An Integral Charge Control Model of Bipolar Transistors

By H. K. GUMMEL and H. C. POON

(Manuscript received December 17, 1969)

We present in this paper a compact model of bipolar transistors, suitable for network analysis computer programs. Through the use of a new charge control relation linking junction voltages, collector current, and base charge, the model includes high injection effects. The performance substantially exceeds that of existing models of comparable complexity. For low bias and with some additional idealization, the model reduces to the conventional Ebers-Moll model.

I. INTRODUCTION

Since its formulation in 1954, the Ebers-Moll model[1] has been the major large-signal model for bipolar transistors. It is based directly on device physics and covers all operating regimes, that is, active, saturated, and cut-off operation. But various approximations limit the accuracy of the model. In the original paper the frequency dependence was described in terms of frequency-dependent current generators. For time-domain analysis, this required use of Laplace transforms.

In 1957 Beaufoy and Sparkes[2] analyzed the bipolar transistor from a charge control point of view. The charge control model[2] or the equivalent[3,4] charge control form of the Ebers-Moll model, is directly useful for transient analysis. It is this model, in various forms, that is presently used in all the major general purpose network analysis programs. More elaborate models[5-7] are useful for detailed device studies, but they have not found widespread use in analysis programs due to their complexity.

As device technology evolved over the years making possible devices of reproducible characteristics, and as deeper understanding was gained in device-theoretical studies, many new effects were identified that are not represented by the basic Ebers-Moll model. Among these are a finite, collector-current-dependent output conductance due to basewidth modulation[8] (Early effect), space-charge-layer generation and recombination[9] (Sah-Noyce-Shockley effect), conductivity modulation in the base[10] (Webster effect) and in the collector[11] (Kirk effect), and emitter crowding.[12,13]

Some of these effects have been included into the Ebers-Moll model in various ways. The common approach has been to specify some parameters of the model as functions of bias and to describe this bias dependence in tabular form or through parametric expressions. Thus, in the NET-1 program, for example,[14] the common-emitter current gain is given by a series expansion in the emitter-base voltage. In the CIRCUS program,[15] forward and reverse current gains and forward and reverse transit times are specified as functions of collector current in tabular form. Such "curve-fitting" modeling tends to require large numbers of parameters or table entries for an accurate description. Also, frequently the parameters are not easily interpretable in terms of the device structure and thus can be obtained only a *posteriori*, from detailed measurements, but cannot be predicted.

The present paper makes use of a new, general charge-control relation[16] which links junction voltages, collector-current, and base charge. This new charge-control relation, used in conjunction with conventional charge-control theory, allows many of the effects not contained in the basic Ebers-Moll model to be incorporated in an integral, physical way, in compact form and with good parameter economy.

In the conventional treatment of the transport equation, as used in the derivation of the Ebers-Moll model, attention is focused on recombination in the base (which in modern transistors is generally *not* the dominant source of base current). Through the use of simplified "boundary conditions" relating carrier concentrations at the edges of the base to the junction voltages, and through neglect of the variation of the electric field in the base with bias, the voltage dependence of emitter and collector currents contains errors, except near zero bias.

The new charge control relation arises from the treatment of the transport equation for the carriers that pass between emitter and collector. Use is made of the fact that recombination has only a very small effect on the junction-voltage dependence of the current passing from emitter to collector (later called the dominant current component). Hence for this dependence, but of course not for the base current, recombination is neglected. A direct, closed-form solution of the transport equation from inside the emitter to inside the collector is then possible. An important feature of the result, quoted in equation (10), is a denominator containing the base charge, that is the charge of all those carriers that communicate with the base terminal.

With the concentration in the base of the carriers carrying the dominant current component known, recombination in the base could be computed, and the small effect that this recombination has on the dominant current component could be applied as a correction.[17] Actually, however, for typical transistors this correction is negligible. The base current is modeled by empirical relations since detailed recombination parameters are not available.

The general features of the new model are presented in Section II. Details of a particular implementation using 21 model parameters are given in Section III.* Table I contains a list of the model parameters. Appendix A describes the treatment of base push-out. Discussion of some aspects of the model, and a set of computed characteristics of a sample transistor are given in Section IV. A compact summary of the model is given in Appendix B. A discussion of the significance of the model parameters and of default values is contained in Appendix C. This paper considers mainly the dc and low-frequency aspects of the model. Use of the model for ac or transient analysis will be presented elsewhere.

II. GENERAL CONSIDERATIONS

As a point of departure we will consider the Ebers-Moll equations.[1] They may be written in the form

$$\begin{pmatrix} I_e \\ I_c \end{pmatrix} = [T] \begin{pmatrix} \exp\,(qV_{eb}/kT) - 1 \\ \exp\,(qV_{cb}/kT) - 1 \end{pmatrix} \tag{1}$$

where T is a symmetric matrix of coefficients that are constant, that is, bias independent. At the time when the Ebers-Moll model was developed, attainable base widths were large by today's standards, and in order that useful current gains could be obtained lifetime in the device had to be long. Reverse saturation currents were used as indicators of lifetime. These circumstances are reflected in the notation used in Ref. 1 for the elements of T:

$$T = \begin{bmatrix} \dfrac{I_{eo}}{1 - \alpha_n \alpha_i} & -\dfrac{\alpha_i I_{co}}{1 - \alpha_n \alpha_i} \\[2ex] -\dfrac{\alpha_n I_{eo}}{1 - \alpha_n \alpha_i} & \dfrac{I_{co}}{1 - \alpha_n \alpha_i} \end{bmatrix} \tag{2}$$

where I_{eo} and I_{co} are emitter and collector reverse saturation currents

---

* A simplified implementation of model, and a comparison with measurements, have been presented in the Digest of Technical Papers, 1970 IEEE International Solid State Circuits Conference, Paper 7.1.

TABLE I—MODEL PARAMETERS

Group 1: Knee parameters and transit times		
	I_k	Knee current*
	V_k	Abs. value of knee voltage, in unit of kT/q
	τ_f	Forward tau (forward transit time)
	r_t	Tau ratio (ratio of reverse to forward transit time)
Group 2: Base Current		
	i_1	Ideal base current coefficient
	i_2	Nonideal base current coefficient
	n_e	Forward base current emission coefficient
	i_3	Reverse base current coefficient
	n_c	Reverse base current emission coefficient
Group 3: Emitter Capacitance		
P_e	v_{oe}	Abs. value of emitter offset voltage, in units of kT/q
	m_e	Emitter grading coefficient
	a_{e1}	Emitter zero bias capacitance coefficient
	a_{e2}	Emitter peak capacitance coefficient
Group 4: Collector Capacitance		
P_c	v_{oc}	Abs. value of collector offset voltage, in units of kT/q
	m_c	Collector grading coefficient
	a_{c1}	Collector zero bias capacitance coefficient
	a_{c2}	Collector peak capacitance coefficient
Group 5: Base Push-out		
	v_{rp}	Abs. value of base push-out reference voltage, in units of kT/q
	r_w	Effective base width ratio
	r_p	Base push-out transition coefficient
	n_p	Base push-out exponent

Auxiliary Quantities

$e_k = \exp(-v_k)$

$e_{ke} = \exp(-v_k/n_e)$

$e_{kc} = \exp(-v_k/n_c)$

---

* Quantity is negative for npn transistor.

and α_n and α_i are forward and reverse common base current gains. A simpler and more appropriate but fully equivalent notation is

$$T = \begin{pmatrix} (1 + 1/\beta_f)I_s & -I_s \\ -\,I_s & (1 + 1/\beta_r)I_s \end{pmatrix}. \tag{3}$$

Here β_f and β_r are forward and reverse common emitter current gain and I_s is the "intercept" current, that is, the current obtained when on a semilog plot of I_c vs. V_{eb} the collector current is extrapolated to $V_{eb} = 0$. The notation of equation (3) is considered more appropriate since the intercept current I_s is nearly independent of current gains. The matrix elements T_{12} and T_{21} in equation (2) show an apparent dependence on lifetime through the forward and reverse current gains.

