CIRCUIT THEORY AND APPLICATIONS, VOL. 2, 133–147 (1974)

A THREE-TERMINAL PIECEWISE-LINEAR MODELLING APPROACH TO DC ANALYSIS OF TRANSISTOR CIRCUITS*

I. N. HAJJ, D. J. ROULSTON AND P. R. BRYANT
Department of Electrical Engineering, University of Waterloo, Ontario, Canada

SUMMARY

The simulation of electronic circuits by computer has become an important part of present-day circuit analysis and design, especially in the area of integrated circuit design. One of the goals in computer simulation of integrated circuits is to have a program 'package' for which the input consists of chip fabrication data (mask dimensions, impurity profiles, material data such as carrier lifetimes) and the output displays the complete circuit response. This requires both an efficient modelling approach and a fast circuit analysis method.

In this paper a simulation method is described which generates dc responses (in the form of operating points or transfer characteristics) of transistor circuits directly from physical parameter data. The basis of the method is a two-dimensional piecewise-linear approach to the dc modelling of bipolar transistors. The model is directly used in a piecewise-linear circuit analysis program to simulate the dc response of a given circuit.

INTRODUCTION

The analysis of electronic circuits by computer has become an important part of present-day circuit analysis and design. The computer-aided simulation of a given electronic circuit offers a cheaper and faster method of testing the performance of a circuit design than the actual building of the circuit in the laboratory and the carrying out of experimental measurements. The correct correlation between the computer simulated results and the experimental results, however, depends primarily on two important factors. The first is the accuracy of the computational method used in the simulation; the second is the accuracy of the modelling technique used in characterizing the devices in the circuit.

The most common approach used up to now in computer-aided circuit analysis and design has been the separation of the circuit problem from the device problem. Thus in almost all cases, the circuit equations are formulated with the assumption that the network is composed of two-terminal elements.[1–5] Many electronic devices, however, have three or more terminals, hence the most commonly used modelling approach has dealt with representing the equivalent circuit models of these devices as interconnections of two-terminal circuit elements.† Instead of treating the circuit analysis problem and the device modelling problem separately, we consider both problems together.

In this paper we confine our study to dc analysis and consider only double diffused bipolar transistors (similar to devices used in integrated circuits). The goal is to have a 'package' for which the input consists of chip fabrication data (e.g., mask dimensions, impurity profiles, material data such as carrier lifetimes) from which output data (in the form of dc operating points or dc transfer characteristics of a complete integrated circuit) is automatically generated.

Our simulation of a transistor is based upon Stern's generalized piecewise-linear characterization.[7] This approach does not require that the model be expressed as an interconnection of two-terminal elements;

* The research reported herein has been supported by the National Research Council of Canada under Grant number A-7113 and Grant number A-4321.

† An example is the Ebers–Moll model of the bipolar transistor.[6] We consider a two-terminal element as an element, the terminal characteristics of which is defined in terms of two variables only; thus a controlled source is considered as a two-terminal element.

Received 25 April 1973
Revised 16 August 1973

rather the behaviour of the transistor is characterized by its terminal description which is then approximated by an appropriate piecewise-linear representation. The terminal description is handled in terms of tabulated sampled points, which may be obtained either experimentally or, as described in this paper, numerically from equations governing the physical behaviour of the transistor. This approach is not new,[8] but the generation of the characteristics directly from the physical parameters of the device is believed to be new. A program has been written which generates in this way a Stern model of a bipolar transistor and this program is used as a part of a larger dc analysis program. Hence, the computation of the dc solutions of transistor circuits directly from a given set of transistor physical parameters is achieved.

In the second section of this paper we briefly review the three-terminal piecewise-linear characterization method. In the third section we present a method of generating the dc characteristics of bipolar transistors given the values of the device physical parameters. In the fourth section we describe an algorithm for computing the dc solutions and the dc transfer characteristics of circuits containing three-terminal elements. In the fifth section we give a short description of a computer program which computes the dc solutions of transistor circuits from a given set of values of device physical parameters and circuit description. Some examples are also given along with a comparison between simulated results and experimental measurements.

PIECEWISE-LINEAR CHARACTERIZATION OF THREE-TERMINAL ELEMENTS

Piecewise-linear characterization of two-terminal elements has been frequently used in the analysis of electrical circuits,[5,9,12,13] but then three-terminal elements, such as transistors, had to be represented as interconnections of two-terminal elements. Only recently has the generalized piecewise-linear characterization of n-terminal elements been used by Ohtsuki and Yoshida to represent the dc characteristics of transistors.[8] The method was first proposed by Stern.[7] The advantage of the method is that it does not require that the n-terminal element be represented as an interconnection of two-terminal elements. However, from the practical point of view, the method becomes numerically involved for more than three-terminal elements. Thus, in this section we review briefly the dc characterization of three-terminal elements only.

Let the terminal characteristics of a three-terminal element be given in the form

$$\begin{aligned} y_1 &= f_1(x_1, x_2) \\ y_2 &= f_2(x_1, x_2) \end{aligned} \tag{1}$$

Where (y_1, y_2) is a set of currents and/or voltages at the two ports formed between each of any two terminals and the third, and (x_1, x_2) is the complementary set of voltages and/or currents, as shown in Figure 1.

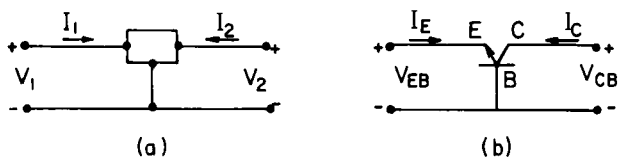

Figure 1. (a) A three-terminal element. (b) A three-terminal NPN transistor

Each of the two equations in (1) represents a surface in three-dimensional space. In order to linearize these two surfaces, a set of sampling points is selected to be regularly spaced throughout the $x_1 - x_2$ space, as shown in Figure 2(a). The problem is to approximate each function $f_i(x_1, x_2)$, $i = 1, 2$, by a continuous, piecewise-linear function $\hat{f}_i(x_1, x_2)$ such that

$$\hat{f}_i(x_1, x_2) = f_i(x_1, x_2), \qquad i = 1, 2 \tag{2}$$

at the selected sampling points. The value of $\hat{f}_i(x_1, x_2)$ at the sampling points will be referred to as the sampled points. As seen in Figure 2(a), the sampling points form grid points on the $x_1 - x_2$ plane. The grid points can be connected by straight lines which are orthogonal to one of the co-ordinates. The grid lines divide the

