THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 4, April, 1972
Printed in U.S.A.

Three-Dimensional Small-Signal Analysis of Bipolar Transistors

By J. L. BLUE

(Manuscript received November 15, 1971)

One-dimensional transistors are well-understood today; computer techniques for detailed large-signal and small-signal analyses are available. Real transistors are three-dimensional, however, and lateral effects are only understood qualitatively. Accurate modeling of lateral effects cannot be accomplished without quantitative analyses of three-dimensional transistors. Unfortunately, even the simplest analysis of lateral effects leads to a partial differential equation. In this paper, a fast and accurate numerical technique is used to solve the partial differential equation. This makes feasible a three-dimensional small-signal analysis of transistors operating in the low-injection regime.

Calculated h-parameters for a high-frequency, double-diffused, silicon transistor are in good agreement with experimental values.

I. INTRODUCTION

One-dimensional transistors are well-understood today; computer techniques for detailed large-signal and small-signal analyses are available.[1] A new charge-control model,[2] which is quite promising for use in circuit analysis programs, has been developed with the aid of insight obtained from the large-scale computer analyses.

However, real transistors are three-dimensional and lateral effects are only understood qualitatively. Accurate modeling of lateral effects cannot be accomplished without quantitative analyses of three-dimensional transistors. Unfortunately, even the simplest analysis of lateral effects leads to a partial differential equation [namely (7)]. Except for certain very simple transistor geometrical configurations, an analytic solution to the partial differential equation is not possible, and up to now, numerical solutions have been at best difficult and expensive in computer time.

In this paper, a fast and accurate numerical technique[3,4] is used to

solve the partial differential equation. This makes feasible a three-dimensional small-signal analysis of transistors operating in the low-injection regime.

It is hoped that with the aid of this technique, quantitative understanding of lateral effects in transistors can be achieved, and an accurate but simple model can be developed.

In Section II, the partial differential equation for the potential in the base region of the transistor is derived. Section III briefly reviews the solution technique. In Section IV, the calculation of h-parameters from the partial differential equation solution is described. Section V presents the results from a sample calculation, and compares them with experimental values.

II. TRANSISTOR MODEL

In this section, a discrete bipolar transistor with collector contact made to the substrate is analyzed. (A planar bipolar transistor could equally well have been chosen.) The small-signal behavior of the transistor in one dimension is modeled and a partial differential equation in the other two dimensions is derived.

Figure 1 is a cross section perpendicular to the surface of a typical discrete transistor with diffused base and emitter. A partial differential equation for φ, the potential in the base region, will now be derived. Let z be the coordinate perpendicular to the surface of the transistor. Let $\varphi(x, y)$ be the potential in the base, $\rho(x, y)$ be the charge density, and $\mathbf{J}(x, y)$ be the current density parallel to the transistor surface; φ, ρ, σ, and $\mathbf{J}$ are average values, and are taken to be independent of z.

Ohm's law is

$$\nabla\varphi = -\frac{1}{\sigma}\mathbf{J}, \tag{1}$$

where σ is the conductivity, and the continuity equation for the charge is

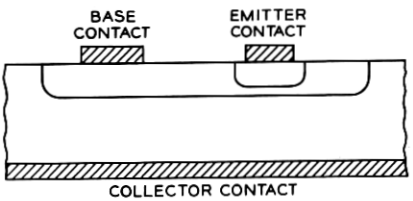

Fig. 1—Cross section of transistor.

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0. \tag{2}$$

The $\partial\rho/\partial t$ term is composed of capacitatively and resistively injected charge into the base, and will be considered shortly. Taking the divergence of (1) and substituting (2), one obtains

$$\nabla^2 \varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} = \frac{1}{\sigma} \frac{\partial \rho}{\partial t}. \tag{3}$$

In order to calculate $\partial\rho/\partial t$, the net rate of the charge injection into the base region, a specific model is necessary. This paper will consider only a simple model, since the purpose is to illustrate the technique rather than to model exactly some particular transistor. Two separate sets of simplifying assumptions will be made, one set in calculating $1/\sigma\ \partial\rho/\partial t$ to set up the partial differential equation, and the second set in calculating small-signal parameters from the solution of the partial differential equation. The second set will be discussed in Section IV.