Actually, this dependence is nearly cancelled by that of I_{eo} and I_{co}. To a very good approximation the intercept current I_s depends only on the total number of impurities in the base.[18]

Equation (1) with matrix T given by equation (3) suggests the following interpretation: The emitter and collector current have a common, dominant component

$$I_{cc} = -I_s[\exp(qV_{eb}/kT) - \exp(qV_{cb}/kT)]. \tag{4}$$

In addition the emitter and collector currents have each a separate component I_{be} and I_{bc}, which is proportional to the reciprocal forward and reverse common emitter current gain:

$$I_{be} = (I_s/\beta_f)[\exp(qV_{eb}/kT) - 1], \tag{5}$$

$$I_{bc} = (I_s/\beta_r)[\exp(qV_{cb}/kT) - 1]. \tag{6}$$

The terminal currents are then given by

$$I_e = -I_{cc} + I_{be}, \tag{7}$$

$$I_c = I_{cc} + I_{bc}, \tag{8}$$

$$I_b = -I_{be} - I_{bc}. \tag{9}$$

So far, we have made no changes in the content of the Ebers-Moll equations; we have only brought them into a form that will facilitate the development of the new model. We shall retain equation (7) through (9)* but we will replace equations (4) through (6) by relations giving an improved representation of the physical processes in the transistor.

The separation of emitter and collector currents into the dominant I_{cc} component and the base current components in equations (7) and (8) allows us to give different voltage dependence to the individual components. For example, at low injection levels, collector current and emitter-base voltage are related through the "ideal" diode law; that is, the collector current is proportional to $\exp(qV_{eb}/nkT)$ where the "emission coefficient" n is very close to unity. The base current at low forward bias, on the other hand, is typically a "nonideal" current,[19] that is it has an emission coefficient n with values typically between 1.5 and 2. This nonideal current results from space charge recombination[9] or surface recombination, or the presence of both effects. At higher forward emitter-base voltages the base current is dominated by an ideal component.[19]

In principle, it is possible to compute the base current as a function

---

* These equations are not independent. Equation (7) will be dropped later.

of V_{eb} (for given V_{cb}) for one-dimensional structures,[17,20–22] provided that the doping profile and the recombination parameters, for example the concentration as a function of distance of the important species of recombination centers, as well as their energy levels and capture cross sections, are known. However, in practice, the recombination properties are not known to the detail required for such calculations. Even the assumption of a constant concentration of one species of recombination centers is a gross oversimplification. In real transistors the lack of lattice perfection in heavily doped regions causes enhanced recombination, and interfaces between substrates and epitaxial layers provide local regions of high recombination.

Nevertheless, detailed studies have confirmed that the base current can be described by a sum of terms exponential in voltage with emission coefficients of the magnitude indicated above. Equations (5) and (6) are thus replaced by more general functions of V_{eb} and V_{cb} that are characterized by pre-exponential coefficients and emission coefficients which become model parameters and which represent the overall recombination properties and influence the dependence of forward and reverse current gain on bias. In the next section we will use a specific set of base current components.

The new charge control relation[16], derived from basic physical considerations, permits the dominant current components I_{cc} to be obtained without use of low-injection approximations:

$$I_{cc} = -I_s Q_{bo} \frac{\exp(qV_{eb}/kT) - \exp(qV_{cb}/kT)}{Q_b}. \tag{10}$$

Here Q_b is the "base charge", that is, the charge of all carriers of the type that communicate with the base terminal: electrons in a pnp transistor and holes in an npn transistor. The base charge is a function of bias; Q_{bo} is the zero-bias value. The bias dependence of Q_b is the subject of the conventional charge control theory[12] which approximates the excess (or stored) base charge $Q_b - Q_{bo}$ as the sum of capacitive contributions and products of collector current components and transit times. The essential feature of the model presented here is the simultaneous use of conventional charge-control theory and relation (10). As shown before,[16] relation (10) is a good approximation for one-dimensional transistors for a wide range of bias conditions, including those leading to high-injection and base push-out effects.

It is of interest to note that the Ebers–Moll equations embody *superposition*, that is that the collector current can be expressed as the sum of a function of the emitter voltage and a function of the collector voltage. For real transistors, violations of the superposition principle are easily observed. Consider, for example, the Early effect[8] that is, the dependence of the low-frequency output conductance on bias. As shown schematically in Fig. 1, a region of bias exists in which the collector current varies approximately linearly with collector-emitter voltage for fixed base current, in such a way that the straight-line sections, when extrapolated, intersect (approximately) at a negative voltage which we shall call the "Early voltage", V_A . For superposition to be valid, the lines would have to be nearly parallel to each other. The base charge Q_b in the denominator of equation (10) through its dependence on collector voltage via the collector capacitance disables superposition and provides a realistic description of the output conductance.

Another point of interest concerns high-injection effects in the base region. The "ideal" voltage dependence of I_{cc} on V_{eb} is caused primarily by the dependence of the minority carrier concentration in the base near the emitter as exp (qV_{eb}/kT). This dependence holds, however, only as long as the minority carrier concentration is small compared to the doping concentration. If the minority carrier (the word minority starts to lose its literal meaning here) concentration is large compared to the doping concentration, then it varies as exp $(qV_{eb}/2kT)$, and so does, approximately, I_{cc} except for additional complications due to

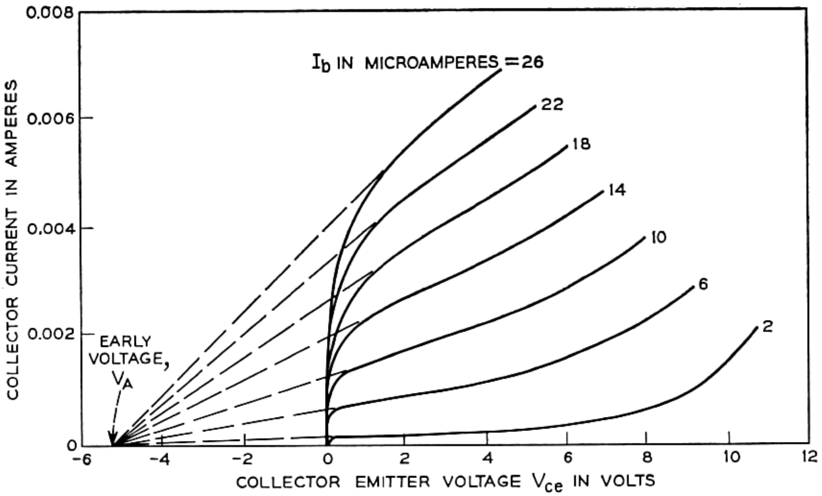

Fig. 1—Definition of Early voltage.

