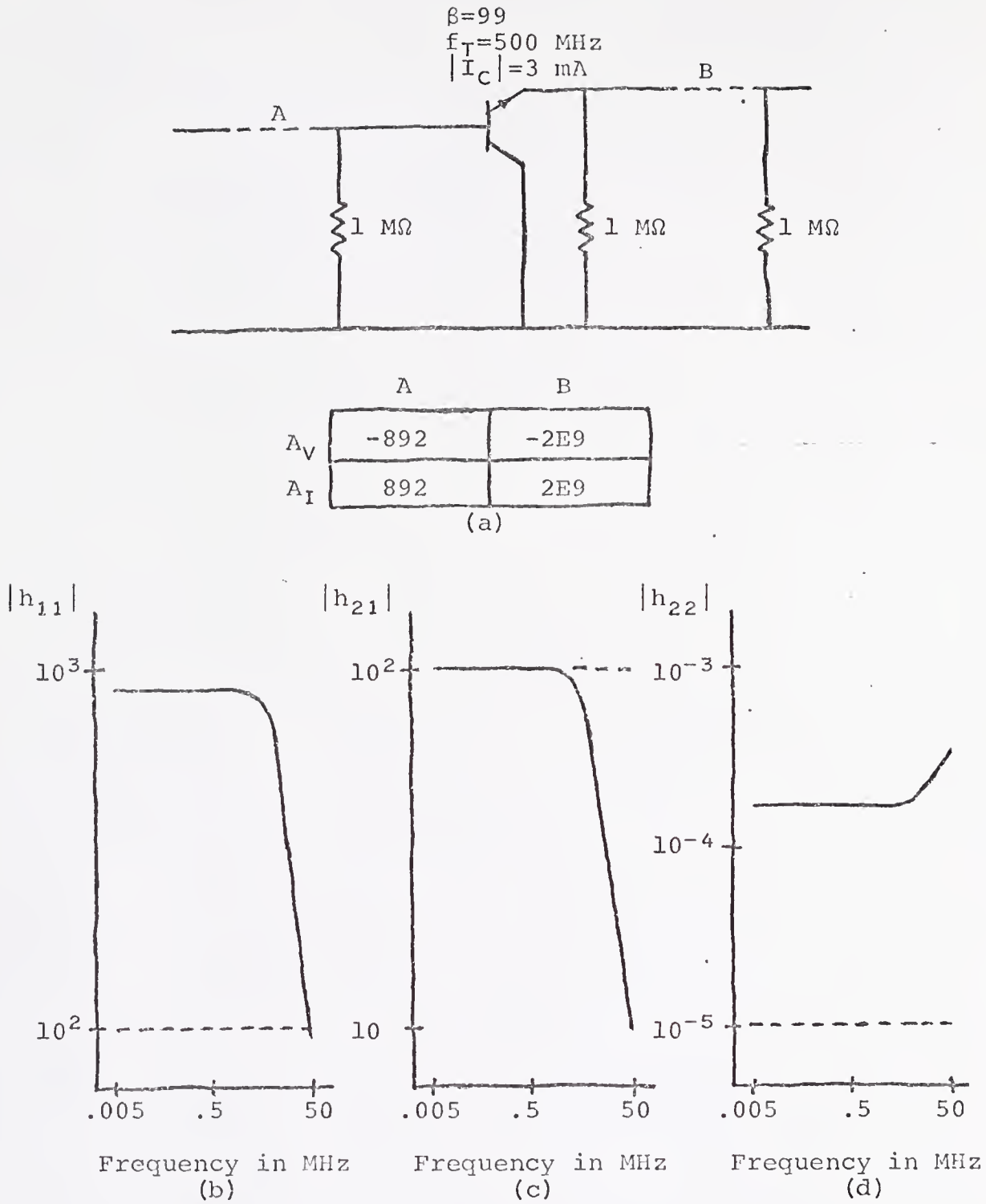
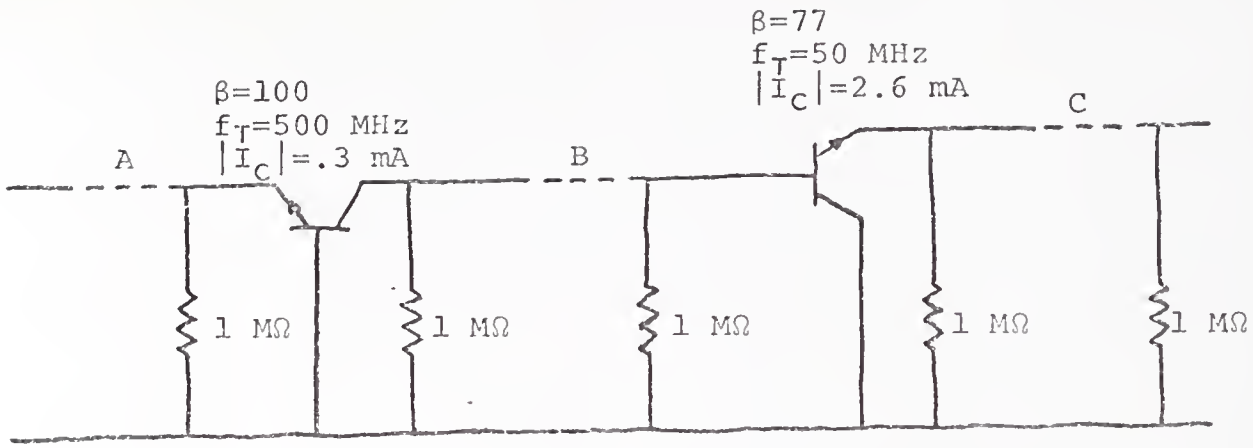


Fig. 4-16

(a) Common-collector stage grew to satisfy gain requirement
(b)-(d) Desired (dashed) and actual (solid) responses



	A	B	C
A_V	-.06	-4E8	-8E8
A_I	-.6	4E8	8E8

(a)

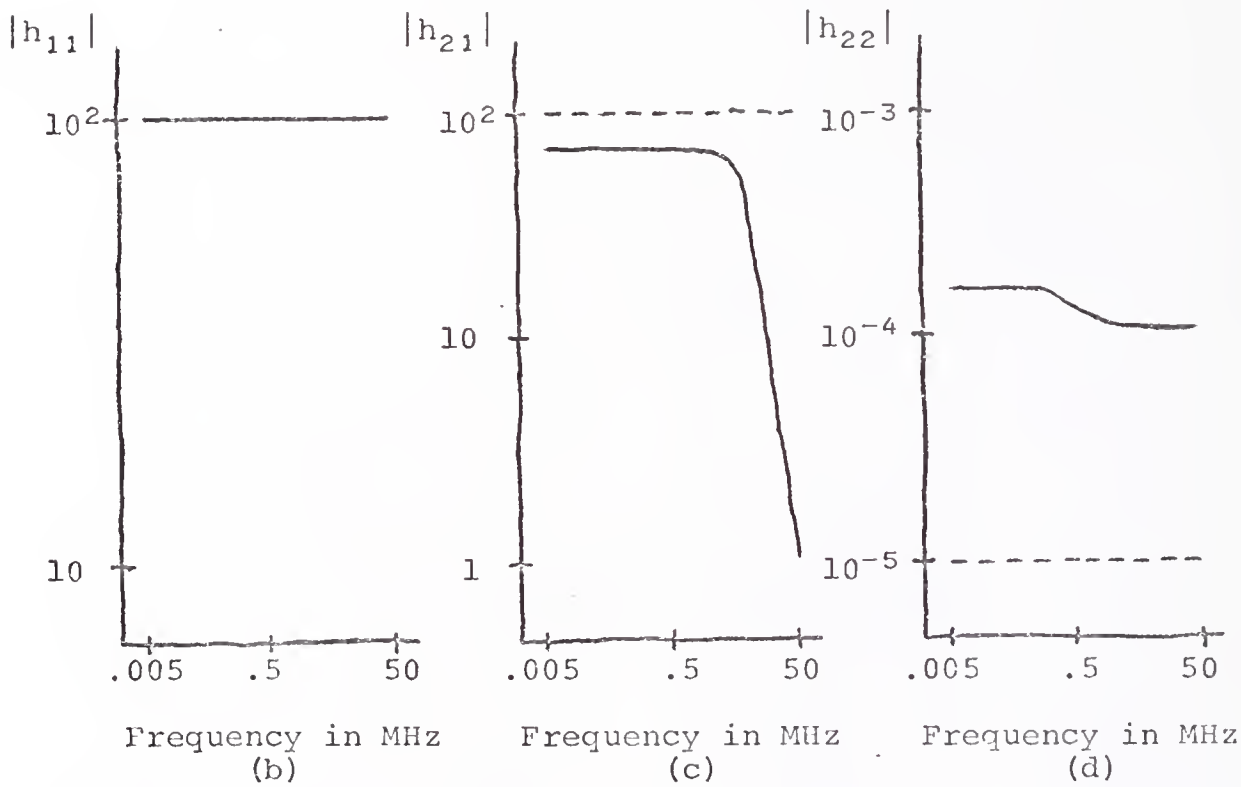


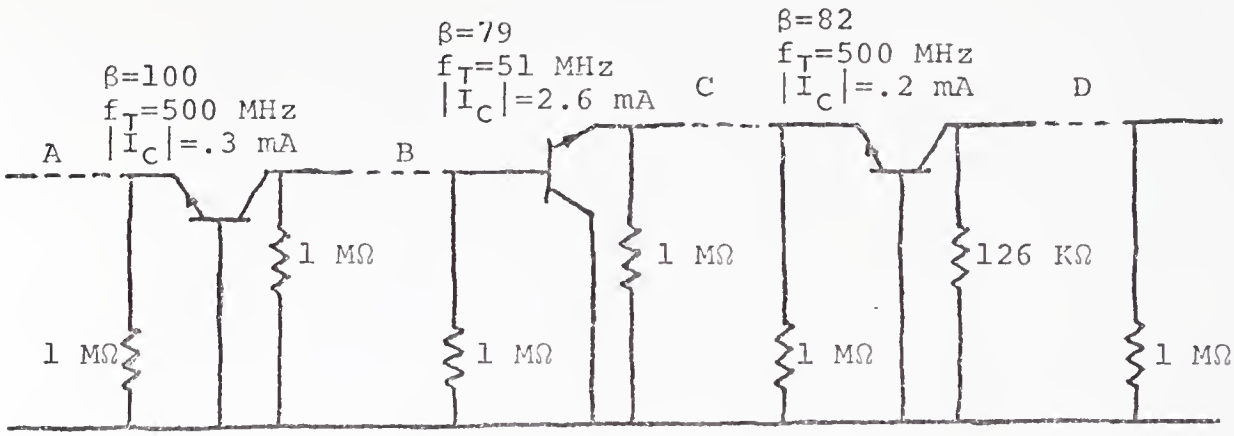
Fig. 4-17

(a) Common-base stage grew to satisfy input impedance requirement
(b)-(d) Desired (dashed) and actual (solid) responses

then yielded the network and responses shown in Fig. 4-18. The current gain specification was then emphasized again to yield the network and responses in Fig. 4-19. To obtain a network that met all specifications equally, weights of one were used for all desired responses to yield the network and responses of Fig. 4-20. The complete design required 19 seconds of IBM 360/65 computer time.

The order in which the input and output specifications were emphasized was reversed for the design of the current amplifier. The network and responses shown in Fig. 4-21 were obtained when output admittance was emphasized first. After which input impedance was emphasized to obtain the network and responses shown in Fig. 4-22. Further emphasis of current gain and a final run with balanced weights yielded the same results that are shown in Fig. 4-20.

Since transistor removal was not exercised in the preceding examples, networks were constructed that required removal to demonstrate the removal indicators. A fifth transistor was inserted into a version of the four-transistor voltage amplifier and a design was performed to yield the network and responses shown in Fig. 4-23. At this point, all transistors were designated removable and removal indicators were calculated. Transistor four was removed and replaced by a short circuit between the base node and the collector node. A design was performed that yielded the network and responses shown in Fig. 4-24. The desired responses are satisfactorily met as would be expected since the transistor that was removed



A_V	-.02	-5578	-11842	-8
A_I	-671041	-665462	-659209	-671033

(a)

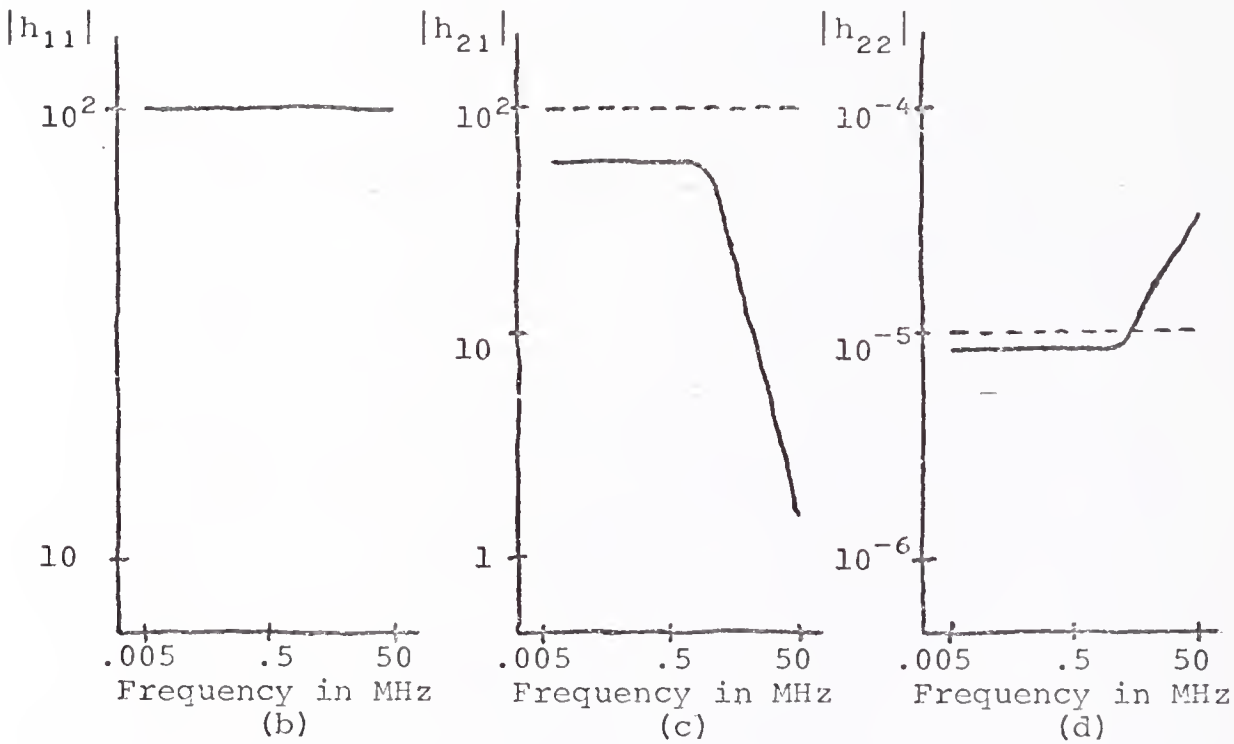


Fig. 4-18

(a) Common-base stage grown to satisfy output admittance requirement
(b)-(d) Desired (dashed) and actual (solid) response

