Report R-910010F

AFOSR-TR- 87-0689

AD-A181 182

NUMERICAL SIMULATION OF THE PERMEABLE BASE TRANSISTOR

M. Meyyappan, J.P. Kreskovsky, B. Morrison, H.L. Grubin
Scientific Research Associates, Inc.
50 Nye Road
Glastonbury, CT 06033

DTIC
ELECTE
JUN 0 9 1987
S D
D

Approved for public release;
distribution unlimited.

May 1987

Final Report for Period 1 September 1984 - 31 December 1986

Prepared for:

Air Force Office of Scientific Research
Bolling Air Force Base
Washington DC 20332

87 5 21 060

ADA181182

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
	Approved for Public Release
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	Distribution Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
R910010	AFOSR-TR- 87-0689

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Scientific Research Associates	8N189	AFOSR/NE

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
50 Nye Road	Bldg 410
Glastonbury, CT 06033	Bolling DC 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Air Force Office of Scientific Research	FQ8671 NE	F49620-84-C-0106

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
Bolling Air Force Base	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
Washington DC 20332	61102F	3005	A1	

11. TITLE (Include Security Classification)

Numerical Simulation of the Permeable Base Transistor

12. PERSONAL AUTHOR(S)
Meyyappan, J.P. Kreskovsky, B. Morrison, H.L. Grubin

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 84/09/01 TO 87/05/04	87/05/04	146

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Permeable Base Transistor, GaAs, Nonequilibrium
			Transport, Breakdown voltage, cutoff frequency, power

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

ABSTRACT

This document describes the results of an SBIR Phase II program to investigate, through numerical simulation, the operational physics of the gallium arsenide permeable base transistor. Two distinct algorithms were used in the study: the semiconductor drift and diffusion equation algorithm and the moments of the Boltzmann transport equation algorithm. The results of the study show that one-micron feature size PBTs are capable of: a) cutoff frequencies in excess of 120 GHz, b) fmax values in the vicinity of 200 GHz, and c) breakdown voltages in excess of 17 volts. The study also demonstrates that the design of the PBT is compatible with hydrodynamic flow concepts, which suggest means to significantly reduce the capacitance of the PBT.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Dr. Gerald Witt	202 767-4931	NE

DD Form 1473, JUN 86 · Previous editions are obsolete.

TABLE OF CONTENTS

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	

By ....
Distribution /

Availability Codes

Dist	Avail and/or Special
A-1	

ABSTRACT

This document describes the results of an SBIR Phase II program to investigate, through numerical simulation, the operational physics of the gallium arsenide permeable base transistor. Two distinct algorithms were used in the study: the semiconductor drift and diffusion equation algorithm and the moments of the Boltzmann transport equation algorithm. The results of the study show that one-micron feature size PBTs are capable of: a) cutoff frequencies in excess of 120 GHz, b) fmax values in the vicinity of 200 GHz, and c) breakdown voltages in excess of 17 volts. The study also demonstrates that the design of the PBT is compatible with hydrodynamic flow concepts, which suggest means to significantly reduce the capacitance of the PBT.

NUMERICAL SIMULATIONS OF THE PERMEABLE BASE TRANSISTOR

A. INTRODUCTION AND RECOMMENDATIONS

This document describes the results of an SBIR Phase II program to investigate, through numerical simulation, the operational physics of the gallium arsenide permeable base transistor (PBT). Two distinct algorithms, developed at Scientific Research Associates, Inc. (SRA), were used in the study. The first algorithm was used to solve the semiconductor drift and diffusion equations (DDE) for electrons and holes, and was the basis for most of the study. The second algorithm was used to solve the moments of the Boltzmann transport equation (MBTE) for electron transport, and enabled an assessment to be made of the influence of non-equilibrium transport (velocity overshoot) on the operation of the PBT.

While study of the operational physics of the PBT was at the core of the Phase II SBIR program, it was SRA's additional intention to develop robust codes, drawn upon its broad experience in both solid state device physics and computational fluid dynamics, to assist in the design and development of the device. In the design and development phase of the study, workers at SRA made frequent visits to Lincoln Laboratories to discuss progress in the study. Additional design and development of the PBT (Phase III) is presently sponsored by DARPA.

In examining the PBT, specific technical goals were sought. Foremost was whether the PBT could operate at frequencies in excess of 94GHz. Here, it was clear to workers at SRA, based upon Monte Carlo calculations in Japan, that submicron scale PBTs would perform at these high frequencies. However, in view of the difficulties in fabricating micron feature size PBTs, it was SRA's view that a detailed study of 0.25 micron emitter-collector spacing PBTs should be delayed until the microwave and millimeter wave capabilities of 1.0 micron emitter-collector spacing PBTs were fully assessed. While additional study is needed, the results of the 1.0 micron study are very encouraging. For example:

1. The results of the non-equilibrium study, through solutions to the MBTE, show cutoff frequencies significantly in excess of 120GHz.

2. The results of a "Y" parameter calculation, through solutions to the DDE, show values of f_{max} near 200GHz.

3. The results of a "breakdown" calculation, through solutions to the DDE, show breakdown voltages far in excess of those achievable with coplanar FET configurations.

4. The implementation of fluid dynamic concepts to reducing the deleterious influence of the protrusive base was shown to result in substantial increases in the cutoff frequency.

Arriving at the above conclusions required a number of complementary approaches.

Initially, the achievement of high frequency performance was sought through scaling. Here, while scaling was shown to provide design directions to high frequency operation, the scaling study demonstrated that:

1. Standard scaling procedures are not suitable for submicron channel PBTs, primarily because the Schottky barrier height of the protrusive base does not scale.

2. Standard scaling procedures resulting in increased donor concentrations are not suitable for PBTs, because qualitative operation of the PBT may change. For example, high donor concentration may lead to dipole formation, whereas low concentration will not.

However, increases in carrier concentration result in reduced Schottky barrier depletion widths. And whereas the initial Lincoln Laboratories PBT calculations emphasized structures with donor concentrations of $N_D = 10^{16}/cm^3$, the largest number of calculations performed in the SRA Phase II study were or donor concentrations of $N_D = 5x10^{16}/cm^3$.

It has been known for a number of years that the protrusive base results in enhanced capacitive contributions. From the point of view of capacitance design, reducing the base penetration to zero would yield the lowest capacitance, and the highest cutoff frequency. We note that if the base penetration was zero, we would not have a PBT. However, let us, for the moment, entertain the possibility of a PBT with zero base penetration. Such a configuration would be expected to yield "short channel" effects, high fields associated with edge effects at the corners of the base contacts, and a reasonable number of carriers reacting to these high fields. For the PBT with finite base penetration, the Schottky barrier depletion extends to three sides of the base contact. While this increased depletion "area" contributes to the capacitance, this same depletion area supports fewer carriers than the zero base penetration structure and may, in a significant way, be responsible for the superior breakdown characteristics of the device. In any case, the base penetration of the PBT is taken as an engineering variable and a large number of calculations were performed with the ratio of base penetration to channel opening of 1:2. This ratio is a departure from standard Lincoln Laboratories design considerations.

In addition to the increased depletion "area" of the protrusive base, the base is also a physical barrier. Aerodynamic studies indicate that protrusions with sharp edges offer unwanted resistance to fluid flow, and that improvements in fluid flow arise from smoothing these edges, and shaping the protrusion. To see if these principles were applicable to the PBT, a region of undoped GaAs (generally half-ellipse) were inserted upstream and downstream of the base contact. The presence of these regions remarkably reduced the device capacitance and increased the cutoff frequency.

In examining the nature of these aerodynamic changes it is recognized that the Schottky barrier depletion layer tends to smooth the effects of the physical protrusion. Consistent with this concept are calculations showing that when the undoped GaAs region was contained within the Schottky barrier depletion layer, its influence was marginal. When it extended beyond the depletion layer its effect was substantial. The presence of the undoped GaAs appears to reduce the number of carriers not contributing to conduction, thereby reducing the net capacitance of the device.

The Schottky barrier PBT is a vertical device and there is no substrate for the carriers to transit across. Calculations indicate that low concentration PBTs lead to low current level devices and low cutoff frequencies, whereas high concentration devices lead to high current levels and high cutoff frequency PBTs. Mixed design PBTs yield results consistent with these observations. For example, PBT structures in which the doping in the base region was below that of a uniformly doped PBT showed lower current and cutoff frequency than the uniformly doped PBT. These latter results are significant from a design viewpoint and tend to support the speculation that the tungsten deposition and subsequent regrowth introduced traps and lowered the donor concentration in the base region. The results also underscore the sensitivity of the PBT to design modifications. While this sensitivity may be regarded as an undesirable feature, the sensitivity can be used advantageously, as the introduction of the semi-insulating region indicates.

The microwave performance of the device was calculated in a number of different ways. For most of the study, the cutoff frequency, transconductance and capacitance were obtained from the dc calculations. These dc calculations were implemented for both the DDE and MBTE studies. The extent to which the dc calculations are applicable to high frequency performance is uncertain. To generalize the calculations, "Y" parameter calculations were performed using the DDE. Here, step changes in potential were applied first to the base and then to the collector. The resulting current transients, which for the most part were completed in under 2 ps, were Fourier analyzed and the "Y" parameters computed. From the Y parameters, the current gain and power gain were computed. Each of these quantities, as expected, was extremely sensitive to circuit loading, a result that cannot be extracted from the dc calculations. Additionally, the maximum current gain was computed. The frequency at which the maximum current gain is unity is the cutoff frequency. The cutoff frequency from the Y parameter calculation was approximately 10% higher than that obtained from the dc parameters. The maximum available gain (MAG) was also computed from the "Y" parameters. The calculations for MAG show 8.5 db gain at 60 GHz (albeit slightly unstable) and 5.4 db gain at 100 GHz (albeit slightly unstable), and $f_{max} = 200$ GHz.

Most of the calculations discussed above were obtained using the drift and diffusion equations, rather than the moments of the Boltzmann transport equation. Fewer calculations were done with the MBTE because there are more equations (nine for the MBTE, two for the DDE) and they are more costly. The difference between the MBTE and DDE are profound, both qualitatively and quantitatively. For example, where there are high field Gunn domains in the DDE calculation, at least one MBTE study displayed no Gunn domains. This result emerges in spite of the fact that the potential distribution for the

two calculations are qualitatively similar. The current levels are nearly four times higher than that obtained using the DDE algorithm, a result that is a direct consequence of velocity overshoot. The MBTE also displays a negative forward conductance that is electronic in origin, a result that is not obtained from any DDE studies. Further, cutoff frequencies in excess of 200 GHz were obtained. SRA strongly recommends continued study of the PBT and other device structures be performed using the MBTE.

There are several new and unique features of the Phase II study:

1. The Phase II study is the first to use the moments of the Boltzmann transport equation to examine transport in the PBT.

2. The Phase II study is the first to include both electron and hole transport (in the drift and diffusion equations) in examining breakdown in the PBT.

3. The Phase II study is the first to implement transient accuracy, in going from one bias level to a second bias level, and to use the output to compute the frequency dependent admittance (or "Y") parameters.

4. The Phase II study is the first to use the "Y" parameter calculation to compute the figures of merit for the PBT, which heretofore were obtained either from dc calculations or analytical approximations. These figures of merit include the cutoff frequency, the current gain, the maximum available power gain, the maximum stable power gain, the stability parameter and f_{max}.

5. The Phase II study is the first to use aerodynamic principles to improve transport in the PBT.

The results of the study are summarized in four self-contained sections. Section B summarizes the scaling, material and design characteristics of the PBT. Section C describes the "Y" parameter calculations. Section D is a summary of the breakdown studies. Section E describes the non-equilibrium MBTE studies.

The broad conclusion of the study is that the PBT is, in principle, capable of achieving high-frequency analog operation with breakdown characteristics superior to that of the conventional FET, but that to achieve this promise, strict attention must be paid to the design of the doping of the channel.

B. SCALING, MATERIAL AND DESIGN CHARACTERISTICS OF THE PERMEABLE BASE TRANSISTOR

I. INTRODUCTION

This section reports the results of a detailed numerical study of the gallium arsenide permeable base transistor. The numerical results arise from solutions to the semiconductor drift and diffusion equations (DDE) and are subject to the limitations imposed by mobility models. A major goal of the study was to stay within current design constraints and to systematically determine those parameters that would most strongly affect device performance. Within this framework, it was of interest to determine whether present design features [1,2] would prevent the attainment of extremely high frequency oscillations, or whether minor design modifications could achieve significantly improved performance. The latter was shown to be the case. A second goal was to determine what design or material characteristics would lead to degraded device performance. This goal was also achieved.

Insofar as use of the drift and diffusion equations prevents an understanding of the effects of overshoot on device performance a necessary adjunct to this paper is a study that includes these effects. Two companion studies,[3,4], using solutions to the moments of the Boltzmann transport equation indicate a factor of three or better in cutoff frequency and close the loop as to whether frequencies in excess of 100 GHz can be achieved.

In examining the PBT to achieve the above goals several issues arose:

I-1. Scaling

What constraints are imposed in reducing the feature size of the PBT, and/or increasing the donor background of the device to improve the current drive and high frequency performance? Thus scaling was addressed from several viewpoints.

First a purely numerical attitude, i.e., the constraints of the differential equation are exposed. This served to highlight the inadequacies of current scaling approaches. The second approach involved constraining the average field, a condition normally imposed in VLSI scaling. The universal limitations imposed upon this type of scaling are exposed in the PBT study and have broad implications for scaling of other devices. The third approach involved scaling the depletion layer. A common feature of all three scaling approaches is the neglect of diffusion. The neglect of diffusion is shown to be the principal limitation of scaling as a direct means of improving the design of a high frequency PBT. Other approaches must be coupled to scaling to achieve improved performance.

I-2. Reconfigured Designs

While the standard approach to scaling is shown to be deficient in significant

ways, manipulation of device parameters is shown to exhibit several clear trends. For example, increases in the carrier density increase the drive current as well at the cutoff frequency. But a very serious limitation arises from the inherent design of the PBT. The base penetration introduces both an obstruction as well as a Schottky barrier with its consequent large parasitic capacitance. This latter conclusion emerged from a study that addressed the relative capacitive contributions from different sections of the device, a procedure originally implemented in reference 5. For example it was found that for a given channel opening the current density was much less sensitive to decreases in base penetration, than were changes in the charge density. The result is an increase in the cutoff frequency. It was also determined that the amount of unwanted stored charge, upstream of the base could be reduced through the introduction of a semi-insulating gallium arsenide region. The latter result is a consequence of direct application of aerodynamic principles, as discussed below and in reference 6. Here, while the transconductance of the structure increases slightly, the overall effect is to increase the cutoff frequency, in some cases by as much as 40%.

A large number of calculations were performed taking advantage of the above reconfigured designs. For example, calculations were performed for the case in which the ratio of base penetration to channel opening was 1:2, and the micron length feature size common to present PBT experiments was retained. It was found that significant improvements in the cutoff frequency resulted. sometimes as much as a factor of two. Companion studies with the moments of the Boltzmann transport equation (MBTE) for this same configuration indicated cutoff frequencies in excess of 100 GHz. Additional studies have also been performed for this reconfigured design in which the bias level was taken to the threshold for impact ionization. The unusual feature is the presence of high breakdown voltages, in excess of 15 volts [7].

I-3. Effects of Material Characteristics

It is known that regrowth of gallium arsenide after deposition of the tungsten base introduces undesirable material properties in the region surrounding the base. This is particularly troublesome in devices, such as the PBT, whose design constrains the transport of carriers to well defined narrow regions. The situation with GaAs FETs with, e.g., semi-insulating substrates are likely to be more forgiving. In the case of the PBT, calculations were performed with a reduced concentration surrounding the base. The results indicated a dramatic drop in the cutoff frequency, and suggest that efforts to tailor the doping profile, with an increased donor concentration under the base region are likely to yield fruitful results.

Each of the above issues are discussed in the ensuing sections. Section II discusses scaling from a computational point of view and introduces the role of scaling on boundary conditions. Section III outlines the computational details. Section IV discusses all of the scaling results. Section V is a description of the PBT with a ratio of base penetration to channel opening of 1:2. Section VI analyzes the role of doping variations, and the influence of the semi-insulating region. Section VII summarizes the results.

II. THE STRUCTURE OF THE PBT, ANALYSIS AND SCALING

II-1. Structure

A schematic cross-section view of the PBT, taken from the first PBT discussion[1], is displayed in figure 1. Majority carriers, electrons, enter the transit region through the metal-emitter contact, leave through the metal-collector contact, and are controlled by the Schottky base-contact. As represented in figure 1, the PBT is assumed to consist of a bank of tungsten gratings that constitute the base contacts. Edge effects are ignored and computations are confined to one unit cell of the PBT. The unit cell is captured between the dotted lines shown in figure 1. The structure of the PBT is strikingly similar to that of a diode having plane electrodes[8]. However, the presence of condensed matter as the media of transport, rather than the vacuum, alters the operating characteristics of the device in fundamental ways.

II-2. Analysis

While some analyses of the PBT have proceeded through solutions to moments of the Boltzmann transport equation [3,4] and through Monte Carlo procedures [9,10] most of the PBT discussion has been within the framework of solutions to the semiconductor drift and diffusion equations (see e.g., references 1,5, 11,12,13,14). The relevant DDE equations, ignoring ionization (see reference 7 for breakdown studies), and assuming transport by one species of carriers (electrons) are the continuity equation and Poisson's equations

$$\frac{\partial N}{\partial T} - \frac{\nabla \bullet J}{q} = 0 \tag{1}$$

$$\nabla^2 \Psi - q[N - N_D] = 0 \tag{2}$$

Here N, $J[= q(N\mu\nabla\Psi + D\nabla N)]$ and Ψ are, respectively, the carrier density, current density and potential within the PBT. N_D is the background doping, q the magnitude of the electron charge, μ the electron mobility and D the electron diffusivity.

For purposes of analysis, physical, intuitive and numerical, the above equations, with the current J, separated into its contingent drift and diffusion components, are recast into dimensionless form.

$$\frac{\partial n}{\partial t} = \frac{1}{Re}\nabla \bullet (D_n \nabla n) - Cn\nabla \bullet (n\mu_n \nabla\psi) \tag{3}$$

$$\nabla^2 \psi = Sn \ (n - n_D) \tag{4}$$

7

where the following denote dimensionless quantities:

$$\psi = \Psi/\Psi_r \quad \text{potential}$$

$$n = N/N_r \quad \text{electron density}$$

$$n_D = N_D/N_r \quad \text{donor density}$$

$$D_n = D/D_r \quad \text{electron diffusivity}$$

$$\mu_n = \mu/\mu_r \quad \text{electron mobility}$$

$$t = T/T_r \quad \text{time}$$

$$x = X/X_r \quad \text{distance}$$

$$Cn = \frac{\mu_r \Psi_r}{X_r V_r} \qquad Sn = \frac{q_r X^2_r N_r}{\epsilon_r \Psi_r} \qquad Re = \frac{X_r V_r}{D_r}$$

In the above Cn is the ratio of the mean velocity of a carrier, under an assumption of constant mobility, to a reference velocity V_r. Sn is the ratio of the square of the reference length, to twice the depletion layer width of a Schottky barrier with a "barrier" of Ψ_r (within the framework of the depletion layer approximation). ϵ_r is the reference permittivity. The quantity Re, which is analogous to the Reynolds number in fluid dynamics, represents the effectiveness of diffusion in carrier transport. For the case when Cn=1.0, and the carrier diffusion is represented by the Einstein relation Re is the ratio of the reference potential energy to the thermal energy:

$$Re = q\Psi_r/kT$$

Note: under zero current conditions, the solution to the governing equations yields $n = A + B\exp+[\psi/Re]$, where A and B are integration constants, and the Einstein relation holds. This simple example highlights the role of the Reynolds number in determining the charge distribution.