base push-out. The intersection of the $n = 1$ and $n = 2$ asymptotes to I_{cc} represents an important characteristic feature of the transistor. We shall call it the "knee point" and use its coordinates, V_k and I_k as model parameters and as a basis for normalizations.*

Consolidating the development up to this point, the model, exclusive of parasitic effects, is described by

$$I_c = -I_s Q_{bo} \frac{\exp(qV_{eb}/kT) - \exp(qV_{cb}/kT)}{Q_b} + I_{bc} , \tag{11}$$

$$I_b = -I_{be} - I_{bc} . \tag{12}$$

As discussed above, the base current components I_{be} and I_{bc} depend strongly on the recombination properties of the structure and are in practice not readily calculated from first principles. By contrast, the base charge as a function of bias depends primarily on the doping profile and is nearly independent of recombination properties. Hence, given the doping profile, Q_b as a function of V_{eb} and V_{cb} can be computed by existing techniques. However, for such a calculation considerable computer resources (memory and time) are required. For network analysis programs it is preferred to approximate Q_b by simple algebraic or algorithmic (implicit functions) representations which, depending on complexity, can give reasonable accuracy. One such representation is given in the next section. Special features are use of a modified representation of junction capacitance[23] which avoids the problem of an infinite capacitance when the junction voltage equals the built-in voltage, and use of a 4-parameter representation of base push-out.

III. SPECIFIC MODEL IMPLEMENTATION

In this section we present a specific implementation of the general model, equations (11) and (12). The bias dependence of base charge and base current will be modeled. The polarity assumed is that of a pnp transistor.

The dominant current component I_{cc} may be separated into an emitter and a collector component, or a forward and reverse component:

$$I_{cc} = -I_s Q_{bo} \frac{\exp(qV_{eb}/kT) - 1}{Q_b} + I_s Q_{bo} \frac{\exp(qV_{cb}/kT) - 1}{Q_b} \equiv -I_f + I_r . \tag{13}$$

---

* The exact definition of these parameters is given in the following Section.

We then model the excess base charge as consisting of emitter and collector capacitive contributions Q_e and Q_c, and of forward and reverse current-controlled contributions*

$$Q_b = Q_{bo} + Q_e + Q_c - \tau_f B I_f - \tau_r I_r \,. \tag{14}$$

Here τ_f and τ_r are forward and reverse transit times. The coefficient B has been included to model the increase of the transit time when base push-out occurs; it has a value of unity in the absence of base push-out.

For modeling these parameters it is convenient to normalize all charges in equation (14) with respect to the zero bias charge Q_{bo} and to denote the normalized charges by lowercase symbols. Also, we replace I_f and I_r according to equation (13). Then

$$\underbrace{1 + q_e + q_c}_{q_1} \underset{+\,\frac{1}{q_b}}{+\,\frac{1}{q_b}} \underbrace{\frac{I_s}{-Q_{bo}} \{\tau_f B[\exp(qV_{eb}/kT) - 1] + \tau_r[\exp(qV_{cb}/kT) - 1]\}}_{q_2\,.} \tag{15}$$

Multiplication of equation (15) by q_b removes q_b from the denominator of the last term on the right side of equation (15) and gives rise to a quadratic equation in q_b. Its solution gives q_b explicitly in terms of the junction voltages, except for a possible q_b-dependence of B:

$$q_b = \frac{q_1}{2} + \left[\left(\frac{q_1}{2}\right)^2 + q_2\right]^{\frac{1}{2}}. \tag{16}$$

The term q_1 represents the sum of the zero-bias charge and the charge associated with the junction capacitances; q_2 represents the excess base charge, or the current-dependent charge associated with diffusion capacitances. The latter charge contains a dependence on the base push-out effect through the parameter B which is treated in Appendix A.

For high forward bias the charge q_2 is the dominant component of the base charge q_b. Except for the base push-out term B, it is characterized by four parameters: I_s, Q_{bo}, τ_f, and τ_r. It will be convenient to normalize these parameters. For this we define the knee voltage V_k as the emitter voltage for which q_2 equals unity (for zero collector voltage and with neglect of terms small compared to the exponential of the emitter voltage):

---

* The minus sign arises because the base charge contains electrons neutralizing the positive charges $\tau_f B I_f$ and $\tau_r I_r$. Q_b, Q_{bo}, Q_e, and Q_c are all negative quantities for positive V_{eb} and V_{cb}.

$$V_k = \frac{kT}{q} \ln \frac{-Q_{bo}}{I_s \tau_f}. \tag{17}$$

The low-injection-extrapolated collector current for $V_{eb} = V_k$ is

$$I_k = \frac{-Q_{bo}}{\tau_f}. \tag{18}$$

It will be convenient to normalize all quantities having dimensions of current with respect to I_k, and to express voltages by their difference from V_k, in units of kT/q. We shall use lower-case symbols for normalized quantities. Thus, we define

$$v_k = \frac{qV_k}{kT}, \tag{19}$$

$$v_e = \frac{qV_{eb}}{kT} - v_k, \tag{20}$$

$$v_c = \frac{qV_{cb}}{kT} - v_k, \tag{21}$$

$$e_k = \exp(-v_k). \tag{22}$$

With these normalizations, equations (11) and (12) become

$$I_c = I_k i_c = I_k\left(-\frac{e^{v_e} - e^{v_c}}{q_b} + i_{bc}\right), \tag{23}$$

$$I_b = I_k i_b = I_k(-i_{be} - i_{bc}). \tag{24}$$

For the base charge we obtain

$$q_2 = B(e^{v_e} - e_k) + (\tau_r/\tau_f)(e^{v_c} - e_k), \tag{25}$$

$$q_b = q_1/2 + [(q_1/2)^2 + q_2]^{1/2}, \tag{26}$$

$$Q_b = -I_k \tau_f q_b . \tag{27}$$

With these normalizations, we can use the set I_k, v_k, τ_f, $r_t \equiv \tau_f/\tau_r$ as the four model parameters describing q_2 for the case $B = 1$. These four parameters constitute Group 1 of the model parameters listed in Table 1.