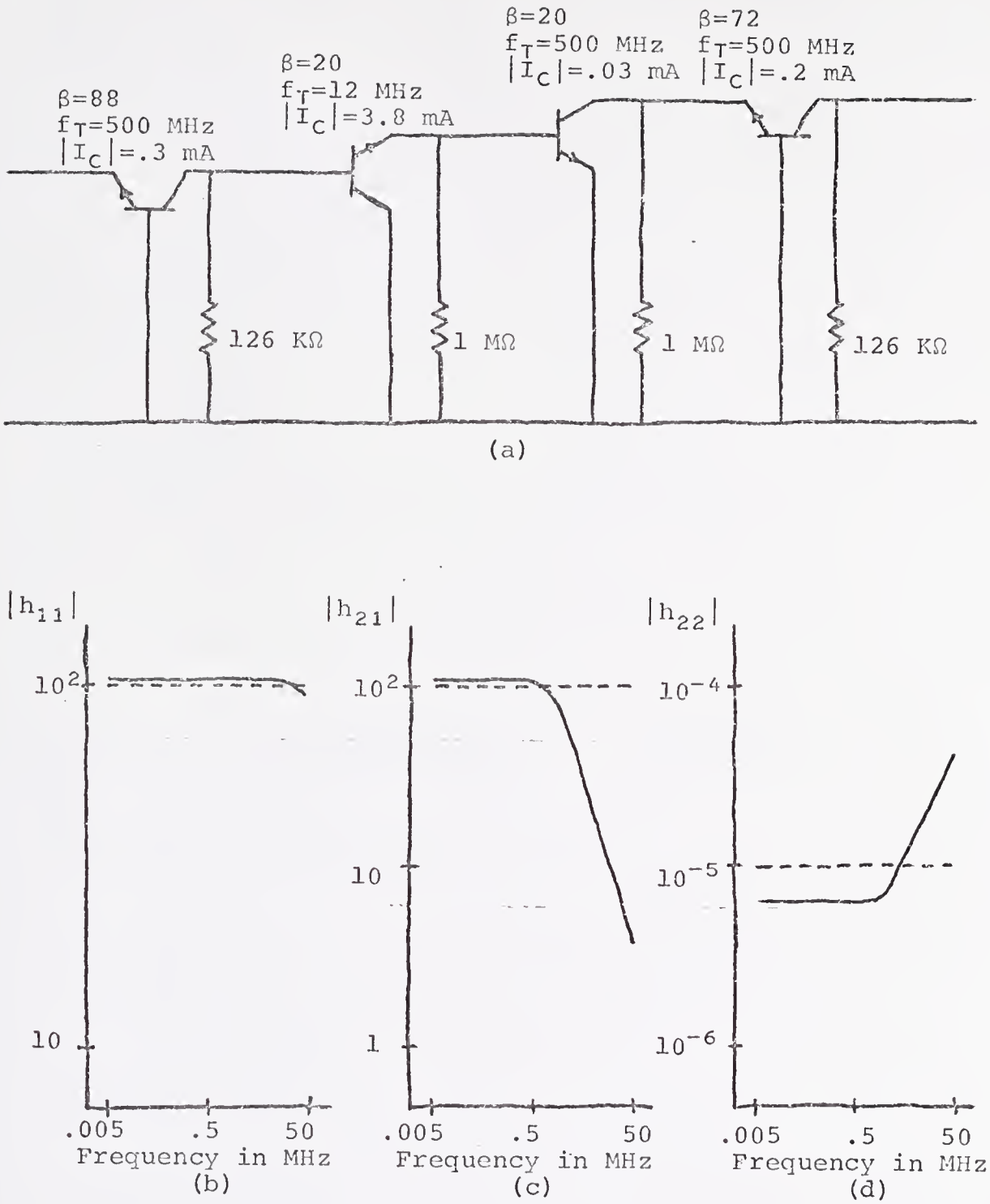
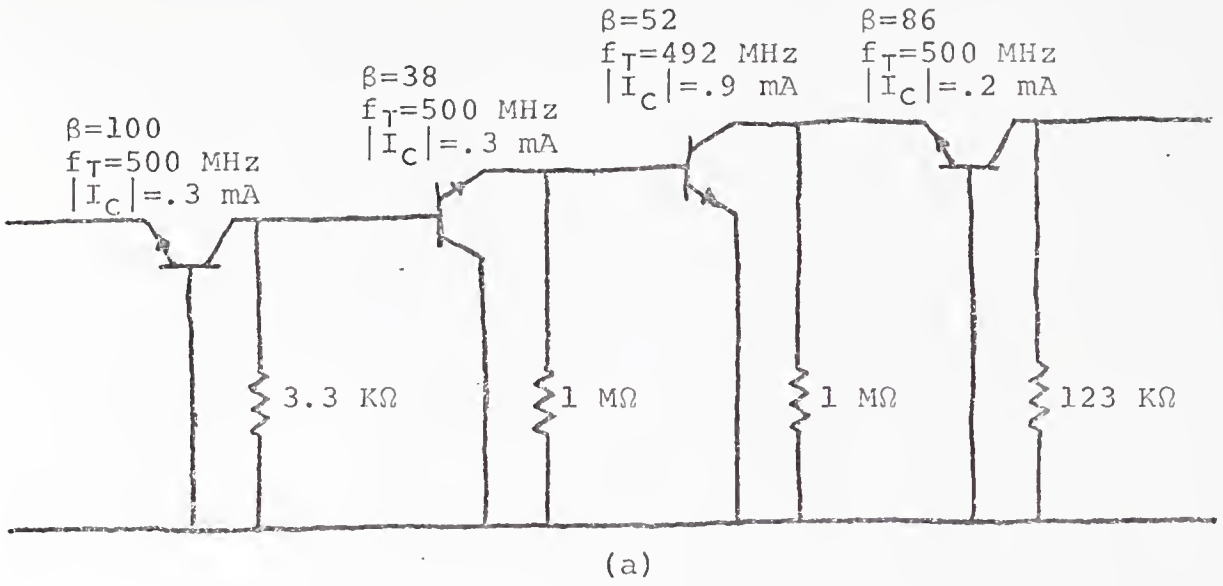
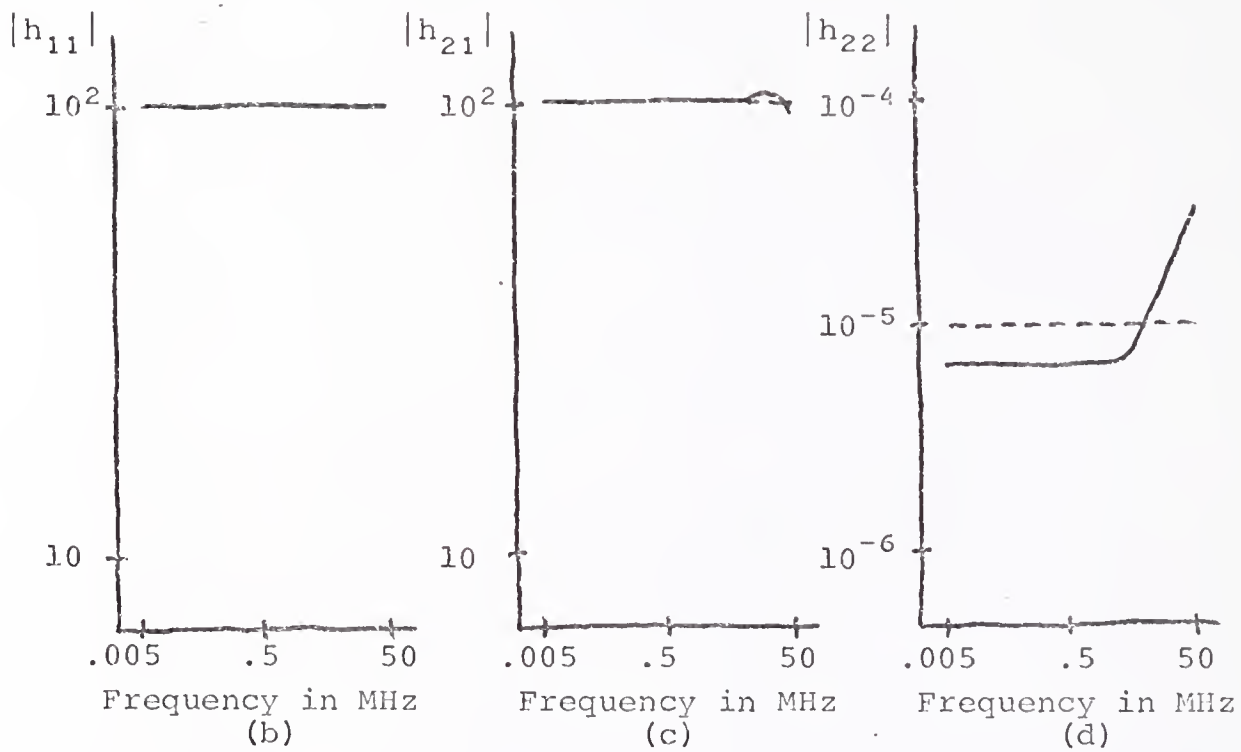


Fig. 4-19

(a) Common-emitter stage grew to satisfy gain requirement
(b)-(d) Desired (dashed) and actual (solid) responses



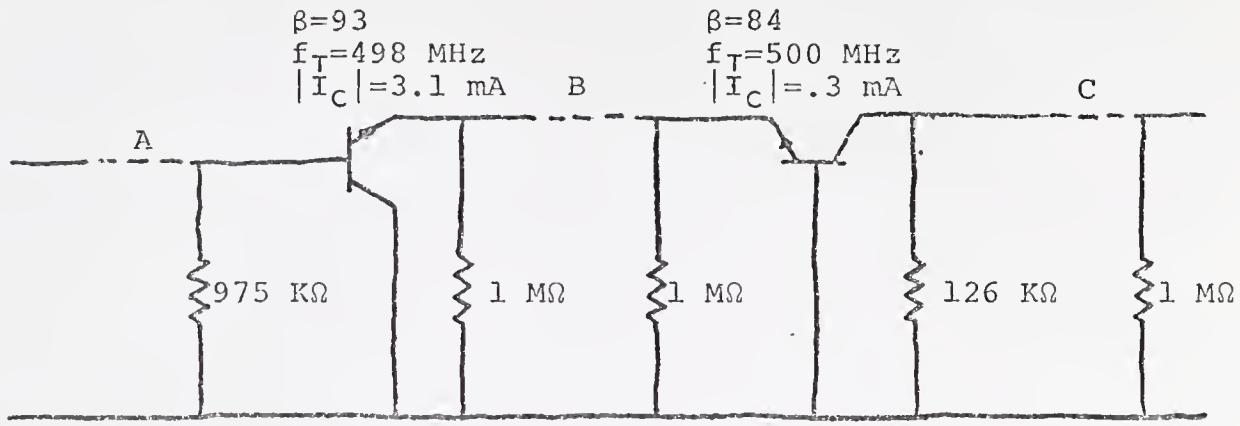
(a)



(b) (c) (d)

Fig. 4-20

(a) Final configuration for current amplifier example
(b)-(d) Desired (dashed) and actual (solid) responses



	A	B	C
A_V	-3E10	-2E10	-5060
A_I	3E10	2E10	2533

(a)

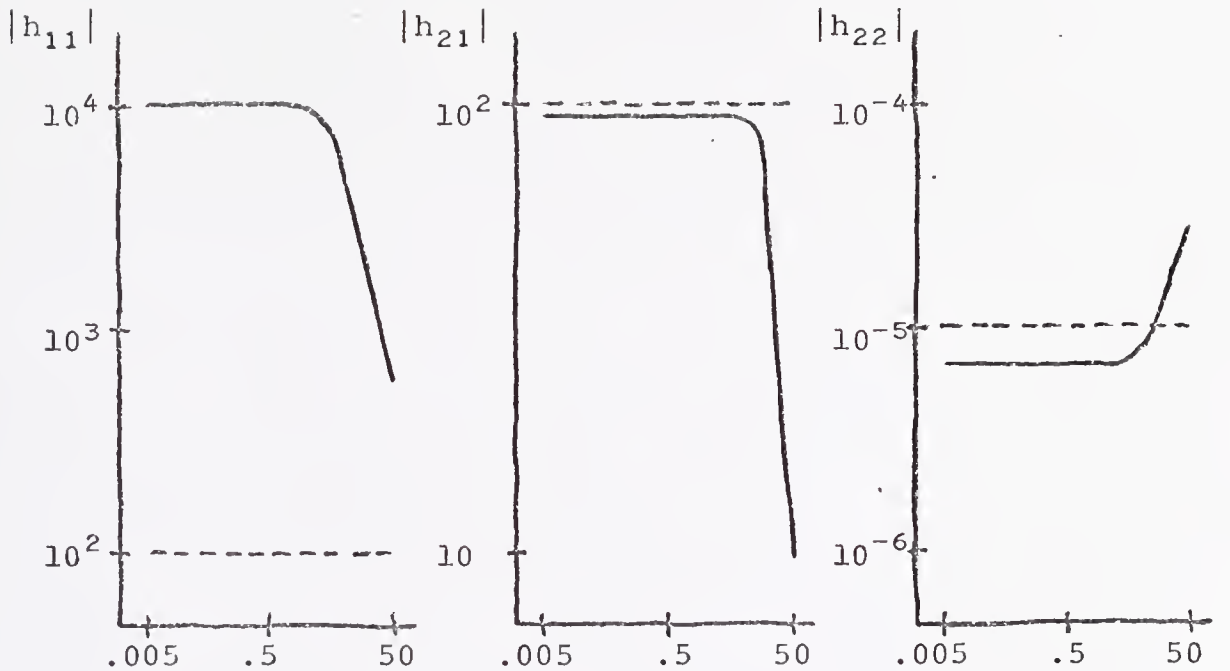
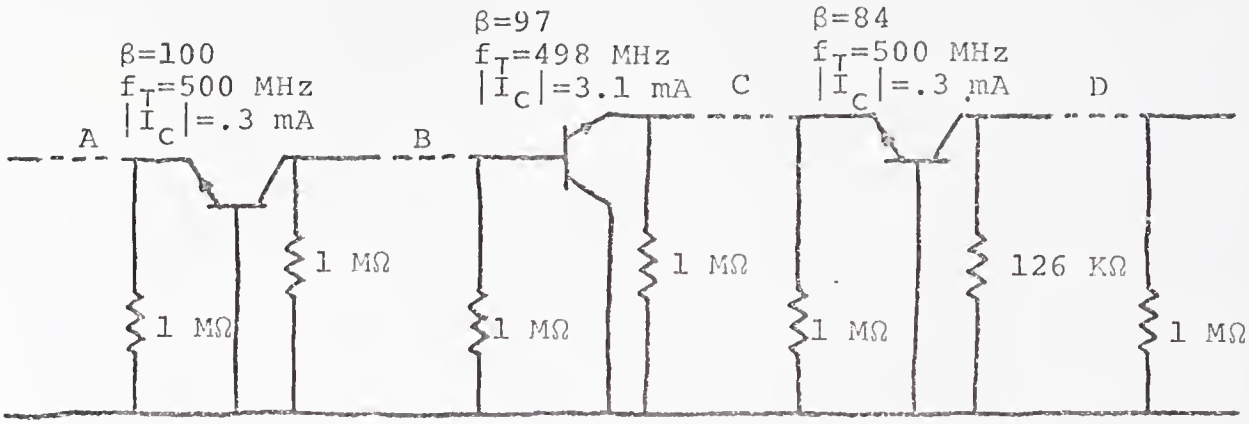


Fig. 4-21

(a) Common-base stage grew to satisfy output admittance requirement
(b)-(d) Desired (dashed) and actual (solid) responses



	A	B	C	D
A_V	-.004	-4588	-9048	-8
A_I	-276274	-271686	-267236	-276267

(a)