II-3. Scaling Choices

The choice of scaling and hence reference quantities is arbitrary. One useful choice is to retain constant values for the dimensionless parameters. This choice is labeled intrinsic scaling as the device material parameters are only marginally altered. To achieve intrinsic scaling a common choice for the reference time is the dielectric relaxation time of the material. Unfortunately, intrinsic scaling is never completely successful if device operation is constrained to a given ambient temperature, as the temperature dependence of the Reynolds number above, suggests. The choices made by others and also below are done to emphasize different contributions of the various components to the transport process. For example, in one of the many

8

calculations performed by Bozler and Alley [1], specifically their figure 15, the base penetration was chosen as the reference length. In these calculations rather than attempt to retain constant intrinsic parameters, an attempt was made to keep the dimensionless parameter Sn constant. In reference 1, three sets of calculations were performed in which the product N_0X^2 was constant. In particular for a nominal doping of $1.6 \times 10^{17}/cm^3$, $1.0 \times 10^{16}/cm^3$, and $4.0 \times 10^{14}/cm^3$, the base penetration was successively increased from 250A, to 1000A and 5000A, respectively. The reference potential was apparently unchanged in this calculation, as was the device length, kept at 1.0 microns.

With reference to equations (3) and (4) if the boundary conditions are unchanged, and the reference time is tied to the reference length as $T_r = X_r/V_r$, where V_r is a reference velocity, then it is anticipated that transients are shorter as the carrier density is increased. Further, if a constant mobility is assumed then in the steady state it is anticipated that the linear current density (current density x periodicity) is a constant, as all dimensionless differential equation parameters are constant (at steady state). Bozler and Alley's[1] results do not sustain this comment, partly because other device parameters were not scaled, in particular the device length. But more specifically, scaling is prevented because of nonlinearities in the material field dependent parameters and because of proximity effects.

In the analysis that follows, the reference quantities are chosen as:

1. X_r, device length as reference length,

2. Ψ_r, the potential applied across the device as reference potential,

3. V_r, saturated drift velocity as reference velocity,

4. μ_r, low field mobility as reference mobility,

5. D_r, low field diffusivity as reference diffusivity,

6. N_r, the nominal doping density as reference carrier density,

7. T_r, the reference time, given by,

$$T_r = X_r/V_r$$

(Note, when X_r is chosen such that Ψ_r/X_r is a constant, we have constant field scaling.)

There is broad significance to the parameter grouping following equation (4), which is highlighted through traditional scaling arguments that often ignore diffusion components[14]. We ignore diffusion contributions in the immediate discussion, although not in the numerical computations. Ignoring diffusion contributions in equation (3) we rewrite equations (3) and (4) as follows:

$$\frac{\partial n}{\partial t} + Cn\nabla\psi(n\mu_n) + CnSn\mu_n n(n-n_D) = 0 \tag{5}$$

Within the context of equation (5), if the boundary conditions are unchanged through parameter variation, then the successful design of a device for one set of parameters will permit equally successful scaling down in size and scaling up in frequency. We will adapt and scrutinize this point of view. We note that this point of view is basic to VLSI scaling, as discussed by Bar-Lev[15]: "The basic idea underlying present scaling theory is to keep the electric field strength invariant with size reduction. Shorter distances will translate into shorter transit times and lower voltages. Vertical, as well as horizontal distances must be scaled down....Substrate doping must be increased so that depletion regions are also scaled down..." It is noted that in the Bozler-Alley study[1], the horizontal distances were not scaled. Within the context of equation (5) scaling is exact, both spatially and temporally, if Cn and Sn are unchanged. This is satisfied if Ψ_r/X_r and $X_r N_r$ are constant, a departure from the study in reference 1. But while we will see that this type of scaling is severely limited as submicron dimensions are approached it is widely implemented within the semiconductor community.

We note that under conditions of constant field scaling $\Psi_r \rightarrow \Psi_r/\lambda(\lambda>1)$, $X_r \rightarrow X_r/\lambda$ and $N_r \rightarrow N_r\lambda$. Also, the Reynolds number, which appears in equation (3) scales as $Re \rightarrow Re/\lambda$. Thus the effect of constant field scaling is to enhance the diffusive contribution through more rapid variation of the electron concentration.

II-4. Boundary Conditions

In general boundary conditions (e.g., contact properties) do not scale in a manner that bears a simple relation to the scaling of the governing equations. This is illustrated for two simple cases. The emitter and collector contacts are ohmic and the base is a Schottky contact. The emitter and collector potentials are predicated on the assumption that the contact carrier density is equal to the nominal donor density. Thus

$$\Psi_E = -(E_C-E_F)/q = -\frac{kT}{q}\ln\frac{N_C}{N_D} \tag{6}$$

and

$$\Psi_C = -\frac{kT}{q}\ln\frac{N_C}{N_D} + V_{CE} \tag{7}$$

where $E_C - E_F$ is the difference between the conduction and Fermi energies. V_{CE} is the applied potential and N_C is the electron density of states. In the discussion below we ignore differences between the conduction and valence band density of states.

For the base contact the free carrier concentration is given by

$$N_B = N_C \exp[-q(\Phi_{BN} - V_{BE})/kT] \tag{8}$$

where we are ignoring image potential contributions, and take ϕ_{BN} to be independent of concentration. The base potential is

$$\Psi_B = -\Phi_{BN} + V_{BE} \tag{9}$$

where Φ_{BN} is the Schottky barrier height. We note:[15] using a value of $\Phi_{BN} = 0.8v$, and $N_C = 4.7 \times 10^{17}/cm^3$ and $V_{BE} = 0.4$ volts, we find $N_B = 8.68 \times 10^{10}/cm^3$. Because the relative size of the barrier concentration is orders of magnitude below the computed concentrations, the barrier concentration, N_B was taken as 'numerically zero'. In one of these calculations this zero boundary condition is physically violated. However, as discussed below, the absence of a bias dependence on the density is at least as weak as the absence of hole transport in the formulation of forward current across a Schottky contact.

Insofar as potential differences are the only relevant quantities, and we are dealing with homojunction devices, the boundary conditions are rewritten as:

$$\Psi_E = 0 \tag{10}$$

$$\Psi_C = V_{CE} \tag{11}$$

$$\Psi_B = -\Phi_{BN} + \frac{kT}{q} \ln \frac{N_C}{N_D} + V_{BE} \tag{12a}$$

Note: in Ref. 1, the density of states of the conduction and valence bands was assumed to be equal. Then,

$$\Psi_B = - \left| \Phi_{BN} - \frac{E_g}{2q} \right| - \frac{kT}{q} \ln \frac{N_D}{N_i} + V_{BE} \tag{12b}$$

where E_g is the band gap, and N_i is the intrinsic concentration. For $V_{BE} = 0.3v$, $N_D = 10^{16}/cm^3$ $N_i = 1.79 \times 10^6/cm^3$, $E_g = 1.424ev$, $\Psi_B(12a) = -0.40$, Ψ_B (12b) $= -0.37v$ for a difusion of the order of kT.

In terms of dimensionless variables:

$$\psi_E = 0 \tag{13}$$

11

$$\psi_C - \frac{V_{CE}}{\Psi_r} \tag{14}$$

$$\psi_B - - \frac{\Phi_{BN}}{\Psi_r} + \frac{kT}{q\Psi_r} \ln \frac{N_C}{N_D} + \frac{V_{BE}}{\Psi_r} \tag{15}$$

In the discussion of Bozler and Alley[1], it appears to these authors that the reference potentials were not scaled, and so the boundary conditions are determined solely by the intrinsic material properties and the applied potential. Under conditions of intrinsic scaling the dimensionless variables as well as the dimensionless boundary conditions are kept constant. However, the transformation back into dimensional units results in significant alteration of the physical boundary conditions. Because of the importance of this scaling we elaborate on the variation of boundary conditions:

a) Constant field scaling (partial), V_{BE} held constant

Under constant field scaling, discussed immediately below, the device reference length is scaled, as is the reference potential. Under this condition of 'constant field scaling, $V_{CE} \rightarrow V_{CE}/\lambda$ and the boundary conditions at the emitter and contacts are unchanged. For the case of a Schottky barrier, while the reference potential scales, the barrier height does not change. The scaled dimensionless base potential thus becomes:

$$\psi_B \rightarrow \lambda\psi_B - \frac{kT\lambda}{q\Psi_r}\ln\lambda \tag{15a}$$

b) Constant field scaling (partial), V_{BE} is determined from constant dimensionless boundary condition

Here, ψ_B is constant, and the unscaled base voltage becomes

$$V_{BE} \rightarrow V_{BE} + \left[\frac{kT\lambda}{q\Psi_r}\ln\lambda + (1-\lambda)\psi_B\right]\frac{\Psi_r}{\lambda} \tag{15b}$$

As seen in figure 2a, as λ increases V_{BE} increases. And for $\lambda>2$, increases in λ result in only marginal changes in V_{BE}

c) Constant field scaling (complete)

Here, $V_{BE} \rightarrow V_{BE}/\lambda$, and the dimensionless base potential becomes

$$\psi_B \rightarrow \lambda\psi_B + (1-\lambda)\frac{V_{BE}}{\Psi_r} - \frac{kT\lambda\ln\lambda}{q\Psi_r} \tag{15c}$$

12

As seen in figure 2b, the dimensionless potential scales, approximately
linearly with λ, for $\lambda > 1$.

d) Concentration Scaling

The situation when only the carrier concentration is scaled ($N \rightarrow \Lambda N$)
results in

$$\psi_B \Rightarrow \psi_B - \frac{kT\ln\Lambda}{q\Psi_r} \tag{15d}$$

and the dimensionless Schottky barrier is more reverse biased.

e) Concentration scaling, constant dimensionless boundary condition.

Keeping ψ_B fixed is equivalent to a slight forward bias increase (see
figure 2c) on the Schottky equal to

$$V_B \Rightarrow V_B + \frac{kT\ln\Lambda}{q} \tag{15e}$$

f) Concentration ($N \rightarrow N\Lambda\lambda$) plus constant field scaling (complete)

Combining equations 15c and 15e for $\lambda \neq 1$

$$\psi_B \Rightarrow \lambda\psi_{BE} + (1-\lambda)\frac{V_{BE}}{q\Psi_r} - \frac{kT\lambda}{q\Psi_r}\ln\lambda\Lambda \tag{15f}$$

g) Concentration plus constant field scaling (partial), V_{BE} is determined
from constant dimensionless boundary condition.

Combining equations 15c and 15e results in

$$V_B \Rightarrow V_{BE} + \left[\frac{kT\lambda}{q\Psi_r}\ln\lambda\Lambda + (1-\lambda)\psi_B\right]\frac{\Psi_r}{\lambda} \tag{15g}$$

Figure 2 displays the dimensionless boundary conditions under the assumption
of constant field scaling and an unscaled $\psi_B = -0.35$ for $V_{BE} = 0.35$ v
(see below). The important point to note here is that as the scaling
parameter increases the dimensionless Schottky barrier boundary becomes more
negative, while that associated with the ohmic contacts remains unchanged. The
change in dimemsionless potential on the Schottky contact appears to be
ignored in the discussion by Bar-Lev[14], the consequences of which are
significant. The first set of calculations on scaling illustrates the

13

consequences of ignoring the potential boundary condition alterations. Other scaling calculations will follow in order.

III. COMPUTATIONAL DETAILS

In the calculations to follow, the carrier velocity is field dependent with region of negative differential mobility. The magnitude of the carrier velocity and diffusivity are represented by the following analytical relations:

$$v = \mu_o F + a(F/F_v)^2 + b(F/F_v)^3 + c(F/F_v)^4$$

(16)

$$D = \frac{kTv}{qF} + \frac{2\tau}{3} v^2 \tag{17}$$

where $F = |F|$, and

$$F = -\nabla \psi \tag{18}$$

The term μ_0 represents the low field mobility of GaAs and is dependent upon the donor density. The term τ in equation (18) is an assigned constant. Velocity overshoot effects are not included in this calculation, but are in a companion study[3][4]. The parameters for equations (17) and (18) are displayed in Table 1, and the velocity is shown in figure 3.

For these calculations the grid structures employed were variable in spacing and employed a more densely packed mesh near the base contact ($X = Z = 40A$) than in the rest of the device. Typically, but not exclusively, for a given device structure and V_{BE} and V_{CE} values, a steady state solution for the governing equations was obtained. Then the base potential was perturbed by a small value ($V_{BE} = 0.05v$) and a new steady state solution was computed. The results were then used to calculate the performance characteristics such as transconductance, g_m, capacitance, C_T, and unity-current-gain frequency, f_T, using the following relations:

$$g_m = \frac{\Delta I_c}{\Delta V_{BE}} \bigg|_{V_{CE} = \text{constant}} \tag{20}$$

$$C_T = \frac{\Delta Q_c}{\Delta V_{BE}} \bigg|_{V_{CE} = \text{constant}} \tag{21}$$

$$f_T = \frac{g_m}{2\pi C_T} \tag{22}$$

14

Additionally, in order to obtain a clearer picture of the flow of carriers within the device, current streamlines were plotted. Its significance is as follows: A current streamfunction, σ, is defined according to the expression

$$\vec{\nabla} \times \vec{\sigma} = J \qquad (23)$$

where $\vec{J}$ is the total current density. Note: conservation of total current is consistent with the following:

$$\vec{\nabla} \cdot \vec{\nabla} \times \vec{\sigma} = \vec{\nabla} \cdot J_0 \qquad (24)$$

Hence,

$$\frac{\partial \sigma}{\partial X} = -J_Z \qquad (25)$$

and

$$\frac{\partial \sigma}{\partial Z} = -J_X \qquad (26)$$

σ can be obtained at every grid point by numerically integrating either equations (25) or (26). A current streamline is then a contour of constant streamfunction in the device, and the total current between any two stream-lines is the same.

IV. SCALING CALCULATION

IV-1. Basic Unscaled Calculation and Constant Field Scaled Device

All of the scaling calculations are done with respect to a reference or unscaled calculation. The unscaled calculation is based upon a Lincoln Laboratory calculation with the parameters listed in Table 2a. The boundary conditions, in dimensionless and dimensional units, are displayed in both Table 2b and figure 2 for $\lambda = 1$. The ac electrical parameters are displayed in Table 2c, where two sets of value are shown. The left column of Table 2c is expressed in units of the unmodeled direction, while the right column is for an assumed 100 micron channel width per finger. The symbols associated with Table 2 are shown in figure 4.

For reference d is the base penetration, h is the channel width, L_B denotes the base thickness. D is one-half of the device periodicity; $D = d+h$. L_{BE} and L_{BC} are the distances of the base from the emitter and collector boundaries, respectively. The dimensions and doping level for the computations of the reference device are displayed in figure 5. The bias levels for these computations were $V_{CE} = 1.0$ volts and $V_{BE} = 0.35$ volts.

The computations for the unscaled PBT with Table 2 parameters are displayed in figure 6. Here contours of constant carrier density (6a), potential (6b) and current streamlines (6c) are shown. The contours for the constant field scaled device with $\lambda = 4$ are shown in figure 7, with the parameters listed in Table 3. Scaling this device with the prescription of equation 15c would yield $V_{BE} = 0.0875$, and $\psi_B = -2.59$. In scaled calculation ψ_B was somewhat lower at 2.44, with $V_{BE} = 0.1276$ (see Table 3b).

For the unscaled device one notes the presence of carrier depletion under the gate and a local potential barrier upstream from the base contact. (The scaled device calculations are discussed below.) The depletion region does not pinch-off the channel opening. Pinch-off occurs for a base voltage of $V_{BE} = 0.0$ volts. Comparison of figure 6 with that of figure 4b and 5b of Bozler and Alley[1] shows very similar results, in spite of the very different representations of the field dependence of velocity and diffusivity. Additionally, the Einstein relation was used in reference 1, whereas a generalized empirical diffusivity is used in the present calculations[17]. Differences that do appear in the calculations are, for the most part, of a minor quantitative nature. There were no qualitative differences.

A few remarks are in order with respect to the current streamlines, as they represent a recent introduction into examining device physics. As drawn they indicate that between each pair of current streamlines the current/micron (in the unmodeled direction) is the same. We note the streamlines in figure 6c are not equally spaced. Also the streamlines show no emitter or collector crowding, as would be expected from coplanar source and drain contacts. There is also a small decrease in the current density along the line of symmetry of the base contact, a result familiar to those versed in vacuum tube technology[8]. We also note that immediately to the left and right of the base there are no current streamlines. The electrons in these regions are "parasitic" to the device in that they do not contribute favorably to the

16

device performance. As will be shown later, their presence results in increased device capacitance which ultimately raises the device response time.

While analogies to vacuum tube transport are significant here, there is also significant analogy to the flow of a hydrodynamic fluid. It is this latter analogy that will provide the insight, as shown later for structural modifications in the PBT that improve device performance. The analogy to hydrodynamic flow has a firm physical basis since the equations governing the flow of electrons, the drift and diffusion equations and, more generally the moments of the Boltzmann transport equation are qualitatively similar to those governing fluid flow. The concepts of mass, momentum and energy transport are fundamental to both electron transport and fluid dynamics, and the governing equations for both processes are derivable from the Boltzmann transport equation.

To carry the hydrodynamic analogy further, the electron flow in a PBT structure is similar to fluid flow between parallel flat plates with an obstruction[6,18] as shown in figure 8. As seen in figure 8 the fluid streamlines shown closely resemble the current streamlines in the PBT. From aerodynamic considerations, it is well known that when the shape of an obstruction is like an airfoil rather than a blunt protrusion, the resistance to flow is reduced. This suggests that if the base region of the PBT could be contoured or shaped the PBT would show improved performance. As discussed in a later section the introduction of a shaped semi-insulating region immediately and downstream of the base reduces the number of stagnant electrons and improves device performance.

We return to the density contours of figures 6 and 7 and supplement them with line drawings along the channel line of symmetry. With respect to figure 6b, we see that the electrons encounter a barrier, as shown by the negative potential contours near the base. This weak potential barrier is present at the center of the channel opening, as shown in figure 9a. The corresponding carrier density is shown in figures 6a and 9b, and displays carrier depletion everywhere except in the vicinity of the boundaries. Additionally, there is local charge accumulation (there is no net charge accumulation) upstream from the base contact. This local carrier accumulation occurs at the point where there is a local change in curvature of the potential (figure 9a). The change in potential curvature is more than compensated by an increasingly negative second derivative of the potential in a direction normal to the channel. The change in curvature implies the existence of a local non-dipole layer high field domain whose origin is two dimensional, rather than a consequence of the region of negative differential mobility. (Note, qualitatively similar results appear from the PBT calculation of reference 1). The high field domain is shown in figure 9c.

As seen in figures 9b and 9c the electron density and electric field indicate steep gradients on either side of the base. Most of the potential drop occurs in the region immediately downstream from the base contact. Inspection of figure 9d, which displays carrier velocity within the channel at the center of the device reveals that in the emitter region of the device the electron flux flux flowing toward the collector is due to opposing drift and diffusion components, and that the net flux is dominated by diffusion. In the base region, the drift and diffusion current augment each other. Note the

17

characteristic 'camel-back' structure associated with the field dependent mobility. The absence of this feature in the drift and diffusion simulation in reference 10 is somewhat puzzling. The current densities in the base region are large compared to the rest of the device (figure 9e) as required by current continuity and the reduction in the cross-sectional area of the base region channel.

The transistor parameters, using equations (20) through (22) and the obtained current levels are shown in Table 2c. The unit/micron reflects the unmodeled direction. It is noted that Bozler and Alley[1] do not report the values of g_m, C_T and f_T for the above structure. However, the density and potential contours in figures 6 are very similar to those reported in reference 1, and the transistor parameters extracted thereof are expected to be very similar. Note also, that the f_T calculated here is similar to that reported by Osman et al[5], who used a two piece field dependent velocity similar to the PLC curve of Bozler and Alley[1].

The results for the constant field scaled device are considerably different and are shown in figure 10. These differences are due in part to the reverse potential on the base which, as seen from figure for constant field scaling is equal to -2.44 which results in a depletion width that greatly exceeds the channel opening. The structure is essentially depleted of carriers with the current at all contacts near zero. The contour plots for this set of calculations are displayed in figure 7, while the line plots are shown in figure 10. Note the streamlines are not shown as the current is nearly zero everywhere. With regard to the line plots, note the enhanced barrier within the channel and the reversal of sign of the electric field. Thus, the carrier velocity changes sign and diffusive effects are pronounced. We note: if in scaling the dimensionless boundary was fixed at the value of the unscaled calculation, the base would be forward biased with respect to both the emitter and collector (see Table 3c and figure 2a). The consequences of this will be discussed later.