The charge contribution from the emitter and collector junction capacitances will be considered next. The conventional representation of junction capacitance is through an expression containing three parameters:

$$C = \frac{\text{const}}{(V - V_{\text{built-in}})^m}. \tag{28}$$

The parameters are $V_{\text{built-in}}$ (which for silicon is typically $\approx .7V$), the grading coefficient m, and the constant in the numerator which can be related to the zero-bias capacitance. This expression causes difficulties when the junction voltage V approaches the built-in voltage and C goes to infinity. In a real transistor, of course, a finite amount of charge is stored for all bias conditions, and the derivatives of charge with respect to junction voltages are finite. As shown in detail in Ref. 23, equation (28) can be modified so as to be free of singularities by the introduction of a fourth parameter which relates to the forward-bias capacitance inferred from measurements of transit time vs. emitter current. For compactness of notation, we express the four parameters as elements p_1, p_2, p_3, and p_4 of a four-dimensional vector $\mathbf{P}$ (see Table I and Appendix C). If we now define a function

$$f(v, \mathbf{P}) = p_3 \left\{ \frac{1}{(1 + p_4)^{p_2}} + \frac{\dfrac{v}{p_1} - 1}{\left[\left(\dfrac{v}{p_1} - 1 \right)^2 + p_4 \right]^{p_2}} \right\}, \tag{29}$$

then the normalized emitter and collector charges are given by:

$$q_e = f\left(\frac{qV_{eb}}{kT}, \mathbf{P}_e \right), \tag{30}$$

$$q_c = f\left(\frac{qV_{cb}}{kT}, \mathbf{P}_c \right), \tag{31}$$

where $\mathbf{P}_e$ and $\mathbf{P}_c$ are the four-parameter vectors describing the emitter and collector junctions, respectively. These two vectors constitute Groups 3 and 4 of the model parameters listed in Table I. As discussed in Appendix C, some of the parameter values can be estimated or approximated in terms of other model parameters. In any case, these parameters are readily amenable to numerical evaluation from the device structure.

As mentioned in the previous section, the recombination in transistors is best handled through a description of the base current as a sum of exponentials in the junction voltages. Pertinent model parameters are pre-exponential factors and emission coefficients. For typical transistors the forward base current is adequately described by two components, one ideal ($n = 1$) and the other nonideal ($n = n_e$). For the reverse base current a single nonideal ($n = n_c$) component is adequate. We define

$$e_k = \exp(-v_k), \tag{32}$$

$$e_{ke} = \exp(-v_k/n_e), \tag{33}$$

$$e_{kc} = \exp(-v_k/n_c). \tag{34}$$

Then we let

$$i_{be} = i_1(e^{v_e} - e_k) + i_2(e^{v_e/n_e} - e_{ke}), \tag{35}$$

$$i_{bc} = i_3(e^{v_c/n_c} - e_{kc}). \tag{36}$$

The quantities i_1, i_2, n_e, i_3, n_c are the five parameters which characterize the recombination behavior of the transistor. They are listed as Group 2 in Table I.

The last set (Group 5) of the model parameters listed in Table I describes the base push-out effect. For its description four parameters are required. The details of the base push-out modeling are given in Appendix A.

We thus have a total of 21 parameters which describe the transistor model. For several of these parameters default values can be used as discussed in Appendix C.

The following features are a consequence of the normalizations used:

(*i*) I_k is proportional to the emitter area. All other parameters are, to first order, independent of area. Area scaling (neglecting complications caused by lateral effects such as emitter crowding) is achieved simply by changing the value of I_k. This feature is particularly convenient for integrated circuit work, where transistors on a given slice differ nominally only in their lateral dimensions.

(*ii*) For pnp transistors all model parameters have positive numerical values. For npn transistors two changes are required: a) I_k must be made a negative quantity; b) the Boltzmann voltage kT/q must be given a negative value (or the Boltzmann voltage is given the sign of I_k for pnp's and npn's). When this is done, the polarity of terminal currents and voltages is in agreement with standard practice (currents positive if flowing *into* the device).

(*iii*) The offset voltages V_{oe} and V_{oc} used in modeling the capacitive charges (see Appendix C) are approximately proportional to the absolute temperature. Hence use of constant, that is, temperature independent, values for the normalized quantities v_{oe} and v_{oc} implements automatically the temperature dependence of the offset voltages. The normalized knee voltage v_k is not, to first order, temperature independent, but varies with temperature approximately as

$$v_k(T) = v_k(T_o) + \frac{qV_g}{kT_o}\left(\frac{T_o}{T} - 1\right) \tag{37}$$

where T_o is a reference temperature (for example, room temperature) and V_g is the band-gap voltage (1.12 V for silicon).

IV. DISCUSSIONS AND EXAMPLES

This Section will illustrate the performance of the new model. The following features are of interest:

(*i*) For low bias so that Q_b is nearly equal to Q_{bo} (or $q_b \approx 1$) and with the choice $n_e = n_c = 1$, the model reduces to the Ebers–Moll model.

(*ii*) Superposition which is operative in the Ebers–Moll equations, is disabled through the base-charge denominator Q_b in equation (11). The dependence of Q_b on collector voltage produces a finite output conductance, that is, the Early effect.

(*iii*) The rapid increase of Q_b when base push-out occurs causes fall-off of current gain and frequency response.

(*iv*) If we define an effective emission coefficient n by

$$\frac{1}{n} \equiv \frac{kT/q}{I_c} \frac{dI_c}{dV_{eb}} \bigg|_{V_{cb}=\text{const}}, \tag{38}$$

then it is seen that for the model, n varies from approximately unity at low currents to approximately two at high currents (and larger values when base push-out occurs). The shift to a value of two represents high injection effects. The $n = 1$ and $n = 2$ asymptotes intersect approximately at $I_c = I_k$ and $V_{eb} = V_k$ (see Fig. 2).

If we define an emitter capacitance C_e by

$$C_e = \frac{dQ_e}{dV_{eb}} \bigg|_{V_{cb}=\text{const}}, \tag{39}$$

then the effective emission coefficient at low current values, where the current contribution to Q_b, equation (14), is negligible, is given by*

$$n = 1 + \frac{(kT/q)C_e}{Q_{bo}}. \tag{40}$$

Thus, small deviations from the ideal exponential law are caused by the emitter capacitance. These deviations are present even at

---

* If the transistor is used in a common emitter configuration it may be useful to define the effective emission coefficient as in equation (39), but with V_{ce} instead of V_{cb} held constant. For this case, shown in Fig. 2, the emitter capacitance C_e in equation (40) is to be replaced by the sum of emitter and collector capacitances.

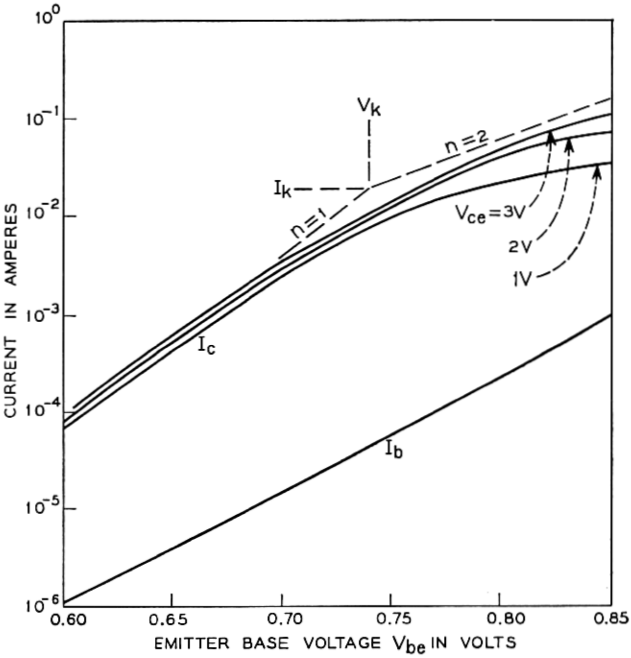

Fig. 2—Collector and base currents for three values of collector-emitter voltage versus base-emitter voltage. Also shown are "knee" point (V_K, I_K), and slopes corresponding to values of 1 and 2 for the emission coefficient n.

low forward currents†. Additional increases of n at intermediate currents are caused by base resistance (not considered in this paper).