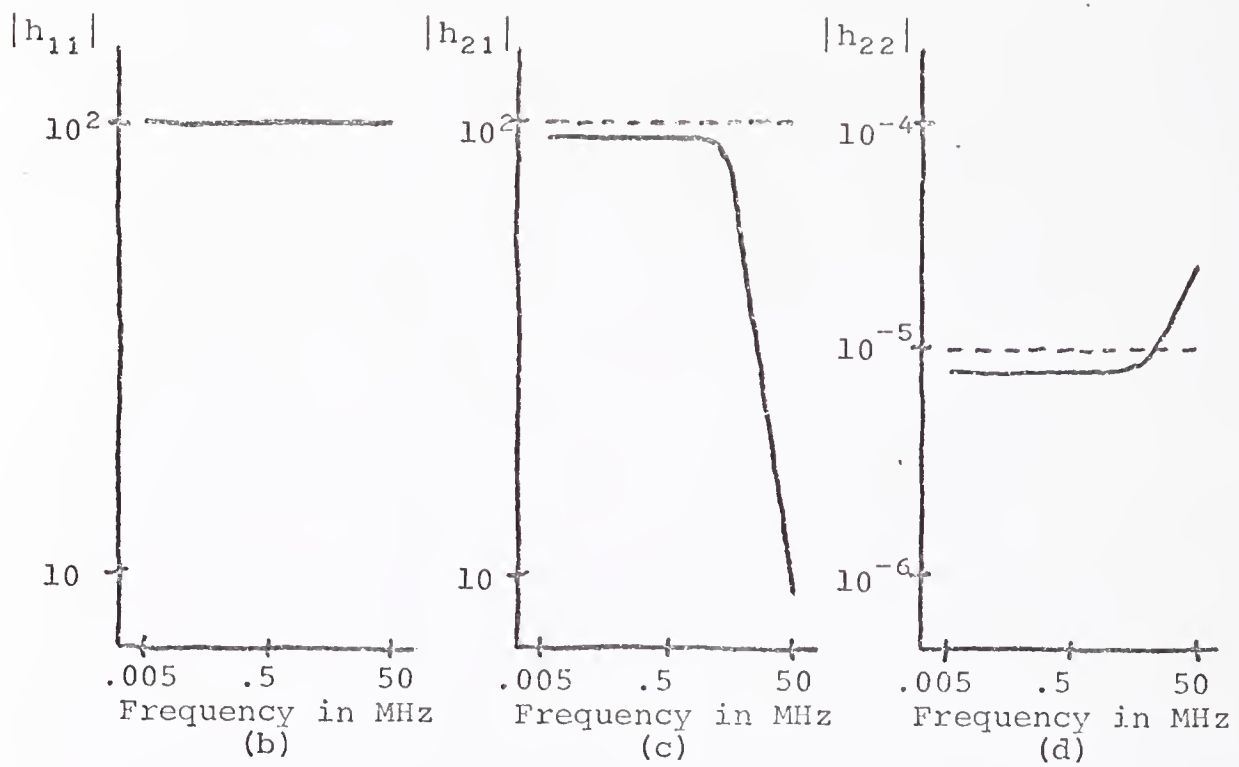
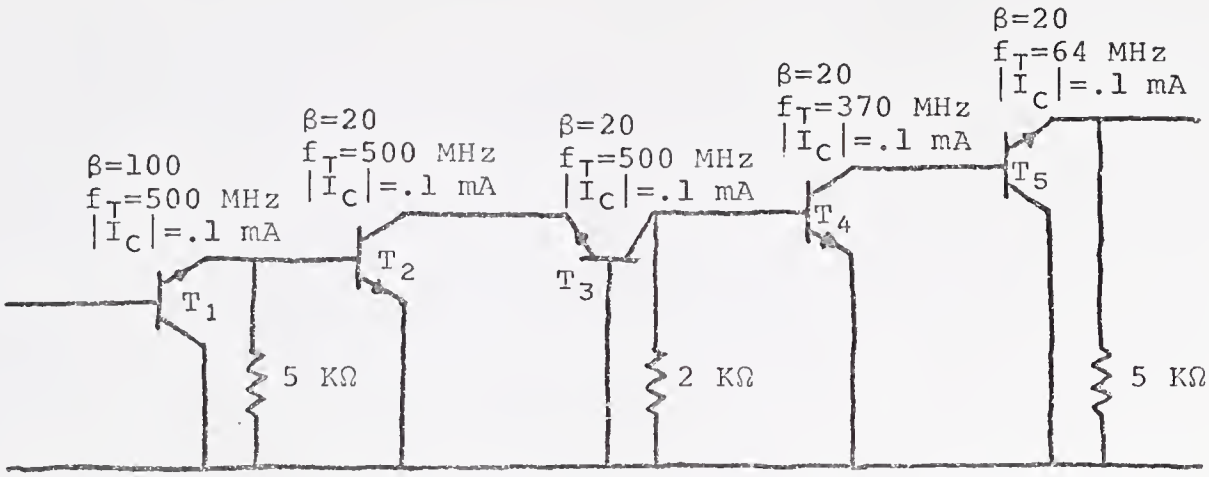


Fig. 4-22

(a) Common-base stage grown to satisfy input impedance requirement
(b)-(d) Desired (dashed) and actual (solid) responses



	T_1	T_2	T_3	T_4	T_5
EB	3E7	-3E8	-4E8	-2E9	-6E9
BC	-7E7	-1E9	-2E9	-9E11	-1E11
CE	-3E8	-4E8	-7E7	-1E11	-6E9

(a)

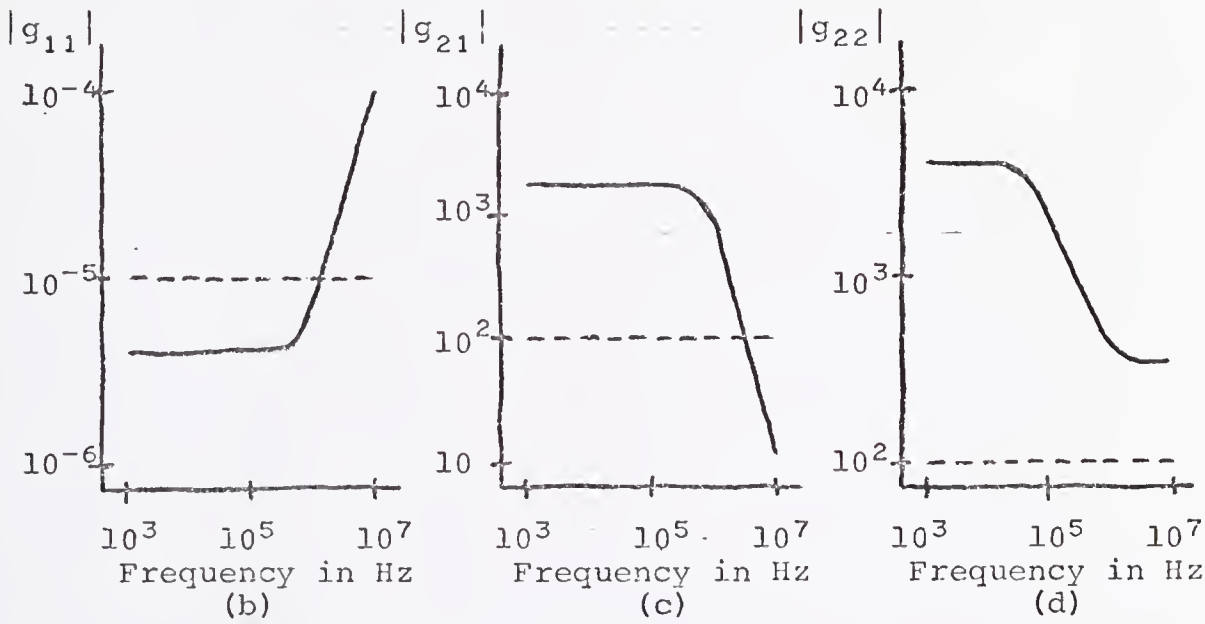


Fig. 4-23

(a) Initial configuration of voltage amplifier with extra transistor
(b)-(d) Desired (dashed) and actual (solid) responses

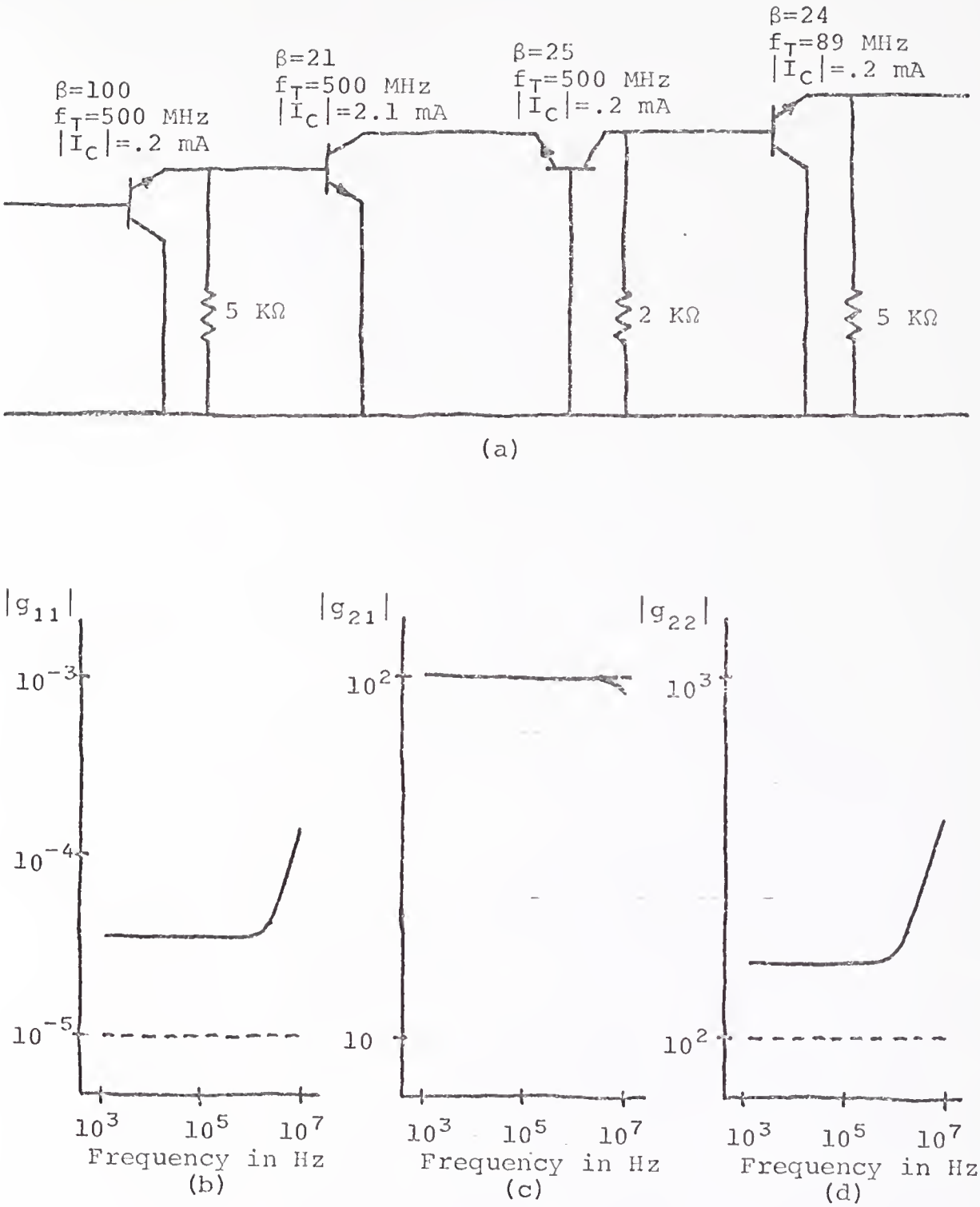


Fig. 4-24

(a) Final configuration of voltage amplifier after transistor had been removed
(b)-(d) Desired (dashed) and actual (solid) responses

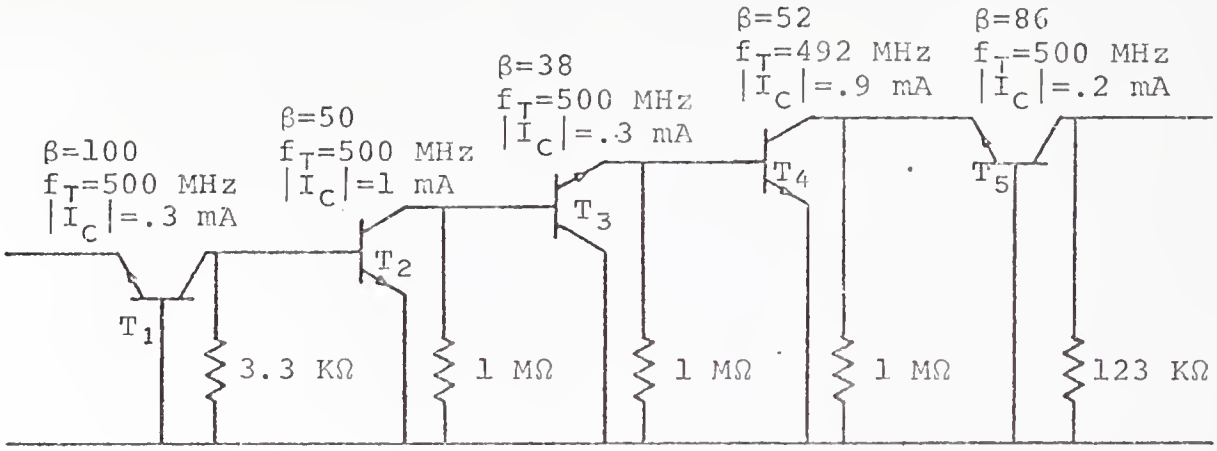
was the one that had been inserted and it was replaced properly with a short-open combination. The total design required 6 seconds of IBM 360/65 computer time.

A transistor was also inserted into a version of the four-transistor current amplifier previously considered. The network, its responses, and removal indicators for each transistor are shown in Fig. 4-25. Removal of transistor two is suggested and performed, after which a design is performed to yield the network and responses in Fig. 4-26. The desired responses are again met satisfactorily. The total design required 4 seconds of IBM 360/65 computer time.

The growth indicators have been shown to accurately predict where to insert transistors and in what configuration to put them. Although the principle upon which the indicators are based, the ability of a transistor to supply gain, is simple, it is used here to solve the complex problem of adding a transistor to a complicated network.

The removal of a transistor from a network is not as simple. A method has been proposed and examples have been described which support its ability. But a problem does exist with the method; it strongly prefers to remove common-emitter stages. This preference is not understood but hopefully an improved method will be available in the future.