It is assumed that the emitter doping is sufficiently high that the potential on the emitter side of the emitter-base junction is constant. This constant is taken equal to zero. It is similarly assumed that the potential on the collector side of the collector-base junction is constant. The base region is divided into two parts—the *active base region*, directly under the emitter diffusion, and the *passive base region*, the remainder of the base. The conductivity of each of the two regions is taken to be constant, but the two conductivities are not equal. Finally, low injection is assumed so that current and charge injected by the emitter-base junction is uniform in the active base region.

With these assumptions, the rate of charge injection into the base may be calculated. Equations (4), (5), and (6) may be interpreted with the aid of the equivalent circuit of Fig. 2, which is essentially a hybrid-π model of a transistor neglecting base resistance corrections. The charge injected into the base at point (x, y) may be obtained by calculating the base current for the circuit in Fig. 2; such an equivalent circuit obtains at each point of the active base. Only the circuit elements which contribute to charge flow in the z-direction are included; the lateral charge flow is calculated by solving the partial differential equation (3). Thus, it is not necessary to include in the circuit of Fig. 2 the usual hybrid-π base resistances. In the equations below, $\varphi(x, y) = v_{BE}(x, y)$ is the only quantity that depends on x and y, for (x, y) in the active base region. The rate of charge injection into the base has three components.

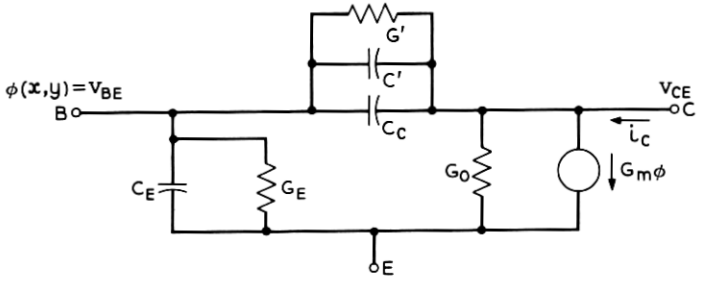

Fig. 2—Equivalent circuit used to calculate charge injection into the base and current injection into the collector, at a point (x, y) of the active base region.

(*i*) The dynamic resistance and capacitance of the emitter-base junction contributes to $\partial\rho/\partial t$ only in the active base region.

$$\left(\frac{\partial\rho}{\partial t}\right)_1 = \frac{1}{W_A}\left(C_E\frac{\partial\varphi}{\partial t} + G_E\varphi\right), \tag{4}$$

where W_A is the width (z-direction) of the active base region; C_E is the dynamic capacitance per unit area of the emitter-base junction, and includes the ordinary junction capacitance; and G_E is the conductance per unit area.

(*ii*) The collector junction capacitance contributes

$$\left(\frac{\partial\rho}{\partial t}\right)_2 = \frac{C_C}{W}\frac{\partial}{\partial t}(\varphi - v_{CE}), \tag{5}$$

where C_C is the capacitance per unit area of the collector-base junction, v_{CE} is the potential of the collector side of the collector-base junction, and W in the active base region is W_A and in the passive base region is W_P, the width of the passive base.

(*iii*) Base-width modulation in the active base region contributes

$$\left(\frac{\partial\rho}{\partial t}\right)_3 = \frac{1}{W_A}\left(G' + C'\frac{\partial}{\partial t}\right)(\varphi - v_{CE}), \tag{6}$$

where the conductance G' and capacitance C' are each per unit area. Generally C' and G' are not important. In addition, base-width modulation is the source of G_0, a conductance per unit area.

In the passive base region, the only charge injected is through the capacitance C_C.

Collecting the three terms, assuming frequency dependence $e^{i\omega t}$,

defining sheet resistance $R_A = 1/W_A\sigma_A$ and $R_P = 1/W_P\sigma_P$, the following partial differential equations for the potential are obtained.

Active base:

$$\nabla^2\varphi = R_A[G_E + G' + i\omega(C_C + C_E + C')]\varphi - R_A[G' + i\omega(C' + C_C)]v_{CE}\,. \tag{7}$$

Passive base:

$$\nabla^2\varphi = i\omega R_P C_C(\varphi - v_{CE}). \tag{8}$$

For typical parameters,

$$\begin{aligned} G_E &\gg G_0, \\ C_E &\gg C_C + C', \\ C_C &\gg C'; \end{aligned} \tag{9}$$

then in the active base the equation simplifies to

$$\nabla^2\varphi = R_A[(G_E + i\omega C_E)\varphi - (G' + i\omega C_C)v_{CE}]. \tag{10}$$

This equation was used for numerical calculations.