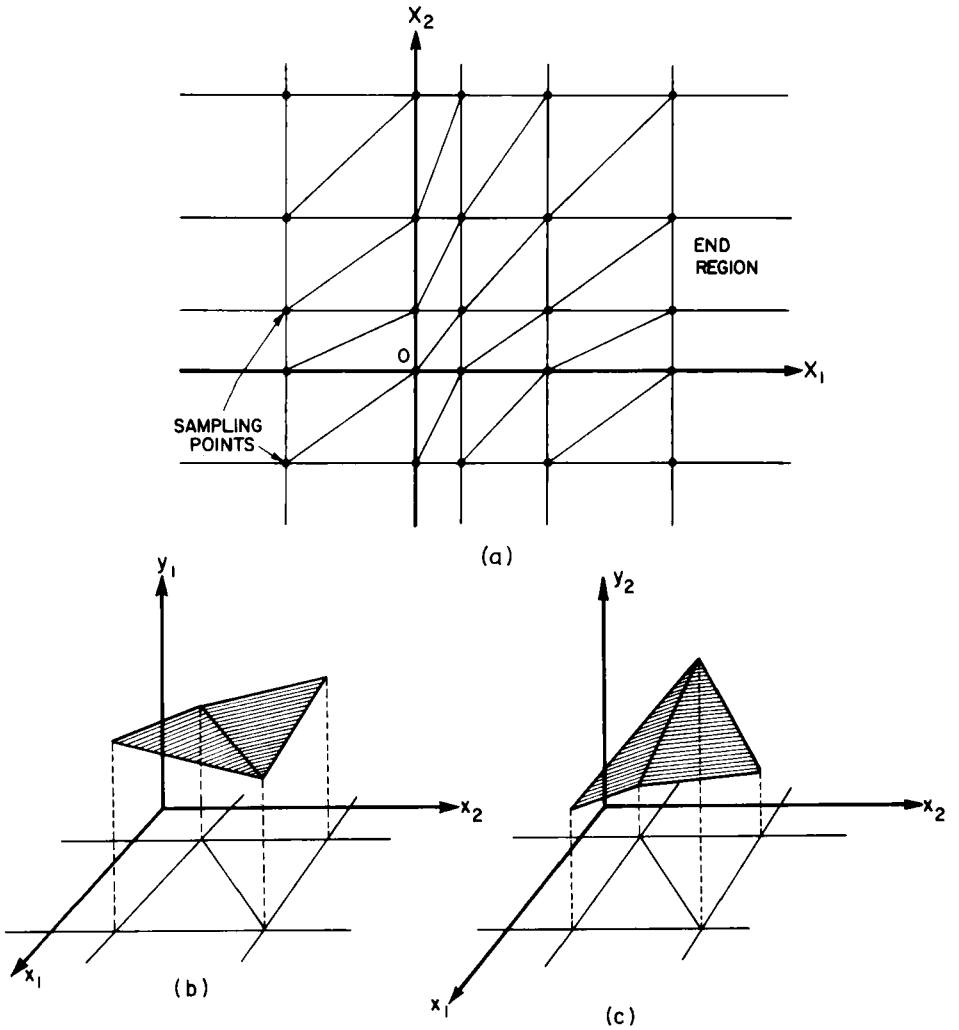

Figure 2. (a) Sampling points and boundary lines in the $x_1 - x_2$ plane. (b, c) Piecewise-linear characteristics in one of the rectangular regions

plane into rectangular regions. Each closed† rectangular region is, in turn, divided into two triangular regions, in each of which the characteristic is linear and is of the form

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + a_{13} \\ y_2 &= a_{21}x_1 + a_{22}x_2 + a_{23} \end{aligned} \tag{3a}$$

or, in matrix form

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \tag{3b}$$

Note that each closed rectangular region is subdivided into two triangular regions because although in general a plane cannot be made to pass through four sampled points, it can be made to pass through three points.

† As seen in Figure 2(a), some regions are not closed. These are referred to as end regions. For computational purposes, the characteristics in the end regions are arbitrarily specified.

It should be pointed out that in order to construct the piecewise-linear approximation, it is not necessary to know the original function provided it is known to be continuous. It is only necessary to obtain tabulated sampled points, which will then determine the piecewise-linear characteristics, except in the end regions. In practice, the tabulated points may be obtained either experimentally from actual measurements, or numerically from equations governing the behaviour of the three-terminal device. The following section describes the numerical technique we use to obtain the tabulated sampled points from the physical parameters of bipolar transistors.

DESCRIPTION OF METHOD USED TO OBTAIN TRANSISTOR DC CHARACTERISTICS

The aim of the modelling method described in this section is to develop a subroutine which can be used as a part of a larger circuit analysis program. The requirements of this subroutine are as follows:

1. The input data must be in-the form of process parameters (mask dimensions, impurity profile parameters, material parameters).
2. The subroutine must be rapid in execution time and economical in core requirements.
3. Accuracy should be compatible with, but not exceed, the limits imposed, on the one hand, by the accuracy with which the final impurity profile and material data (such as carrier lifetime data) can be related to known process parameters (surface concentrations, oven temperatures, dopants, etc.); and, on the other hand, by the accuracy to which the final response is computed by the piecewise-linear method described in the next section.
4. The method must be capable of characterizing the transistor at any bias levels encountered in the final circuit. This will normally mean collector-base voltages between breakdown and saturation and collector currents from low values (where the gain falls off due to space charge recombination effects) to high values at which the characteristics are very degraded (gain fall-off due to high level injection effects, etc.), or the heat dissipation rating (a function of the thermal resistance) is exceeded.
5. The subroutine must have built-in algorithms to select voltage or current ranges, but it should be possible to override these automatic choices if desired.

It was decided that the most economic way of interfacing such a 'device characterization' program with a 'circuit analysis' program was to generate the complete I–V characteristic tables (V_{CB}, I_C, V_{EB}, I_E), as shown in Table I. These tables are similar to those used in Reference 8. Since the bipolar transistor is essentially a current-controlled device, and since there is almost a linear relationship between injected carrier concentrations and current densities, it was initially decided to define the collector current density J_C as one independent variable. The second obvious choice is collector-base voltage V_{CB}. The base-emitter voltage V_{EB} and base current I_B then become dependent variables. Unfortunately, the transistor is a two-dimensional device when operating close to, or beyond 'nominal operating' conditions. It does not appear to be possible to specify V_{CB} *and* I_C independently without recourse to costly iteration procedures. Consequently we were free to choose which variables we designated as independent variables and for reasons of convenience a voltage controlled description was chosen.

Before we describe the modelling method, it should be pointed out that the major source of inaccuracies in practical fabrication facilities lies in device profile and lifetime data. It is felt however, that this can be taken care of in an established calibrated system by defining 'effective' parameters for the basic profiles (erfc, Gaussian) whereby effects such as 'base push out' can be included by simple empirical formulae. The controlled variables are generally the diffusion times, surface concentrations of dopants, and temperatures. In practice, however, an oven is calibrated for a given dopant, given surface concentration, and fixed oven temperature; the only variable which remains is the diffusion time. It is possible by using simple techniques such as bevelling and staining combined with sheet resistivity, or using more sophisticated methods such as depletion layer capacitance-voltage plots, ionic probing, or radioactivity methods to precalibrate the system and thus to predict the complete double diffused profile in the form of sums of simple analytic

Table I. Characteristic tables for NPN (CSF) transistor. The top table gives values of I_E, the lower table gives values of I_C. For heavily saturated conditions, an approximate solution has been used in order to reduce execution time.