For a complete design we still need a suitable dc bias network to supply the necessary dc collector currents. This design typically requires the ability to grow additional transistors to meet dc specifications. The present algorithm



	T_1	T_2	T_3	T_4	T_5
EB	-5E9	-4E10	-1E11	-5E10	-7E9
BC	-4E10	-7E13	-2E12	-3E11	-6E6
CE	-4E8	-2E12	-5E10	-7E9	9E7

(a)

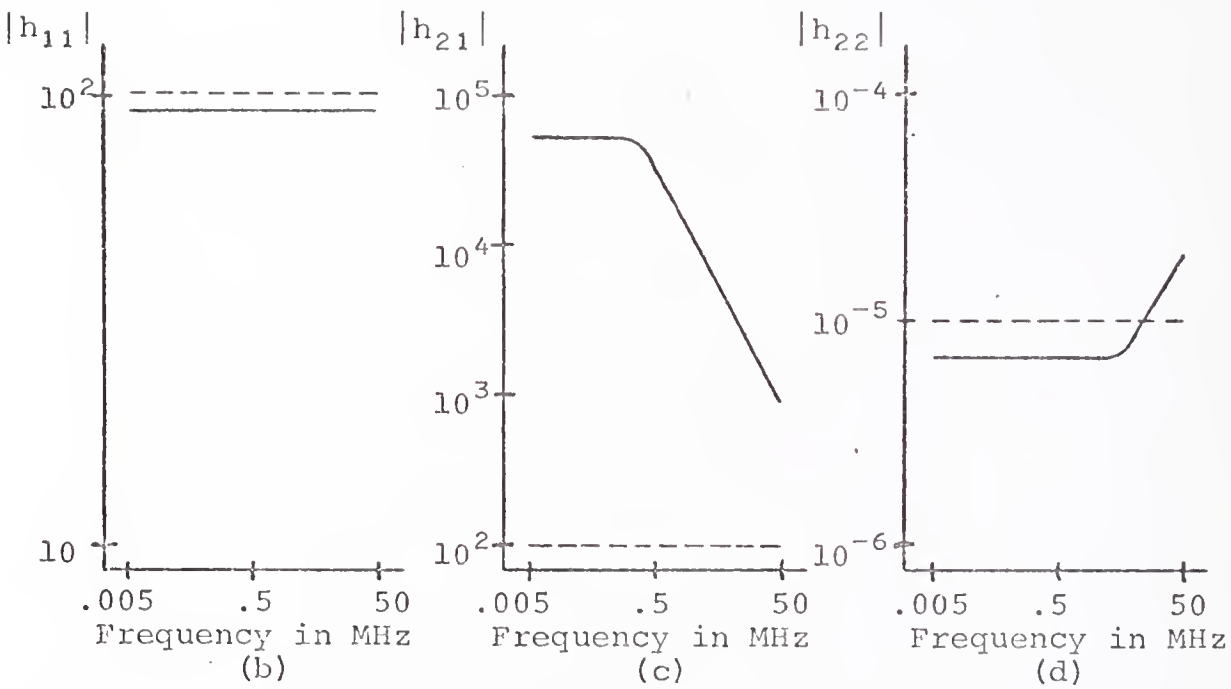
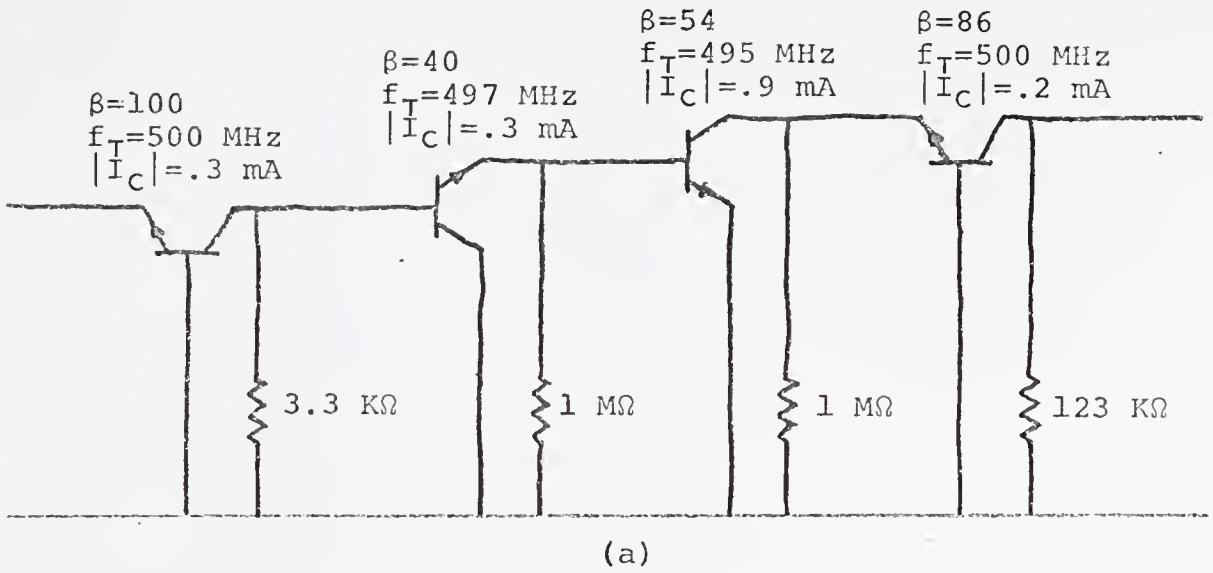
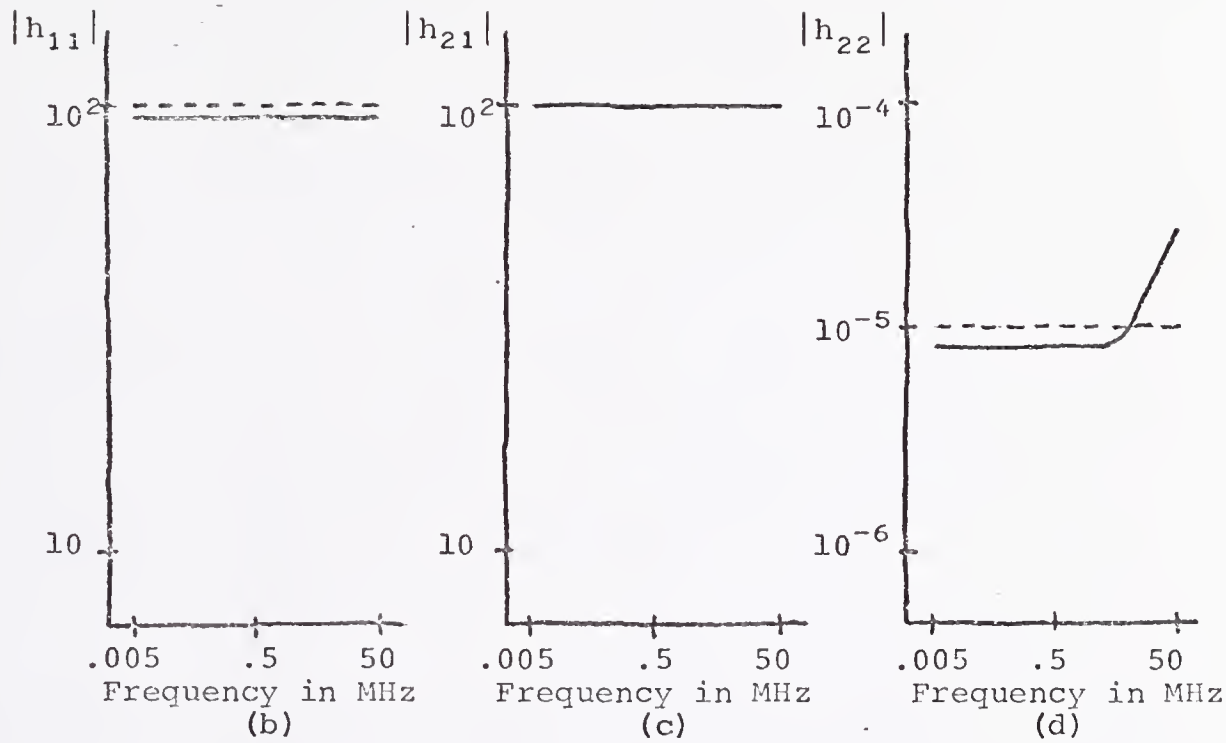


Fig. 4-25

(a) Initial configuration of current amplifier with extra transistor
(b)-(d) Desired (dashed) and actual (solid) responses



(a)



(b) (c) (d)

Fig. 4-26

(a) Final configuration of current amplifier after transistor had been removed
(b)-(d) Desired (dashed) and actual (solid) responses

is capable only of adjusting parameter values to realize the dc specifications. Hence the circuit designer must decide on the location of any additional transistors. Further investigations will hopefully eliminate the necessity of this step. These amplifiers can be realized in differential configurations in which the dc collector currents can be supplied by simple current sources.

CHAPTER 5

SIMULTANEOUS AC AND DC DESIGN

Although separate ac and dc analyses of linear networks are efficient and informative, separate ac and dc designs are misleading and have uncertain value. Some network elements influence both the ac and dc performance of a network so that adjustment of such elements to meet ac specifications alone could disturb dc performance. The ac models for transistors and diodes are dependent upon their dc operating conditions. It is therefore readily seen that parameter adjustment to meet ac or dc specifications alone could cause deviation from the remaining specifications.

An advantage of simultaneous ac and dc design is the ability of the algorithm to reach compromises between ac and dc specifications if necessary. Some specifications require parameter adjustments that would keep other specifications from being met. Quite often a compromise between the competing specifications is acceptable or even desirable.

Methods have been discussed in the previous chapters that were selected because they perform ac and dc design efficiently and accurately. We now describe how these methods may be combined and show the results of some typical designs.

Methods Required for Simultaneous AC and DC Design

Many of the methods used to perform simultaneous ac and dc design are identical to those used for separate ac and dc designs. There are two areas that require revision; these revisions accomodate the design algorithm that will be described later in this chapter. Specifications are included in the dc design that temperature desensitize the network under consideration. In order to implement these desensitizing specifications, additions are made to the performance function. The performance function and the parameter gradients become more general and more useful due to the modifications that are explained in this section.

Each transistor's operation is dependent on two dc variables. For the Ebers-Moll model, the voltages across the two diodes are considered, and for the hybrid-pi model, the collector current and the collector-to-emitter voltage are considered. Each diode's operation is also dependent on a dc variable, the junction voltage. In general, these dc variables are the most important dc variables of a linear amplifier. If other dc variables are selected as representative variables to be desensitized, procedures similar to those that follow can be developed. In the discussion that follows, a general dc variable, $X_j(\underline{r}(T),T)$, will be considered as it is representative of all possible dc variables.

There are two basic approaches to desensitization of a variable with respect to temperature. The first method is to desire that the variable assume a specific value at all tem-

peratures. To achieve this specification, a term is added to the performance function for each dc variable. The term is

$$E_T = \sum_{J=1}^{n_T} \int_{T_I}^{T_F} 1/p\{W_J(T)[X_J(\underline{r}(T),T)/\hat{X}_J(T)-1]\}^p dT \qquad (5.1)$$

where E_T is the contribution to the performance function from temperature desensitization,

n_T is the number of dc variables specified,

T_I is the initial temperature of interest,

T_F is the final temperature of interest,

p is an even positive integer,

$W_J(T)$ is a nonnegative weight function,

$X_J(\underline{r}(T),T)$ is the actual value of the dc variable,

$\underline{r}(T)$ is the vector of network parameters, and

$\hat{X}_J(T)$ is the desired value of the dc variable.

The total performance function can be written

$$E = E_{AC}+E_{DC}+E_T.$$

Although this approach is an acceptable method of desensitizing variables, it requires the user to select a value for each dc variable before the final network is known. Thus the course of the design is altered to achieve the specified values of the dc variables. Even if this method is useful for some design situations, an alternate method is required that does not detract from the other objectives of the network design.

The second approach to temperature desensitization is to select a standard temperature at which a value of the variable

will be determined and used as the desired value for the variable at other temperatures. Thus a specific value of the dc variable is not desired and the desensitization process can be accomplished with minor influence on the other objectives of the design. This method is implemented through the performance function by adding a term of the form

$$E_T = \sum_{J=1}^{n_T} \int_{T_I}^{T_F} 1/p\{W_J(T)[X_J(\underline{r}(T),T)/X_J(\underline{r}(T_0),T_0)-1]\}^p dT \qquad (5.2)$$

where T_0 is the standard temperature, 298° Kelvin. This approach has proven to be very satisfactory.

In Chapter 3 we allowed the dc bias parameters of diodes and transistors as well as the nonbias parameters of the network to vary to meet the ac specifications. For simultaneous consideration of both ac and dc specifications, the dc bias parameters cannot be adjusted arbitrarily but rather assume values which are dependent on conditions existing in the network. This apparent stumbling block can be removed by making a simple modification to the ac gradients of the nonbias parameters which incorporates the gradient information for the bias parameters.

We must realize that many of the nonbias parameters of the network directly influence the values of the bias parameters. For instance, if the value of a resistor which is connected to the emitter of a transistor is altered, the value of the emitter-base junction voltage is altered. The interdependence between the emitter resistor and the emitter-

base junction voltage is described by the partial derivative

$$\partial V_{EB}/\partial R_E .$$

In general, a bias parameter is related to a nonbias parameter by the partial derivative

$$\partial r_B/\partial r_N$$

where r_B is a bias parameter and r_N is a nonbias parameter.

Using this partial derivative, a general ac gradient expression can be determined for each nonbias parameter that includes gradient information for itself and a portion of the gradient information for each bias parameter. Thus, the ac gradient component for each nonbias parameter is of the form

$$\partial E_{AC}/\partial r_N = [\partial E_{AC}/\partial r_N]' + \sum_{I=1}^{n_T} [\partial E_{AC}/\partial r_{BI}][\partial r_{BI}/\partial r_N] \qquad (5.3)$$

where $[\partial E_{AC}/\partial r_N]'$ is the gradient expression heretofore used to directly adjust the nonbias parameter, n_T is the number of bias parameters, and $\partial E_{AC}/r_{BI}$ is the gradient expression heretofore used to directly adjust each bias parameter. This expression can also be obtained by application of the chain rule of differentiation to the ac performance function,

$$E_{AC}(\underline{r}_N, \underline{r}_B(\underline{r}_N)).$$

The only components of (5.3) which are not already known are the partial derivatives of the bias parameters with respect to the nonbias parameters. We can again use Tellegen's theorem to obtain the necessary partial derivatives if we require that

bias currents be measured through voltage sources and bias voltages be measured across current sources. If we refer back to (3.32), rewritten here for convenience

$$-\sum_{V}\Delta I_{V}\psi_{V}+\sum_{I}\Delta V_{I}\phi_{I} = -\sum_{J=1}^{n} G_{J}\Delta r_{J},$$

we notice that the partial derivatives are easily obtained after an analysis of the original network and n_T analyses of the adjoint network. For the partial of a voltage bias parameter, we set all voltage and current sources of the adjoint network to zero, except the current source corresponding to the parameter. Setting the remaining current source to unity yields

$$\Delta V_{K} = \Delta r_{B} = -\sum_{J=1}^{n} G_{J}\Delta r_{J}$$

where V_K is the bias voltage being considered. In the limit as each Δr_J approaches zero, the desired partials are

$$\partial r_{B}/\partial r_{J} = -G_{J}, \quad J = 1,2,3,\ldots n. \tag{5.4}$$

A similar result can be obtained for a current bias parameter if all adjoint sources are set to zero except the voltage source corresponding to the parameter which is set to minus one.

It should be clear that the dc variables used to temperature desensitize a network are identical to the dc bias parameters. Although this correspondence is not necessary, it is convenient for what follows.

Since we have added a temperature desensitization term to the performance function, gradient components for that term must be derived. When specific values are desired for each dc variable as shown in (5.1), a procedure identical to that described for E_{DC} is used. When the value of the variable at T_0 is desired, an alternate procedure must be developed. This new procedure is required because the desired response is now variable also. Previously, all desired responses were constant and differentiation of the performance function yielded an expression in terms of partials of port voltages and currents. Differentiation of (5.2) yields

$$\partial E_T/\partial r_K = \sum_{J=1}^{n_T} \int_{T_I}^{T_F} \{W_J(T)[X_J(\underline{r}(T),T)/X(\underline{r}(T_0),T_0)-1]\}^{p-1}$$

$$\cdot W_J(T)\{[\partial X_J(\underline{r}(T),T)/\partial r_K]/X(\underline{r}(T_0),T_0)$$

$$-X(\underline{r}(T),T)/X(\underline{r}(T_0),T_0)^2 \partial X(\underline{r}(T_0),T_0)/\partial r_K\}dT \qquad (5.5)$$

where

$$\partial X(\underline{r}(T_0),T_0)/\partial r_K = \partial r_B/\partial r_K.$$

The only quantity of (5.5) that needs to be calculated is the partial of each fixed desired response with respect to the network parameter of interest. These partials can be determined in a procedure identical to that described for obtaining (5.4).