The basic conclusion that can be drawn from the above is that constant field scaling is inadequate for device prediction if submicron scales are approached. For example, if current density levels are to scale then the device cross-sectional area will have to be increased. Now if we look at these problems from a dimensionless coefficient point of view, and here we refer the reader to equations (3) and (4) we see that in going from the unscaled device to the constant field scaled device, two alterations were made. First, the relative contribution of the diffusive component of transport was increased by a factor of four. Second, the boundary conditions were changed. In the following calculation, referred to as depletion width scaling, only the doping level is altered.

IV-2. Basic Unscaled Calculation/Depletion Layer Scaling

For the basic unscaled calculation the depletion width calculated from the

18

depletion layer approximation exceeds the channel opening and also displays a local barrier within the channel. In the next set of calculations the relative conductive and diffusive components are the same as in the unscaled calculation. The only change is obtained by increasing the donor concentration, in this case to $1.25 \times 10^{17}/cm^3$. The parameters for this calculation are displayed in Table 4 where we note the increase in forward bias (as obtained from equation 15e, also see figure 2c). For this calculation the depletion layer thickness, as computed from the depletion layer approximation is 600A which is smaller than the channel opening. There are several additional points of note. First, the dimensionless boundary conditions are the same as in the unscaled device, yielding a slight increase in the base potential. Second, since the depletion layer thickness is smaller here, than for the unscaled device, there is an anticipated increase in the f_T.

The results of the calculation are displayed in figures 11 and 12. Figure 11 displays contour plots of carrier density, potential and current streamlines. Note, as for the unscaled device, negligible current flows through the base. However, unlike the results of the unscaled device (figure 6a) the potential contours display significant clustering within the vicinity of the base region. Further, a cursory examination of the unscaled contours (figure 6) and the depletion layer scaled contours (figure 11) indicates an average field up and downstream of the base that is lower for the latter. While this is indeed true, the details of the distribution indicate a totally different picture. For example, the potential contours associated with figure 6 exhibit a saddle-point structure. No such structure is present in the figure 11 contours, although the strong monotonic increase in potential (figure 12a) demonstrates the presence of a high field domain in the channel (figure 12c). The high field domain within the channel is strong enough to create a dipole layer as figure 12c attests. In addition, it is important to note that the field near the collector in the figure 12c calculation is relatively flat and equal to 680v/cm. Indeed, most of the collector current is conductive. For the figure 6 calculations, the field near the collector is lower and equal to 130v/cm. The current for the unscaled device is lower than the depletion-layer scaled device and is dominated by diffusive rather than convective transport. The consequence of this difference are profound. For example, the f_T for the depletion-layer-scaled device is equal to 37.1 GHz and is dominated by the properties of the base contact and transport through the base region. As a result marginal alterations in the emitter-collector spacing are expected to have only a marginal effect on f_T. Such is not the case in the basic structure with lower density, figure 6 calculation, where an alteration in the emitter-collector spacing may be expected to drastically alter the field profile at the collector contact with the result that f_T is altered.

A further point of interest is that an increase in the donor density by slightly over an order of magnitude results in an increase in the collector current from 0.015 A/cm to 1.27 A/cm. The increase is greater than one order of magnitude and arises from a synergistic combination of a quantitative

alteration in the relative convective/diffusive contributions (even though these alterations do not arise from the dimensionless variables) and the increased donor concentration.

IV-3. Basic Unscaled Calculation/Depletion Layer and Constant Field Scaling

The depletion layer scaling of the previous section resulted in two qualitative differences from both the unscaled and constant field scaled device. First, there was the absence of a channel potential barrier. Second, there was the presence of a high field domain and net carrier accumulation within the channel. There was also an achievement of higher f_T. In this section we exploit the two scaling procedures. First, we retain a high carrier concentration, in this case, a doping density of $2 \times 10^{17}/cm^3$. Second, we reduce the reference length to 0.5 microns, scale all the other dimensions accordingly. For example, if the average field between the contacts is the same as in the unscaled device, then $V_{CE} - 0.5$ volts and $V_{BE} - 0.175$. This has the effect of reducing the dimensionless boundary to $\psi_B - - 1.21$. In the calculation below $\psi_B - 1.15$, for $V_{BE} - 0.20$, and $V_{CE} - 0.5v$. (See Table 5b)

It is recalled that in the constant field scaled device the dimensionless parameters indicated a increase in the diffusive contribution through the decrease of Reynolds number by a factor of four. In the present calculation the dimensionless Reynolds number is decreased by a factor of two, and the carrier density variations are more pronounced. The parameters of the calculation are listed in Tables 5a and 5b.

The contours and current streamlines are displayed in figure 13, while the line plots are shown in figure 14. Figure 13a and 14b show the absence of charge accumulation within the channel, although there is a local minima. The potential contours show a local barrier, but it is not within the channel, rather it is upstream from the base contact. The potential contours when combined with the line plots show the presence of a saddle point within the vicinity of the base contact as in the case of the unscaled device. But significantly, at and near emitter and collector contacts the field is relatively uniform, indicating that transport to the contacts is primarily conductive. The transistor parameters indicate an $f_T - 43.8GHz$, which is approximately three times that of the unscaled device, and about 30% higher than that of the depletion layer scaled device. Further the current level for the depletion layer/constant field scaled device is 0.39 A/m which is approximately an order of magnitude greater than that of the unscaled device, although it is lower than that of the depletion layer scaled device. The latter is a consequence the fact that the dimensionless potential is more reverse bias for the depletion layer/constant field device, than for the depletion layer scaled device.

There is another point that should be recognized. For this calculation there was some leakage current through the base contact. As displayed in the current streamlines (figure 13c) approximately 10% of the current flows from the emitter to the base contact. This will be discussed in more detail below.

IV-4. Basic Unscaled Calculation/Depletion Layer and Constant Boundary Condition Scaling

The calculation in Section IV-2 involved constant field scaling. The numerical boundary conditions for the constant field scaling device, as shown in Table 5b indicate that the base was more reverse bias then the unscaled device. If the numerical boundary conditions are kept constant, as shown in Table 3c, the only parameter that differs between the scaled and unscaled calculation is the contribution of the diffusivity. Further, according to Table 5a, the dimensional units for the potential (from equation 15c) indicate that the base is forward bias with respect to both the collector and the emitter, when all dimensionless boundary conditions are kept constant. The consequences of this as shown in figure 15 display an unacceptable base leakage current. A very similar result holds for the 0.5 micron device, with the same dimensionless boundary conditions (see equation 15c) as the unscaled device. This is shown in figure 16. The question of relevance is: why is there such a large leakage current?

The most obvious place to look for an answer is in the assumptions of the calculations. First the calculations neglect hole transport, a feature which renders the physics weak in the forward bias regime. A second assumption lies in the inclusion diffusion the calculation. Also note, that the diffusivity used here is generally above that determined from the Einstein relation. The origin of our choice lies in the measurements of diffusivity which show strong departures from the Einstein relation[17]. The third and most important contribution is associated with proximity effects. Figure 17 illustrates, where one-dimensional calculations were performed for two structures, $X_r = 1.0\mu m$ and $X_r = 0.25\mu m$. The dimensionless potentials for both calculations was -0.349 and 0.0, respectively. For the $1.0\mu m$ calculation, the carrier concentration at the classical (diffusionless) depletion length is $X_D = X_r/2\sqrt{5}Sn = 0.09N_r$. For the $0.25\mu m$ calculation, the carrier concentration at X_D is $0.95 N_r$. The points X_D are indicated in Figure 17. Thus, we must conclude that the enhanced forward base bias and proximity effects violate the depletion layer approximation and may thereby contribute to the current.

IV-5. Provisional Scaling Conclusions/ The Role of the Embedded Base

The question that arises now is how we assess the above scaling procedures. Certainly, scaling can only be regarded as successful if the scaled device parameters lead to a consistent set of output results with simple design rules. For example, constant field scaling should lead to the following scaled results: $I_C \rightarrow I_C/\lambda$, where I_C is the is the collector current; $I_C V_C \rightarrow I_C V_C/\lambda^2$, where V_C is the collector voltage. The capacitance should scale as $C_T \rightarrow C_T/\lambda$, whereas the transconductance should be unchanged, and the cutoff frequency should scale as $f_T \rightarrow f_T\lambda$. The results of this study, as well as those of others indicate, that these scaling rules are only qualitatively followed. Thus in the strict sense, scaling rules fail as submicron dimensions are approached.

The above discussion indicates that a direct straightforward way to scaling is

21

not likely to lead to confidence that the scaled device will perform as simple design rules anticipate. Thus other approaches are needed to improve the structure of the PBT. In examining these approaches we introduce a diagnostic procedure first discussed by Osman, et al.[5] for examining the capacitance contributions of different regions of the device. The capacitance contributions are obtained by integrating the change in the mobile charge due to an incremental change in the base voltage. The regional contributions involve integrating the regional charge distribution due to incremental base voltage changes. The study below also includes an analysis of the regional transconductance contributions. For example, figure 18 shows the distribution of capacitance for the depletion layer scaled device (Section IV-2). As seen about 50% of the device capacitance arises from the distribution of space charge immediately upstream and downstream from the embedded base contact. Further, since the current streamlines (see figure 11c) indicate that the electron current flux from the emitter to the collector is typically not flowing immediately upstream and downstream of the embedded base region, these capacitance contributions are regarded as largely parasitic.

There are at least two approaches to reducing the these parasitic contributions. The first is through a reduction in the penetration of the base. This is considered next. The second approach is through the alteration of the material characteristics. This is considered in Section VI. Before proceeding it is worthwhile noting that in the Bozler and Alley study[1] the suggestion was made that devices in which the ratio of channel opening to base penetration, h/d is less then unity would have cutoff frequencies greater than those structures in which h/d > 1. Thus there is added incentive for this approach.

The calculations reported below involve varying the ratio h/d by varying the base penetration. In these calculations h is constant and equal to 1000A. Note the periodicity of the structure is D − 2(h+d). In these calculations it is found that f_T increases as the ratio h/d is increased. These results are consistent with the calculations of reference 5, rather than reference 1.

The variable base penetration calculations are displayed in figure 19 (density), figure 20 (potential) and figure 21 (current streamlines). The calculations were performed for the same material and device parameters as the depletion layer scaled device of figure 11. The added feature here is that the base penetration was taken as 1500A, 1000A, 500A, 250A and 0A. Figure 19 establishes the essential features of the calculation. Here for the 1500A, 1000A, 500A and 250A calculations, it is visually apparent that the distribution of charge within the channel for all four structures is approximately independent of the base penetration. Changes occur only for the zero penetration structure. The remarks remain the same when the potential distribution and the current streamlines are examined. It may therefore be expected that the relative contribution of the parasitic capacitance has been reduced as the base penetration is reduced. Indeed as figure 22 indicates the total capacitance contribution decreases as the base penetration decreases while the transconductance undergoes more moderate changes. The cutoff frequency increases from under 30 GHz to values in excess of 80 GHz. It is worthwhile noting that the variation in the base penetration, with the exception of the zero base penetration, does not alter the current density through the PBT.

22

This result is displayed in Table 6. But it must also be recalled that these base penetration studies are being performed for a structure that must be regarded as operating in an on-state, as the current density is sufficiently high to warrant the presence of a negative differential mobility dipole layer. Calculations of this type were not performed for devices that were normally-off, and the generalization to them is not in order.

V. IMPROVED DEVICE DESIGN

The above calculations suggest that improvements in device performance can be obtained with modest alterations in design. For example, increases in carrier density, narrower base penetration, shorter emitter collector spacing, will increase the cutoff frequency. We have chosen to stay within the constraints of present technology to offer improvement in design. The structure chosen for a more detailed evaluation has a narrow base penetration, in this case 500A, a channel opening of 1000A and a nominal doping level of $5 \times 10^{16}/cm3$. This level of doping was chosen because it is one commonly being used [2]. Further, since most of the structures being treated experimentally have a half-channel opening of 1000A, this particular design is highly relevant to today's studies. The parameters of the calculation are displayed in Table 7.

The first set of calculations is displayed in figure 23, which shows the collector current versus collector voltage. Several features should be noted. First, on the scale of figure 23, the current voltage characteristics show a relatively soft saturation. In a companion study for these device parameters, but where the moments of the Boltzmann transport equation are solved the current-voltage characteristics show harder saturation[3,4]. Second, we note, the absence of any negative forward conductance. Experimentally a negative forward conductance is observed at the higher base bias levels[2]. At least part of this observed negative conductance is attributable to heat transport, which lowers the mobility of the semiconductor. However, in the companion study, a negative forward conductance that is electronic in origin is also observed[3,4]. Third, the soft characteristic does not continue at all bias levels. In another companion study[7] in which breakdown effects are being studied, the collector current was calculated at a gate bias level of 0.3 volts. In that study, it was observed that saturation is harder at the higher bias levels. Fourth, there is an apparent increase in transconductance as the base potential is increased. Now as discussed in the earlier sections, the transistor parameters may be obtained through an incremental change in voltage on the relevant contacts. For the dc parameters an alternative method is to determine the collector current and charge as shown in figures 24 and 25. By eliminating voltage between figures 24 and 25 the relation between current and stored charge is obtained. This is shown in figure 26, the slope of which yields the bias dependence of the cutoff frequency, also shown in figure 27a. An important point to note is that the cutoff frequency increases with increasing base voltage (the maximum increase is nearly 30%), and the increase is smaller as the collector voltage increases. Further, the bias dependence of the cutoff frequency is smaller than that reported by Alley[12] and Osman, et al[5]. Here, however, we note that the dimensions, doping and field dependent velocity curve used were different in the three calculations.

The sensitivity of the cutoff frequency to design parameters was discussed in

reference 5. To expand on this study, and to throw some light on the structure-to-structure variation a series of calculations were performed, at one collector bias level, for a PBT with the same parameters as those of the calculation corresponding in figure 23, but with a base penetration of 1000A. This calculation is displayed in figure 27b, with results that were closer to those reported by reference 5, indicating a more sensitive dependence on bias. The most sensitive bias dependence occurred for the unscaled device, where increasing the base bias resulted in a transition from a low current to a high current drive device. Figure 27c results indicate that the percent change in cutoff frequency is reduced as the base penetration is reduced (see also reference 5), and that the percent change in the cutoff frequency increases as the donor concentration decreases.

The trend of the percent change in the cutoff frequency is clear from the above set of calculations, and also through the discussion in reference 5. For the $5 \times 10^{16}/cm^3$ and $1.0 \times 10^{16}/cm^3$ calculations the percent increase in cutoff frequency decreases as the base penetration decreases (for a given channel opening). For the same device structure the percent change in f_T decreases as the concentration increases. But in this latter case there is a qualitative change in the distribution of charge. For the higher donor concentration a high field domain exists within the channel, as displayed in figure 28.

Figure 28 displays contours of charge density, potential and current streamlines for the figure 23 calculation with $V_{CE} = 1.0$ volt and $V_{BE} = 0.3$ volts. We note the presence of a high field domain whose origin lies in the region of negative differential mobility associated with the device. This is displayed in figure 28. The result is of interest because it suggests the presence of a high field Gunn domain in the devices reported in [2]. There is however, a caveat. The calculations with the moments of the Boltzmann transport equation show a reduced susceptibility to high field Gunn domain formation[3,4].

There are several other figures of merit for the PBT. For example Bozler and and Alley[1] compared the PBT to other transistors and demonstrated that the high bias collector characteristics followed a square law dependence. Their study also demonstrated that for the low doped $1.0 \times 10^{16}/cm^3$ that the extrapolated threshold voltage was slightly less than 0.3 volts. Their calculation is repeated here and shown in Figure 29.. Additionally, the calculation was repeated for the case when the nominal carrier density is a higher $5.0 \times 10^{16}/cm^3$, and when the base penetration is reduced to 500A. It is anticipated that a larger (in magnitude) reverse bias on the base is required to turn the device off. As shown in figure 29, the current follows a square law dependence at high bias levels and extrapolates to a threshold voltage of less than -0.8volts for the 500A base and less than -1.0 volts for the 1000A base.

VI. MATERIAL DEPENDENCE AND DEVICE CHARACTERISTICS

The discussion in the above sections suggested that material properties strongly affect device performance, some of which are deleterious and some beneficial. We address the former issue first.

24

VI-1. The Effects of Channel Material Variations on Device Performance

It is known that the effects of regrowth on the GaAs PBT often introduce deleterious modifications in the material characteristics of the PBT. Some of these deleterious results arising from regrowth are mitigated through alternative growth procedures. The question of interest is why the PBT is extremely sensitive to material properties, whereas the standard configured GaAs coplanar FET while material dependent does not appear to be as sensitive to material variations. While the answer to this question is speculative it must be remembered that the current paths available to carriers in the standard GaAs FET configuration are broader than that of the PBT, with the FET permitting injection into the substrate. As a result small changes in the material characteristics of the PBT are likely to have a greater influence than in the case of the FET. Perhaps the clearest role of the influence of the material characteristics was in a discussion of Gopinath et al [13]. In the study below the device whose characteristics are displayed in figure 23 was altered through a reduction in the donor concentration under the base contact. Here the donor concentration was decreased to $1.0 \times 10^{16}/cm^3$ over a distance of approximately 2000A. The altered structure is shown in figure 30. Detailed calculations were performed at four pairs of base and collector voltage levels, as shown in Figure 20. Several results of this study are described below.

First, the effect of the increased resistance lowers the current (See Figure 31). Second, the effect of the reduced donor concentration is to increase the effective depletion layer width (See Figure 32). Hence, there is an expected decrease in cutoff frequency. Indeed the cutoff frequency has been reduced to 22 GHz, which is approximately half of that for the unmodified structure. But there was also a major qualitative alteration. For the high concentration flat profile device, we observed the presence of a dipole layer (See Figure 28). For the structure in which the concentration was reduced, the dipole layer was absent. For a direct comparison, see Figure 32b.

VI-2. The Influence of a Semi-Insulating Region on Device Performance

As discussed earlier, a variety of tools are required to redesign the PBT to minimize parasitics. Aerodynamic considerations have suggested that the blunt shape of the embedded base increases the resistance of the structure to current flow. Insofar as shaping the embedded base is an unlikely scenario, a design procedure was introduced to minimize the presence of blunt edges on current flow. The procedure involved the placement of 'contoured barriers' immediately upstream and downstream of the base. For these barriers to be effective it is necessary that they extend a distance away from the base that nominally exceeds the electrical depletion width of the base. The 'contoured barriers' were conceived as as regions of semi-insulating gallium arsenide, thereby eliminating the occurrance of unwanted surface states. The realization of the contoured barriers is displayed in figure 33 for a structure with an embedded base of 1000A.
Quarter-circle regions each of 1000A radius are placed immediately upstream and downstream of the embedded base. As seen in figure (34) and in comparison

to figure 18, there is a dramatic drop in the distribution of capacitance. In figure (35) we display the channel density and potential contours. It is apparent that the channel dynamics are only marginally modified. However, the protrusive sharp corner effect of the penetrating base is reduced, and as indicated in figure (34) the cutoff frequency increased by 33% to 49.6GHz. Improvements in device performance is not restricted to one design configuration.

Earlier calculations showing the effects of base penetration on device performance were repeated with the quarter-circle embedded region. The results are displayed in figure (36) and show general broad improvement, over the entire range of variable base penetration although the current level is reduced (see Table 8).

The importance of designing the semi-insulating region to exceed the depletion width associated with the tungsten Schottky barrier is illustrated in figure (37) for a device with a donor concentration of 2×10^{17} and a base penetration of 500A. Here, all device structural parameters are the same as discussed above, but because of the increased carrier concentration, quarter circle regions of radius equal to the base penetration do not exceed the electrical depletion width. This is illustrated in figure (37). Rather, an elliptical region of semi-insulating gallium arsenide is required, as shown in figure (7). The latter result strongly suggests the broad result that means of altering electronic flow patterns through static means may provide a powerful design tool for improving device performance.

VII: CONCLUSIONS

The PBT is a structure capable of achieving high frequency, high speed operation. Because of its size, and the constraint imposed on current paths, the PBT is highly sensitive to material variations. While at one end of the scale this sensitivity to material variation might be regarded as deleterious, at the other end of the scale dramatic performance improvements are possible, as the introduction of the semi-insulating region attests to. Additionally, the ultra-submicron features of the device permit the testing of scaling concepts. As is apparent from the scaling discussion, direct scaling is inadequate, primarily because the Schottky barrier does not scale. As a result, full two-dimensional numerical simulations are necessary for analyzing the usefulness of design changes.