(*v*) For currents low enough such that base widening effects are negligible, the emitter-collector delay time τ_d for common emitter operation is given by

$$\tau_d = \left.\frac{dQ_b}{dI_c}\right|_{V_{ce}=\text{const}},$$

$$= \frac{\tau_f + \dfrac{kT}{q}\dfrac{(C_e + C_c)}{I_c}}{1 - \dfrac{kT}{q}\dfrac{(C_e + C_c)}{Q_b}}. \tag{41}$$

---

† In principle, equation (40) could be used to obtain the emitter capacitance from a dc semilog plot of I_c vs. V_{eb}. However, very accurate temperature control would be required.

At low current values, and for $C_c \ll C_e$, the denominator of equation (41) is approximately $1/n$, equation (40). Then the value of the emitter capacitance C_e may be obtained from the slope $n(kT/q)C_e$ of a plot of delay time τ_d versus reciprocal collector current (see Fig. 6). It is this forward-bias emitter capacitance that may be used to set the parameter a_{e2} (see equation 1 of Appendix C).

As a demonstration of the proposed model, various characteristics for a double diffused silicon npn transistor at room temperature ($kT/q = 0.02585\ V$) were computed with the parameter values listed in Table II. The results are shown in Figs. 2 through 7. We treat the intrinsic transistor model, without inclusion of parasitics, such as base resistance and inactive base region. Figure 2 is a semilog plot of collector and base current vs. emitter-base current for $V_{ce} = 1, 2$ and $3\ V$. Note the transition from $n = 1$ to $n = 2$ near $I_k = 19$ mA and $V_k = 0.74$ volts. Figure 3 shows common-emitter low-frequency current gain β vs. collector current for various collector voltages. The frequency dependence is conveniently characterized by the low-frequency approximation to the unity-gain frequency,[24] f_L. (For high-current-gain transistors in the active region, f_L is synonymous with the conventional cut-off frequency f_T.) Figure 4 shows f_L vs. collector current for various collector voltages.

TABLE II—MODEL PARAMETER VALUES

$I_k = -1.875 \times 10^{-2}$ amps.
$v_k = 28.7$
$\tau_f = 3.2 \times 10^{-10}$ secs.
$r_t = 10.0$

$i_1 = 2.35 \times 10^{-4}$
$i_2 = 2.19 \times 10^{-3}$
$n_e = 1.5$
$i_3 = 7.65 \times 10^{-3}$
$n_c = 1.5$

$v_{oe} = 27.1$
$m_e = 0.24$
$a_{e1} = 0.337$
$a_{e2} = 1.03 \times 10^{-2}$

$v_{oc} = 27.1$
$m_c = 0.1265$
$a_{c1} = 0.187$
$a_{c2} = 7.17 \times 10^{-3}$

$v_{rp} = 18.2$
$r_w = 10.0$
$r_p = 4.55$
$n_p = 3.0$

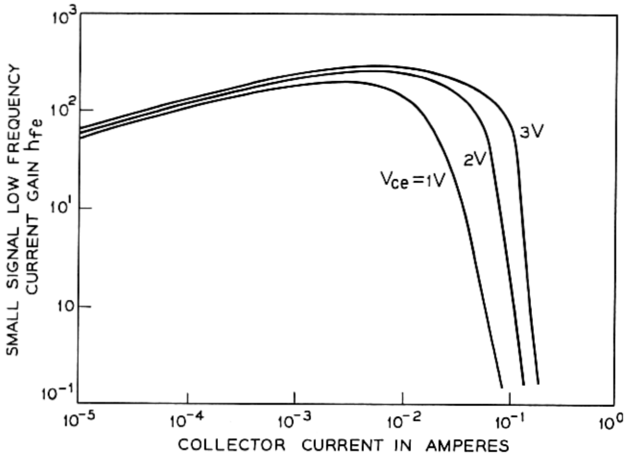

Fig. 3—Small-signal low-frequency current gain h_{fe} versus collector current for three values of collector-emitter voltage.

Figure 5 presents a family of I_c vs. V_{ce} characteristics, with I_b as parameter. Figure 6 gives results of simulated τ_d vs. $1/I_c$ measurements of the model. The same information is displayed as f_L contour plots in Fig. 7.

V. CONCLUSION

With the aid of a new charge control relation, a compact bipolar transistor model, suitable for network analysis computer programs, has been developed. This model is an "integral model" in that parameters important for the ac response also shape the dc characteristics. The model includes high-injection effects; for low bias and idealized emission coefficients it is equivalent to the Ebers–Moll model. Figures 2–7 exhibit many features of real transistors that conventional models do not contain, or can give only by curve-fitting with the use of many parameters. These figures were obtained for an implementation of the model using 21 parameters (Table I, Appendix B). For several of these parameters, values can be estimated *a priori*. Such values may be used as default values. Sometimes users may be willing to make some sacrifice in the accuracy of the model in return for having to specify only few key parameters. It appears feasible to require that only I_k (proportional to emitter area), the Early voltage V_A, maximum β (at some collector voltage, for example, $V_{ce} = 5$ V), maximum f_L (at $V_{ce} = 5$ V) and the collector current at which maximum f_L occurs, need be specified and

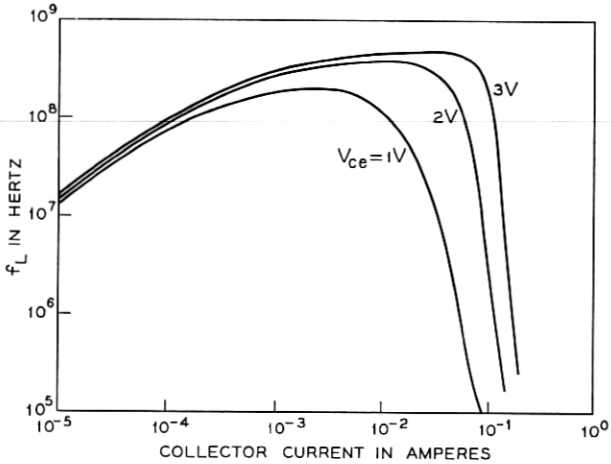

Fig. 4—Low-frequency approximation to unity-gain frequency f_L versus collector current for three values of collector-emitter voltage.