Equations (8) and (10) are of the form

$$\nabla^2\varphi = \gamma^2(\varphi - v_{CE}), \tag{11}$$

where γ^2 is frequency-dependent and takes on different values γ_A^2 and γ_P^2 in the active and passive base regions, but otherwise is independent of x and y.

Any other modeling of the z-direction of the transistor is equally usable, provided that it leads to equations of the type

$$\nabla^2\varphi = a\varphi + b, \tag{12}$$

where a and b are independent of x and y in each of several regions.

Figure 3 shows a typical double-base-stripe discrete bipolar transistor from the top. The active base region is under the emitter diffusion, marked E; this region is represented by eq. (10). Under the two regions marked B, φ is constant and $\varphi = v_{BE}$. The remainder of the region shown is the passive base region; in this region, eq. (8) holds. Around the passive base region is the boundary of the base diffusion. A boundary condition is that no current flows through this boundary; the current flowing through the boundary is proportional to the normally-directed derivative of the potential, so on this boundary

$$\mathbf{n}\cdot\nabla\varphi \equiv \frac{\partial\varphi}{\partial n} = 0, \tag{13}$$

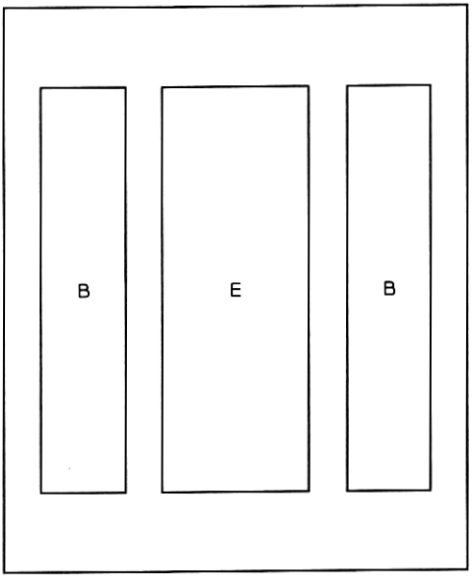

Fig. 3—Top view of transistor used in calculations.

where $\mathbf{n}$ is the outer-directed unit normal to the boundary. At the edge of the emitter diffusion, the dividing line between the active and passive base regions, the boundary conditions are that the potential be continuous and that the total current flow be continuous. These conditions are

$$\varphi_A = \varphi_P, \tag{14}$$

$$\mathbf{n}\cdot\left(\frac{1}{R_A}\nabla\varphi_A\right) = \mathbf{n}\cdot\left(\frac{1}{R_P}\nabla\varphi_P\right), \tag{15}$$

where $\mathbf{n}$ is the unit normal outward from the emitter diffusion, φ_A refers to the potential just inside the active base, and φ_P refers to the potential just inside the passive base. Equation (14) neglects the edge injection of current from the emitter-base junction; such injection could easily be included. It leads to a boundary condition, similar to eq. (15), but involving both φ and $\partial\varphi/\partial n$.

III. SOLUTION TECHNIQUE

This section describes an economical method for solving the coupled partial differential equations (8) and (10) with boundary conditions (13) through (15). The usual way to solve such equations is by finite dif-

ferences. A square or rectangular grid is superimposed on the x, y plane and derivatives at the grid vertices are approximated by finite differences. A large number of linear algebraic equations are generated and solved iteratively. While very powerful, finite difference methods have several disadvantages for the problem at hand.

It may not be easy to get the transistor boundaries (emitter diffusion, base contacts, and base diffusion) to fall on grid lines, so that more complicated programming is necessary. The areas in which the equations must be solved are frequently odd-shaped (see Fig. 4). In such cases, the usual iteration schemes are not guaranteed to converge and it may be necessary to develop an efficient scheme.