VIII: ACKNOWLEDGEMENT

26

REFERENCES

1. C.O. Bozler and G.D. Alley, IEEE Trans. Electron Devices 27, 1128 (1980).

2. C.O. Bozler, M.A. Hollis, K.B. Nichols, S. Rabe, A Vera and C.L. Chen, IEEE Electron Dev. Lett. 6, 456 (1985).

3. J.P. Kreskovsky, M. Meyyappan, and H.L. Grubin, Proceedings of NUMOS Symposium, Oct. 1985.

4. To be published by the authors.

5. M.A. Osman, D.H. Navor, T.W. Tang and L. Sha, IEEE Trans. Electron Devices, 30, 1348 (1983).

6. SRA Report R-910010-2.

7. To be published by the authors.

8. E.L. Chaffee, Theory of Thermionic Vacuum Tubes, McGraw-Hill (1933), Cf figures 75 and 76.

9. Y. Awano, K. Tomizawa and N. Hashizume, IEEE Trans. Electron Devices.

10. C.G. Hwang, D.H. Navon, and T.W. Tang, IEEE Trans. Electron Dev. Lett. 6, 114 (1985).

11. D.E. Snyder and R.L. Kubena, IEDM Technical Digest, 612 (1981).

12. G.D. Alley, IEEE Trans. Electron Devices, 30, 52 (1983).

13. A. Gopinath and J.B. Jenkin, Proceedings of the Cornell Conference, 1985.

14. A. Bar-Lev, Semiconductors and Electronic Devices, Prentice Hall, 1984.

15. S.M. Sze, Physics of Semiconductor Devices, John Wiley & Sons, 1981.

TABLE 1. Values of Constants in Equation (17) and (18)

μ^0 cm²/volt	Density Dependent
a, cm/s	1.0×10^6
b, cm/s	6.0×10^6
c, cm/s	7.5×10^6
F_V, volts/cm	4000
τ, s^{-1}	2.25×10^{-13}

TABLE 2a. Unscaled Device

Dimensions (microns)	Reference Values	Dimensionless Parameters
L - 1.02	$D_r - \mu kT/q$	Sn - 16.3
L_{BE} - 0.40	X_r - $1.0 \mu m$	Cn - 7.5
L_B - 0.02	ψ_r - 1.0V	Re - 5.2
L_{BC} - 0.60	N_r - $10^{16}/cm^3$	
d - 0.10	t_r - 10.0ps	
h - 0.10		

TABLE 2b. Potential in Dimensionless and Dimensional Units

$$\psi_B - - 0.349, \qquad \frac{V_E}{\Psi_r} - 0, \qquad \frac{V_{CE}}{\Psi_r} - 1.0$$

Potential in Volts

$$V_{BE} - +0.35V \qquad V_{CE} - 1.0V$$

TABLE 2c. Transistor Parameters Per Half Cell

g_m (mS/μm) -	0.02	g_m (mS) -	2.0
C_T (pf/μm) -	2.0×10^{-4}	C_T (pF) -	0.02
f_T(GHz) -	16.1	f_T(GHz) -	16.1
I_D (mA/μm) -	0.15	I_D (mA) -	15
I_G (mA/μm) -	0.003	I_G (mA) -	0.3

TABLE 3a. Constant Field Scaling

Dimensions (microns)	References Values	Dimensionless Parameters
$L = 1.02/4 = 0.255$	$D_r = \mu_r kT/q$	$Sn = 16.3$
$L_{BE} = 0.100$	$X_r = 0.25\mu m$	$Cn = 7.5$
$L_B = 0.005$	$\psi_r = 0.25V$	$Re = 5.2/4 = 1.3$
$L_{BC} = 0.150$	$N_r = 4 \times 10^{16}/cm^3$	
$d = 0.025$	$t_r = 2.5ps$	
$h = 0.025$		

TABLE 3b. Potential in Dimensionless Units

$$\psi_B = -2.44, \qquad \frac{V_E}{\psi_r} = 0, \qquad \frac{V_{CE}}{\psi_r} = 1.0$$

Potential in Volts

$$V_{BE} = +0.128 \qquad V_{CE} = 0.25V$$

TABLE 3c. Potential in Dimensionless Units (see figure 15)

$$\psi_B = -.349, \qquad \frac{V_E}{\psi_r} = 0.0, \qquad \frac{V_{CE}}{\psi_r} = 1.0$$

Potential in Volts

$$V_{BE} = 0.649 \qquad V_{CE} = 0.25$$

TABLE 4a. Depletion Layer Scaling

Dimensions (microns)	Reference Values	Dimensionless Parameters
L = 1.02	$D_r = \mu_r kT/q$	Sn = 204.1
L_{BE} = 0.40	X_r = 1.0μm	Cn = 7.5
L_B = 0.02	ψ_r = 1.0V	Re = 5.2
L_{BC} = 0.60	N_r = 1.25 x 10^{17}/cm^3	
d = 0.10	t_r = 10.0ps	
h = 0.10		

TABLE 4b. Potential in Dimensionless Units

$$\psi_B = -0.349, \qquad \frac{V_E}{\Psi_r} = 0.0, \qquad \frac{V_{CE}}{\Psi_r} = 1.0$$

Potential in Volts

$$V_{BE} = +0.417V \qquad V_{CE} = 1.0V$$

TABLE 4c. Transistor Parameters Per Half Cell

Per Unmodeled Direction		Per 100 Microns Base Width	
g_m (mS/μm) = 0.111		g_m (mS) = 11.1	
C_T (pF/μm) = 4.75 x 10^{-4}		C_T (pF) = 4.75 x 10^{-2}	
f_T (GHz) = 37.1		f_T (GHz) = 37.1	
I_D (mA/μm) = 0.128		I_D (mA) = 12.8	
I_G (mA/μm) = 0.0036		I_G (mA) = 0.36	

TABLE 5a. Depletion Layer and Constant Field Scaling

Dimensions (microns)	Reference Values	Dimensionless Parameters
$L = 0.5$	$D_r = \mu_r kT/q$	$Sn = 163$
$L_{BE} = 0.02$	$X_r = 0.5\mu m$	$Cn = 7.5$
$L_B = 0.01$	$\psi_r = 0.5V$	$Re = 5.2$
$L_{BC} = 0.30$	$N_r = 2 \times 10^{17}/cm^3$	
$d = 0.05$	$t_r = 5.0ps$	
$h = 0.05$		

TABLE 5b. Potential in Dimensionless Units

$$\psi_B = -1.15, \qquad \frac{V_E}{\psi_r} = 0, \qquad \frac{V_{CE}}{\psi_r} = 1.0$$

Potential in Volts

$$V_{BE} = 0.20 \qquad V_{CE} = 0.50$$

TABLE 5c. Potential in Dimensionless Units

$$\psi_B = -0.349, \qquad \frac{V_E}{\psi_r} = 0.0, \qquad \frac{V_{CE}}{\psi_r} = 1.0$$

Potential in Volts

$$V_{BE} = +0.60V \qquad V_{CE} = 0.5V$$

TABLE 6. Effect of Base Penetration on Collector Current (No I Region)

Current, mA/μm

Base Penetration	Without I Region Drain	Base
0	0.1417	3.5×10^{-3}
250	0.1282	4.3×10^{-3}
500	0.1284	4.6×10^{-3}
1000	0.1289	5.5×10^{-3}
1500	0.1297	6.5×10^{-3}

TABLE 7. Device Parameters Used to Obtain Figure 23
(In these calculations, dimensionless parameters are obtained from equation 12b)

L (μm) — 1.02	$D_r = \mu_r kT/q$	Sn — 81.6
L_{BE} (μm) — 0.40	X_r — 1.0μm	Cn — 7.5
L_B (μm) — 0.02	ψ_r — 1.0v	Re — 5.2
L_{BC} (μm) — 0.60	N_r — 5×10^{16}/cm^3	
d (μm) — 0.05	t_r — 10.0ps	
h (μm) — 0.10		

TABLE 8. Effect of Base Penetration Channel Opening 1000°Å With I Region

Current, mA/μm

Base Penetration	With I Region Drain	Base
0		
250	0.1184	3.5×10^{-3}
500	0.1175	3.6×10^{-3}
1000	0.1192	4.1×10^{-3}
1500	0.1203	4.7×10^{-3}
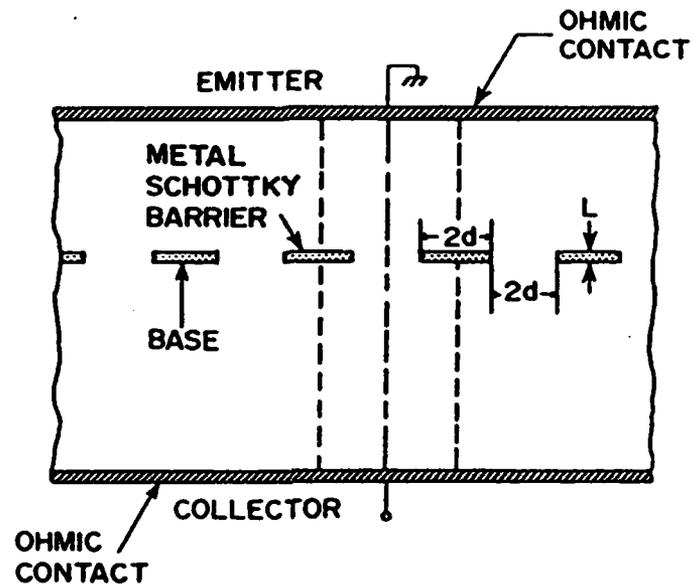
Fig. 1 — Schematic of the PBT.
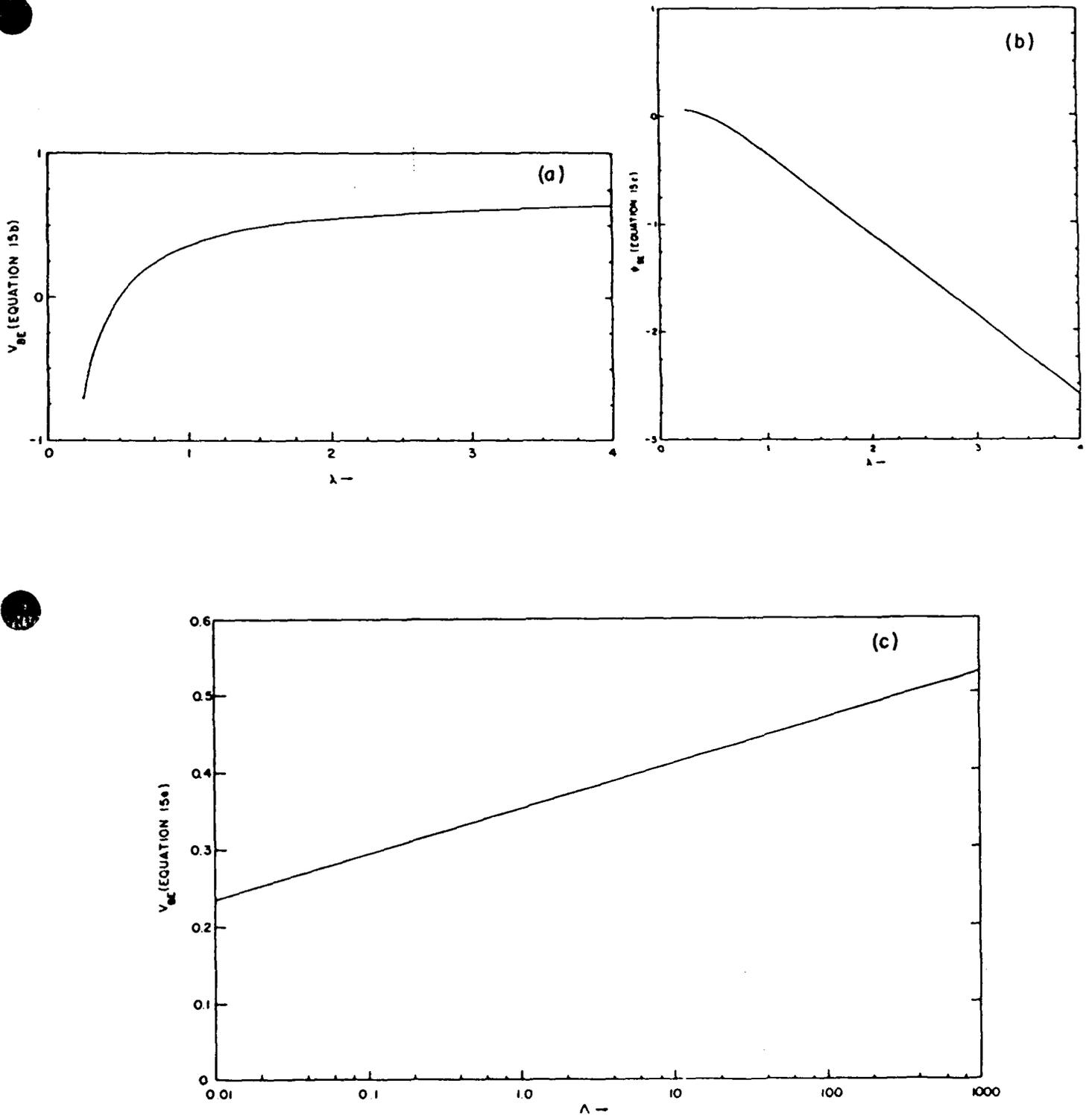
Fig. 2 — Parameter dependence of the scaling equation (equation 15).

35
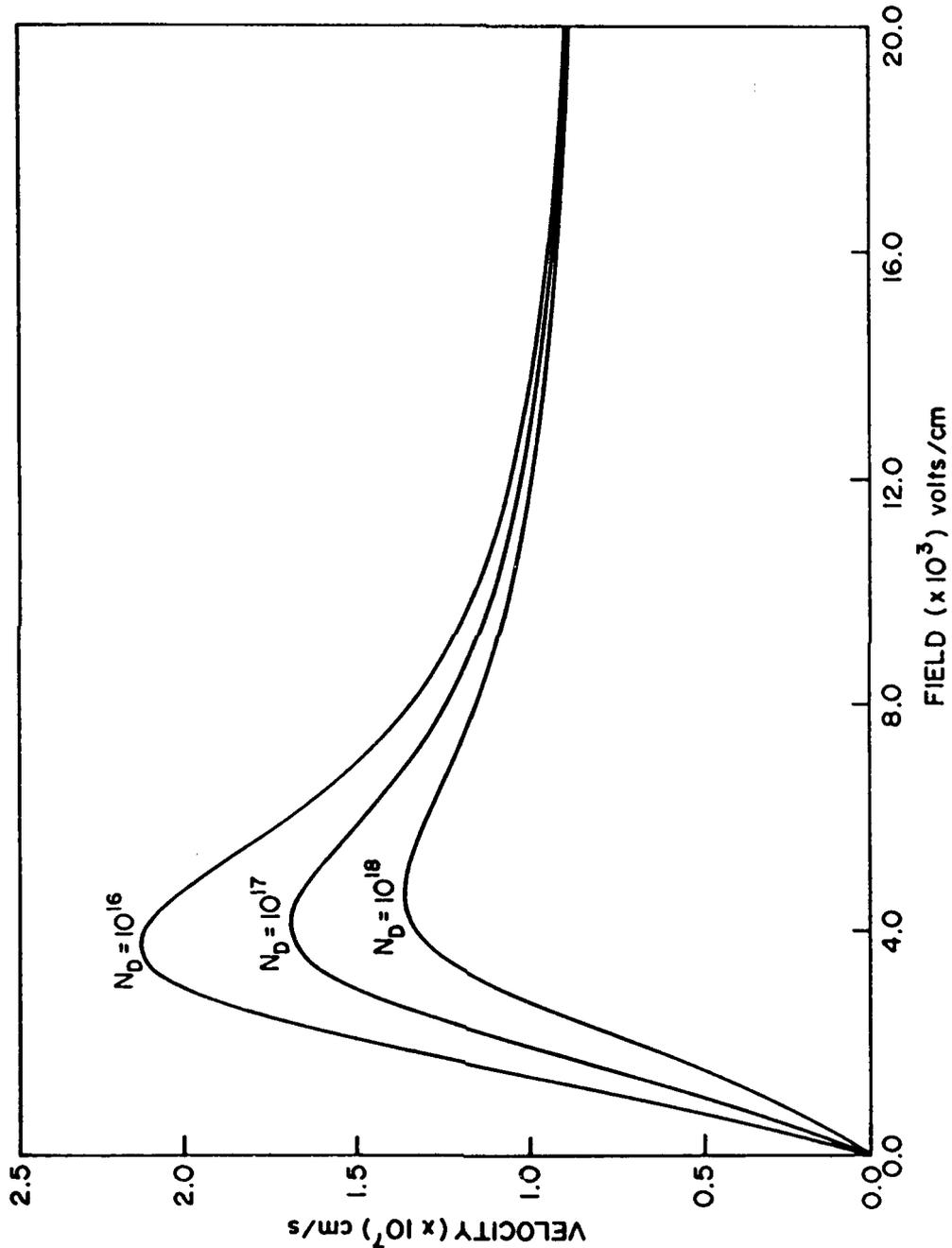
Fig. 3 - Velocity-field curve used in this study.
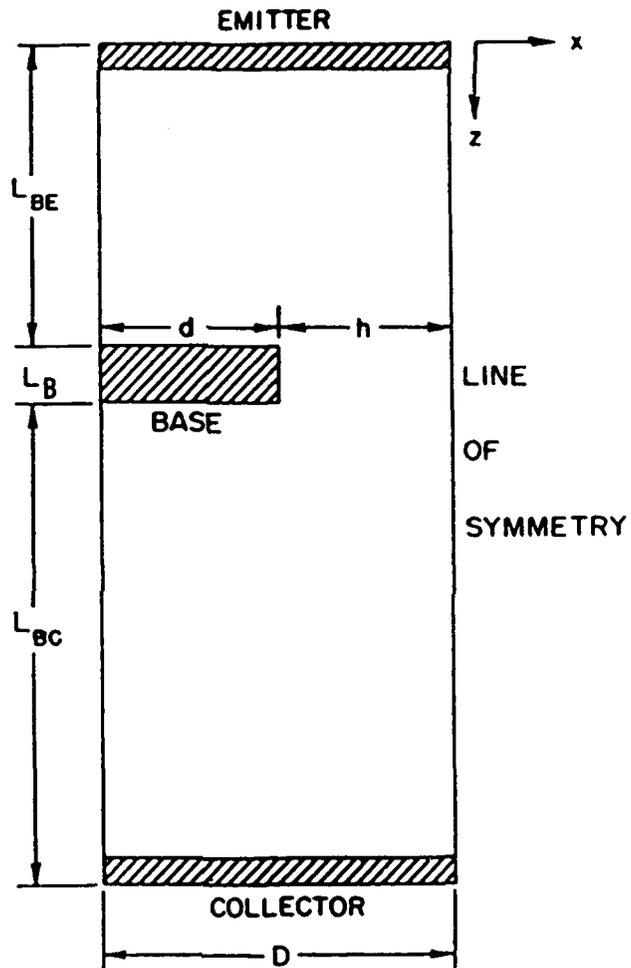
Fig. 4 - Basic device as proposed by Lincoln Laboratory.