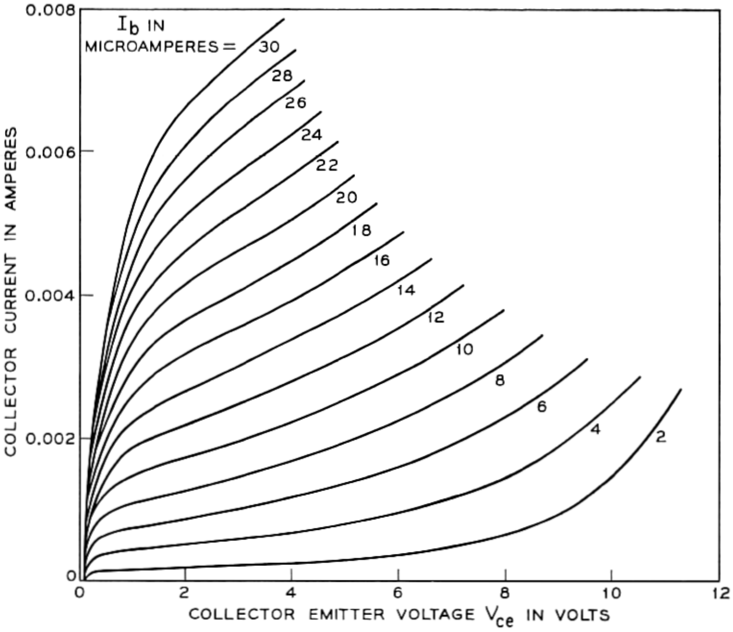

Fig. 5—Collector current versus collector-emitter voltage for various values of base current.

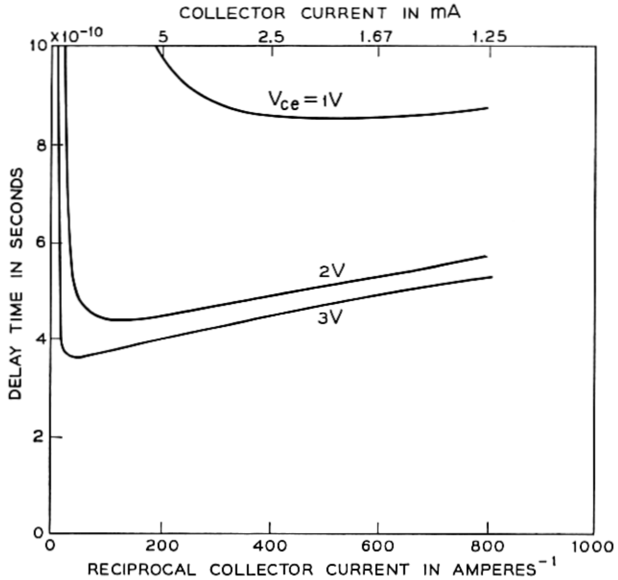

Fig. 6—Emitter-collector delay time versus reciprocal collector current for three values of emitter-collector voltage.

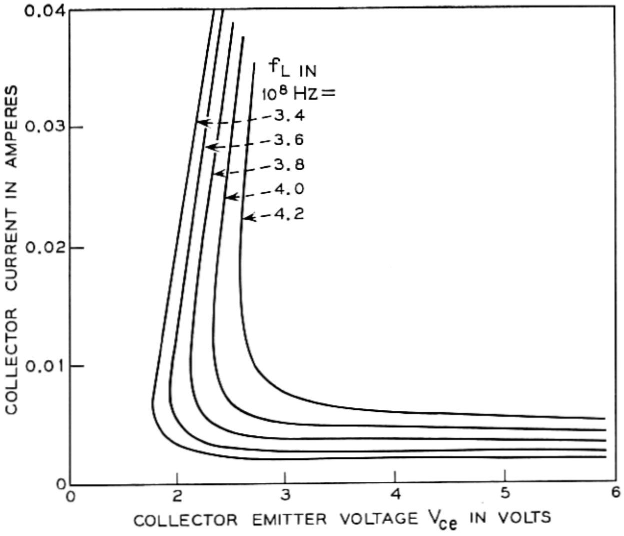

Fig. 7—Contours of constant f_L in collector-emitter voltage, collector-current plane.

that the full set of model parameters is automatically computed with the use of default values.

Not included in this paper, but to be included in careful modeling, are parasitics such as base resistance, inactive base region, collector resistance, and so on. The base resistance may be modeled as consisting of a constant, exterior resistance incurred in the inactive base and an interior resistance. To reflect conductivity modulation of the base, the value of the interior resistance may be made inversely proportional to the base charge Q_b. It appears feasible to extend the intrinsic model, for example to include carrier generation by impact ionization.

VI. ACKNOWLEDGMENT

The authors are indebted to F. M. Smits for encouragement and for substantial improvements in the presentation of the material.

APPENDIX A

Base Push-Out Effect

The approach towards modeling the base push-out effect was guided by results obtained in a detailed analysis of this effect.[25] The discussion is based on pnp polarity (I_c is a negative quantity). Assuming constant resistivity ρ in the collector region adjacent to the base (epitaxial region), the base push-out effect starts approximately at a collector current value

$$I_1 = \frac{A_e(V_{oc} - V_{cb})}{\rho W_c} \tag{42}$$

where A_e is the emitter area and W_c is the width of the lightly doped collector region. Let W_{eff} be the effective width of the base. For $-I_c < I_1$, the effective base width is equal to the metallurgical base width, W_b,

$$W_{\text{eff}} = W_b . \tag{43}$$

For $-I_c > I_1$, the effective base width is approximately given by

$$W_{\text{eff}} = W_b + W_c\left(1 - \frac{I_1}{I_c}\right). \tag{44}$$

Equations (43) and (44) can be written as

$$W_{\text{eff}} = W_b + \frac{W_c}{2} \frac{[(I_c + I_1)^2]^{\frac{1}{2}} - (I_c + I_1)}{I_c}. \tag{45}$$

Equation (45) gives much too abrupt a transition for W_{eff} from W_b to

$(W_b + W_c)$, and becomes numerically unstable when I_c approaches zero. Hence, we modify equation (45) with an additional parameter I_2, such that

$$W_{\mathrm{eff}} = W_b + \frac{W_c}{4} \frac{\{[(I_c + I_1)^2 + I_2^2]^{\frac{1}{2}} - (I_c + I_i)\}^2}{I_c^2 + I_2^2}. \tag{46}$$

The low-current forward transit time τ_f is to be modified by the square of the ratio of the effective base width to metallurgical base width, to give the total base transit time τ_{fb}:

$$\tau_{fb} = \tau_f \left(\frac{W_{\mathrm{eff}}}{W_b}\right)^2 \equiv \tau_f B \tag{47}$$

with

$$B = \left\{1 + \frac{W_c}{4W_b} \frac{\{[(I_c + I_1)^2 + I_2^2]^{\frac{1}{2}} - (I_c + I_1)\}^2}{I_c^2 + I_2^2}\right\}^2. \tag{48}$$

The quantity B may be expressed in terms of model parameters and the normalized collector current $i_c = I_c/I_K$

$$B = \left\{1 + \frac{r_w}{4} \frac{\{[(i_c + i_1)^2 + r_p]^{\frac{1}{2}} - (i_c + i_1)\}^2}{i_c^2 + r_p}\right\}^2 \tag{49}$$

with

$$r_w = \frac{W_c}{W_b}, \qquad r_p = \left(\frac{I_2}{I_K}\right)^2,$$

and

$$i_1 = \frac{(V_{oc} - V_{cB})A_e}{\rho W_c I_k} = \frac{v_{oc} - v_k - v_c}{v_{rp}}. \tag{50}$$

So far, equation (49) models effects in a one-dimensional transistor. Emitter crowding and carrier storage in the inactive base cause the transit time at high currents to increase more strongly than given by equation (49). For a first-order modeling of emitter crowding we replace the exponent 2 outside the square brackets by an adjustable model parameter n_p ("push-out exponent"):

$$B = \left\{1 + r_w \frac{\{[(i_c + i_1)^2 + r_p]^{\frac{1}{2}} - (i_c + i_1)\}^{n_p}}{4(i_c^2 + r_p)}\right\}^2. \tag{51}$$

The quantities r_w, r_p, v_{rp}, n_p and v_{oc} are model parameters (groups 5 and 4).