The method[3,4] developed for solving equations like (11) does not employ a grid and does not require iteration to solve the set of linear algebraic equations. The method will be briefly described for a single equation,

$$\nabla^2\varphi = \gamma^2\varphi, \tag{16}$$

holding in some area A with boundary Γ. The area A may be any shape; it may even be multiply connected. For (16), Ref. 4 derives the exact expression,

$$\pi\varphi(s') = P\int_\Gamma \left[K_0(\gamma r)\frac{\partial\varphi(s)}{\partial n} + (\mathbf{n}\cdot\mathbf{r}/r)K_1(\gamma r)\varphi(s)\right] ds. \tag{17}$$

Here s' is any point on the boundary (except at a corner point), P denotes the Cauchy principle value, K_0 and K_1 are modified Bessel functions of the first kind,[5] and $\mathbf{r}$ is the vector from s' to s. Points inside the region A are not involved; however, if s' is inside A, (17) can be

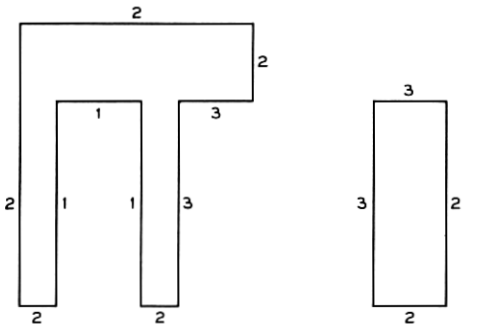

Fig. 4—"Exploded" view of one-fourth of transistor. Left—passive base region; right—active base region (under emitter). Numbers indicate boundary conditions applying (see text).

used if π is replaced by 2π. Thus, if φ and $\partial\varphi/\partial n$ are known on the boundary, φ may be found inside the boundary.

To find $\varphi(s)$ and $\partial\varphi(s)/\partial n$ on the boundary, (17) is regarded as an integral equation. This and the boundary conditions for the original partial differential equation suffice to determine the solution. In principle, an accurate numerical solution is quite simple, although the computer programming is rather complex. Various methods may be used to solve (17) plus the boundary conditions; the method used in Ref. 4 will be described here. The boundary Γ is divided into a number of straight-line segments; on each segment $\varphi(s)$ and $\partial\varphi(s)/\partial n$ are approximated by polynomials of any desired order with unknown coefficients. Using geometrical information about Γ, and the expansion of K_0 and K_1 in terms of logarithms and polynomials,[5] the integrals in (17) may be done symbolically.[4] Equation (17) cannot in general be satisfied exactly at each point s' along the boundary; N points are chosen at which to satisfy (17) exactly, where N is the total number of unknown coefficients in all of the polynomials approximating φ and $\partial\varphi/\partial n$, on every part of Γ. Then (17) is replaced by N linear algebraic equations for the polynomial coefficients. For typical transistors, N is no more than 50 or so, and therefore the equations may be solved directly without iteration by Gaussian elimination; $N = 50$ gives an accuracy of about one percent for the numerical example considered later.

When the polynomial coefficients have been found, φ and $\partial\varphi/\partial n$ are known on the boundary and φ may be found inside A if it is desired. For the transistor analysis, however, φ is not needed inside A; it is only needed on the boundary. Although the program is quite complex, the computer time is quite short. For Laplace's equation, the boundary integral equation method is considerably faster than finite difference methods.[4]

The two coupled partial differential equations, (8) and (10), may be solved in a similar manner. Fig. 4 shows the areas of Fig. 3 in which each of the two equations can be used. Because of the fourfold symmetry of Fig. 3, only one-fourth of the transistor need be considered, and only one-fourth is shown. On the left is the passive base region, in which (10) holds; on the right is the active base region, in which (8) holds. The numbers on the line segments indicate the boundary conditions on each segment. On segments numbered 1, the base contacts, $(\varphi - v_{CE})$ is constant; on the exterior of the base diffusion, segments numbered 2, and symmetry lines, $\partial\varphi/\partial n = 0$; and on segments numbered 3, the boundary of the passive and active base regions, (14) and (15) hold.

IV. CALCULATION OF h-PARAMETERS

With a fast and accurate method of solving the coupled partial differential equations (8) and (10), finding the small-signal parameters of the transistor is relatively easy. This section describes the method for determining common-emitter h-parameters.