General Algorithm for Simultaneous AC and DC Design

The basis of simultaneous ac and dc design is to determine the value of a performance function that reflects the ac and dc specifications of a linear amplifier and minimize it by means of an optimization algorithm. As indicated earlier, the optimization algorithms employed require determination of partial derivatives of the performance function with respect to the designable network parameters. The performance function and parameter gradients are determined by the appropriate use of responses of the original network to specified excitations and responses of an adjoint network to excitations specified above. A systematic order in which to perform the necessary operations will now be discussed.

Assume a linear amplifier is to be designed and there are both ac and dc specifications. Transistors are used to supply the amplifier's gain. Since the partial derivatives of the bias parameters with respect to the nonbias parameters are required for both the evaluation of the ac gradients and the evaluation of the dc temperature desensitization portion of the performance function, these partials are determined first. A single dc analysis of the original network at T_0 is performed and the values of all of the dc bias parameters are defined. The dc adjoint network is analyzed n_T times and the appropriate responses are used to calculate the partial derivatives.

We choose to evaluate the ac contributions to the performance function and gradients next, even though the dc

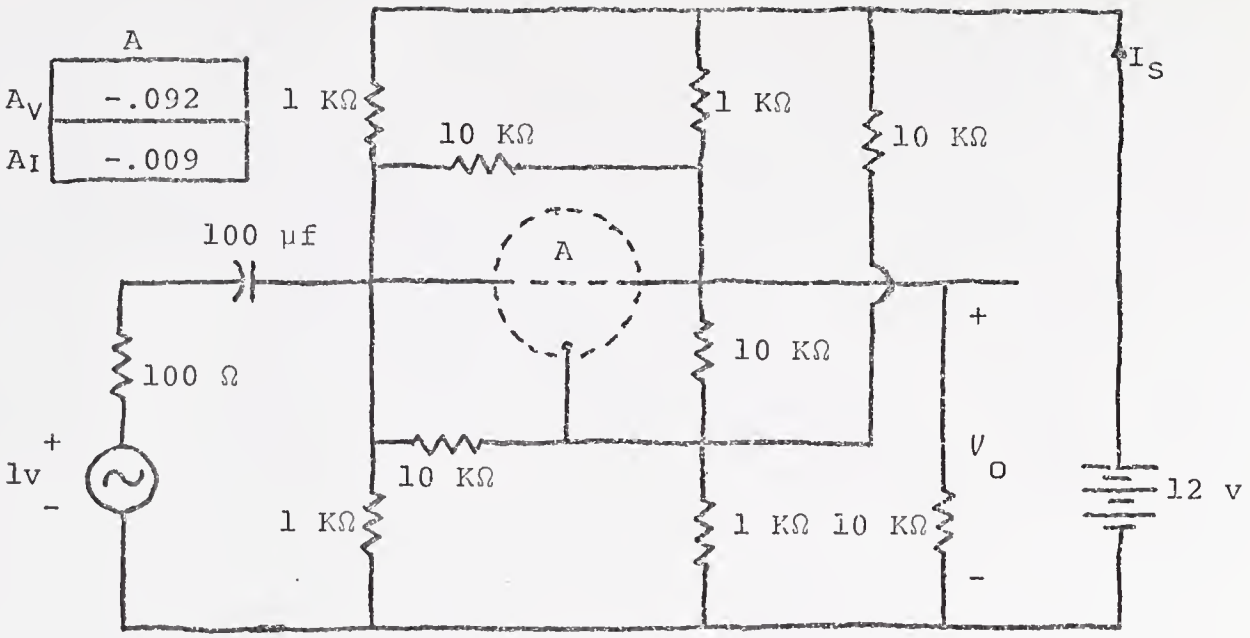
contributions could be calculated first. For the ac contributions, one analysis of the original network and one analysis of the adjoint network must be performed at each frequency. The contribution to the performance function and excitations for the adjoint network are calculated immediately after the analysis of the original network. The adjoint analysis is then performed and is followed by the gradient calculations. Before going to the next frequency, the partial integrals are augmented with the appropriately weighted contributions.

Upon completing the ac calculations, the dc calculations are performed. For each temperature, one analysis on the original network and n_T+1 analyses on the adjoint network are performed. The contribution to the performance function from the regular specifications and the corresponding excitations to adjoint network are calculated after the analysis of the original network. The adjoint network is then analyzed and the gradient contributions related to the regular specifications are calculated. For each dc bias parameter, an additional contribution to the performance function is determined and a corresponding adjoint analysis is performed. Gradient contributions related to each bias parameter are evaluated after its associated adjoint analysis. Before going to the next temperature, the partial integrals are augmented with the appropriately weighted contributions.

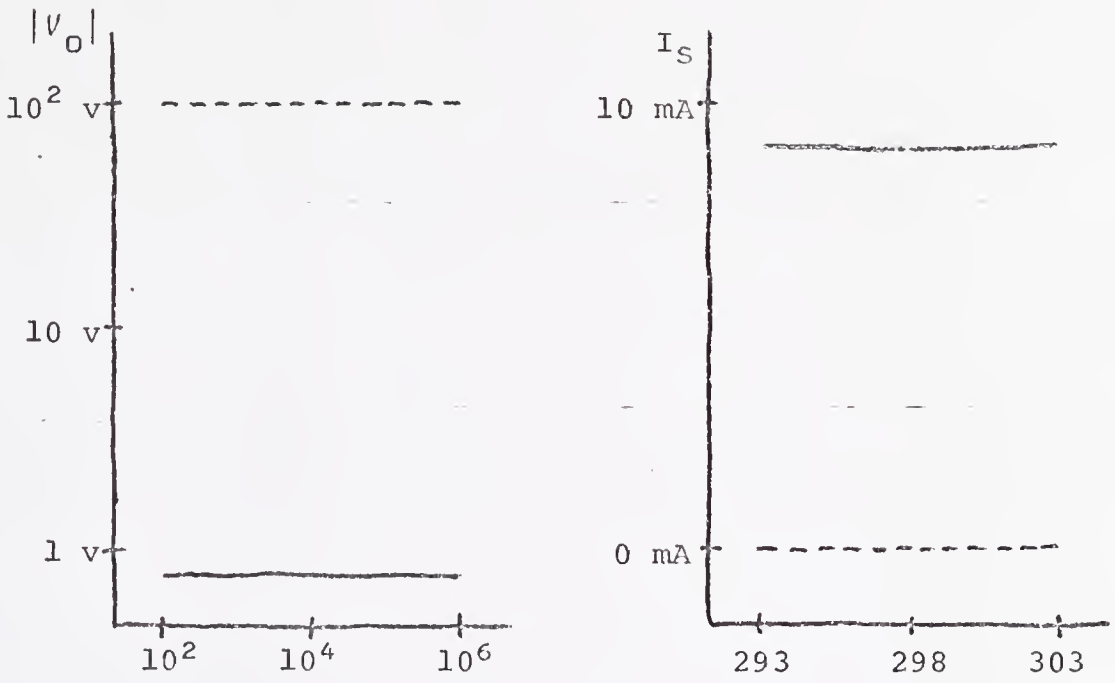
The performance function and its parameter gradients have now been evaluated completely. Control is passed to the

minimization subroutine which uses the gradient information to adjust the network parameters. Both minimization algorithms, conjugate gradients and steepest descent, employ a linear search followed by cubic interpolation. This approach requires multiple evaluations of the performance function and its parameter gradients for each linear search. After determining new parameter values, the minimization subroutine returns control to the subroutines that evaluate the performance function and gradients. The procedure again starts with the evaluation of the partials of the bias parameters with respect to all network parameters. The order of operations is identical to that of the initial calculations.

The examples that will now be described have been run on FROLIC which is depicted through flow charts in the appendix. A simple example that demonstrates the features of this design algorithm has been run. The desired network was an amplifier stage biased by a resistive network. The initial network was an RC network, shown in Fig. 5-1(a), with a centrally located growth site. The four standard bias resistors were initially one kilohm and the remaining resistors were ten kilohms each. Performance functions (3.3), (3.6), and (5.2) were used with p=2. The conjugate gradients minimization algorithm was used. The desired responses, shown in Fig. 5-1(b)-(c), represent a voltage gain of 100 and minimum dc power dissipation. Growth indicators were determined and a common-emitter stage was grown. Upon the insertion of the transistor into the network, two bias parameter temperature desensitizing specifications



(a)



(b) (c)

Fig. 5-1

(a) Initial configuration for possible amplifier stage
(b) Output ac voltage desired (dashed) and actual (solid) responses
(c) Dc supply current desired (dashed) and actual (solid) responses

were invoked. After 41 iterations of the optimization algorithm, 158 function evaluations had been performed and 35 seconds of IBM 360/65 computer time consumed. The resulting network and its responses are shown in Fig. 5-2. The parameters and constants of the final network are shown in Table 5-1. The desired specifications had been met satisfactorily.

A more complicated example involving a differential pair of transistors and a transistor current source was also designed. The initial network and its desired responses are shown in Fig. 5-3. All resistors were initially one kilohm. The desired responses, performance function, and minimization algorithm were basically identical to those used in the previous example. The differential pair was grown upon calculation and investigation of the growth indicators. In this example, two transistors were allowed to grow at once as the network required a balanced pair. Since the network contained four transistors, eight temperature desensitizing specifications were invoked. After 15 iterations, 32 function evaluations had been performed and 17 seconds of IBM 360/65 computer time had been consumed. The resulting network and responses are shown in Figs. 5-4 and 5-5. The parameters and constants of the final network are shown in Table 5-2. Again the desired results were obtained.

An eleven-transistor video amplifier, the Fairchild μA751C, is a good test for the design algorithm. Since this network incorporates feedback, the automatic frequency selection algorithm was used to guarantee that all integrals were evaluated accurately. Transistor growth was not exercised

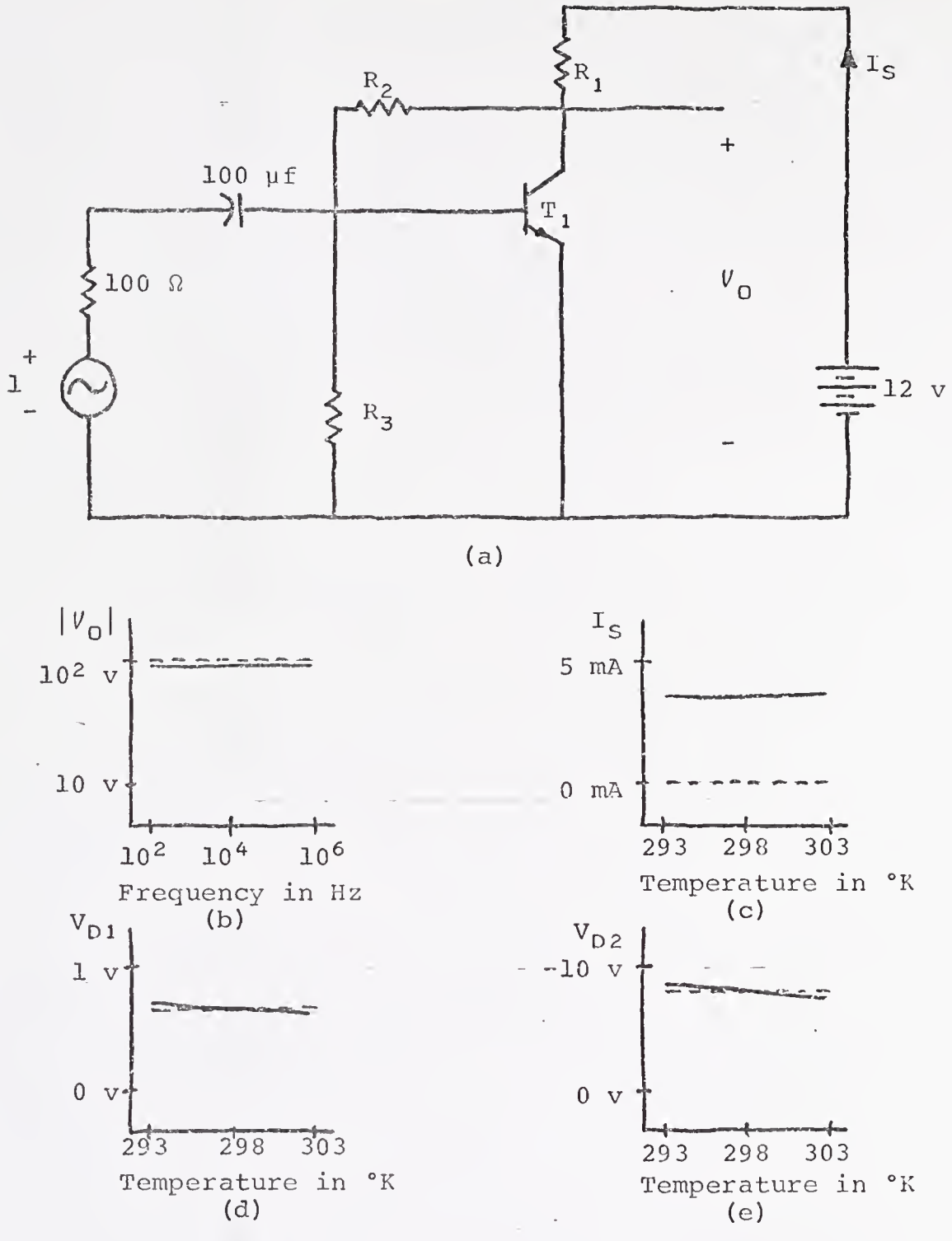


Fig. 5-2

(a) Final configuration of one-stage transistor amplifier
(b)-(e) Desired (dashed) and actual (solid) responses

Table 5-1

Parameters and Constants of Final Configuration of One-Stage Amplifier

	T_1
β_{F0}	100
f_{T1}	500 MHz
V_{D1}	.68 v
V_{D2}	-8.15 v
R_B	0 Ω
R_C	0 Ω
R_E	0 Ω
I_{CS0}	2E-14 A
I_{ES0}	1E-14 A
η	1E-4
C_{0j}	1 pf
C_{02}	1 pf
β_R	1
f_{T2}	500 MHz
n_1	1
n_2	1
V_{01}	.75 v
V_{02}	.75 v
m_1	2.5
m_2	2.5
α_1	.15
α_2	.15
α_3	.008