*** TERMINAL CHARACTERISTICS WITH EXTENSION ***

IE TABULATION (1ST COLUMN IS VEB, 1ST ROW IS VCB)

0.000E 00	-0.100E 05	-0.500E 00	0.000E 00	0.100E 01	0.158E 01	0.251E 01	0.398E 01	0.631E 01	0.100E 02	0.500E 03
-0.100E 05	-0.100E 03	-0.100E 03	-0.100E 03	-0.100E 03	-0.100E 03	-0.100E 03	-0.100E 03	-0.100E 03	-0.100E 03	-0.100E 03
-0.780E 00	-0.416E 01	-0.416E 01	-0.416E 01	-0.122E 01	-0.154E 01	-0.154E 01	-0.233E 01	-0.176E 01	-0.231E 01	-0.231E 01
-0.760E 00	-0.129E 01	-0.129E 01	-0.129E 01	-0.470E 00	-0.115E 01	-0.116E 01	-0.126E 01	-0.109E 01	-0.138E 01	-0.138E 01
-0.743E 00	-0.448E 00	-0.448E 00	-0.449E 00	-0.249E 00	-0.886E 00	-0.885E 00	-0.939E 00	-0.847E 00	-0.100E 01	-0.100E 01
-0.728E 00	-0.177E 00	-0.177E 00	-0.178E 00	-0.178E 00	-0.814E 00	-0.806E 00	-0.831E 00	-0.722E 00	-0.833E 00	-0.833E 00
-0.718E 00	-0.863E-01	-0.863E-01	-0.864E-01	-0.154E 00	-0.667E 00	-0.636E 00	-0.660E 00	-0.458E 00	-0.578E 00	-0.578E 00
-0.696E 00	-0.430E-01	-0.430E-01	-0.431E-01	-0.143E 00	-0.298E 00	-0.297E 00	-0.296E 00	-0.224E 00	-0.277E 00	-0.277E 00
-0.672E 00	-0.303E-01	-0.303E-01	-0.304E-01	-0.607E-01	-0.132E 00	-0.132E 00	-0.133E 00	-0.991E-01	-0.126E 00	-0.126E 00
-0.651E 00	-0.268E-01	-0.268E-01	-0.269E-01	-0.272E-01	-0.607E-01	-0.612E-01	-0.616E-01	-0.455E-01	-0.591E-01	-0.591E-01
-0.631E 00	-0.126E-01	-0.126E-01	-0.126E-01	-0.128E-01	-0.286E-01	-0.288E-01	-0.291E-01	-0.214E-01	-0.280E-01	-0.280E-01
-0.611E 00	-0.576E-02	-0.576E-02	-0.580E-02	-0.588E-02	-0.139E-01	-0.140E-01	-0.141E-01	-0.102E-01	-0.135E-01	-0.135E-01
0.000E 00	0.000E 00	0.000E 00	0.000E 00	0.000E 00	0.000E 00	0.000E 00	0.000E 00	0.000E 00	0.000E 00	0.000E 00
0.100E 05	0.100E 01	0.100E 01	0.100E 01	0.100E 01	0.100E 01	0.100E 01	0.100E 01	0.100E 01	0.100E 01	0.100E 01

IC TABULATION (1ST COLUMN IS VEB,1ST ROW IS VCB)

0.000E 00	-0.100E 05	-0.500E 00	0.000E 00	0.100E 01	0.158E 01	0.251E 01	0.398E 01	0.631E 01	0.100E 02	0.500E 03
-0.100E 05	-0.100E 05	0.253E-01	0.254E-01	0.139E 00	0.981E 00	0.990E 00	0.185E 01	0.153E 01	0.179E 01	0.100E 05
-0.780E 00	-0.100E 05	0.253E-01	0.254E-01	0.139E 00	0.981E 00	0.990E 00	0.185E 01	0.153E 01	0.179E 01	0.100E 05
-0.760E 00	-0.100E 05	0.253E-01	0.254E-01	0.139E 00	0.981E 00	0.990E 00	0.113E 01	0.995E 00	0.126E 01	0.100E 05
-0.743E 00	-0.100E 05	0.253E-01	0.254E-01	0.139E 00	0.807E 00	0.808E 00	0.876E 00	0.803E 00	0.950E 00	0.100E 05
-0.728E 00	-0.100E 05	0.253E-01	0.254E-01	0.139E 00	0.777E 00	0.770E 00	0.799E 00	0.702E 00	0.806E 00	0.100E 05
-0.718E 00	-0.100E 05	0.253E-01	0.254E-01	0.139E 00	0.648E 00	0.619E 00	0.643E 00	0.447E 00	0.563E 00	0.100E 05
-0.696E 00	-0.100E 05	0.253E-01	0.254E-01	0.139E 00	0.290E 00	0.290E 00	0.289E 00	0.218E 00	0.269E 00	0.100E 05
-0.672E 00	-0.100E 05	0.253E-01	0.254E-01	0.582E-01	0.128E 00	0.128E 00	0.128E 00	0.956E-01	0.122E 00	0.100E 05
-0.651E 00	-0.100E 05	0.253E-01	0.254E-01	0.257E-01	0.582E-01	0.587E-01	0.592E-01	0.434E-01	0.565E-01	0.100E 05
-0.631E 00	-0.100E 05	0.116E-01	0.116E-01	0.118E-01	0.271E-01	0.273E-01	0.275E-01	0.201E-01	0.264E-01	0.100E 05
-0.611E 00	-0.100E 05	0.530E-02	0.534E-02	0.542E-02	0.129E-01	0.129E-01	0.131E-01	0.938E-02	0.125E-01	0.100E 05
0.000E 00	-0.100E 05	0.530E-02	0.534E-02	0.542E-02	0.129E-01	0.129E-01	0.131E-01	0.938E-02	0.125E-01	0.100E 05
0.100E 05	-0.100E 05	0.530E-02	0.534E-02	0.542E-02	0.129E-01	0.129E-01	0.131E-01	0.938E-02	0.125E-01	0.100E 05

functions. In this paper, therefore, we shall assume that the impurity profile can be varied at will, in a calibrated fashion, and that the lifetime τ_{p0} is known.†

Description of analysis

The *variable boundary regional* method[10] was found to be a fast, efficient way of generating device characteristics over a wide range of currents and voltages. The method involves an approximate two-dimensional numerical analysis of the transistor. It is based on using a set of differential equations describing the current flow in the longitudinal (emitter-collector) x-direction to generate a table of values of current density, base conductance, current gain, and injected carrier concentration. These values are then used in a separate differential equation in the neutral base region to describe the current flow in the lateral y-direction and to solve for the terminal characteristics. The method described here is designed for an NPN double-diffused transistor.

The equations used in describing the current flow in the x-direction are:

$$[\mathbf{J}_p(x, y)]_x = -qV_t\mu_p(x, y)\frac{\partial p(x, y)}{\partial x} + qp(x, y)\,\mu_p(x, y)[\mathbf{E}(x, y)]_x \tag{4}$$

$$[\mathbf{J}_n(x, y)]_x = qV_t\mu_n(x, y)\frac{\partial n(x, y)}{\partial x} + qn(x, y)\mu_n(x, y)[\mathbf{E}(x, y)]_x \tag{5}$$

$$\frac{\partial[\mathbf{E}(x, y)]_x}{\partial x} = \frac{q}{\varepsilon}[p(x, y) - n(x, y) - N(x)] \tag{6}$$

$$\frac{\partial[\mathbf{J}_p(x, y)]_x}{\partial x} = \frac{-q}{\tau_{p0}}[p(x, y) - |n_i^2/N(x)|] \tag{7}$$

Where $\mathbf{J}_p(x, y)$ and $\mathbf{J}_n(x, y)$ are hole and electron current densities; $\mu_p(x, y)$ and $\mu_n(x, y)$ are the concentration and field dependent mobilities; $V_t \approx 0{\cdot}026$ volts at 300°K; $\mathbf{E}(x, y)$ is the electric field; q is the electron charge; ε is the permittivity; $n(x, y)$ and $p(x, y)$ are the free-electron and hole concentrations; τ_{p0} is the hole lifetime in the emitter region; and $N(x)$ is the net impurity profile which is defined in terms of process parameters such as surface concentration, background doping and junction depths. It should be noted that μ_p and μ_n are functions of both $N(x)$ and $\mathbf{E}$; convenient analytic (empirical) expressions, given in Reference 11, are used in the program. Similar equations to (4)–(7) apply in the y-direction.