Common-emitter h-parameters are defined by

$$v_{BE} = h_{11}i_B + h_{12}v_{CE} = h_{ie}i_B + h_{re}v_{CE}, \tag{18a}$$

$$i_C = h_{21}i_B + h_{22}v_{CE} = h_{fe}i_B + h_{oe}v_{CE}. \tag{18b}$$

In order to find all four h-parameters at a given frequency, two potential solutions are necessary. The first solution gives h_{11} and h_{21} and assumes v_{CE} to be zero. With $v_{CE} = 0$, $v_{BE} = h_{11}i_B$ and $i_C = h_{21}i_B$. The base contacts are taken to be at unit potential, $v_{BE} = 1$. The boundary integral equation method is used to solve (8) and (10) with boundary conditions (13), (14), (15), plus $\varphi = 1$ on the base contact (lines numbered 1 in Fig. 4). In Fig. 4, the output is $\partial\varphi/\partial n$ on lines numbered 1, where $\varphi = 1$; φ on lines numbered 2, where $\partial\varphi/\partial n = 0$; and both φ and $\partial\varphi/\partial n$ on lines numbered 3. The current flowing into the base contact is just

$$i_B = \frac{1}{R_P}\int \frac{\partial\varphi(s)}{\partial n}\,ds, \tag{19}$$

where the integration is over lines numbered 1 in Fig. 4. Since $\partial\varphi/\partial n$ is approximated by a polynomial on each of the sides, the integral in (19) is trivial. This gives $h_{11} = 1/i_B$.

To find h_{21}, the collector current must be found, which involves the modeling of the transistor in the z-direction. Again, the simple model used may be interpreted with the aid of Fig. 2. Transit-time effects in base and collector will be neglected. The z-directed collector current flowing through a small area $\Delta x\ \Delta y$ has three components:

(*i*) Diffusion of injected charge across the active base region. G_m is a conductance per unit area.

$$(\Delta i_C)_1 = G_m\varphi\ \Delta x\ \Delta y. \tag{20}$$

(*ii*) Injected current from collector-base junction capacitance.

$$(\Delta i_C)_2 = -C_C\frac{\partial}{\partial t}(\varphi - v_{CE})\ \Delta x\ \Delta y. \tag{21}$$

(*iii*) Injected current from base-width modulation in the active base region.

$$(\Delta i_C)_3 = G_0 v_{CE}\,\Delta x\,\Delta y - \left(G' + C'\frac{\partial}{\partial t}\right)(\varphi - v_{CE})\,\Delta x\,\Delta y. \tag{22}$$

From the passive base region, current flows to the collector only through the capacitance C_C .

The total collector current is

$$\begin{aligned} i_C = \iint_{\text{Active Base}} &\{[(G_m - G') - i\omega(C_C + C')]\varphi \\ &+ [(G_0 + G') + i\omega(C_C + C')]v_{CE}\}\,dx\,dy \\ &- \iint_{\text{Passive Base}} i\omega_C(\varphi - v_{CE})\,dx\,dy. \end{aligned} \tag{23}$$

In the calculations, G' was neglected, $G_0 + G' \equiv G_1$ were combined, and C' was neglected.

Thus it is necessary to do an area integral of the potential. Since the factors multiplying φ are independent of x and y, a two-dimensional numerical integration may be avoided. Consider one of Green's boundary-value formulas,

$$\iint_A (G\nabla^2\varphi - \varphi\nabla^2 G)\,dx\,dy = \oint_\Gamma \left(G\frac{\partial\varphi}{\partial n} - \varphi\frac{\partial G}{\partial n}\right) ds, \tag{24}$$

where the line integral is over Γ, the boundary of the area A. This is true for any well-behaved functions $\varphi(x, y)$ and $G(x, y)$. Suppose that φ obeys (16). In the area integral, substitute $\gamma^2\varphi$ for $\nabla^2\varphi$ to obtain

$$-\iint_A \varphi(\nabla^2 G - \gamma^2 G)\,dx\,dy = \oint_\Gamma \left(G\frac{\partial\varphi}{\partial n} - \varphi\frac{\partial G}{\partial n}\right) ds. \tag{25}$$

Now G may be chosen so that

$$\nabla^2 G - \gamma^2 G = 1, \tag{26}$$

or

$$G = -1/\gamma^2,$$

to give

$$\iint_A \varphi\,dx\,dy = -\frac{1}{\gamma^2}\oint_\Gamma \frac{\partial\varphi}{\partial n}\,ds. \tag{27}$$

Thus the area integral in (23) may be replaced by a line integral;

since $\partial\varphi/\partial n$ is a known polynomial, the integral is again trivial. In this way i_C may be found, and $h_{21} = i_C/i_B$.