APPENDIX B

A Summary of the Model

Terminal quantities are: collector current I_c, base current I_b, emitter-base voltage V_{eb}, collector base voltage V_{cb}, and base charge Q_b. Define the following normalized quantities: $i_c = I_c/I_k$, $i_b = I_b/I_k$, $v_e = (qV_{eb}/kT) - v_k$, $v_c = (qV_{cb}/kT) - v_k$, $q_b = Q_b/(I_k\tau_f)$, and define the function

$$f(v, P) = p_3\left\{\frac{1}{(1 + p_4)^{p_2}} + \frac{(v/p_1 - 1)}{[(v/p_1 - 1)^2 + p_4]^{p_2}}\right\}.$$

Then the following relations hold between the normalized quantities

$$i_{be} = i_1(e^{v_e} - e_k) + i_2(e^{v_e/n_e} - e_{ke}), \tag{52}$$

$$i_{bc} = i_3(e^{v_c/n_c} - e_{kc}), \tag{53}$$

$$i_b = -i_{be} - i_{bc}, \tag{54}$$

$$i_4 = i_c + (v_{oc} - v_c - v_k)/v_{rp}, \tag{55}$$

$$B = \left\{1 + r_w \frac{\{[i_4^2 + r_p]^{\frac{1}{2}} - i_4\}^{n_p}}{4(i_c^2 + r_p)}\right\}, \tag{56}$$

$$q_1 = 1 + f(v_e + v_k, P_e) + f(v_c + v_k, P_c), \tag{57}$$

$$q_2 = B(e^{v_e} - e_k) + r_t(e^{v_c} - e_k), \tag{58}$$

$$q_b = q_1/2 + [(q_1/2)^2 + q_2]^{\frac{1}{2}}, \tag{59}$$

$$i_c = -\frac{(e^{v_e} - e^{v_c})}{q_b} + i_{bc}. \tag{60}$$

APPENDIX C

Discussion of Model Parameters

The knee parameters I_k and v_k, and τ_f are related to the conventional intercept current I_s (saturation current), the base charge at zero bias Q_{bo}, and the nominal value of the low frequency approximation to the unity-current-gain frequency f_L, by

$$I_k = \frac{-Q_{bo}}{\tau_f}, \tag{61}$$

$$v_k = \ln\left(\frac{-Q_{bo}}{\tau_f I_s}\right), \tag{62}$$

and

$$\tau_f = \frac{1}{2\pi f_L}. \tag{63}$$

In terms of structural parameters τ_f is given approximately by

$$\tau_f = \frac{w_b^2}{2\eta D} \tag{64}$$

where w_b is the base width and where η represents the drift effect in the base.[26] η is unity for uniform base doping and has typical values between two and ten for diffused-base transistors.

The zero-bias base charge is approximately given by

$$Q_{bo} = -A_e q N_b \tag{65}$$

where A_e is the emitter area, q the electronic charge and N_b the number of impurities per unit area in the base. Typical values for N_b are a few times 10^{12} per cm^2 (lower values would cause premature punch through and high values cause low injection efficiency and/or low f_L). The intercept current is given by

$$I_s = A_e \frac{qn_i^2 D}{N_b} = -\frac{A_e^2 q^2 n_i^2 D}{Q_{bo}} \tag{66}$$

where n_i is the intrinsic carrier concentration and D is the effective diffusivity of carriers in the base. From equations (62) and (64) through (66), one obtains

$$v_k = 2 \ln \frac{N_b/w_b}{n_i} + \ln 2\eta. \tag{67}$$

The parameters i_1, i_2, i_3 determine the current gain of the transistor. The variation of forward gain with current at low currents is governed by n_e and the inverse gain variation by n_c.

The group 3 parameters model the emitter junction capacitance. The offset voltage V_{oe} is approximately the conventional "built-in" voltage, which has a typical value near 0.7 volts for silicon at room temperature, or $v_{oe} = V_{oe}/(kT/q) = 27$. The grading coefficient m_e depends on the type of doping transition: it is one-fourth and one-sixth for ideal step and linearly graded junctions respectively. Typical values for emitter junctions are in the neighborhood of 0.2. Parameter a_{e1} is related to the zero-bias capacitance C_{oe} by[23]

$$a_{e1} = \frac{C_{oe} V_{oe}}{Q_{bo}(1 - 2m_e)}. \tag{68}$$

If one assumes the emitter junction to be a step junction and the base doping to be uniform, and if one expresses base doping and base width in terms of the number of impurities per cm^2 in the base, N_b, and in terms of f_L, then an approximate formula for a_{e1} may be derived to be

$$a_{e1} = \frac{1}{(1 - 2m_e)} \left(\frac{\pi \epsilon^2 V_{eo}^2}{4q^2 D} \right)^{\frac{1}{4}} \frac{(f_L)^{\frac{1}{4}}}{(N_b)^{\frac{1}{2}}}, \tag{69}$$

$$= A \frac{(f_L/10^9 \text{ Hz})^{\frac{1}{4}}}{(N_b/5 \times 10^{12} \text{ cm}^{-2})^{\frac{1}{2}}}, \tag{70}$$

where ϵ is the permitivity and D the diffusivity of electrons (holes) in an npn (pnp) transistor. For a silicon npn transistor, the numerical value for A based on equation (69) is 0.147. For an actual double diffused transistor of $f_L = 400$ MHz, the value of A obtained from parameter fitting was found to be 0.202.

The last parameter a_{e2} in this group is related to the forward bias capacitance, C_{ef}, deduced from the slope of delay time versus reciprocal emitter current by[23]

$$a_{e2} = \left[\frac{(1 - 2m_e) r C_{ef}}{C_{oe}} \right]^{-m_e} \tag{71}$$

where r is a numerical coefficient approximately equal to unity, the exact value depending on the doping profile. Typical values for a_{e2} range from 10^{-2} to 10^{-3}. If emitter capacitance effects are not of importance to the users, and if the user does not want to specify group 3 parameter values, the following default values are suggested:

$$v_{oe} = 27, \quad \text{silicon};$$

$$m_e = 0.2;$$

$$a_{e1} = 0.2 \frac{(f_L/10^9 \text{ Hz})^{\frac{1}{4}}}{(N_b/5 \times 10^{12} \text{ cm}^{-2})^{\frac{1}{2}}};$$

$$a_{e2} = 3 \times 10^{-3}.$$

The group 4 parameters model the collector junction capacitance. The parameters v_{oc}, m_c, a_{c1} and a_{c2} have similar meanings as their counterparts in the emitter junction capacitance. Typical default values (for silicon transistors) for v_{oc}, m_c, and a_{c2} are

$$v_{oc} = 27, \quad \text{silicon};$$

$$m_c = 0.15;$$

$$a_{c2} = 10^{-3}.$$

We shall relate a_{c1} to the output characteristics of the transistor. The Early voltage as defined in Section II is of magnitude comparable to that of the punch-through voltage V_T, defined as that voltage for which the charge associated with collector capacitance, Q_{cc}, equals minus Q_{bo}. We denote the coefficient relating the Early voltage and the punch-through voltage by r_A

$$V_T = r_A V_A . \tag{72}$$

Then a_{c1} is given in terms of the Early voltage by

$$a_{c1} = \frac{1}{\left(\dfrac{r_A V_A - V_{oc}}{V_{oc}}\right)^{1-2m_c} - 1} . \tag{73}$$