A. *Longitudinal analysis*

In the longitudinal x-direction a cross-section of the transistor is divided into five regions (see Figure 3(a), (b)): (I) emitter neutral region, $0 \leqslant x \leqslant x_{en}$, (II) emitter-base space-charge region, $x_{en} \leqslant x \leqslant x_{eb}$, (III) neutral-base region, $x_{eb} \leqslant x \leqslant x_{bc}$, (IV) base-collector space-charge region, $x_{bc} \leqslant x \leqslant x_{cn}$, and (V) collector neutral (ohmic) region, $x_{cn} \leqslant x \leqslant W_{epi}$. The boundaries between these regions are current and voltage dependent. The analysis in the x-direction is carried out at an arbitrary value of y, and the results are assumed to apply at any y where $-L/2 \leqslant y \leqslant L/2$.

The steps carried out in the longitudinal analysis are explained in conjunction with the flow chart given in Figure 4. To start with, a table of values of the impurity profile, $N(x)$, is generated with a variable step in the x-direction giving a roughly constant ratio between two successive values of $N(x)$. These steps, which are determined by the shape of the curve $N(x)$, will be used as integration steps when solving equations (4)–(7) in the different regions of the transistor. For an NPN double-diffused transistor, a typical impurity profile, $N(x)$, is shown in Figure 3(b).

†Recently using effective n_i^2 values for high doping levels communicated by DeMan and Overstraeten we have modified the effective emitter profile using the results in Ref. 17 and found that in many transistors the maximum value of h_{FES} can be accurately predicted by setting $\tau_{p0} = \infty$. The exact value of τ_{p0} must still be known however for prediction of low current fall off.

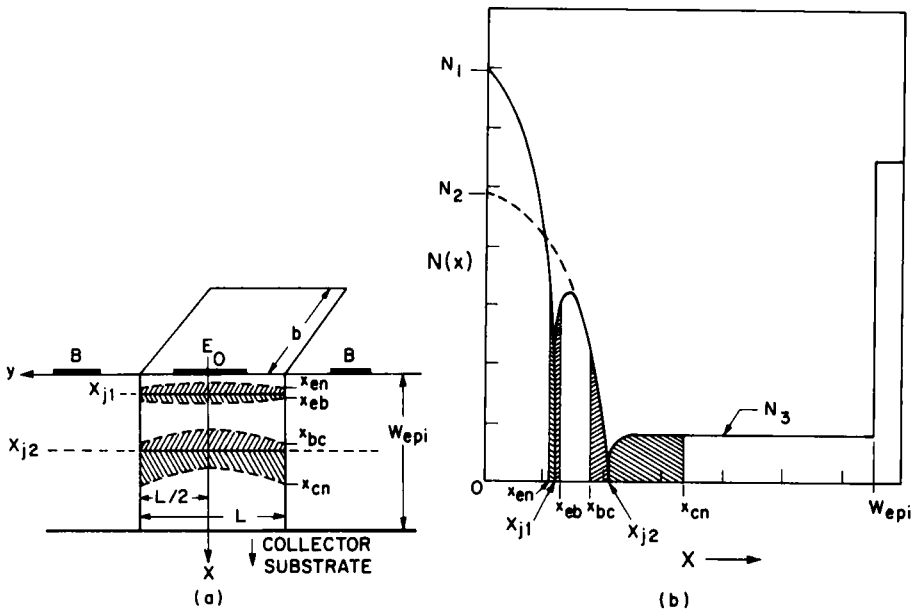

Figure 3.(a) Transistor structure. L is the width of the region in which the analysis is carried out; b is the length of all stripes added together in a multistripe structure. (b) Impurity profile; space-charge regions are shown shaded.

The analysis is divided into five different parts which correspond to the five regions mentioned above. We shall discuss the analysis in each of these regions in the sequence given in the flow chart. Hence, we start with region II.

(a) *Emitter-base space-charge region (region II).* The purpose of the analysis in this region is to obtain tabulated values of region boundaries ($x_{en}(y)$ and $x_{eb}(y)$), and carrier concentrations ($p(x_{en}, y)$ and $n(x_{eb}, y)$) at selected values of $[\mathbf{E}(X_{j1}, y)]_x$, where X_{j1} is the physical junction between the emitter and the base. $x_{en}(y)$ and $x_{eb}(y)$ are determined by finding the points where $[\mathbf{E}(x, y)]_x = 0$. Starting with a selected value of $[\mathbf{E}(X_{j1}, y)]_x$, and assuming that $n(x, y)$ and $p(x, y)$ are both negligible compared with $N(x)$ throughout this region, Poisson's equation (6) is integrated in both directions to determine $x_{en}(y)$ and $x_{eb}(y)$ and to obtain the total voltage V_{jeb} across the region. By assuming that $[\mathbf{J}_p(x, y)]_x$ and $[\mathbf{J}_n(x, y)]_x$ are each small compared with the respective diffusion and drift components in (4) and (5) the Boltzmann relations are obtained and we can write, for a given y:

$$n(x_{eb}, y) = \frac{N(x_{eb}) - N(x_{en}) \exp(V_{jeb}/V_t)}{\exp(2V_{jeb}/V_t) - 1} \tag{8}$$

and

$$p(x_{en}, y) = -\frac{N(x_{en})}{2} \cdot \left\{ \left[1 + \frac{4n(x_{eb}, y)[N(x_{eb}) + n(x_{eb}, y)]}{N(x_{en})^2} \right]^{\frac{1}{2}} - 1 \right\} \tag{9}$$

In addition, the excess hole charge, Q_{eb}, in this region is computed by integrating the excess carrier concentration. Q_{eb} is needed to find the low level current gain at a later stage in the computation.

A user specified nominal value V_{CBn} of V_{CB} is now defined. A value $J1$ for the initial current density $[\mathbf{J}_n(x, y)]_x$ is selected and is assumed constant[10] for $x \geqslant x_{eb}$.

(b) *Neutral base (region III).* In this region, space charge neutrality is assumed and (6) gives an elementary relation between $p(x, y)$, $n(x, y)$ and $N(x)$. Moreover the hole current may safely be set equal to zero in (4). Combining with (5) the result of setting (4) and (6) equal to zero gives a simple first order differential equation relating $\partial n(x, y)/\partial x$ to $n(x, y)$, $N(x)$, $\mu_n(x, y)$. The resulting equation is integrated to determine n as a function