In order to find h_{12} and h_{22} , i_B is assumed to be zero. With $i_B = 0$, $v_{BE} = h_{12}v_{CE}$ and $i_C = h_{22}v_{CE}$. The potential at the base contacts, v_{BE} , is unknown; since $i_B = 0$, $\partial\varphi/\partial n = 0$ at the base. For the second solution, the boundary integral equation method is used to solve (8) and (10) for $(\varphi - v_{CE})$ with boundary conditions (13), (14), (15), plus $\varphi = v_{BE} =$ unknown constant and $\partial\varphi/\partial n = 0$ on the base contact (lines numbered 1 in Fig. 4). v_{CE} is assumed to be 1. The output is φ on lines numbered 1 and 2, where $\partial\varphi/\partial n = 0$, and both φ and $\partial\varphi/\partial n$ on lines numbered 3. Then $v_{BE} = h_{12}$ is found directly from φ on the base contact, and i_C may be found as before to give $h_{22} = i_C$.

V. RESULTS FROM SAMPLE CALCULATION

Experimental h-parameters were available for a double-diffused, double-base-stripe silicon transistor, similar to that shown in Fig. 3. The data were taken with $v_{CE} = 1$ V, and $I_E = 10$ mA (approximately 230 A/cm^2). In the theoretical calculations, there are seven parameters: R_P , C_C , R_A , G_m , G_E , C_E , and G_1 . Two of these, R_P and G_m , were not adjusted. R_P , the sheet resistance of the passive base, was given its nominal fabrication value. G_m , which is essentially a transconductance per unit area, was set equal to $I_E/(A_E kT/q)$. A_E is the emitter area. The other five parameters were chosen so that the experimental and calculated h-parameter matched at low frequencies (below 1 MHz). No additional adjustments were made to fit the high-frequency data. The parameter values used in the calculation are given in Table I. For the calculations, the nominal fabrication geometry was used.

Complex-plane plots of the four common-emitter h-parameters are given in Figs. 5 through 8. The solid lines are drawn from the experimental data, only a few points of which are shown (solid circles). Fre-

TABLE I—PARAMETERS USED IN FITTING h-PARAMETERS

R_P,	sheet resistance in passive base, 170 ohms/□
R_A,	sheet resistance in active base, 6000 ohms/□
G_E,	conductance per unit area of emitter-base junction, 0.00051 mho/mil^2 = 79 mho/cm^2
C_E,	capacitance per unit area of emitter-base junction, 21 pF/mil^2 = 330 μF/cm^2
G_1,	low-frequency limit of h_{oc}, 0.0003 mho
C_C,	capacitance per unit area of collector-base junction, 0.226 pF/mil^2 = 0.0350 μF/cm^2
G_m,	"transconductance" per unit area, 0.058 mho/mil^2 = 8900 mho/cm^2

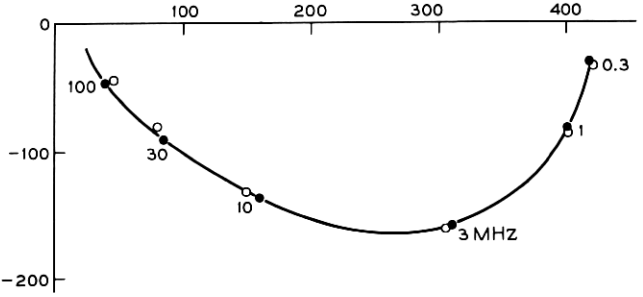

Fig. 5—Input impedance, h_{11} or h_{ie}, in ohms. The solid line and circles are experimental, and the open circles theoretical. Frequency is a parameter along the curve, and selected frequencies are given in MHz.

quency is a parameter along each curve; selected frequencies are indicated in MHz. At these frequencies, the calculated h-parameters are plotted as open circles.

At the higher frequencies, the discrepancy between experimental and calculated results becomes significant. There are several reasons why this is to be expected. In the first place, the experimental data may be in error, since h-parameter measurements are considerably more difficult to obtain at high than at low frequencies. However, it is not necessary to invoke this explanation (the traditional one for theorists) for the discrepancies to be understood.