EMITTER

0.4 μm

0.02 μm

10^{16} N

0.6 μm

COLLECTOR

0.2 μm

x_r = 1.0 μm

ψ_r = 1.0 volt

N_r = 10^{16}/cm^3

V_{CE} = 1.0 volt

V_{BE} = 0.35 volt
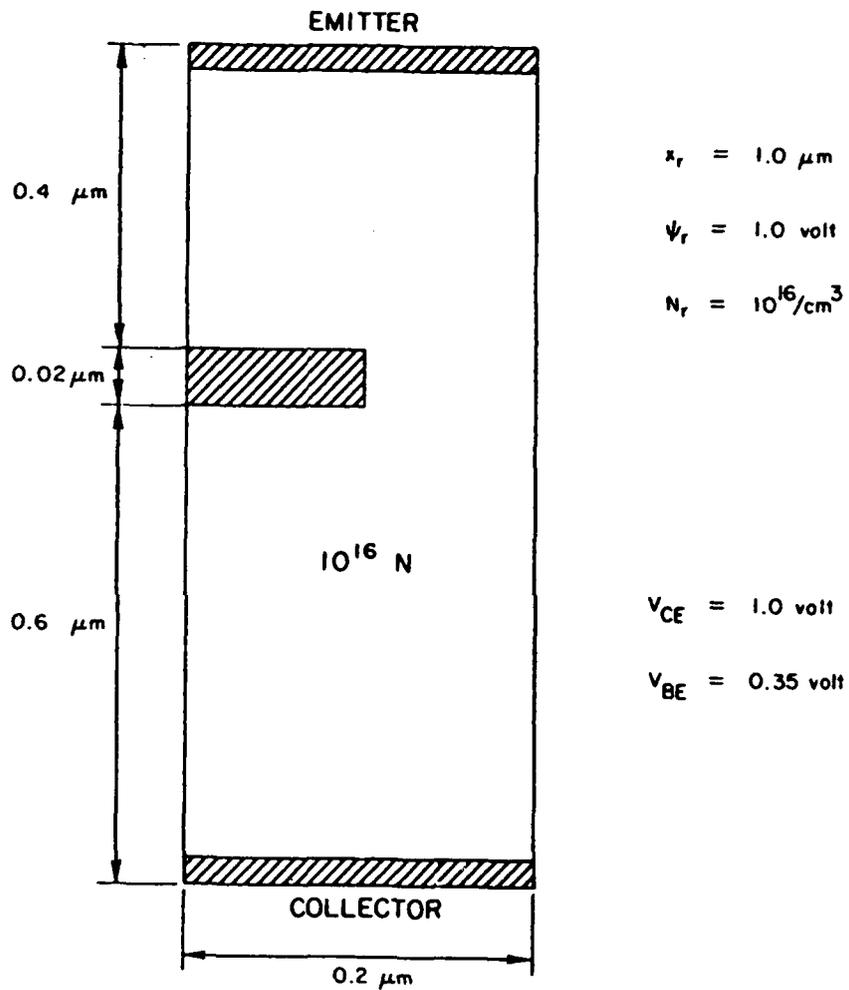
Fig. 5 – Reference quantities and dimensionless parameters for the computation referred to in Fig. 4.
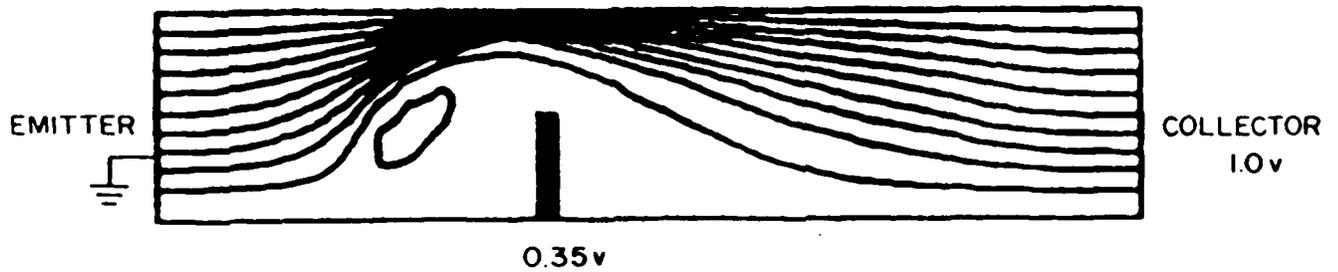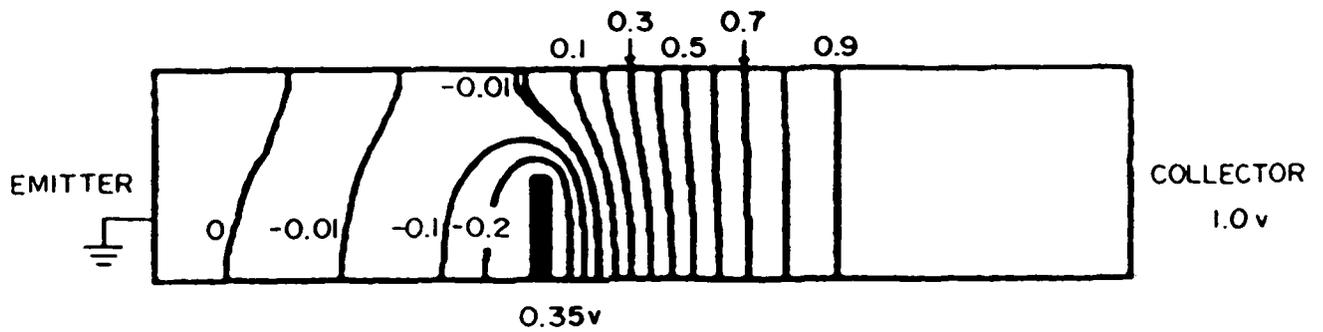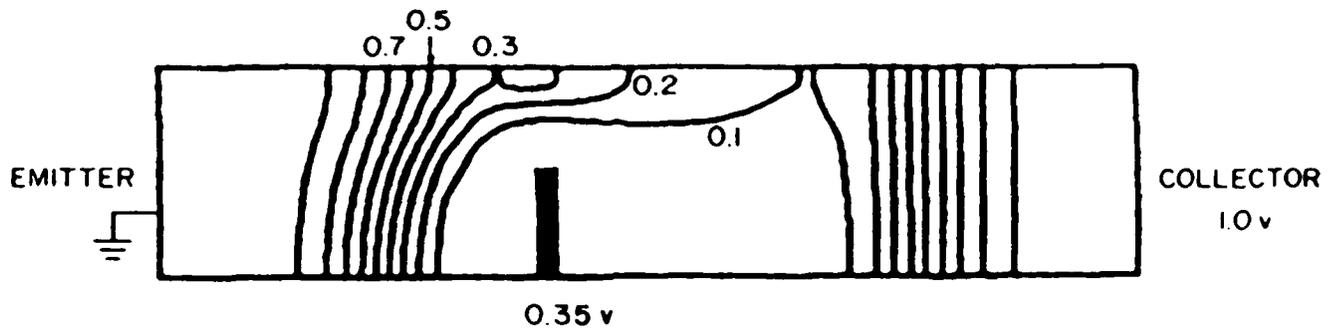
Fig. 6a –   Electron density, potential and streamline
to 6c   contours for the basic unscaled device.
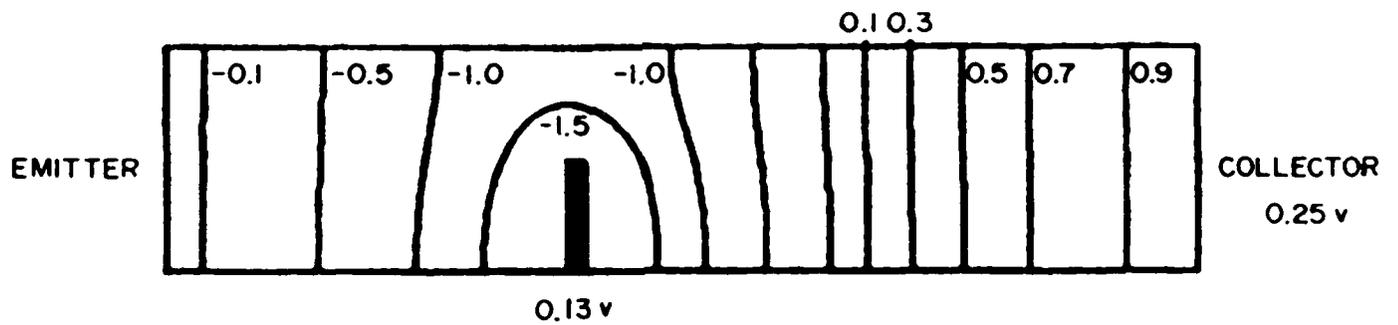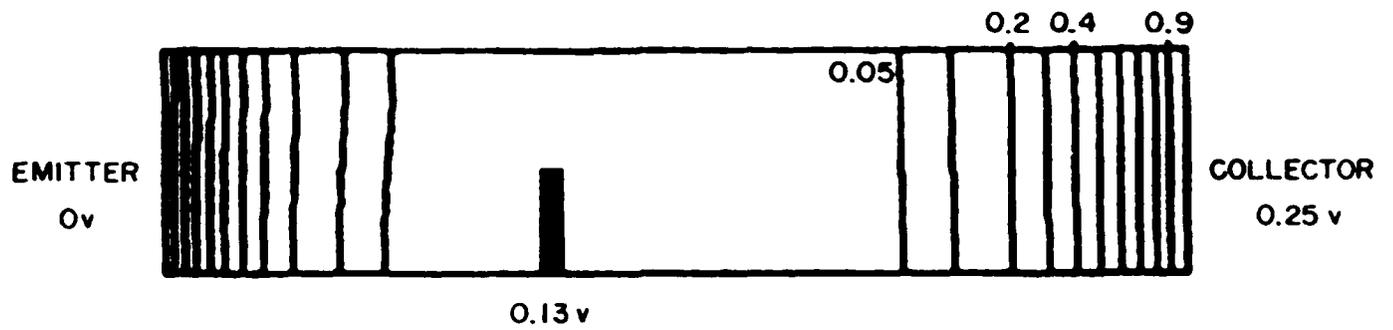
Fig. 7a – Electron density and potential contours for the
and 7b   constant field scaled device at V_{BE} = 0.13 v.

40
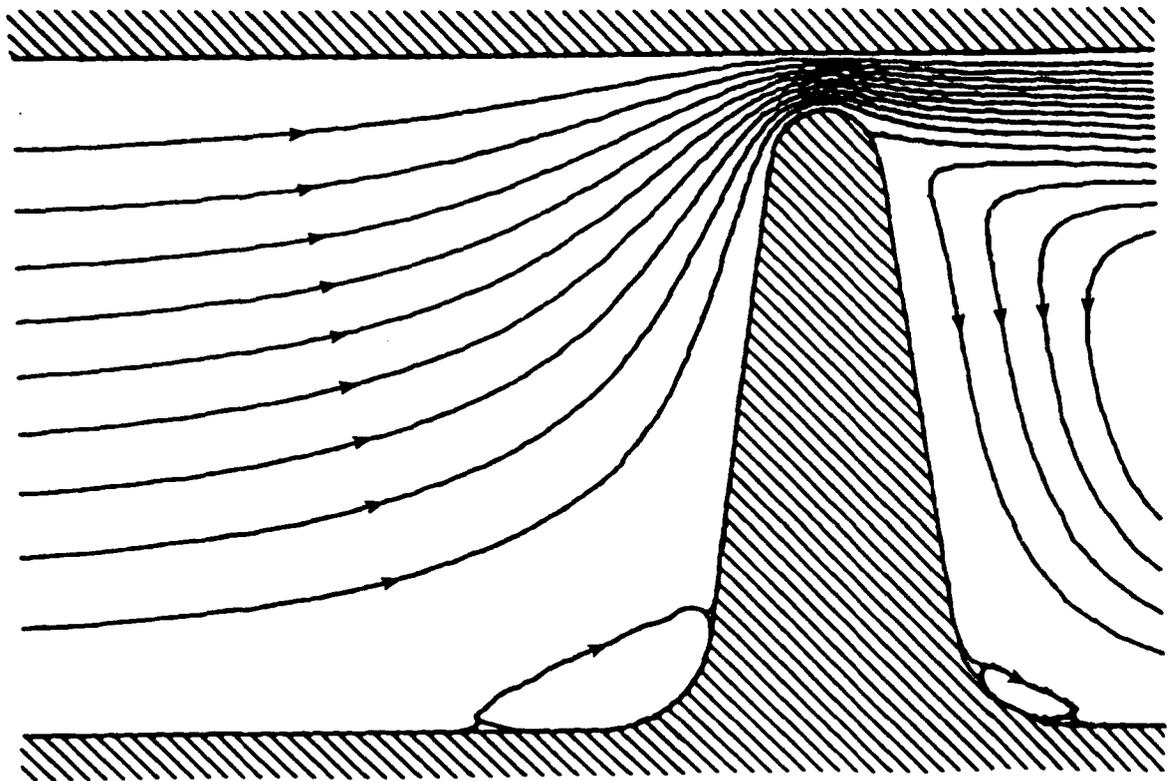
Fig. 8 — Fluid streamlines in a channel with
an obstruction. Embedded base is an
obstruction as well as a Schottky Contact.
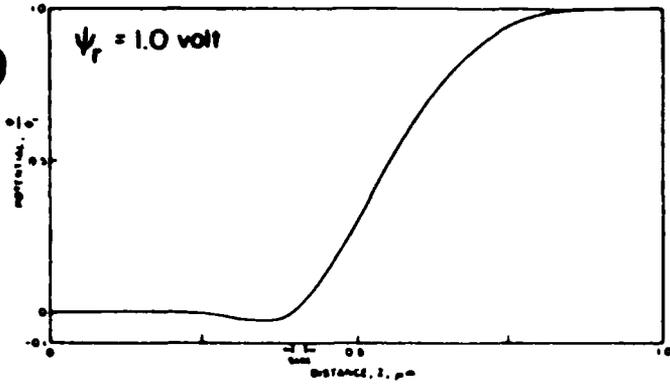
(9a)
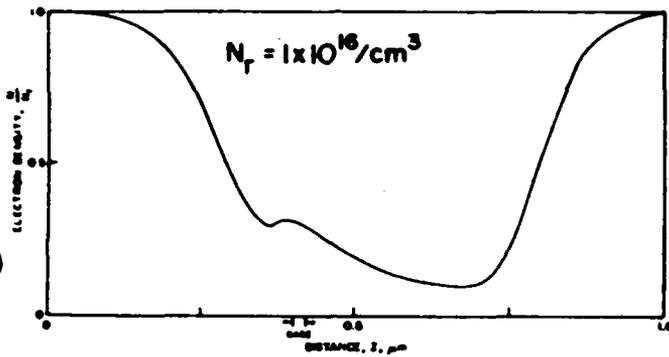
(9b)
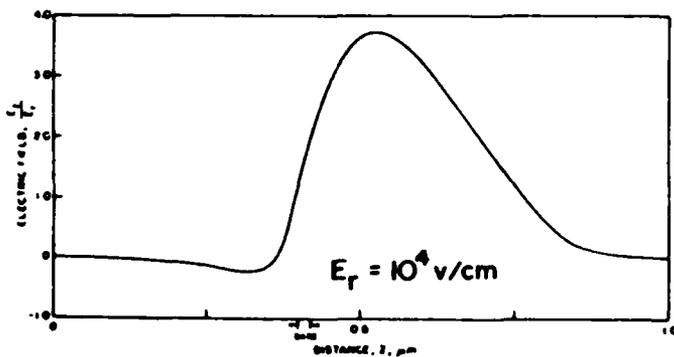
(9c)
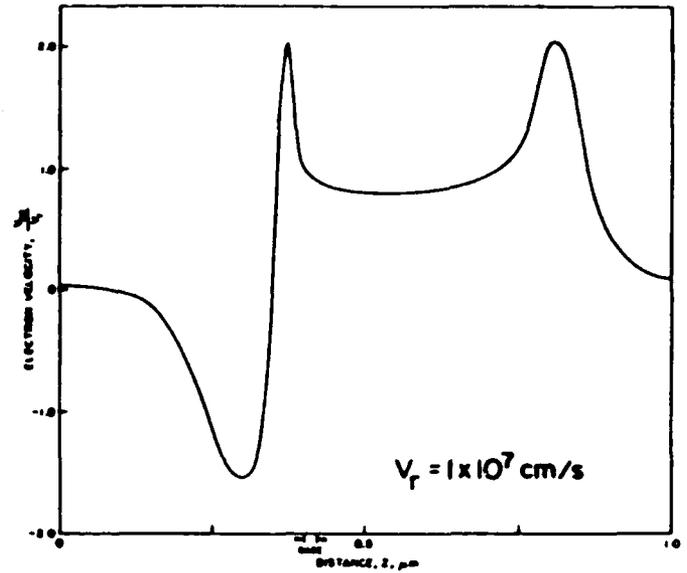
(9d)
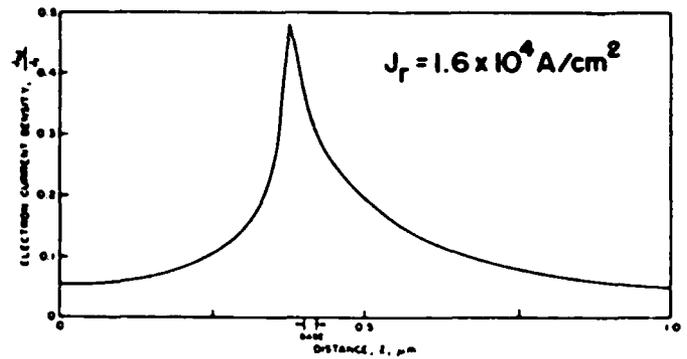
(9e)

Fig. 9a — Variation of potential, electron density, electric field,
to 9e   velocity and current density as a function of distance along
          the channel line of symmetry. Basic unscaled device.

42

(10a)

(10b)

(10c)
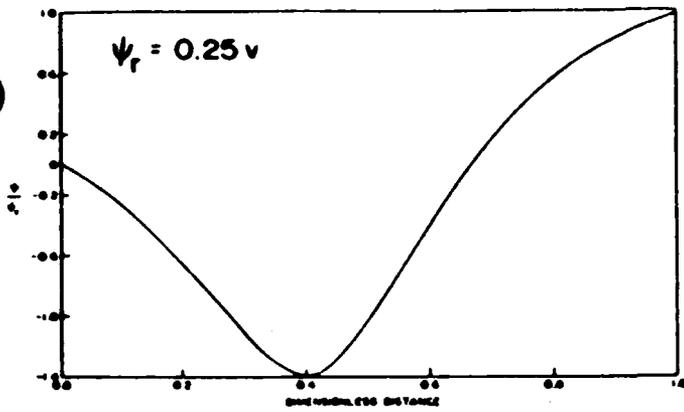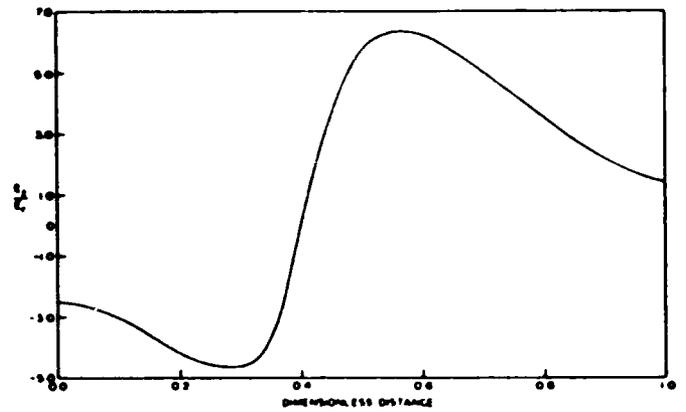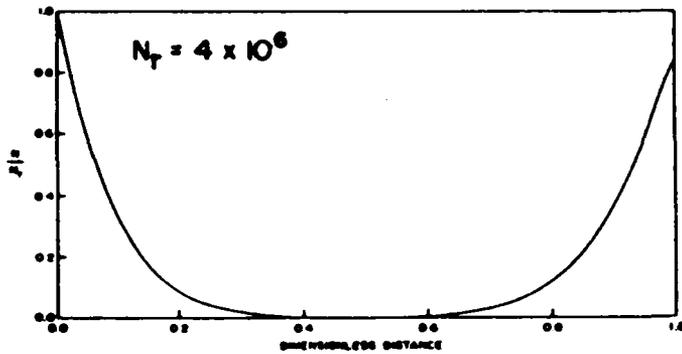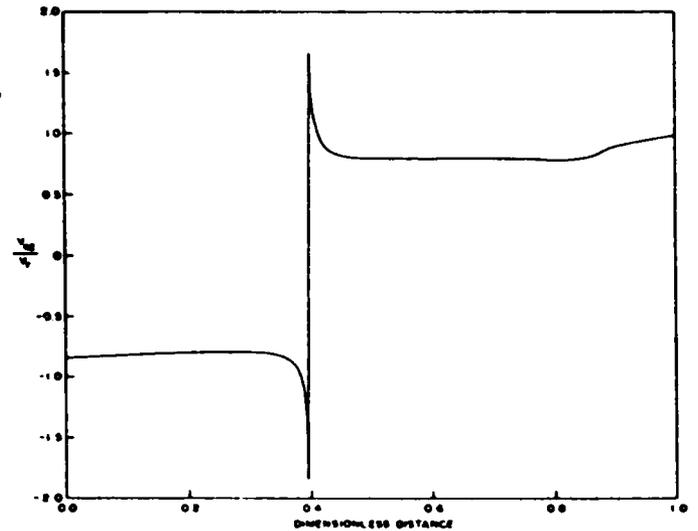
(10d)

Fig. 10a – Variation of potential, electron density, electric field and
to 10d   velocity at $V_{BE} = 0.13$ v as a function of distance along the
channel line of symmetry: Constant field scaled device.