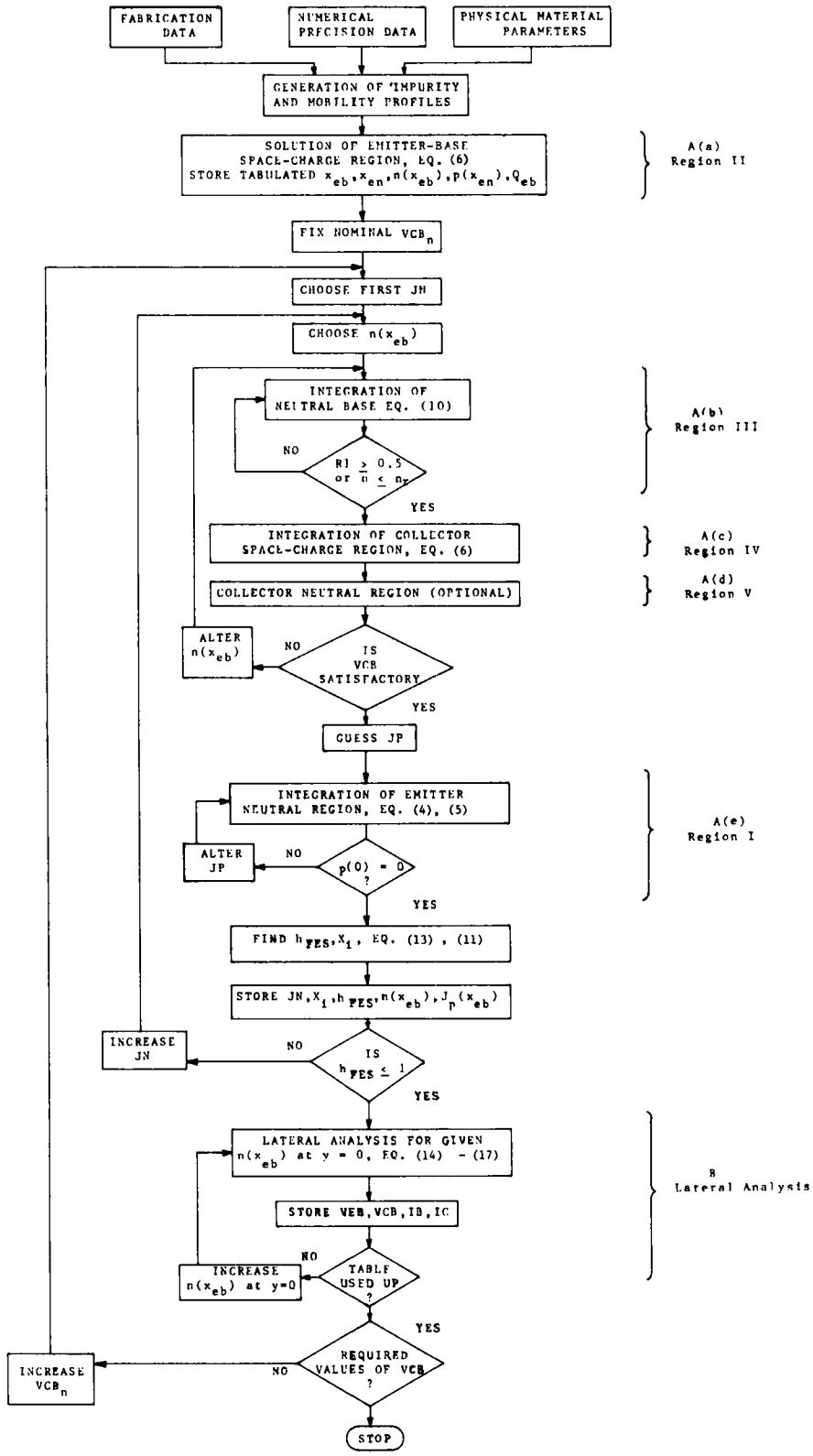

Figure 4. Flow chart of BIPOLAR

of x. The boundary $x_{bc}(y)$ is determined by one of the two conditions

$$\left.\begin{aligned} & n(x, y) \leqslant 2[\mathbf{J}_n(x, y)]_x/qv_{th} \triangleq n_r \\ \text{or} \\ & R_1 \triangleq \frac{\varepsilon}{q[n(x, y)+N(x)]} \frac{\partial[\mathbf{E}(x, y)]_x}{\partial x} \geqslant 0{\cdot}5 \end{aligned}\right\} \tag{10}$$

where v_{th} is the thermal velocity of the carriers ($\simeq 10^7$ cm/sec for electrons in silicon). The variables in (10) are computed point by point and when either inequality is satisfied the integration is stopped. In addition, an effective base width, X_i, which is essential to the lateral analysis, is computed at this stage, where

$$X_i(y) = \frac{1}{[N(x_{eb})+n(x_{eb}, y)]\mu_p(x_{eb}, y)} \int_{x_{eb}}^{x_{bc}} [N(x)+n(x, y)]\mu_p(x, y)\,\mathrm{d}x \tag{11}$$

(c) *Base-collector space-charge layer (region IV).* In this region, starting at $x_{bc}(y)$ as defined above, $p(x, y)$ and $[\mathbf{J}_p(x, y)]_x$ are assumed to be zero; moreover, it is assumed that the electron current is dominated by the drift component and the first term on the right hand side of (5) is set equal to zero. Combining (5) and (6) gives a simple first order differential equation for $\partial[\mathbf{E}(x, y)]_x/\partial x$ in terms of $[\mathbf{E}(x, y)]_x$, $N(x)$, and $\mu_n(x, y)$. Integration is continued until the boundary $x_{cn}(y)$ defined by $n(x, y)/N(x) \approx 0{\cdot}95$ is reached.

(d) *Collector neutral region (region V).* In this region we have only majority carrier drift current $[\mathbf{J}_n(x, y)]_x$ to be taken into account; moreover the field is low and the mobility $\mu_n(x, y)$ is therefore constant. The voltage drop across this region is therefore proportional to $[\mathbf{J}_n(x, y)]_x$. At this point V_{CB} is compared with the user specified value V_{CBn}. If V_{CB}/V_{CBn} is not within a user specified bound, $n(x_{eb}, y)$ is modified and steps (b), (c) and (d) are repeated. After a satisfactory V_{CB} is obtained, the analysis moves back to region I.

(e) *Emitter neutral region (region I).* In this region equations (4), (5) and (7) are solved to determine the hole current density $[\mathbf{J}_p(x_{en}, y)]_x$, given the boundary condition $p(x_{en}, y)$ and given $p(0, y) = 0$. The hole current density in the base region, $[\mathbf{J}_p(x_{eb}, y)]_x$, is given by

$$[\mathbf{J}_p(x_{eb}, y)]_x = [\mathbf{J}_p(x_{en}, y)]_x + \frac{Q_{eb}(y)}{\tau_{p0}} \tag{12}$$

It may be pointed out at this stage that the current gain, h_{FES}, can be obtained from†:

$$h_{FES} = \frac{[\mathbf{J}_n(x_{eb}, y)]_x}{[\mathbf{J}_p(x_{eb}, y)]_x} \tag{13}$$

So far, for a given value of V_{CBn} and $[\mathbf{J}_n(x, y)]_x$, values of $x_{eb}(y)$, $N(x_{eb})$, $n(x_{eb}, y)$, $[\mathbf{J}_p(x_{eb}, y)]_x$, $X_i(y)$ have been obtained and stored. Next, $[\mathbf{J}_n(x, y)]_x$ is increased by a user specified factor and steps (b), (c), (d) and (e) are repeated for the specified V_{CBn}, until a table of values of $x_{eb}(y)$, $N(x_{eb})$, $n(x_{eb}, y)$, $[\mathbf{J}_p(x_{eb}, y)]_x$, $X_i(y)$, is obtained. This table is used in the lateral analysis to compute the terminal variables: V_{EB}, I_C, and I_B, as will be explained below. Computation for new (higher) values of $[\mathbf{J}_n(x, y)]_x$ is stopped when the gain, $h_{FES} \leqslant 1$.

B. *Lateral analysis*

An important approximation is made in order to represent the two dimensional behaviour of the transistor by two separate sets of one dimensional current flow equations. Referring to Figure 3(a) and considering the base current (hole current in an NPN device since $[\mathbf{J}_n(x, y)]_y = 0$), we can use the equivalent of equations (4) and (5), written in the y-direction. However, in order to determine the base current I_p we must integrate $[\mathbf{J}_p(x, y)]_y$ from x_{eb} to x_{bc}. Referring to equation (11) we see that the parameter X_i is an effective base width taking into account conductivity variations between x_{eb} and x_{bc}. The total base current

† See footnote on page 134.

$I_p(y)$ between x_{eb} and x_{bc} flowing in the y-direction can be related in a simple manner through (4) and (5) to X_i, μ_p, n, and N, all of which are functions of y through x_{eb} (see Figure 3(a).