	R_1	R_2	R_3
R_0	861 Ω	173 KΩ	62 KΩ
α_R	.002	.002	.002

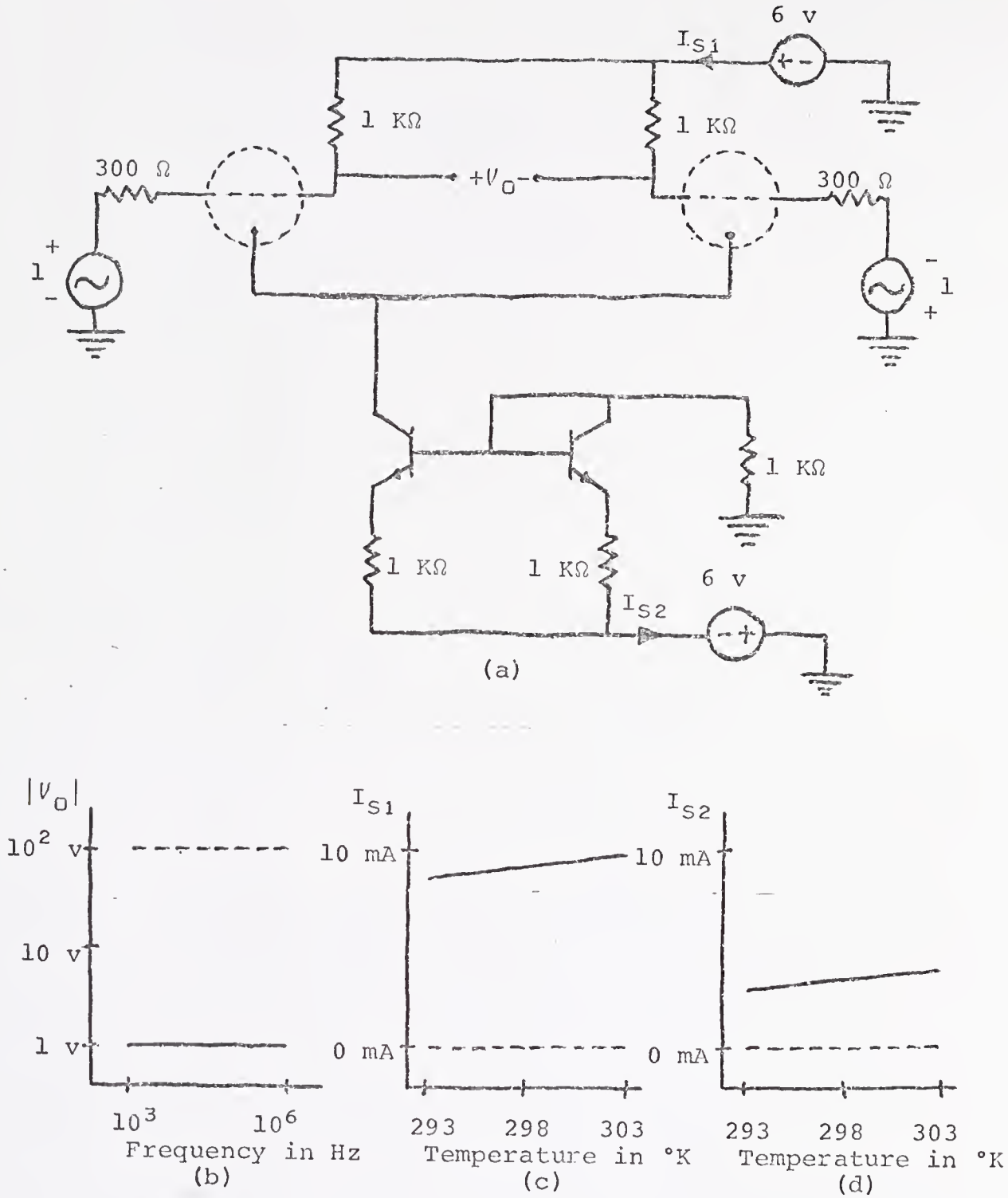
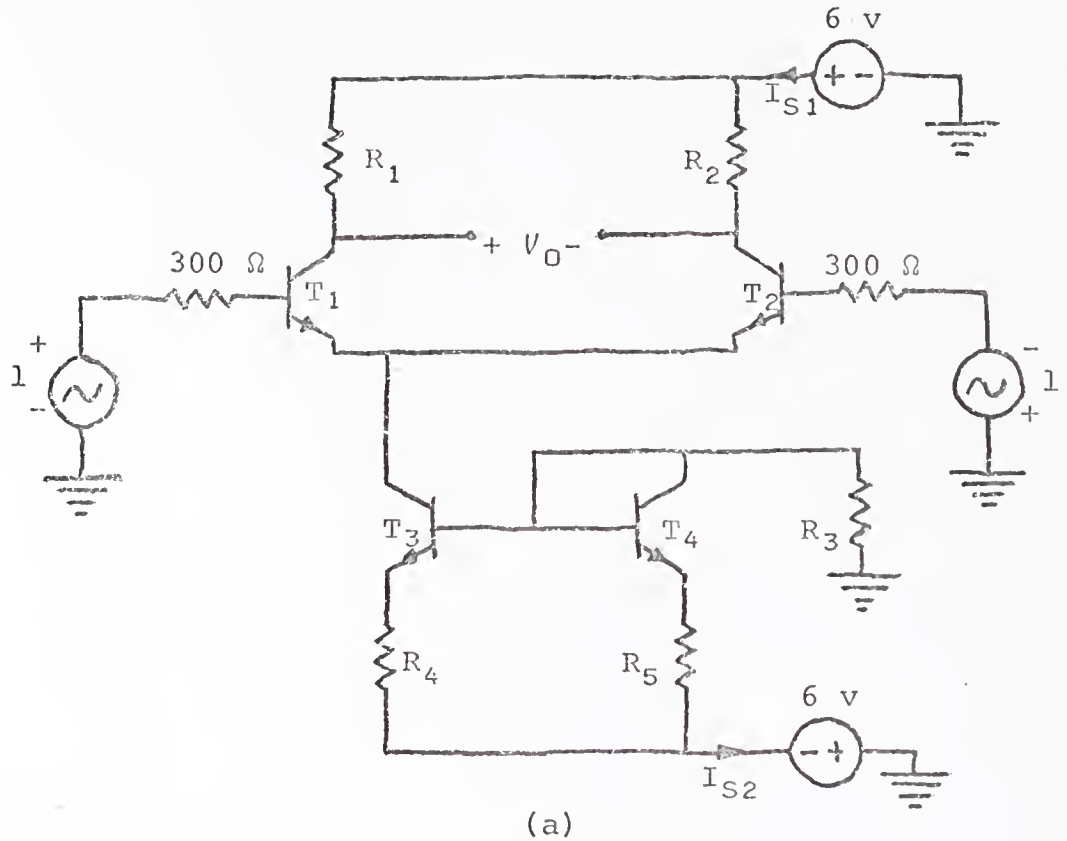


Fig. 5-3

(a) Initial configuration for differential amplifier example
(b)-(d) Desired (dashed) and actual (solid) responses



(a)

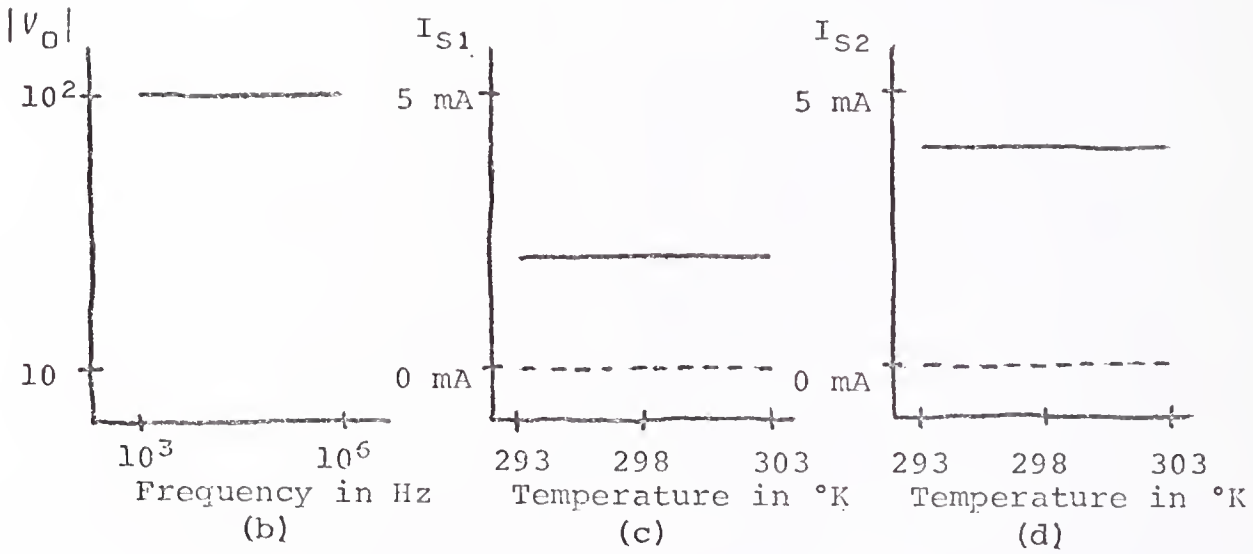


Fig. 5-4

(a) Final configuration for differential amplifier example
(b)-(d) Desired (dashed) and actual (solid) responses

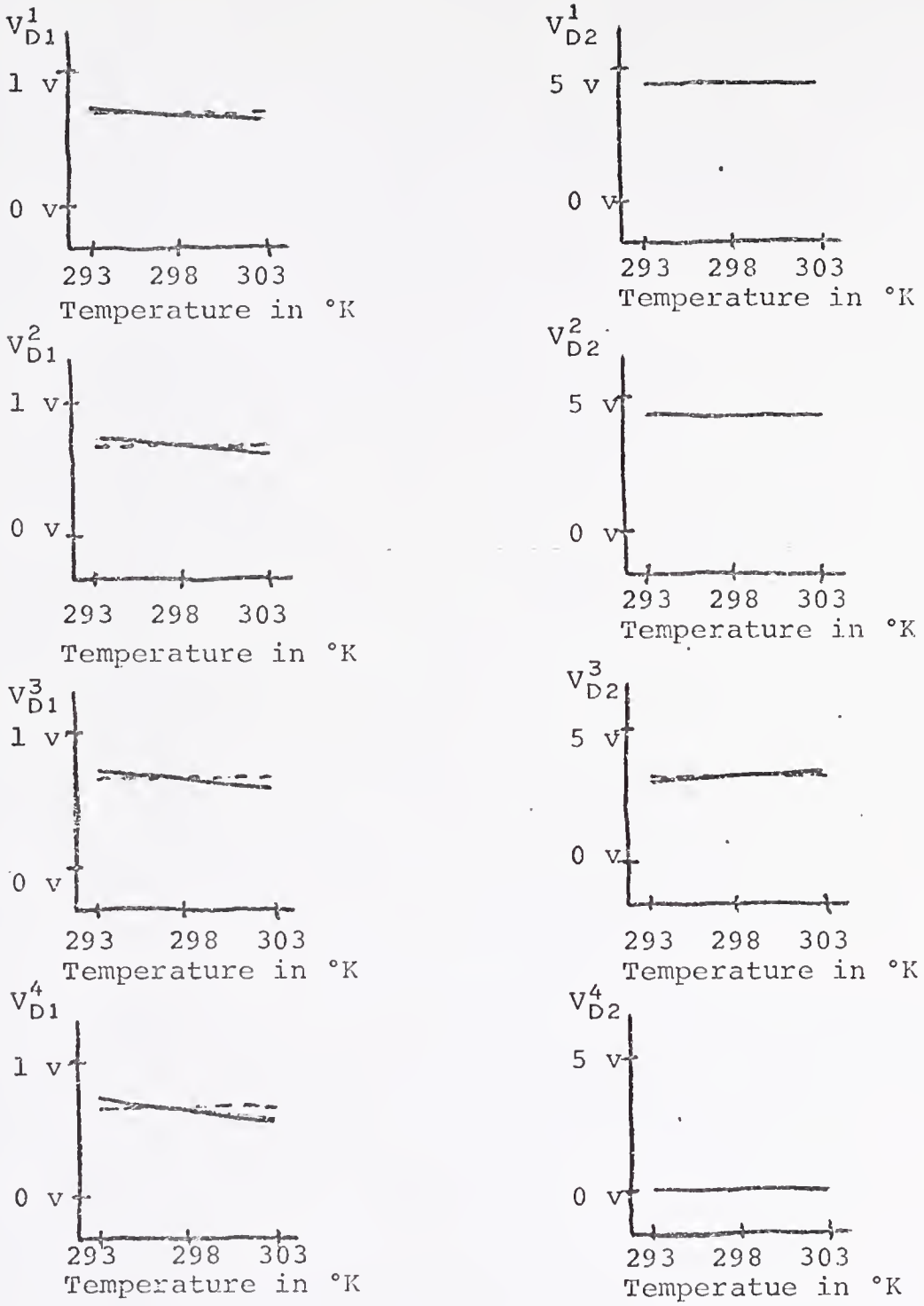


Fig. 5-5

Desired (dashed) and actual (solid) responses for temperature desensitization specifications for differential amplifier example

Table 5-2

Parameters and Constants of Final Configuration of Differential Amplifier

	T_1	T_2	T_3	T_4
β_{FO}	100	100	50	50
f_{T1}	500 MHz	500 MHz	175 MHz	175 MHz
V_{D1}	.65 v	.65 v	.67 v	.65 v
V_{D2}	-4.5 v	-4.5 v	-3.7 v	0.0 v
R_B	0 Ω	0 Ω	0 Ω	0 Ω
R_C	0 Ω	0 Ω	0 Ω	0 Ω
R_E	0 Ω	0 Ω	0 Ω	0 Ω
I_{CSO}	2E-14 A	2E-14 A	2E-14 A	2E-14 A
I_{ESO}	1E-14 A	1E-14 A	1E-14 A	1E-14 A
η	1E-4	1E-4	1E-4	1E-4
C_{O1}	1 pf	1 pf	1 pf	1 pf
C_{O2}	1 pf	1 pf	1 pf	1 pf
β_R	1	1	1	1
f_{T2}	500 MHz	500 MHz	500 MHz	500 MHz
n_1	1	1	1	1
n_2	1	1	1	1
V_{O1}	.75 v	.75 v	.75 v	.75 v
V_{O2}	.75 v	.75 v	.75 v	.75 v
m_1	2.5	2.5	2.5	2.5
m_2	2.5	2.5	2.5	2.5
α_1	.15	.15	.15	.15
α_2	.15	.15	.15	.15
α_3	.008	.008	.008	.008

	R_1	R_2	R_3	R_4	R_5
R_O	1.25 KΩ	1.25 KΩ	3.5 KΩ	418 Ω	875 Ω
α_R	.002	.002	.002	.002	.002

in this example. The desired responses represent an ac voltage gain of 800 and a minimum dc power dissipation. All resistors were started at one kilohm and transistor betas and f_Ts were started at 50 and 500 MHz, respectively. Associated with the 11 transistors were 22 bias parameters, each of which was used as a temperature desensitizing specification. Performance functions (3.3), (3.6), and (5.2) with p=2 and the conjugate gradients minimization algorithm were used. After 44 linear search iterations, 117 function evaluations had been performed and 1,018 seconds of IBM 360/65 computer time had been consumed. The final network and its responses are shown in Figs. 5-6 and 5-7. The parameters and constants of the final network are shown in Table 5-3. The resultant network's resistors are larger than those of the μA751C because of the minimum dc current specification. If a parasitic capacitance were added to each resistor and its value was proportional to the resistance, the resistor values would have been smaller and closer to those of the μA751C.

A design procedure has been proposed and demonstrated that simultaneously achieves ac and dc specifications on a linear amplifier. The procedure has been implemented in a computer program, FROLIC, which incorporates techniques that are efficient, accurate, and versatile.