43
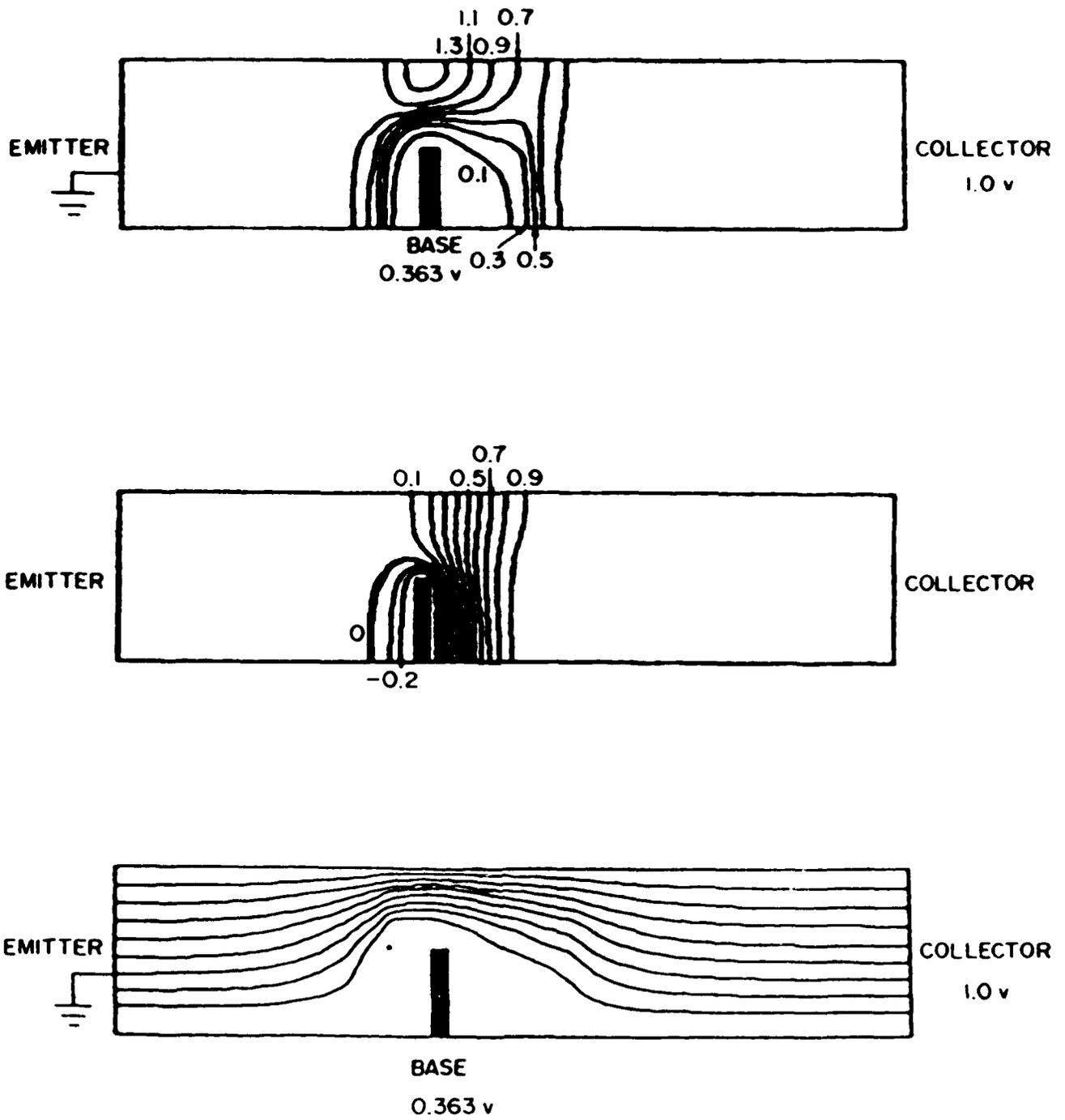
Fig. 11a - Contours of electron density, potential and current
     to 11c   streamlines for the depletion layer scaled device.

(12a)

(12b)

(12c)

(12d)

(12e)
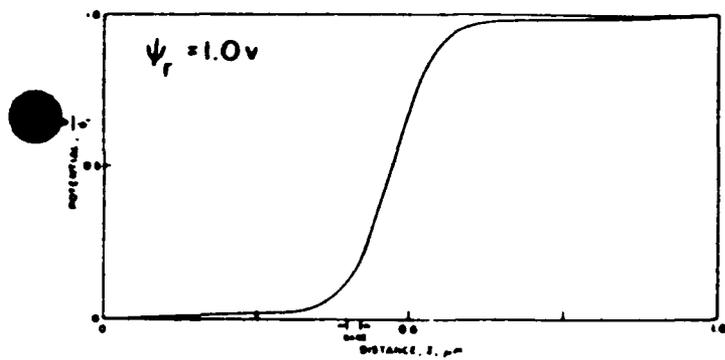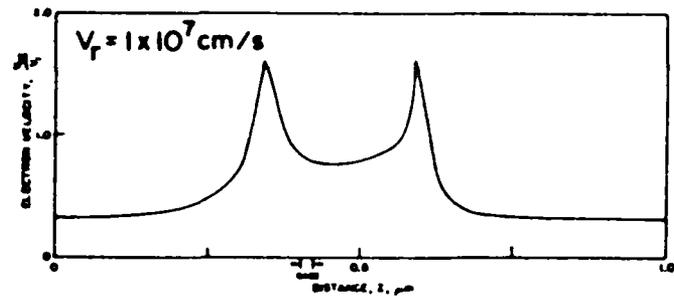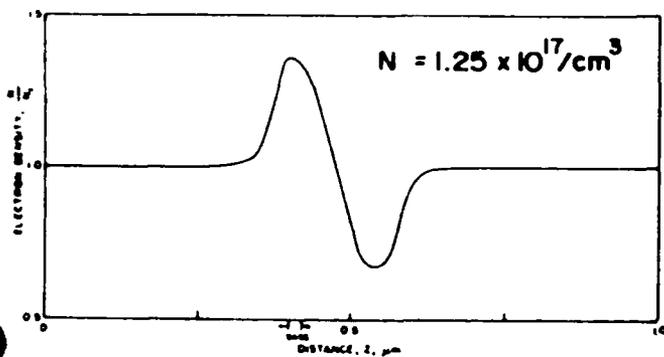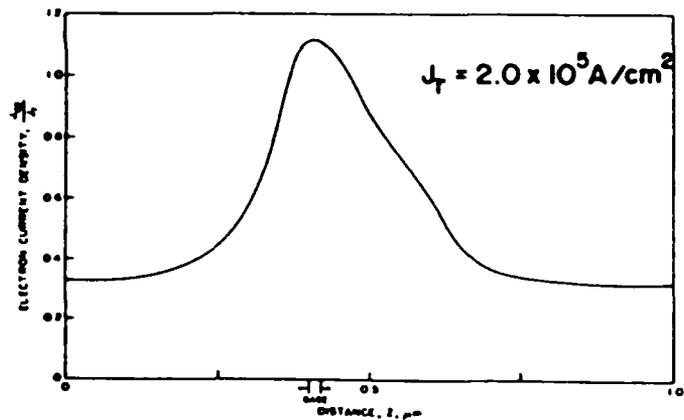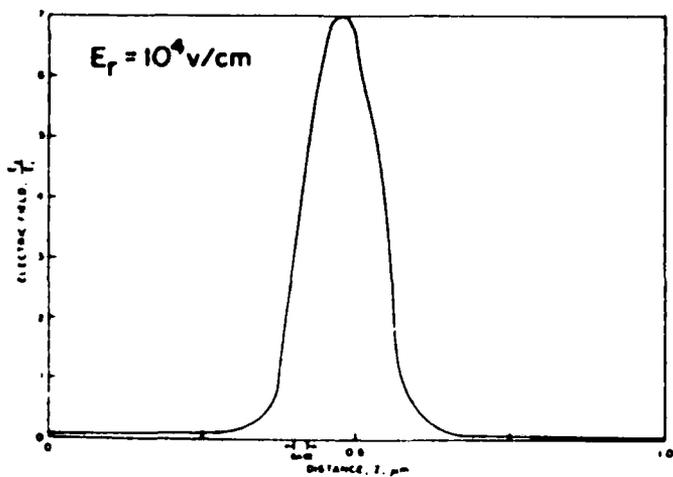
Fig. 12a — Variation of potential, electron density, electric field and
to 12e   velocity at V_{BE} = 0.42 v as a function of distance along the
channel line of symmetry. Depletion layer scaled device.

45
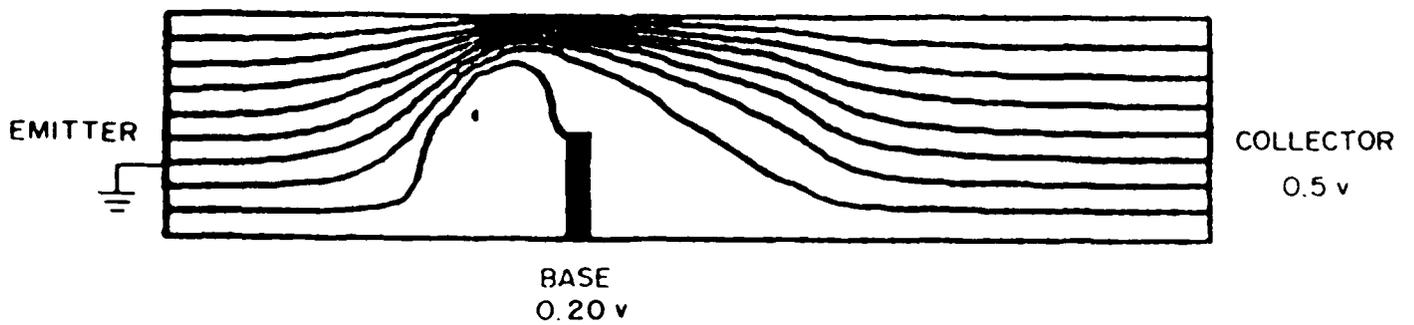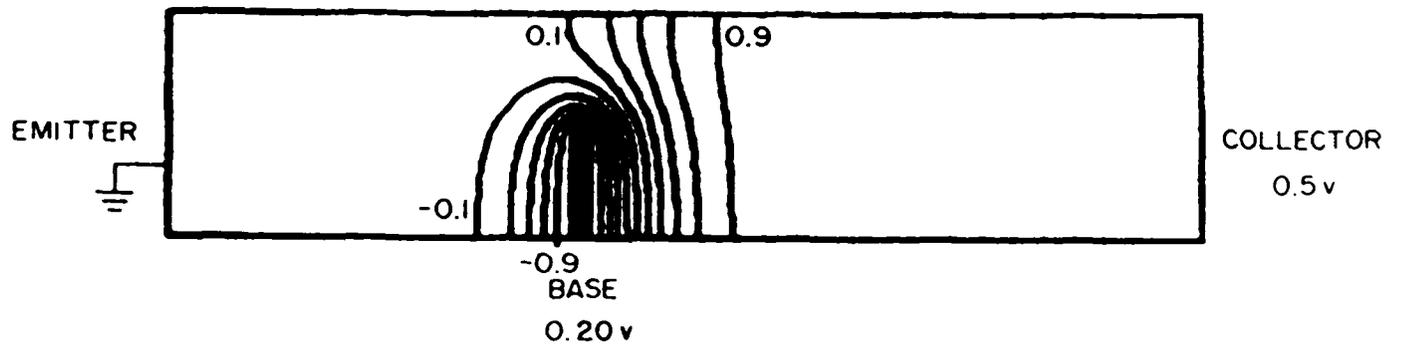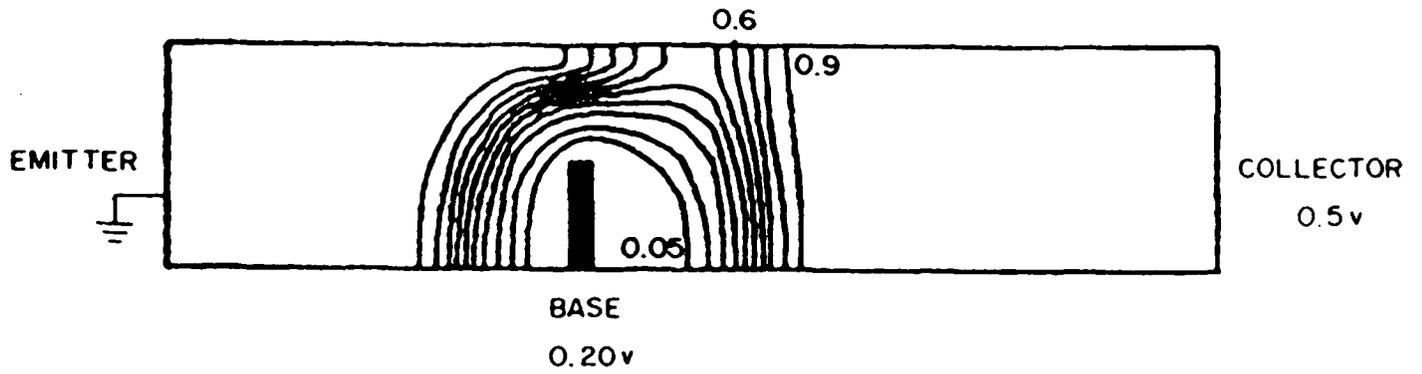
Fig. 13a – Contours of electron density, potential and streamlines for the
to 13c   device with combined depletion layer and constant field scaling.

46
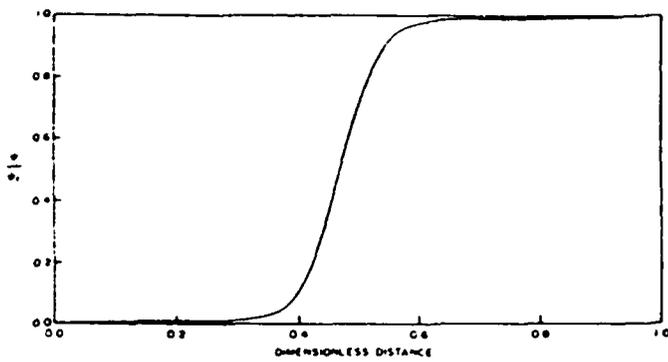
(14a)
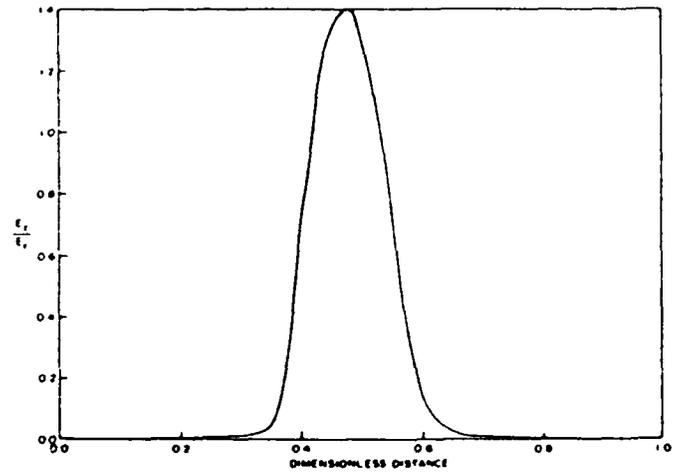
(14c)
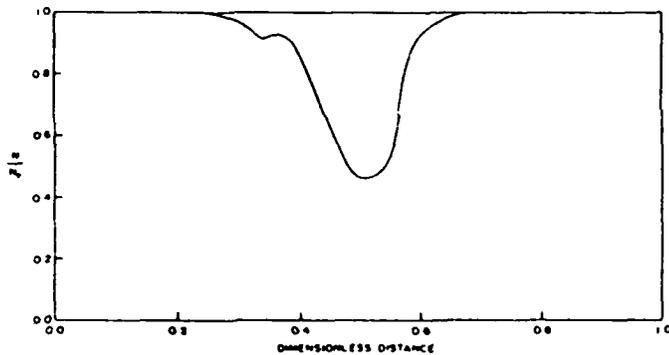
(14b)
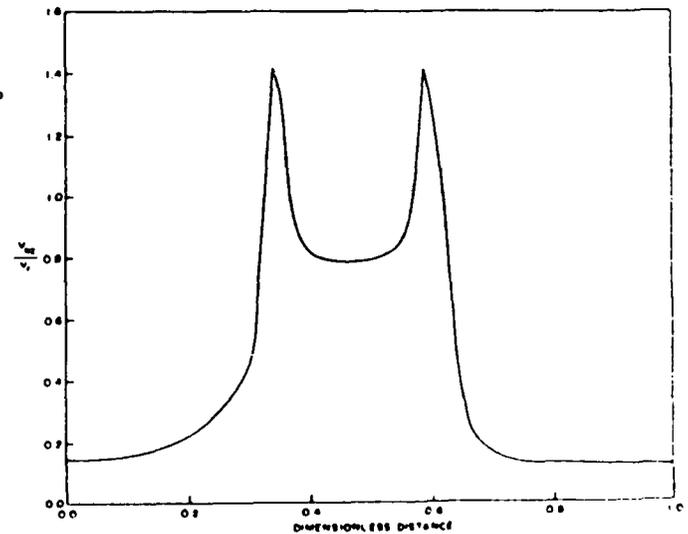
(14d)

Fig. 14a – Variation of potential, electron density, electric field
to 14d   and velocity at $V_{BE} = 0.20$ v as a function of distance
          along the channel line of symmetry. Case with combined
          depletion layer and constant field scaled device.
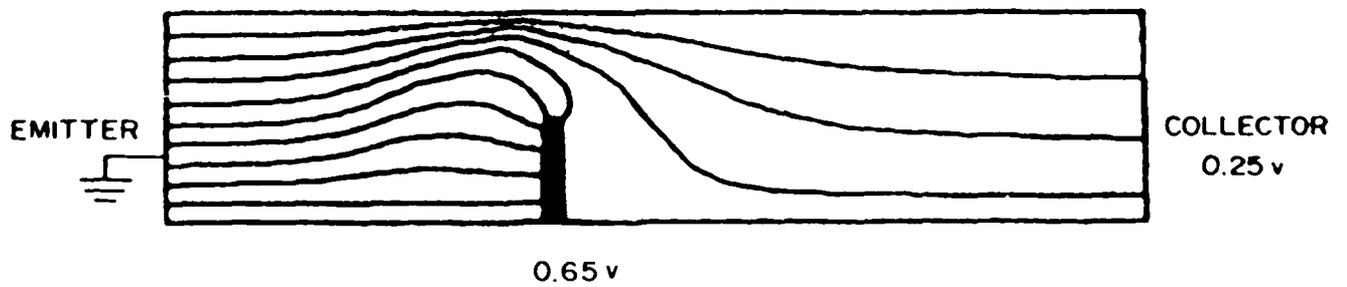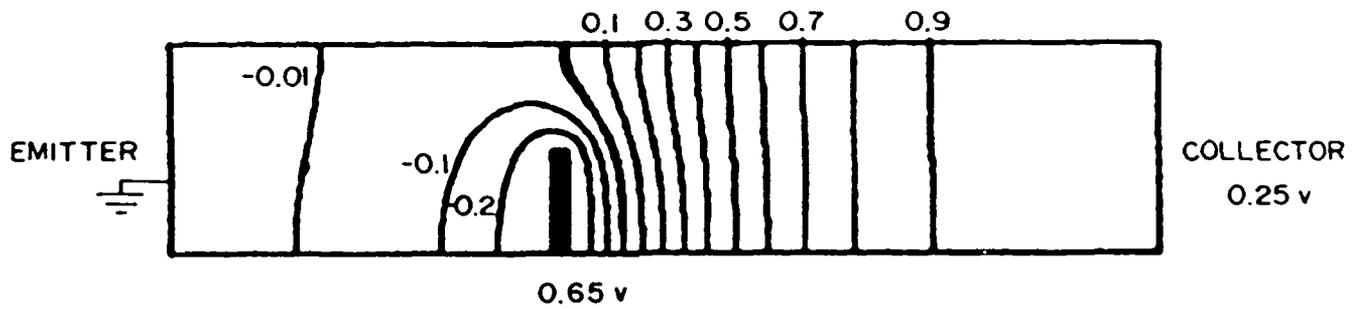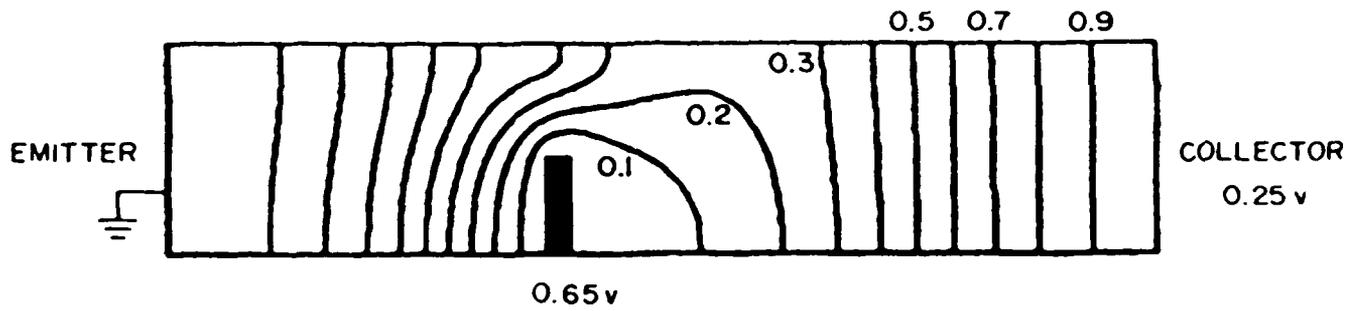
Fig. 15a to 15c – Contours of density, potential and current streamlines.
Case: Depletion layer scaled device at V_{BE} = 0.65 v.