Recognizing that the hole current density $[\mathbf{J}_p(x_{eb}, y)]_x$ calculated in the x-direction in the emitter and emitter-base space-charge regions, flows in the y-direction in the base region, it can be shown[10] that the variation of $n(x_{eb}, y)$ with respect to y is given by

$$\frac{\mathrm{d}n(x_{eb}, y)}{\mathrm{d}y} = \frac{n(x_{eb}, y)}{qV_tX_i(y)\mu_p(x_{eb}, y)[2n(x_{eb}, y) + N(x_{eb})]} \int_0^y [\mathbf{J}_p(x_{eb}, y)]_x \, \mathrm{d}y \tag{14}$$

Equation (14) is solved starting from the centre of the emitter stripe (the point at which $y = 0$), until the point $y = L/2$. In integrating (14), an initial value of $n(x_{eb}, 0)$ is chosen and subsequently, the values of $x_{eb}(y)$, $X_i(y)$, $\mu_p(x_{eb}, y)$, $N(x_{eb})$, and $[\mathbf{J}_p(x_{eb}, y)]_x$ at each integration step, are obtained from the table of values generated in the longitudinal analysis. The total collector current I_C is given by:

$$I_C = 2b \int_0^{L/2} [\mathbf{J}_n(x, y)]_x \, \mathrm{d}y \tag{15}$$

and the base current is defined by

$$I_B = 2b \int_0^{L/2} [\mathbf{J}_p(x_{eb}, y)]_x \, \mathrm{d}y \tag{16}$$

The base-emitter voltage is given by:

$$V_{EB} = -V_t \ln \left[n\left(x_{eb}, \frac{L}{2}\right) N(x_{eb}(L/2)) \Big/ n_i^2 \right] \tag{17}$$

where $x_{eb}(L/2) \triangleq$ the value of x_{eb} at $y = L/2$. A table of values of I_C, I_B, V_{EB} is generated by increasing the values of $n(x_{eb}, 0)$ successively by a user specified ratio. The computation stops when the table of values generated in the longitudinal analysis is 'used up', i.e. when $n(x_{eb}, L/2)$ exceeds the highest tabulated value (corresponding to a gain h_{FES} less than unity).

In order to obtain another set of terminal variables for another user specified value of V_{CB}, steps A(b), (c), (d), (e), and B are repeated.

It should be noted that the above method automatically includes nonlinear variations in the current gain due to space charge recombination (at low currents), base widening, conductivity modulation (at high currents), and emitter current crowding. In its present form the method covers the normal active and saturation regions (the latter is omitted from the flow chart to preserve clarity). The cut-off and inverse regions as well as breakdown will be added at a later date. Recent work which is not described by the above includes the effect of band-gap variation in the emitter[17] and the effect of charge injected into the extrinsic base region.

Second order (Euler–Cauchy) integration techniques are used in the present version of this program. The step size dx in the longitudinal analysis is related to the impurity profile gradient. A variable (current dependent) step routine is used in the lateral analysis to allow for increased precision at high currents when the emitter current crowding effect is severe.

DC ANALYSIS METHOD

DC analysis of piecewise-linear circuits has been the subject of many computer-aided circuit investigations.[8,9,12,13,15] Basically, there are two methods of obtaining the solutions. One is the Newton–Raphson method,[8,13] and the other is the Katzenelson method.[9,12,15] In both methods the network equation is expressed in the form

$$\mathbf{g}(\mathbf{w}) = \mathbf{u} \tag{18}$$

where $\mathbf{g}$ is the piecewise-linear function, $\mathbf{w}$ represents the network variables, and $\mathbf{u}$ the input vector. For a

continuous, piecewise-linear $\mathbf{g}$, the real Euclidean space R^n is divided into a finite number of polyhedral regions by a finite number of boundary hyperplanes.[15] Equation (18) becomes a set of linear equations of the form

$$[\mathbf{M}]_m\mathbf{w}+\mathbf{c}_m = \mathbf{u} \qquad (m = 1, 2, \ldots, r) \tag{19}$$

where $[\mathbf{M}]_m$ is a constant $n \times n$ matrix, $\mathbf{c}_m$ is a constant n-vector, and r is the total number of regions in R^n.

In the Newton–Raphson method an initial guess, $\mathbf{w}_0$, is chosen in an initial region m, and equation (19) is used to obtain

$$\mathbf{w}_1 = \mathbf{w}_0+[\mathbf{M}]_m^{-1}(\mathbf{u}-\mathbf{u}_0) \tag{20}$$

where

$$\mathbf{u}_0 = [\mathbf{M}]_m\mathbf{w}_0+\mathbf{c}_m \tag{21}$$

If $\mathbf{w}_1$ falls in the same region as $\mathbf{w}_0$, it is a true solution of (19); if not, $\mathbf{w}_1$ is then chosen as a new initial guess and the network equation is updated into the region in which $\mathbf{w}_1$ is located, and the process is repeated. However, a Newton–Raphson iterative sequence may not converge even if the network has a unique solution.[13]

A modification of the Newton–Raphson method can be made by changing equation (20) to become

$$\mathbf{w}_1 = \mathbf{w}_0+\lambda[\mathbf{M}]_m^{-1}(\mathbf{u}-\mathbf{u}_0) \tag{22}$$

where λ is chosen such that[8]

$$\|\mathbf{u}-\mathbf{u}_1\| < \|\mathbf{u}-\mathbf{u}_0\| \tag{23}$$

where $\|\cdot\|$ indicates the Euclidean norm. The choice of λ requires a line search. With this modification, convergence becomes faster, but is still not guaranteed.

In Katzenelson's method the scalar λ in equation (22) is chosen such that $0 \leqslant \lambda \leqslant 1$ is the largest value for which $\mathbf{w}_1$ lies on the boundary of region m. In this case it can be easily shown that[9]

$$(\mathbf{u}-\mathbf{u}_i) = (1-\lambda)(\mathbf{u}-\mathbf{u}_{i-1}) \tag{24}$$

where $\mathbf{u}_i = \mathbf{g}(\mathbf{w}_i)$ at the ith iteration. Equation (24) indicates that the vector $(\mathbf{u}-\mathbf{u}_i)$ always lies on the straight line joining $\mathbf{u}_0$ to $\mathbf{u}$. In addition, from equation (24) we get

$$\|\mathbf{u}-\mathbf{u}_i\| = (1-\lambda)\|\mathbf{u}-\mathbf{u}_{i-1}\| \tag{25}$$

Hence $\|\mathbf{u}-\mathbf{u}_i\|$ is always less than $\|\mathbf{u}-\mathbf{u}_{i-1}\|$ except when $\lambda = 0$. In this respect, Katzenelson's method can be viewed as a modified Newton–Raphson's.

In our circuit analysis program we use Katzenelson's method for the following reasons:

1. The method is guaranteed to converge to a solution if the network determinant, $\det [\mathbf{M}]_m$, does not change sign between regions.[15]
2. The method can be extended to find input–output characteristics of networks with unique solutions, and, in some cases, of networks with multiple solutions, such as flip-flops.[9]
3. At every iteration point the new region entered is adjacent to the previous one, and in most cases one boundary only is crossed; the network matrix, $[\mathbf{M}]_{m+1}$, in the new region is then different from the network matrix, $[\mathbf{M}]_m$, in the previous region by only a dyad‡; hence sparse matrix solution techniques with dyad updating of the LU factors can be used.[14,15]

COMPUTER PROGRAM AND EXAMPLES

A program has been developed in FORTRAN and run on IBM 360/75. The program consists of two major parts: a transistor modelling part and a dc analysis part. The transistor modelling program generates the transistor characteristic tables for a particular set of transistor physical parameters. Since only the value

‡ A dyad is a matrix of rank unity which can be expressed as a row vector premultiplied by a column vector, $\mathbf{ab}'$.