The exact value of r_A depends on details of the doping profile and on the region in the I_c vs V_{ce} domain from which the Early voltage is extrapolated; a typical value of r_A is 1.7. Thus, given the Early voltage and the other group 4 parameters, equation (73) is convenient for estimating a_{c1}. It should be noted that a_{c1} refers to the collector capacitance of the intrinsic transistor. The terminal collector capacitance will be dominated by that of the inactive base region.

The group 5 parameters model the base push-out effects. $V_{rp} = (kT/q)v_{rp}$ is the resistive voltage drop across the collector, caused by a current of magnitude I_K. The ratio of the width of the collector epitaxial region to the width of the metallurgical base is designated r_w. The parameter r_p determines the steepness of the variation of the forward delay time as a function of a collector current in the current range where base push-out is incipient. The base push-out exponent n_p determines the fall-off of f_L for high currents. For $n_p = 2$, f_L has a tendency to level off after it has decreased from its maximum value by a factor of $(1 + r_w)^2$. For $n_p > 2$, the decrease continues beyond this level. In the example given in Section IV of this paper, the following values for v_{rp}, r_w, r_p and n_p are used.

$$v_{rp} = 18.2,$$

$$r_w = 10.0,$$

$$r_p = 4.55,$$

$$n_p = 3.0.$$

REFERENCES

1. Ebers, J. J., and Moll, J. L., "Large-Signal Behaviour of Junction Transistors," Proc. IRE, *42*, No. 12 (December 1954), pp. 1761–1772.
2. Beaufoy, R., and Sparkes, J. J., "The Junction Transistor as a Charge-Controlled Device," ATE J. (London), *13*, No. 4 (October 1957), pp. 310–324.
3. Hamilton, D. J., Lindholm, F. A., and Narud, J. A., "Comparison of Large Signal Models for Junction Transistors," Proc. IEEE, *52*, No. 3 (March 1964), pp. 239–248.
4. Koehler, D., "The Charge-Control Concept in the Form of Equivalent Circuits, Representing a Link Between the Classic Large Signal Diode and Transistor Models," B.S.T.J., *46*, No. 3 (March 1967), pp. 523–576.
5. Linvill, J. G., "Lumped Models of Transistors and Diodes," Proc. IRE, *46*, No. 6 (June 1958), pp. 1141–1152.
6. Murphy, B. T., "Diode and Transistor Self-Analogues for Circuit Analysis," B.S.T.J., *47*, No. 4 (April 1968), pp. 487–502.
7. Ohtsuki, T., and Kani, K., "A Unified Modeling Scheme for Semiconductor Devices with Applications of State Variable Analysis," IEEE Transactions on Circuit Theory, *17*, No. 1 (February 1970), pp. 26–32.
8. Early, J. M., "Effects of Space Charge Layer Widening in Junction Transistors," Proc. IRE, *40*, No. 11 (November 1952), pp. 1401–1406.
9. Sah, C. T., Noyce, R. N., and Shockley, W., "Carrier Generation and Recombination in P-N Junctions and P-N Junction Characteristics," Proc. IRE *45*, No. 9 (September 1957), pp. 1228–1243.
10. Webster, W. M., "On the Variation of Junction Transistor Current Gain Amplification Factor With Emitter Current," Proc. IRE, *42*, No. 6 (June 1954), pp. 914–920.
11. Kirk, Jr., C. T., "A Theory of Transistor Cutoff Frequency (f_T) Falloff at High Current Densities," IRE Trans. on Electron Devices, *ED-9*, No. 2 (March 1962), pp. 164–174.
12. Pritchard, R. L., and Coffey, W. N., "Small Signal Parameters of Grown-Junction Transistors at High Frequencies," 1954 IRE Convention Record, Pt. 3, pp. 89–98.
13. Fletcher, N. H., "Self-Bias Cutoff Effect in Power Transistors," Proc. IRE, *43*, No. 11 (November 1955), p. 1669.
14. Malmberg, A. F., and Cornwell, F. L., "NET-1 Network Analysis Program," Los Alamos Scientific Laboratory of the University of California, TID-4500 (19th Ed.), April 30, 1963.
15. Milliman, L. D., Massena, W. A., and Dickhaut, R. H., "CIRCUS, A Digital Computer Program for Transient Analysis of Electronic Circuits," Harry Diamond Lab. User's Guide, HDL 346-1, Program Manual HDL 346-2, January 1967.
16. Gummel, H. K., "A Charge Control Relation for Bipolar Transistors," B.S.T.J., *49*, No. 1 (January 1970), pp. 115–120.
17. Gummel, H. K., "A Self-Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculations," IEEE Trans. on Ed, *11*, No. 10 (October 1964), pp. 455–465.
18. Gummel, H. K., "Measurement of the Number of Impurities in the Base Layer of a Transistor," Proc. IRE, *49*, No. 4 (April 1961), p. 834.
19. Iwersen, J. E., Bray, A. R., and Kleimack, J. J., "Low Current Alpha in Silicon Transistors," IRE Trans. on ED, *9*, No. 6 (November 1962), p. 474–478.
20. Gwyn, C. N., Scharfetter, D. L., and Wirth, J. L., "The Analysis of Radiation Effects in Semiconductor Junction Devices," IEEE Trans. Nuclear Science, *NS-14*, No. 6 (December 1967), pp. 153–169.
21. Scharfetter, D. L., and Gummel, H. K., "Large-Signal Analysis of a Silicon Read Diode Oscillator," IEEE Trans., *ED-16*, No. 1 (January 1969), pp. 64–77.
22. Caughey, D. M., "Simulation of UHF Transistor Small-Signal Behaviour to 10 GHz for Circuit Modeling," Conference on Computerized Electronics, Cornell University, Ithaca, N. Y., August 1969.

23. Poon, H. C., and Gummel, H. K., "Modeling of Emitter Capacitance," Proc. of IEEE Letter, *57*, No. 12 (December 1969), pp. 2181–2182.
24. Gummel, H. K., "On the Definition of the Cutoff Frequency f_T,"Proc. IEEE, *57*, No. 12 (December 1969), p. 2159.
25. Poon, H. C., Gummel, H. K., and Scharfetter, D. L., "High Injection in Epitaxial Transistors," IEEE Trans., *ED-16*, No. 5 (May 1969), pp. 455–457.
26. Moll, J. L., and Ross, I. M., "The Dependence of Transistor Parameters on the Distribution of Base Layer Resistivity," Proc. IRE, *44*, No. 1 (January 1956), pp. 72–78.