48
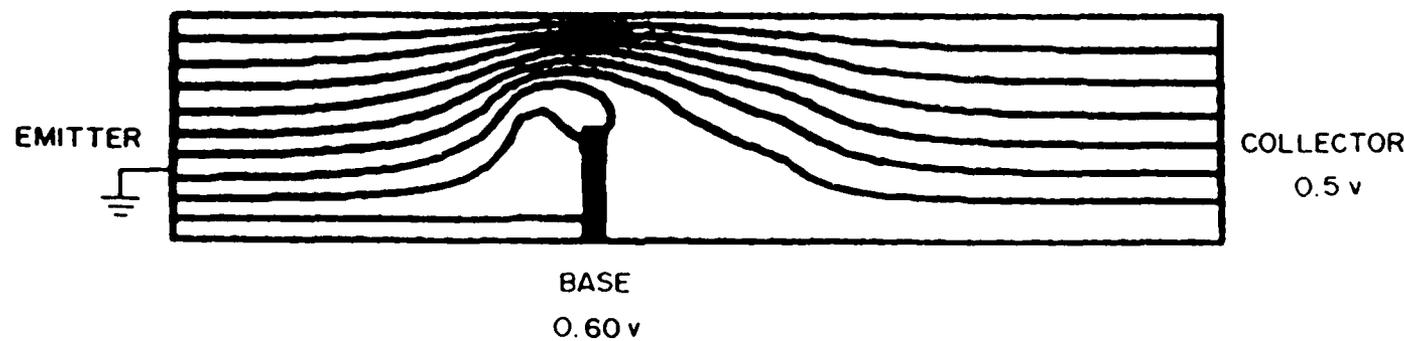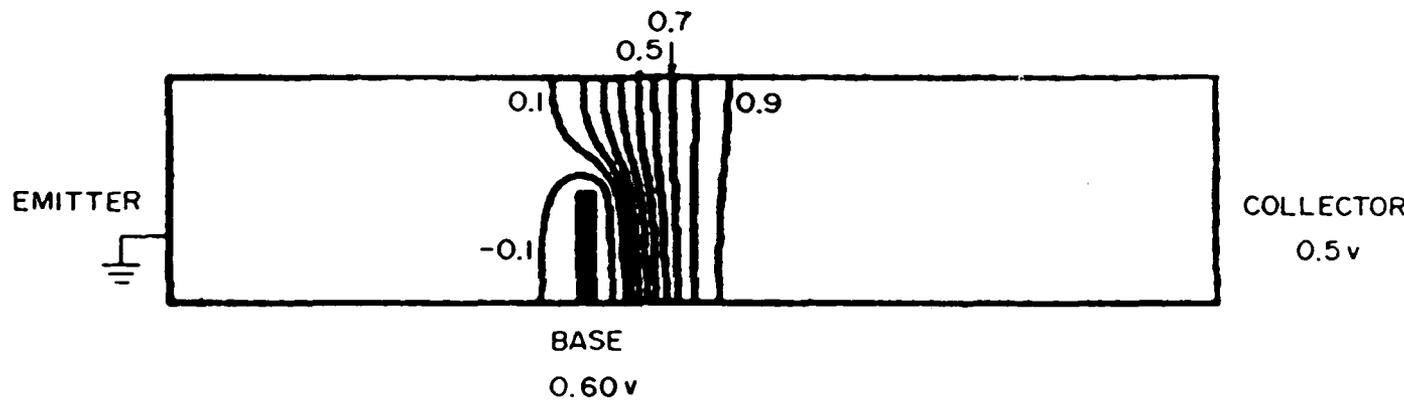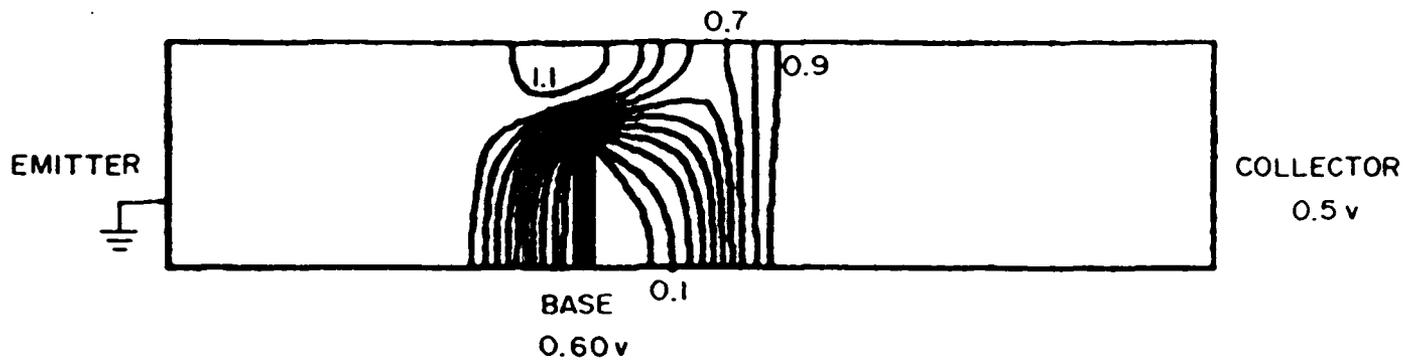
Fig. 16a – Contours of electron density, potential and current streamlines.
to 16c Case with combination of depletion layer and constant field scaling
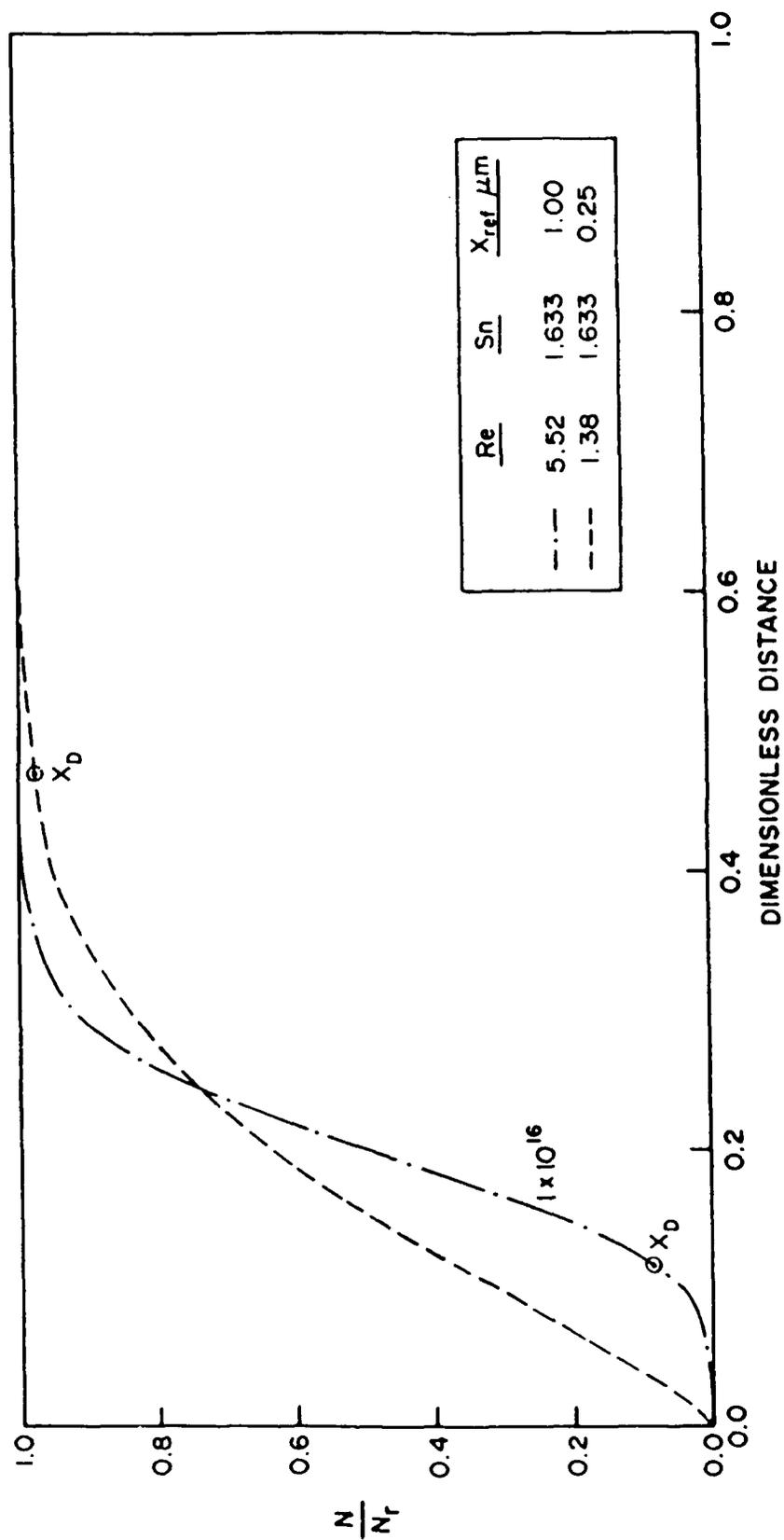
Fig. 17 - Variation of electron density with distance normal to a Schottky barrier.

EMITTER

0.4 μm

35 %

I

0.02 μm

III

II

48 %

0.6 μm

17 %

COLLECTOR

0.1 μm   0.1 μm
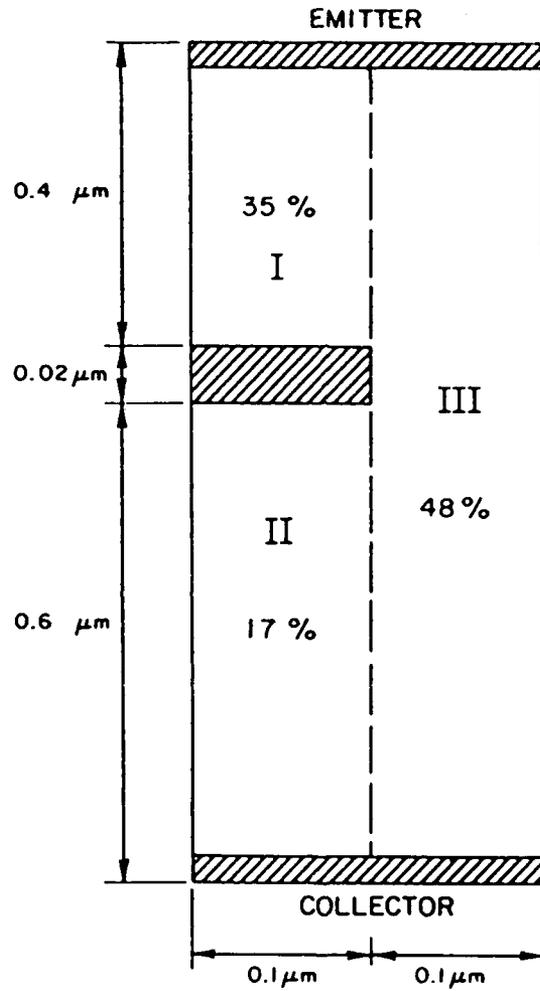
$N = 1.25 \times 10^{17}$

$f_T = 37 \text{ GHz}$

$V_{CE} = 1.0 \text{ volt}$

$V_{BE} = 0.36 \text{ volt}$

Fig. 18 – Capacitance distribution of the depletion-layer-scaled device.
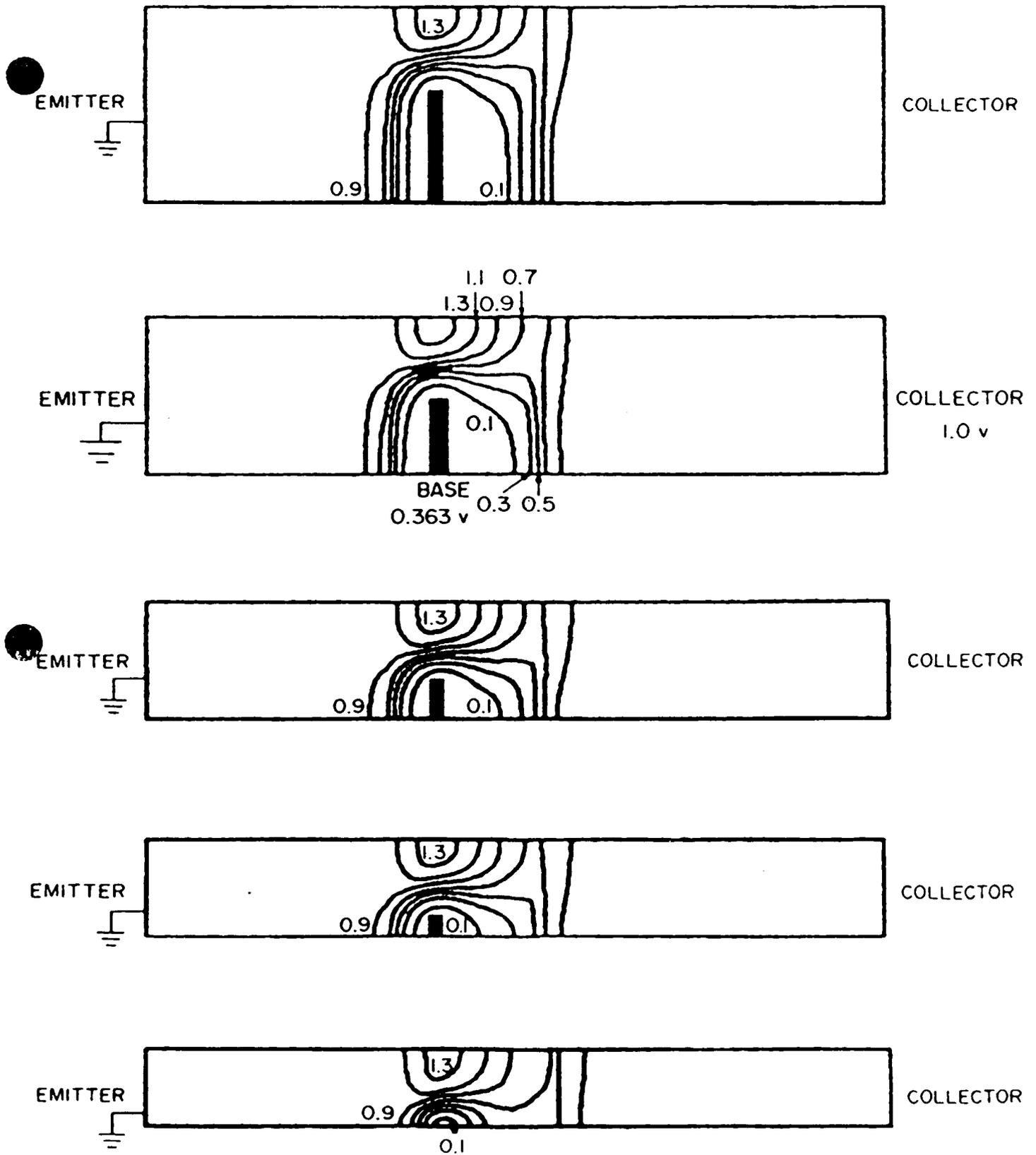
Fig. 19a - Comparison of electron density contours for
to 19e   cases d = 1500, 1000, 500, 250 and 0 Å.
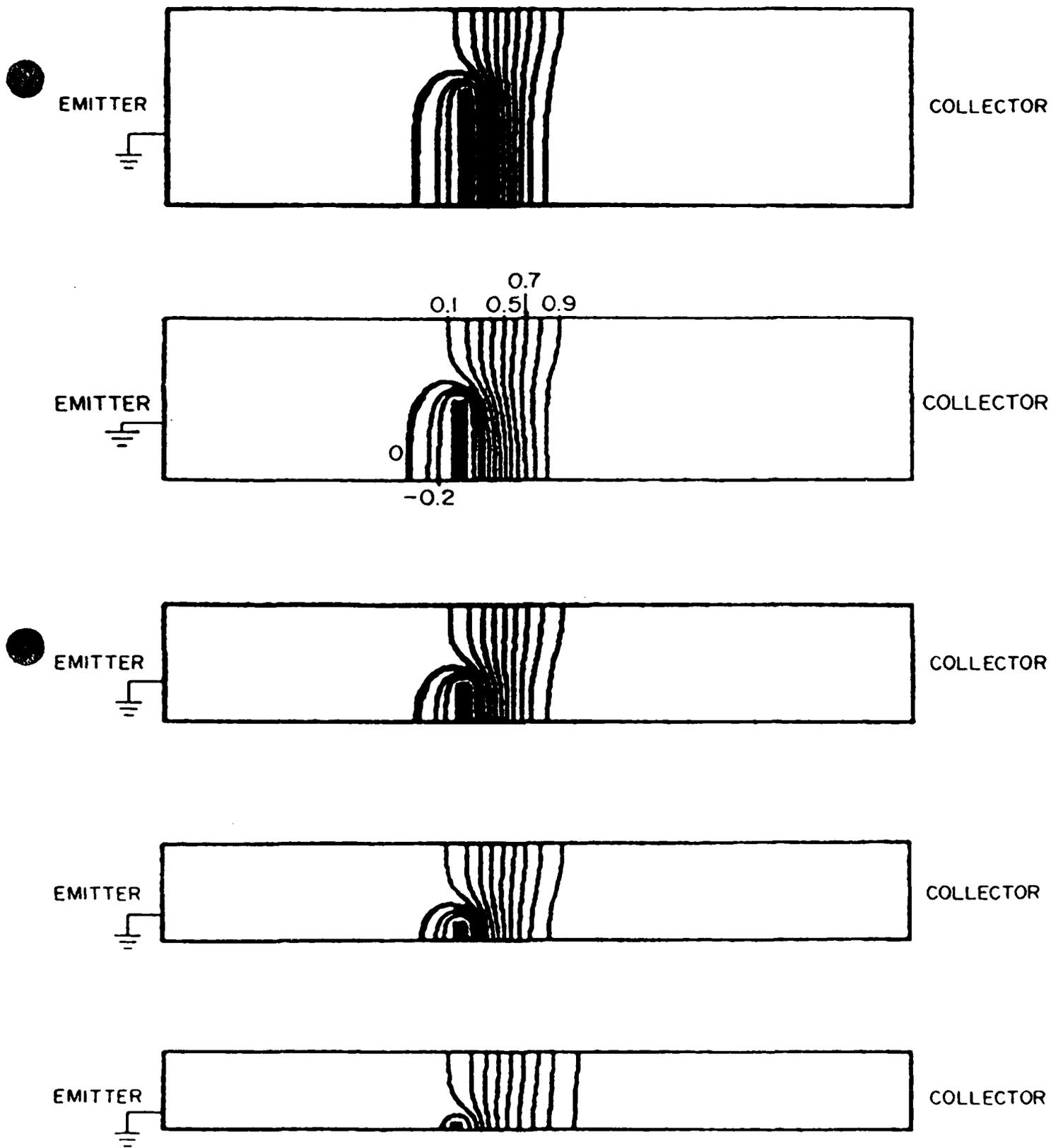
Fig. 20a – Comparison of potential contours for
to 20e   cases d = 1500, 1000, 500, 250 and 0 Å.
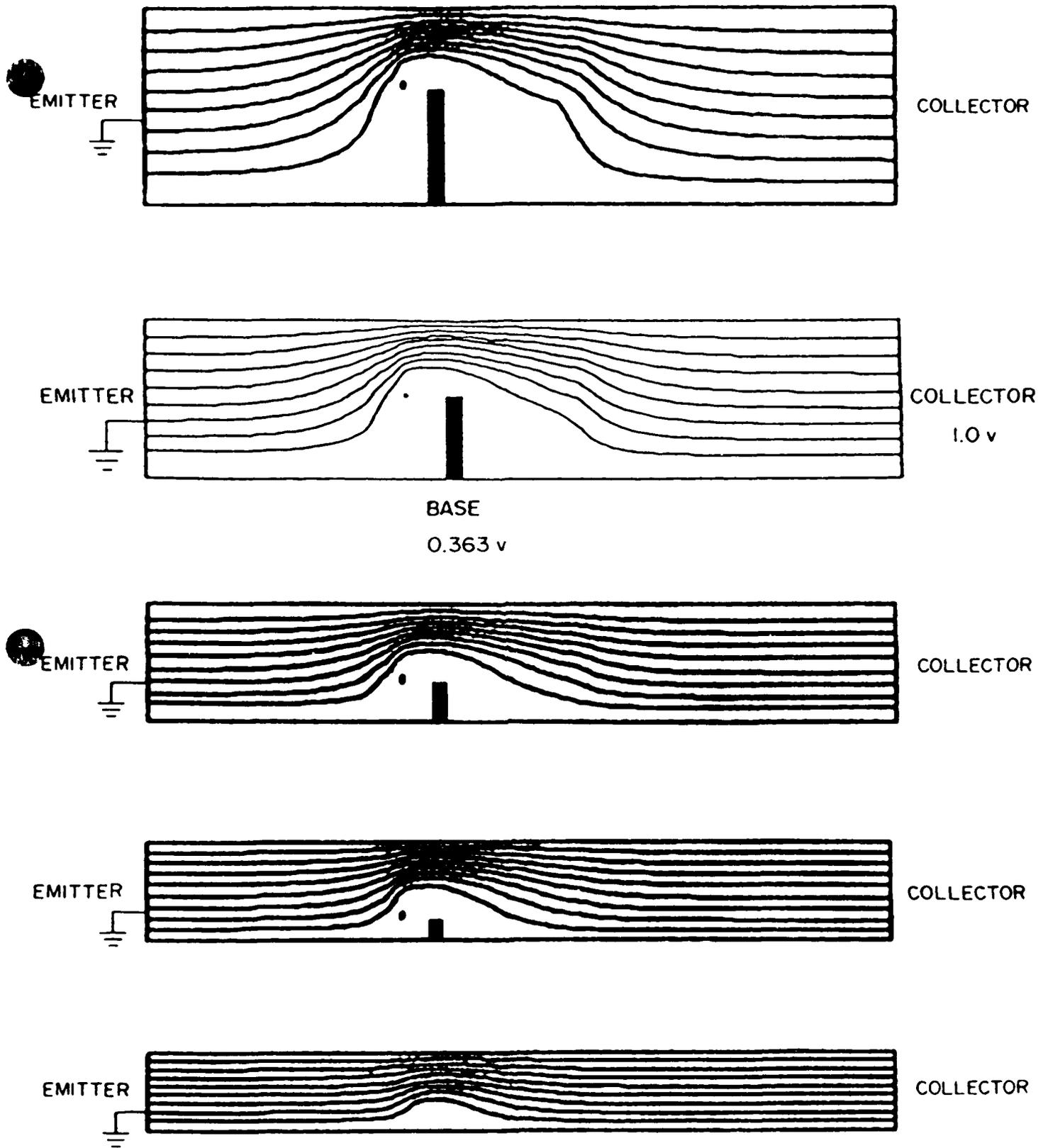
Fig. 21a – Comparison of streamline contours for
to 21e   cases d = 1500, 1000, 500, 250 and 0 Å.