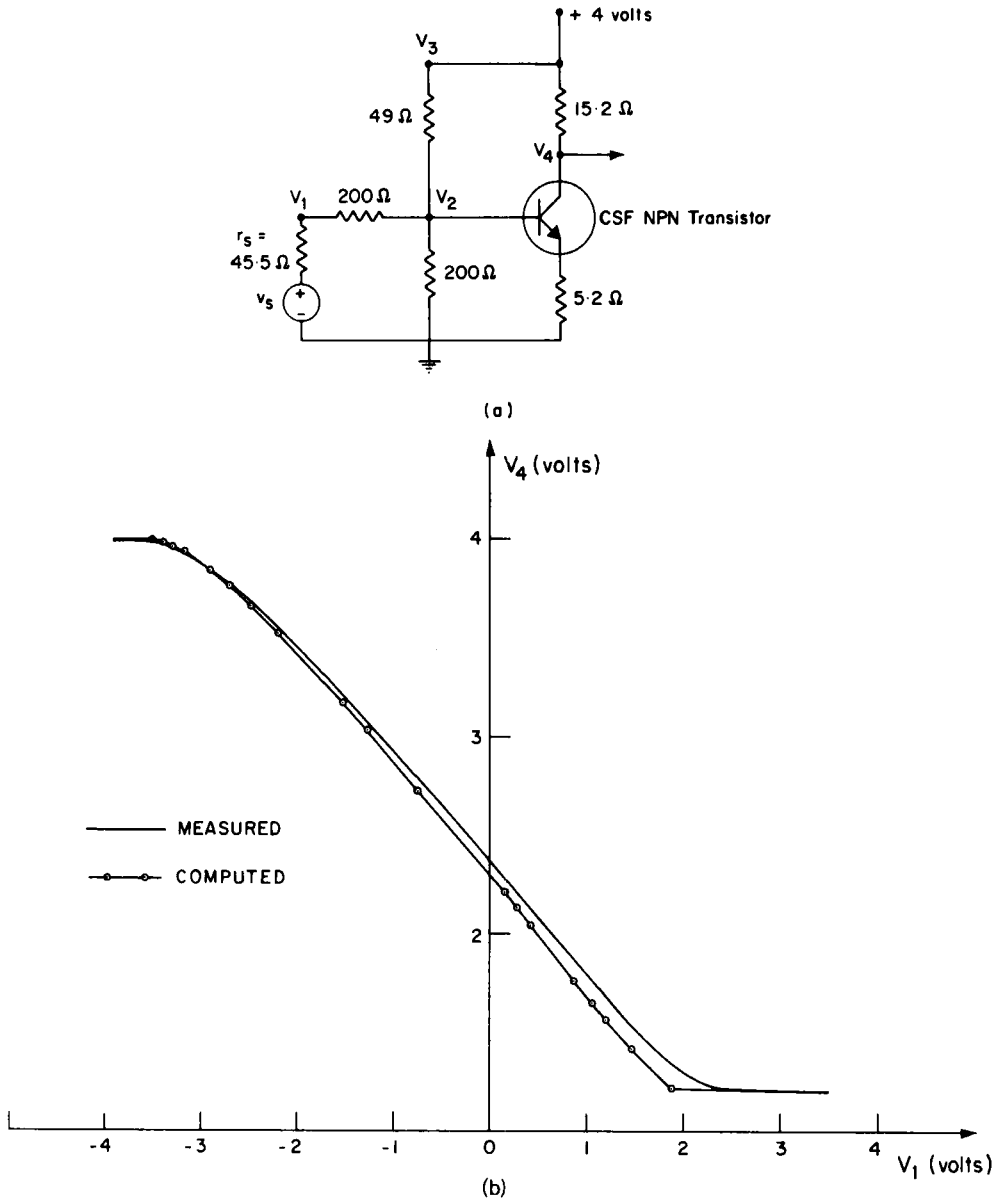

Figure 5. NPN (CSF) transistor amplifier circuit

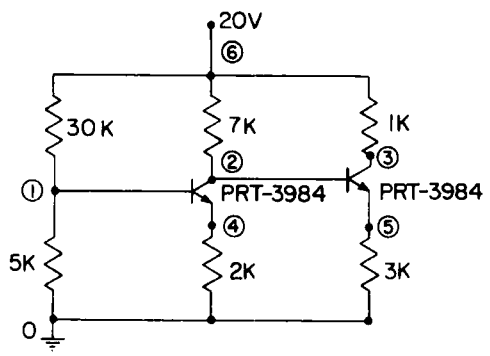

Figure 6. 2N1900 (TRW PRT 3984) two-transistor circuit

of V_{CB} can be externally specified and since for each specified V_{CB} a set of values of V_{EB}, I_E, I_C is computed, exponential interpolation is used to adjust the values of V_{EB}, I_E, I_C, so that V_{EB} is regularly spaced for all values of V_{CB}, as required by the dc analysis program. The output is in the form of two tables per transistor. A typical program requires about 37 kbytes or core memory on an IBM360/75, and a typical set of characteristic tables of $10 \times 10 \times 2$ points requires about 22 seconds using FORTRAN G. The dc program[16] uses nodal analysis and the output is given as node voltages or if required as dc transfer characteristics. It uses dynamic storage and sparse matrix techniques. A typical dc solution of a circuit containing 14 nodes and 12 three-terminal bipolar transistors with $10 \times 10 \times 2$ points in the characteristic tables requires about 4 seconds and a core memory of about 80 kbytes.

Example 1

One-transistor circuit. The simple circuit shown in Figure 5(a) was simulated and also built and tested in the laboratory. The transistor used was a Thomson CSF NPN double diffused bipolar transistor the physical parameters of which are available. The purpose of this example was to construct the DC input–output transfer characteristic between the input v_1 and the output v_4, and thus test the validity of the transistor model. The characteristic tables generated are shown in Table 1; the computed points are enclosed by rectangles. A 'frame' of points is added in order to extend the characteristics into the end regions. The frame is automatically added by an interfacing subroutine which links the transistor modelling program to the dc analysis program. The simulated results and the experimental measurements are plotted together in Figure 5(b). The results show good agreement between the simulated results and the experimental measurements.

Example 2

A second example which involves two transistors is shown in Figure 6. Here we use two 2N1900 (PRT 3984) transistors (the profile and geometry of which were supplied by TRW). The computer simulation output and measured results are both shown in Table II.

Table II. Computed and measured dc solution of amplifier circuit shown in Figure 6

Node No.	Computed voltage	Measured voltage
1	2·812	2·8
2	11·839	10·8
3	16·271	16·9
4	2·321	2·3
5	11·326	10·0
6	20·0	20·0

CONCLUSIONS

A method has been described which generates two-dimensional piecewise-linear dc characteristics of bipolar transistors from a given set of values of physical device parameters. Similar methods may be developed to generate the characteristics of other three-terminal electronic devices. The piecewise-linear model is used in a dc analysis simulation program, thus enabling a circuit designer to simulate the dc performance of circuits directly from the values of the physical parameters of the devices. It should be pointed out however, that the characteristics of the devices may also be generated experimentally and used in the analysis program. One drawback of the method is the large amount of storage allocated to the transistor models, especially when many transistors of different characteristics are included in a given circuit. However, one may save on the storage requirements by reducing the number of sampling points per table. This reduction speeds up the computation, but gives less accurate results. It should also be pointed out that the tables do not

include very low and high values of V_{EB}. This is done in order to reduce the time requirements when generating the table, while maintaining realistic boundaries.