There are two other sources of discrepancy. The first is that the simple model used for the transistor is not adequate at high frequencies. "Second-order" effects such as transit-time delay in the base and collector were ignored; other effects which are small at low frequencies could also be important. At high frequencies, the exact shape and size

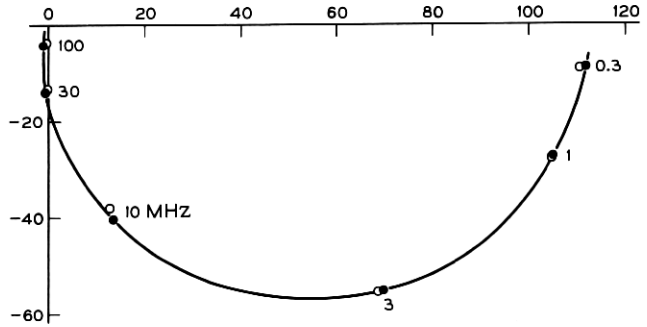

Fig. 6—Forward current gain, h_{21} or h_{fe}, dimensionless.

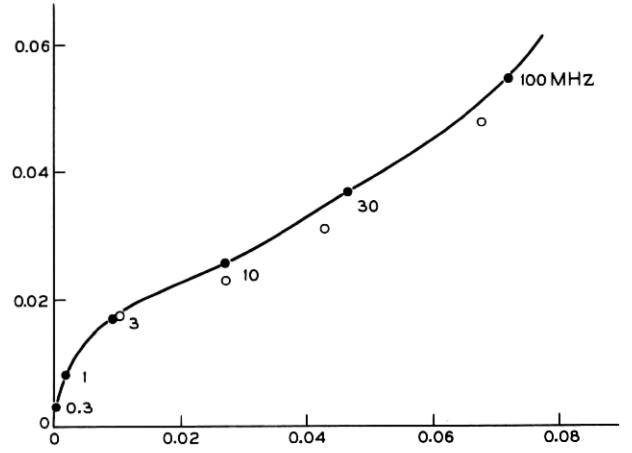

Fig. 7—Reverse voltage transfer factor, h_{12} or h_{re}, dimensionless.

of the transistor is important and the measured transistor may certainly differ somewhat in geometry from the nominal. The effect of geometry may be seen in the plot for h_{22} . The loop in h_{22} at high frequencies is due to the geometry of the transistor. Calculations done with other geometries show that, as the base and emitter stripes are made longer, the horizontal width of the loop shrinks; the loop is absent for a transistor with infinitely long stripes.

Secondly, at high frequencies the parasitic effects of the header and the encapsulation become important and need to be accurately modeled if the experimental data are to be matched. Sample calculations with

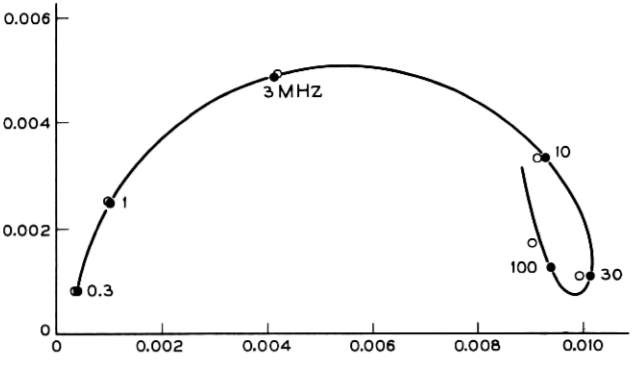

Fig. 8—Output conductance, h_{22} or h_{oe}, in mhos.

representative values showed that header parasitics had at least as large an effect on the h-parameters as the high-frequency discrepancy.

REFERENCES

1. Gwyn, C. N., Scharfetter, D. L., and Wirth, J. L., "The Analysis of Radiation Effects in Semiconductor Junction Devices," IEEE Trans. Nucl. Sci., *NS-14*, (December 1967), pp. 153–169.
2. Gummel, H. K., and Poon, H. C., "An Integral Charge Control Model of Bipolar Transistors," B.S.T.J., *49*, No. 5 (May–June 1970), pp. 827–846.
3. Blue, J. L., "Two-Dimensional Distributed Base Resistance Effects in Bipolar Transistors," talk presented at the 1969 International Electron Device Meeting, October 31, 1969, Washington, D. C.
4. Blue, J. L., "On Finding the Admittance Matrix of a Thin-Film Network by Solving the Reduced Wave Equation in Two Dimensions," J. Comp. Phys., *7*, (April 1971), pp. 327–345.
5. Abramowitz, M., and Stegun, I., Eds., *Handbook of Mathematical Functions*, Washington, D. C.: National Bureau of Standards, 1964, pp. 374–379.