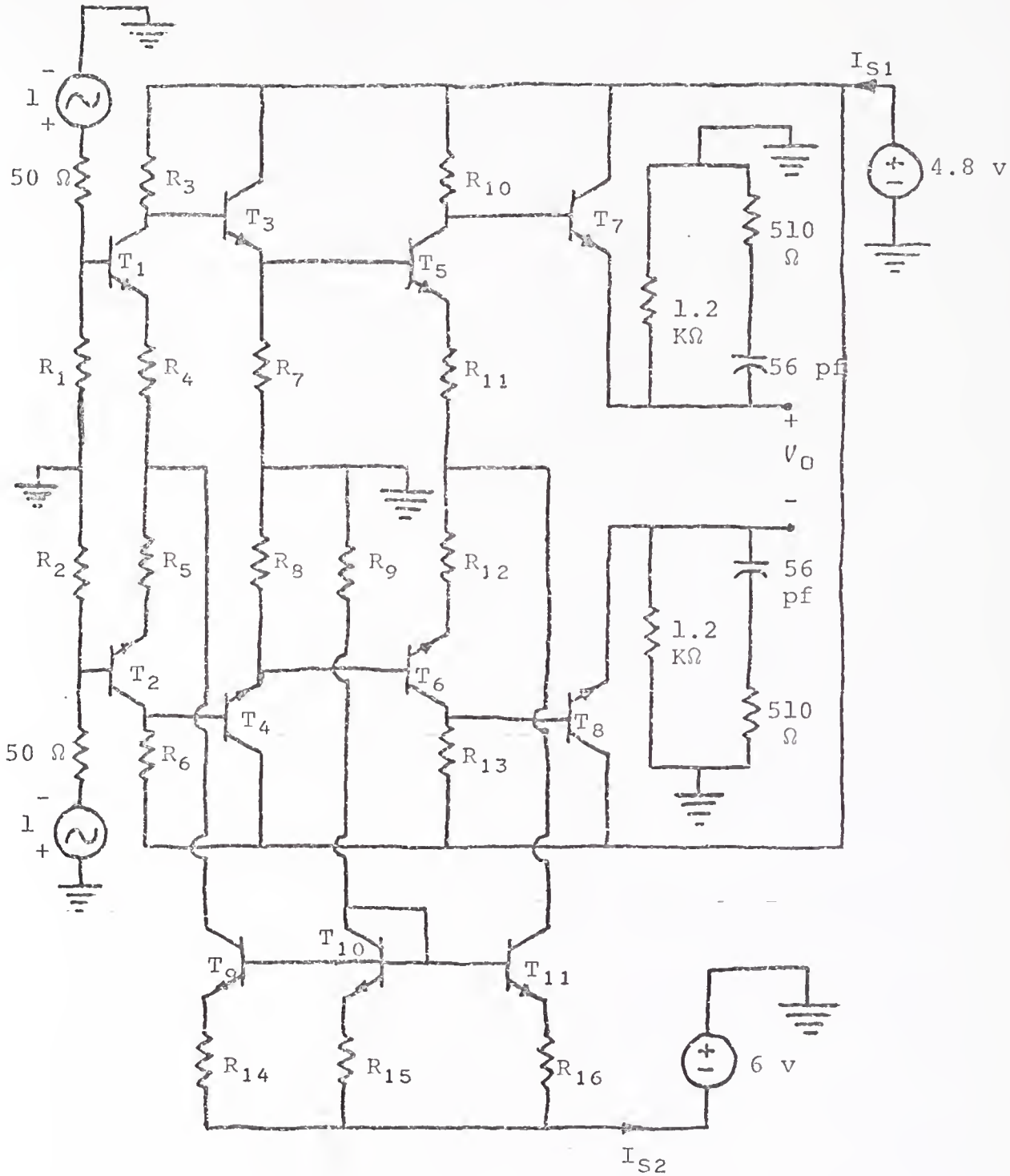


Fig. 5-6

Network configuration of μA751C video amplifier

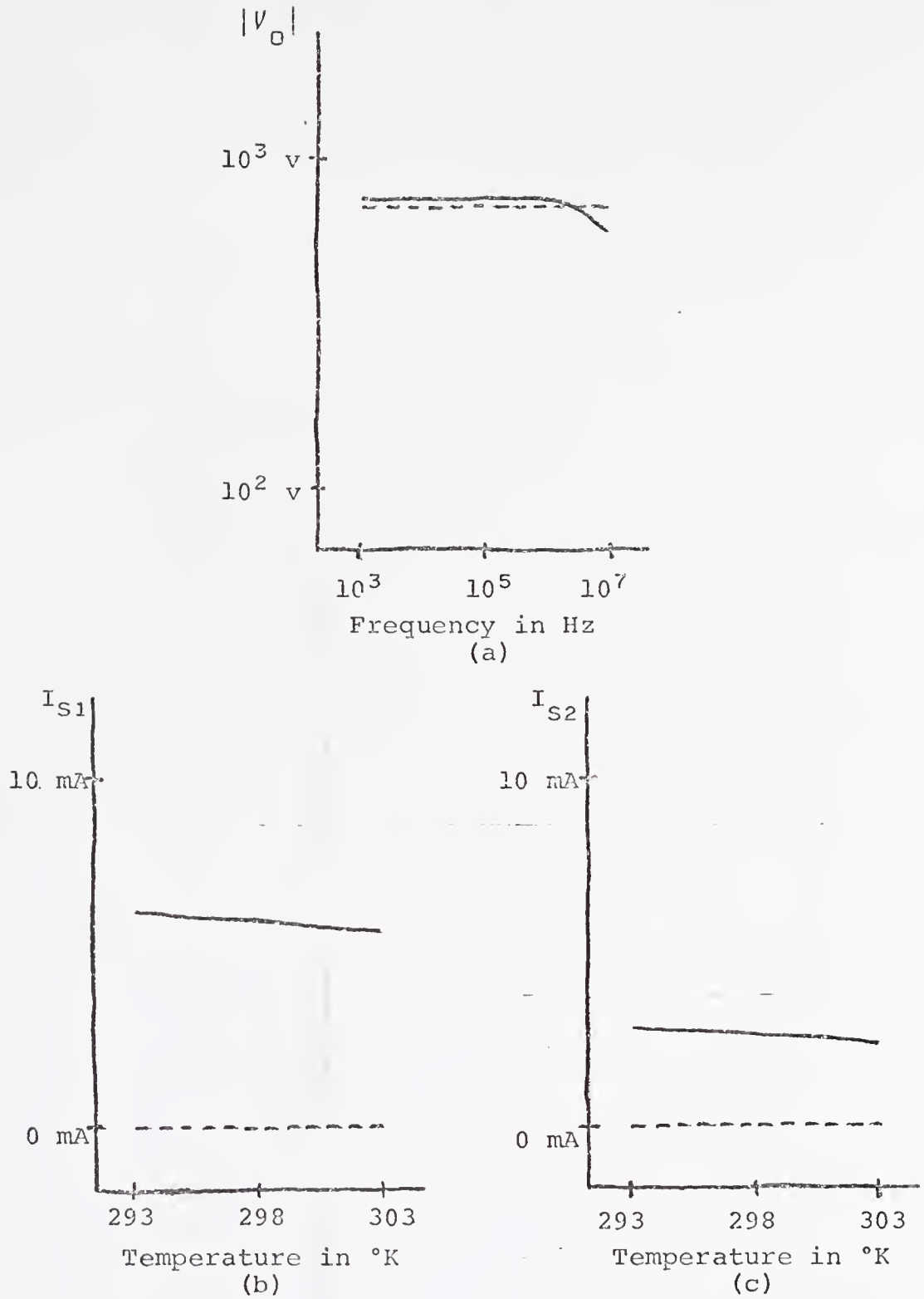


Fig. 5-7

(a)-(c) Desired (dashed) and actual (solid) responses for video amplifier

Table 5-3

Parameters and Constants of Final Configuration of Video Amplifier

	T_1&T_2	T_3&T_4	T_5&T_6	T_7&T_8	T_9	T_{10}	T_{11}
β_{FO}	52	58	53	71	49	50	48
f_{T1}	500 MHz	500 MHz	500 MHz	500 MHz	500 MHz	500 MHz	500 MHz
V_{D1}	.63 v	.58 v	.60 v	.68 v	.65 v	.65 v	.62 v
V_{D2}	-1.35 v	-3.45 v	-3.10 v	-.92 v	-3.40 v	0.0 v	-4.24 v
R_B	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω
R_C	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω
R_E	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω	0 Ω
I_{CSO}	2E-14 A	2E-14 A	2E-14 A	2E-14 A	2E-14 A	2E-14 A	2E-14 A
I_{ESO}	1E-14 A	1E-14 A	1E-14 A	1E-14 A	1E-14 A	1E-14 A	1E-14 A
η	1E-4	1E-4	1E-4	1E-4	1E-4	1E-4	1E-4
C_{Oi}	1 pf	1 pf	1 pf	1 pf	1 pf	1 pf	1 pf
C_{O2}	1 pf	1 pf	1 pf	1 pf	1 pf	1 pf	1 pf
β_R	1	1	1	1	1	1	1
f_{T2}	500 MHz	500 MHz	500 MHz	500 MHz	500 MHz	500 MHz	500 MHz
n_1	1	1	1	1	1	1	1
n_2	1	1	1	1	1	1	1
V_{O1}	.75 v	.75 v	.75 v	.75 v	.75 v	.75 v	.75 v
V_{O2}	.75 v	.75 v	.75 v	.75 v	.75 v	.75 v	.75 v
m_1	2.5	2.5	2.5	2.5	2.5	2.5	2.5
m_2	2.5	2.5	2.5	2.5	2.5	2.5	2.5
α_1	.15	.15	.15	.15	.15	.15	.15
α_2	.15	.15	.15	.15	.15	.15	.15
α_3	.008	.008	.008	.008	.008	.008	.008

	R_1&R_2	R_3&R_4	R_5&R_6	R_7&R_8	R_9	R_{10}&R_{11}	R_{12}&R_{13}	R_{14}	R_{15}	R_{16}
R_O	1.1 KΩ	8.3 KΩ	162 Ω	12.7 KΩ	3.4 KΩ	4.8 KΩ	191 Ω	1.4 KΩ	1.1 KΩ	4.0 KΩ
α_R	.002	.002	.002	.002	.002	.002	.002	.002	.002	.002

CHAPTER 6

CONCLUDING REMARKS AND SUGGESTIONS FOR FURTHER RESEARCH

An automated network design algorithm has been described which simultaneously considers ac and dc specifications. Advanced techniques have been assembled to best accomplish the task of designing the large, complex, integrated circuits in use today. Examples have demonstrated the ability of the algorithm to efficiently achieve desired specifications. Some additional work which would enable the algorithm to be more flexible, more accurate, and more efficient is described below.

Improved models for the physical circuit elements which account for parasitics are necessary. In particular, resistor models need to include capacitive and inductive effects. Bipolar transistor models require a collector current dependent beta.

Consideration of additional types of circuit elements would broaden the scope of the algorithm. Elements such as field-effect transistors and RC structures are often used in integrated circuits. Attention should also be given to often-used combinations of elements such as current sources and differential amplifiers. As new elements and groups of elements are considered, models and gradient expressions for these components must be determined.

Growth and removal of bipolar transistors to meet ac specifications have been investigated. Further work should be done in this area with consideration given to the dc characteristics of transistors. Growth and removal indicators should be determined that consider the total performance of transistors. Possible growth and removal of diodes and other elements that may be added should be explored.

Tolerances related to the various steps of the realization of an integrated circuit make it difficult to obtain the exact network that was specified by the design algorithm. These tolerances should be accounted for during the design process so that resulting networks are capable of meeting the desired specifications without obtaining the exact values designated for all of the parameters.

Temperature desensitized ac gain, a desirable quality of an amplifier, should be incorporated in the design specifications. Complete consideration of this specification would require ac analysis of the amplifier at a set of frequencies which represents the frequency range for each temperature of a set that represents the temperature range. This time-consuming process would have to be carried out for each evaluation of the ac contributions to the performance function and gradients. A condensed method which may closely approximate the above process is to analyze the amplifier at a set of frequencies for a standard temperature and then analyze at a set of temperatures for a standard frequency.

As new options are added to the automated network design algorithm, the ability to produce new and improved circuits is increased. The rate at which these circuits are designed can be increased significantly if results of previous designs are available to the algorithm. A good initial network is important in many cases and a helpful hint is often needed in the midst of the design process. This feature can be realized in two ways: devote large amounts of computer memory to storing all significant circuit configurations or allow interaction by a human circuit designer. Since large amounts of time would be required to maintain and search a storage area full of circuit configurations, the more practical approach is to develop an interactive system that allows a designer to monitor the progress of a circuit design.

APPENDIX

FLOW CHARTS OF "FROLIC"