More recent work has extended the above modelling approach to include the dynamic behaviour of transistors and obtains transient solutions of transistor circuits.[18]

ACKNOWLEDGMENTS

The authors wish to thank K. K. Hau for his work on linking the transistor modelling program to the dc analysis program and for analyzing and testing the examples given above; they also wish to thank Dr. S. G. Chamberlain for valuable discussions. The authors also gratefully acknowledge impurity profile and geometry data supplied by T.R.W., California and Thomson, C.S.F., Orsay, France.

APPENDIX

Notation

a_{ij}	matrix elements of piecewise-linear terminal characteristics
b	length of the emitter-stripe
$\mathbf{c}_m$	constant vector representing equivalent sources in region m
$\mathbf{E}(x, y)$	electric field
$f_1(.,.), f_2(.,.)$	terminal characteristic functions of a three-terminal element
$\hat{f}_1(.,.), \hat{f}_2(.,.)$	piecewise linear approximations to f_1 and f_2
$\mathbf{g}(.)$	piecewise-linear function occurring in the network equations
h_{FES}	d.c. current gain $= \lvert I_C/(I_E - I_C)\rvert$
I_E	emitter current
I_C	collector current
$\mathbf{J}_p(x, y)$	hole current-density
$\mathbf{J}_n(x, y)$	electron current-density
$J1$	initial value of current density of $[\mathbf{J}_n(x, y)]_x$
L	emitter stripe width
$[M]_m$	network matrix in region m
m	enumeration variable of the piecewise-linear regions of $\mathbf{g}(.)$ in R^n
$n(x, y)$	free electron concentration
n_r	$\triangleq 2[\mathbf{J}_n(x, y)]_x/qv_{th}$
n_i	intrinsic carrier concentration
$N(x)$	net impurity profile
$p(x, y)$	free hole concentration
q	electron charge
Q_{eb}	excess hole charge in region II
r	total number of piecewise linear regions of $\mathbf{g}(.)$ in R^n
R_1	$\triangleq \dfrac{\varepsilon}{q[n(x, y)+N(x)]} \dfrac{\partial[E(x, y)]_x}{\partial x}$
R^n	real n-dimensional Euclidean space
$\mathbf{u}$	vector of input quantities in the network equations
$\mathbf{u}_i$	successive values of $\mathbf{u}$
V_t	volt equivalent of temperature $= kT/q$
v_{th}	thermal drift velocity of electrons $\approx 10^7$ cm/sec.
V_{CB}	collector-base voltage
V_{EB}	emitter-base voltage
V_{CBn}	nominal value of V_{CB}
V_{jeb}	total voltage across region II
W_{epi}	thickness of collector epitaxial layer

$\mathbf{w}$	vector of network variables in the network equations
$\mathbf{w}_i$	successive values of $\mathbf{w}$
x_1, x_2	port-variables at ports 1 and 2 complementary to y_1 and y_2 respectively
x	space variable in the longitudinal direction of a transistor cross-section
$x_{en}(y)$	boundary between emitter-neutral and emitter-base space-charge regions
$x_{eb}(y)$	boundary between emitter-base space-charge and neutral-base regions
$x_{bc}(y)$	boundary between neutral-base and base-collector space-charge regions
$x_{cn}(y)$	boundary between base-collector space-charge and collector-neutral regions
X_{j1}	emitter junction depth
X_{j2}	collector junction depth
$X_i(y)$	effective base width
y_1, y_2	port variables at ports 1 and 2 complementary to x_1 and x_2 respectively
y	space variable in the lateral direction of a transistor base region
ε	permittivity
λ	scalar parameter determining region boundary crossings
$\mu_n(x, y), \mu_p(x, y)$	electron and hole mobilities
τ_{p0}	hole lifetime in the emitter region
$\|\cdot\|$	Euclidean norm
$[\cdot]_x, [\cdot]_y$	components of $[\cdot]$ in the x and y directions, respectively

REFERENCES

1. L. Nagel and R. Rohrer, 'Computer analysis of nonlinear circuits, excluding radiation (CANCER)', *IEEE J. Solid-State Circuits*, SC-**6**, 166–182 (1971).
2. F. S. Jenkins and S.-P. Fan, 'TIME—a nonlinear dc and time-domain circuit-simulation program', *IEEE J. Solid-State Circuits*, SC-**6**, 182–188 (1971).
3. F. H. Branin, G. R. Hogsett, R. L. Lunde, and L. E. Kugel, 'ECAP-II, a new electronic circuit analysis program', *IEEE J. Solid-State Circuits*, SC-**6**, 146–166 (1971).
4. L. O. Chua, 'Modelling of three terminal devices: a black box approach', *IEEE Trans. Circuit Theory*, CT-**19**, 555–562 (1972).
5. L. O. Chua, *Introduction to nonlinear network theory*, McGraw-Hill, New York, 1969.
6. J. J. Ebers and J. L. Moll, 'Large-signal behaviour of junction transistors', *Proc. IRE*, **42**, 1761–1772 (1954).
7. T. E. Stern, 'Piecewise-linear network theory', M.I.T. Tech. Rep. 315 (1956).
8. T. Ohtsuki and N. Yoshida, 'DC analysis of nonlinear networks based on generalized piecewise-linear characterization', *IEEE Trans. Circuit Theory*, CT-**18**, 146–152 (1971).
9. E. S. Kuh and I. N. Hajj, 'Nonlinear circuit theory: resistive networks', *Proc. IEEE*, **59**, 340–355 (1971).
10. D. J. Roulston, S. G. Chamberlain, and J. Sehgal, 'Simplified computer-aided analysis of double-diffused transistors including two-dimensional high-level effects', *IEEE Trans. Electron Devices*, ED-**19**, 809–820 (1972).
11. D. M. Caughey and R. F. Thomas, 'Carrier mobilities in silicon empirically related to doping and field', *Proc. IEEE*, **55**, 2192–2193 (1967).
12. J. Katzenelson, 'An algorithm for solving nonlinear resistive networks', *Bell. Syst. Tech. J.*, **44**, 1605–1620 (1965).
13. L. O. Chua, 'Efficient computer algorithms for piecewise-linear analysis of resistive nonlinear networks', *IEEE Trans. Circuit Theory*, CT-**18**, 73–85 (1971).
14. I. N. Hajj, 'Updating method for LU factorisation', *Electronics Letters*, **8**, 186–188 (1972).
15. T. Fujisawa, E. S. Kuh and T. Ohtsuki, 'A sparse matrix method for analysis of piecewise-linear resistive networks', *IEEE Trans. Circuit Theory*, CT-**19**, 571–584 (1972).
16. I. Hajj, K. Singhal, J. Vlach and P. Bryant, 'WATAND—a program for the analysis and design of linear and piecewise-linear networks', *16th Midwest Symposium on Circuit Theory*, Waterloo, Ontario, Canada (1973).
17. R. J. Van Overstraeten, H. J. DeMan, and R. P. Mertens, 'Transport equations in heavy doped silicon', *IEEE Trans. Electron Devices*, ED-**20**, 290–298 (1973).
18. I. Hajj, D. J. Roulston and P. R. Bryant, 'Efficient transient analysis of transistor circuits from device fabrication data', *Proc. IEEE* (correspondence): to be published.