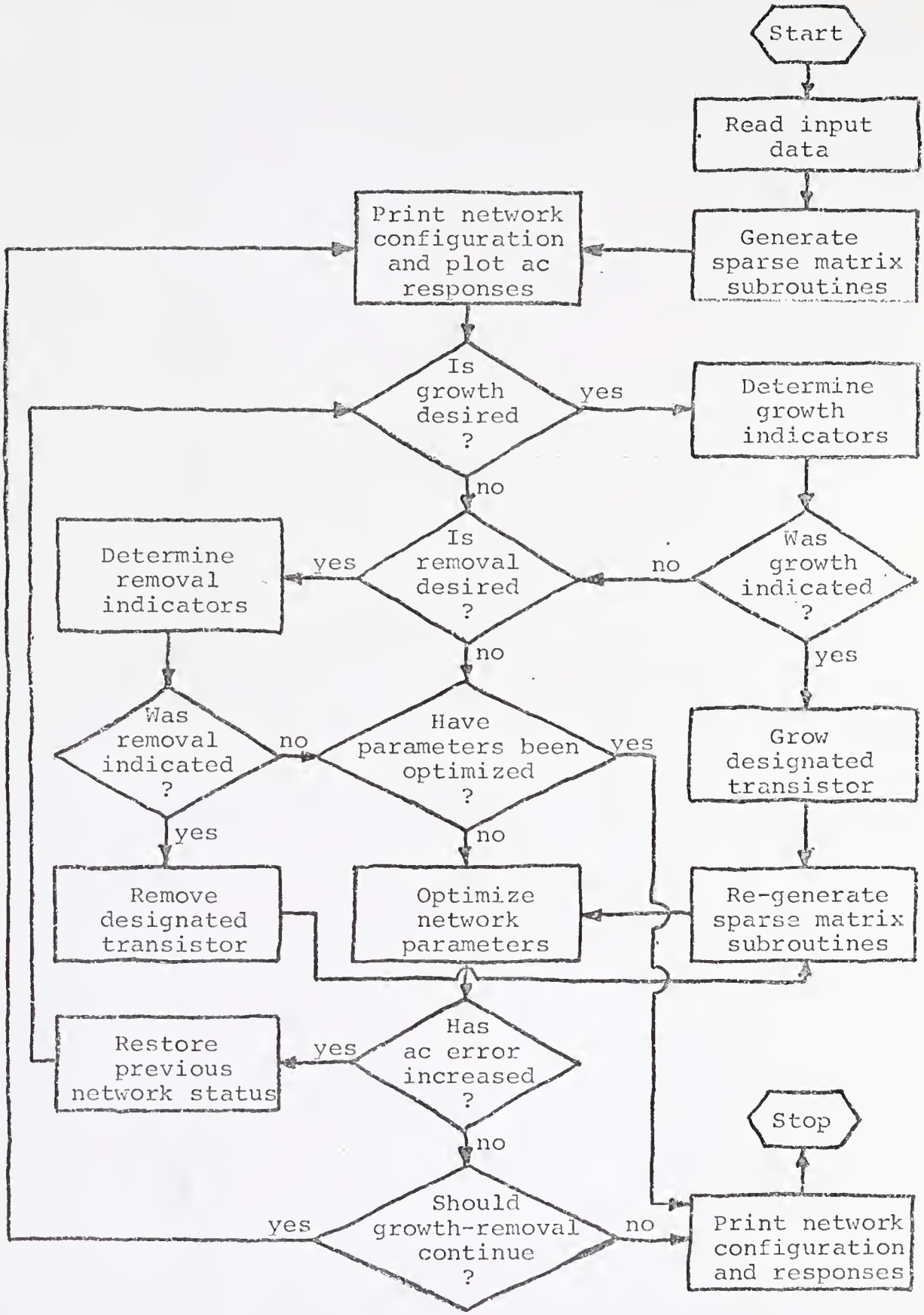


Fig. A-1

Automated network design algorithm

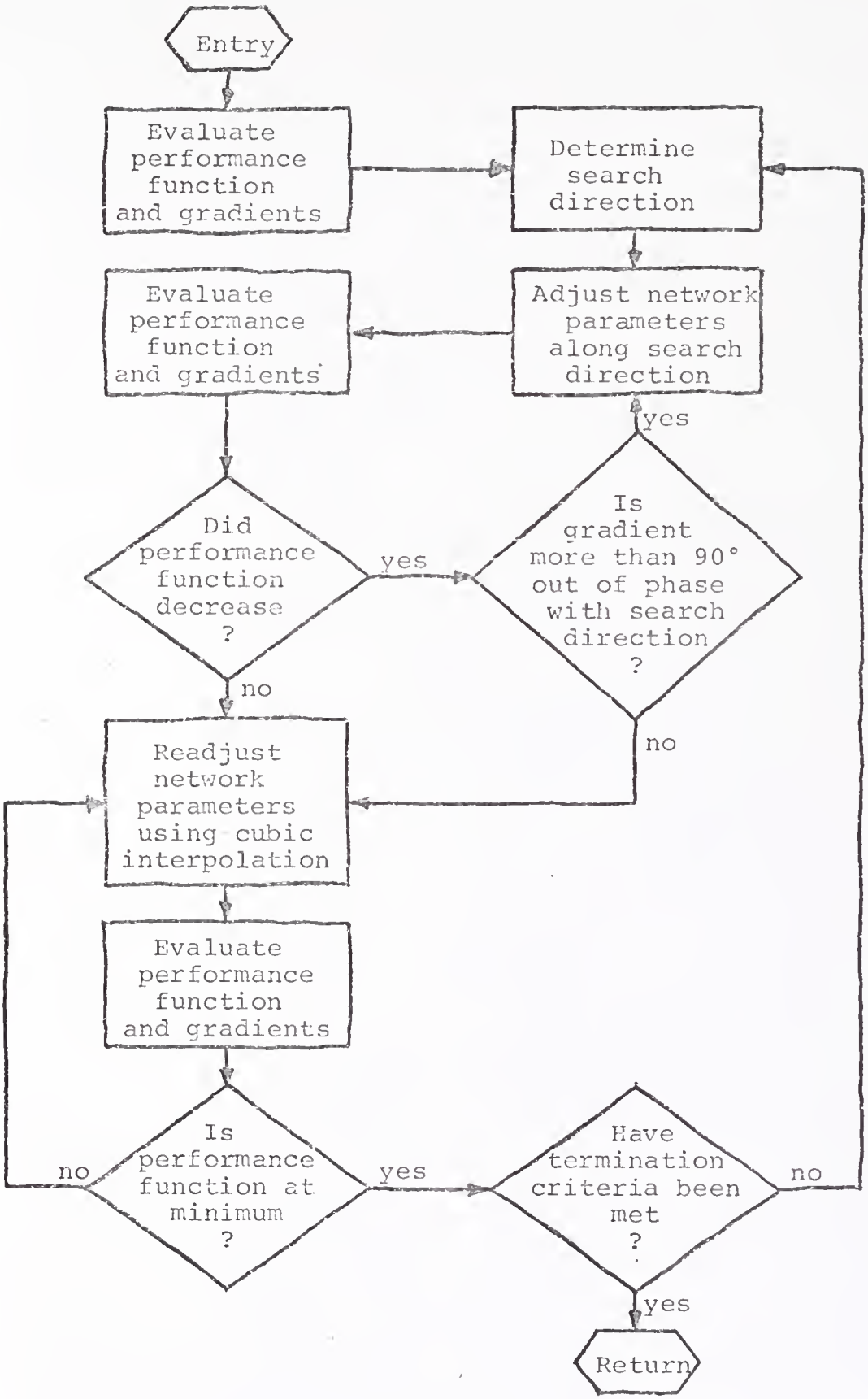


Fig. A-2

Network parameter optimization

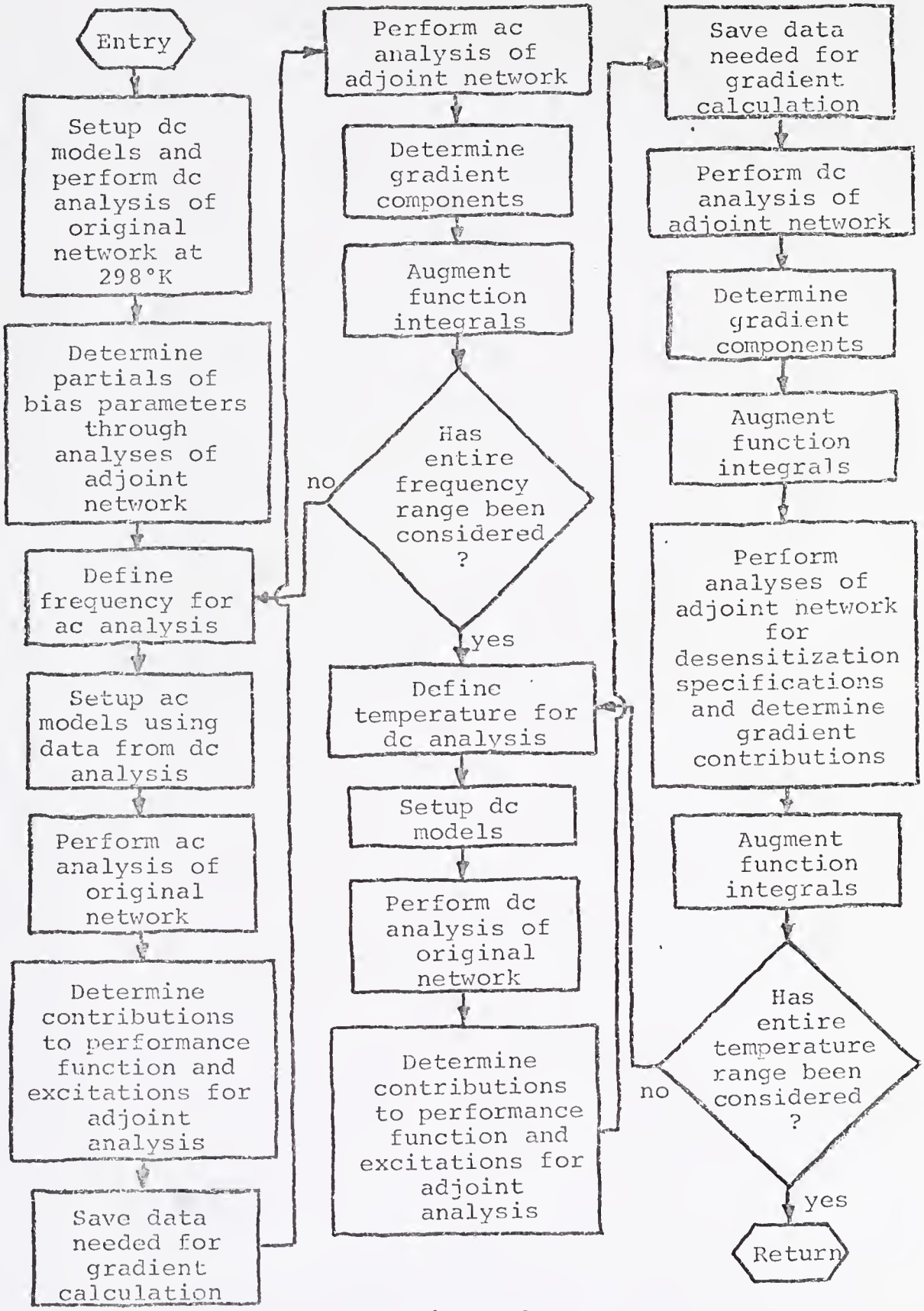


Fig. A-3

Evaluation of performance function and gradients

REFERENCES

1. A. F. Malmberg, F. L. Cornwell, and F. N. Hofer, "NET-1 network analysis program," Los Alamos Scientific Laboratory, Los Alamos, N. Mex., Rept. LA-3119, 7090/94 version, August 1964.

2. D. A. Calahan, "Linear network analysis and realization digital computer programs: An instruction manual," University of Illinois Bulletin, vol. 62, no. 58, February 1965.

3. H. W. Mathers, S. R. Sedore, and J. R. Sents, "Automated digital computer program for determining responses of electronic circuits to transient nuclear radiation (SCEPTRE)," vol. I, IBM Space Guidance Center, Owego, N. Y., IBM File 66-928-611, February 1967.

4. F. H. Branin, "Computer methods of network analysis," Proceedings of the IEEE, vol. 55, pp. 1787-1801, November 1967.

5. R. A. Rohrer, "Fully automated network design by digital computer: Preliminary considerations," Proceedings of the IEEE, vol. 55, pp. 1929-1939, November 1967.

6. S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," IEEE Transactions on Circuit Theory, vol. CT-16, pp. 318-323, August 1969.

7. S. W. Director and R. A. Rohrer, "Automated network design: The frequency-domain case," IEEE Transactions on Circuit Theory, vol. CT-16, pp. 330-337, August 1969.

8. S. W. Director and R. A. Rohrer, "On the efficient computation of first-order network sensitivities (for frequency domain studies of lumped, linear, time-invariant networks)," presented at the second IFAC Conference on System Sensitivity and Adaptivity, 1968, Dubrovnik, Yugoslavia.

9. R. I. Dowell and R. A. Rohrer, "Automated design of biasing circuits," IEEE Transactions on Circuit Theory, vol. CT-18, pp. 85-89, January 1971.

10. G. D. Hachtel, R. K. Brayton, and F. G. Gustavson, "The sparse tableau approach to network analysis and design," IEEE Transactions on Circuit Theory, vol. CT-18, pp. 101-113, January 1971.

11. C. W. Ho, "Time-domain sensitivity computation for networks containing transmission lines," IEEE Transactions on Circuit Theory, vol. CT-18, pp. 114-122, January 1971.

12. S. W. Director, "Network design by mathematical optimization," 1969 Wescon Technical Papers, vol. 23, pp. 3.1-3.9, August 1969.

13. J. W. Bandler and R. E. Seviora, "Current trends in network optimization," IEEE Transactions of Microwave Theory and Techniques, vol. MTT-18, pp. 1159-1170, December 1970.

14. B. A. Wooley, "Automated design of dc-coupled monolithic broad-band amplifiers," IEEE Journal of Solid-State Circuits, vol. SC-6, pp. 24-34, February 1971.

15. D. A. Wayne, "Harmonic analysis and design of nonlinear electronic circuits," Ph.D. dissertation, University of Florida, 1971.

16. P. M. Russo and R. A. Rohrer, "Computer optimization of the transient response of an ECL gate," IEEE Transactions on Circuit Theory, vol. CT-18, pp. 197-199, January 1971.

17. J. J. Ebers and J. L. Moll, "Large-signal behavior of junction transistors," Proceedings of the IRE, vol. 43, pp. 1761-1772, December 1954.

18. F. A. Lindholm, S. W. Director, and D. L. Bowler, "Assessing model adequacy and selecting model complexity in integrated-circuit simulation," IEEE Journal of Solid-State Circuits, vol. SC-6, pp. 213-222, August 1971.

19. W. G. Howard, "Device modelling for computer-aided circuit analysis," 1969 Wescon Technical Papers, vol. 23, pp. 2.1-2.6, August 1969.

20. J. M. Early, "Effects of space-charge layer widening in junction transistors," Proceedings of the IRE, vol. 40, pp. 1401-1406, November 1952.

21. C. L. Searle et al., Elementary Circuit Properties of Transistors, SEEC, vol. 3. New York: Wiley, 1964.

22. F. G. Gustavson, W. Liniger, and R. Willoughby, "Symbolic generation of an optimal Crout algorithm for sparse systems of linear equations," Journal of the ACM, vol. 17, pp. 87-109, January 1970.

23. S. W. Director, "LU factorization in network sensitivity computations," IEEE Transactions on Circuit Theory, vol. CT-18, pp. 184-185, January 1971.

24. W. F. Tinney and J. W. W. Walker, "Direct solutions of sparse network equations by optimally ordered triangular factorization," Proceedings of the IEEE, vol. 55, pp. 1801-1809, November 1967.

25. J. Dickens, "A nonlinear circuit analysis program," Master's Thesis, University of Florida, 1971.

26. T. E. Grapes, "A generating program for the explicit solution of sparse matrix equations," to be published.

27. J. W. Bandler and C. Charalambous, "On conditions for optimality in least pth approximation with $p \to \infty$," Proceedings 9th Annual Allerton Conference on Circuit and System Theory, Urbana, Illinois, pp. 404-413, October 1971.

28. R. Fletcher, "A new approach to variable metric algorithms," The Computer Journal, vol. 13, pp. 317-322, 1970.

29. R. Fletcher and M. J. D. Powell, "A rapidly convergent descent method for minimization," The Computer Journal, vol. 6, pp. 163-168, 1963.

30. R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," The Computer Journal, vol. 7, pp. 149-153, 1964.

31. R. Fletcher, "Function minimization without evaluating derivatives: A review," The Computer Journal, vol. 8, pp. 33-41, 1965.

32. M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," The Computer Journal, vol. 7, pp. 155-161, 1964.

33. J. A. Nelder and R. Mead, "A simplex method for function minimization," The Computer Journal, vol. 7, pp. 308-313, 1964.

34. M. J. Box, "A comparison of several current optimization methods, and the use of transformations in constrained problems," The Computer Journal, vol. 9, pp. 67-77, 1966.

35. G. C. Temes and D. A. Calahan, "Computer-aided network optimization: The state of the art," Proceedings of the IEEE, vol. 55, pp. 1832-1863, November 1967.

36. D. H. Jacobson and W. Oksman, "An algorithm that minimizes homogeneous functions of n variables in n+2 iterations and rapidly minimizes general functions," Division of Engineering and Applied Physics, Harvard University, Technical Report No. 618, October 1970.

37. R. M. Warner and J. N. Fordemwalt, Eds., Integrated Circuits: Design Principles and Fabrication. New York: McGraw-Hill, 1965.

38. B. D. H. Tellegen, "A general network theorem, with applications," Phillips Research Reports, vol. 7, pp. 259-269, August 1952.

39. C. W. Gear, "The automatic integration of ordinary differential equations," Communications of the ACM, vol. 14, pp. 176-179, March 1971.

40. A. J. Brodersen, S. W. Director, and W. A. Bristol, "Simultaneous automated ac and dc design of linear integrated circuit amplifiers." IEEE Transactions on Circuit Theory, vol. CT-18, pp. 50-58, January 1971.

BIOGRAPHICAL SKETCH

William Arthur Bristol was born August 1, 1944, at Chicago, Illinois. In June, 1962, he was graduated from Miami Senior High School. In December, 1966, he received the degree of Bachelor of Science in Electrical Engineering with high honors from the University of Florida. In January, 1967, he enrolled in the Graduate School of the University of Florida where he received the degree Master of Science in Engineering. During this time he worked as a graduate assistant at the University Computing Center. He continued in the Graduate School of the University of Florida and pursued work toward the degree Doctor of Philosophy. While working for this degree he has held a Graduate Assistantship and a Radiation Fellowship.

William Arthur Bristol is married to the former Susan Louise Van Netta. He is a member of Phi Kappa Phi, Tau Beta Pi, Sigma Tau, Eta Kappa Nu, The Institute of Electrical and Electronics Engineers, and the Association for Computing Machinery.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

S. W. Director

Dr. Stephen W. Director, Chairman
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Arthur Brodersen

Dr. Arthur J. Brodersen
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

T. E. Bullock

Dr. Thomas E. Bullock
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Frank D. Vickers

Dr. Frank D. Vickers
Associate Professor of Computer and
Information Sciences

This dissertation was submitted to the Dean of the College of Engineering and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December, 1972

Robert E. Uhrig

Dean, College of Engineering

Dean, Graduate School