54
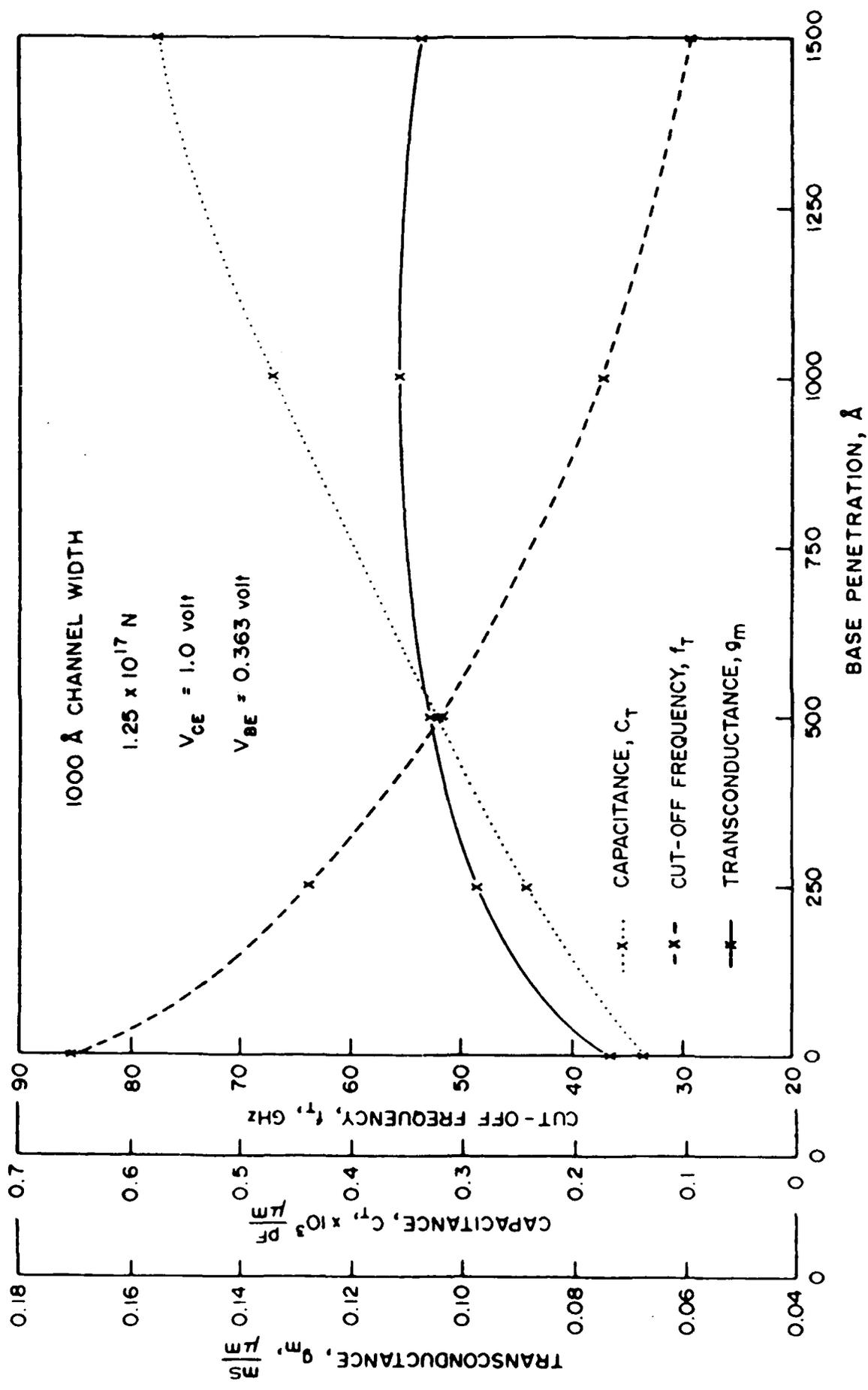
Fig. 22 - Variation of small signal parameters with base penetration.

55
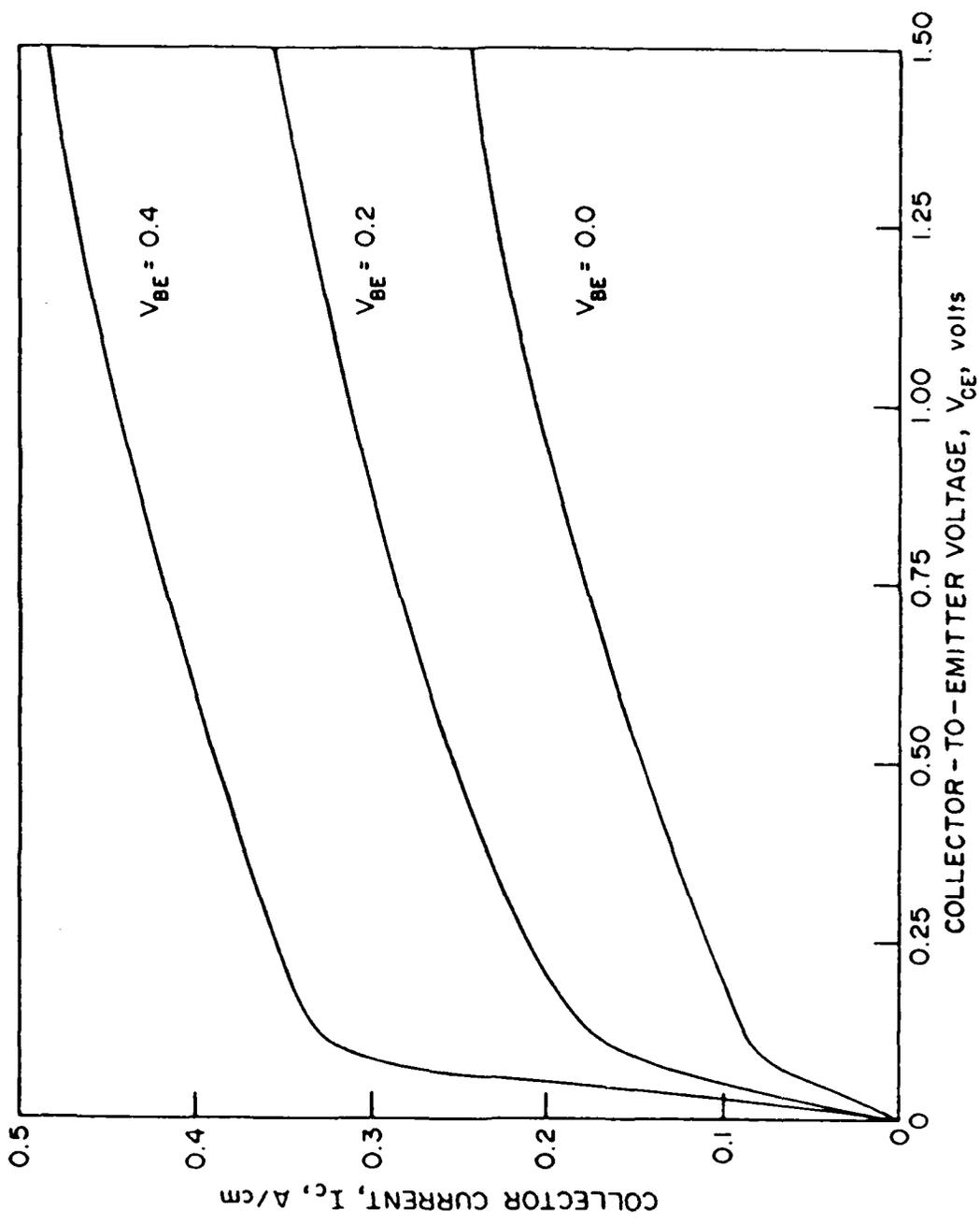
Fig. 23 - I-v characteristics for the device with parameters listed in Table 8.
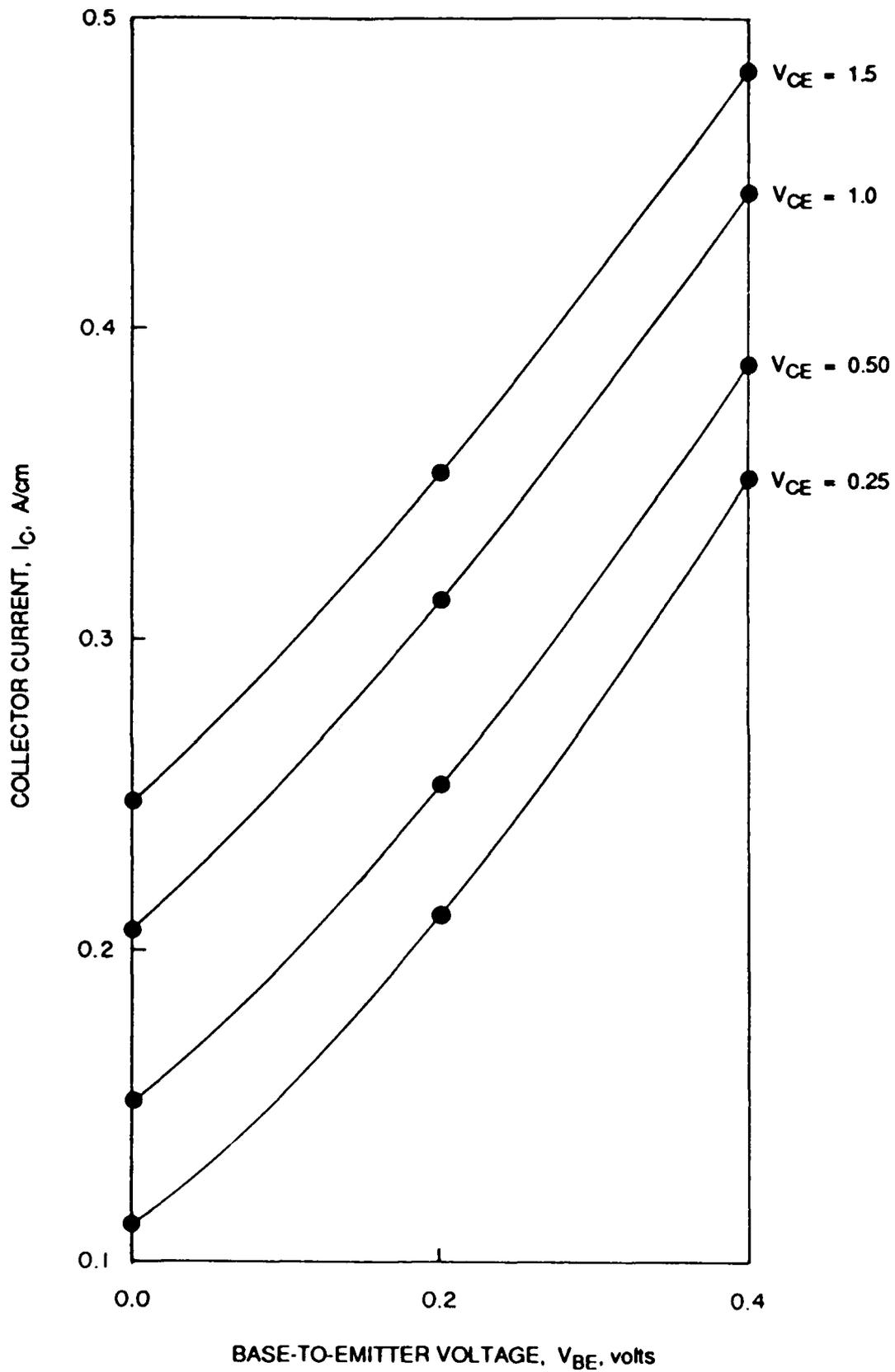
Fig. 24 - Collector current vs. V_{BE}.
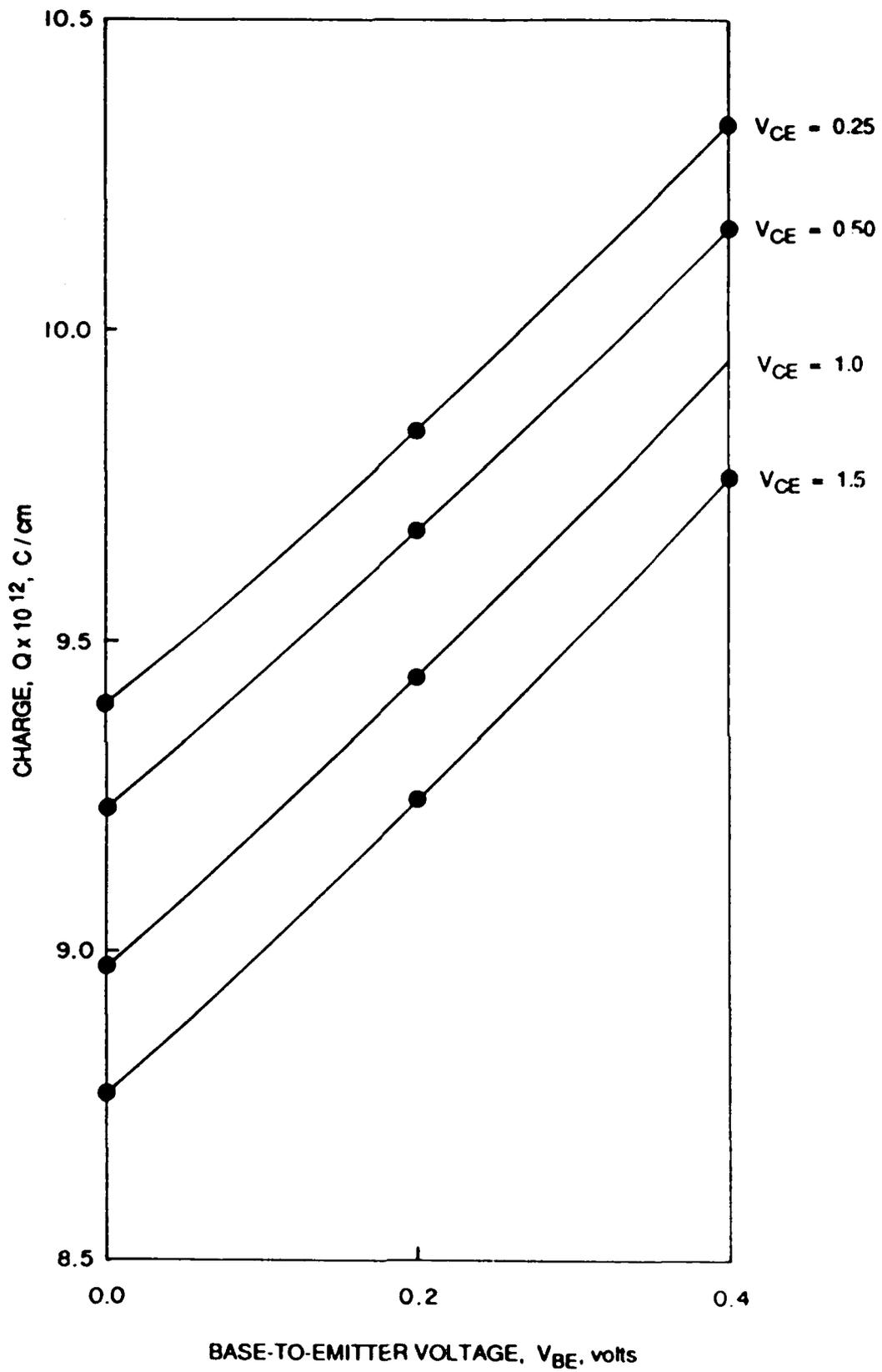
Fig. 25 - Charge vs. V_{BE}.

CHARGE, Q×10^{12}, C/cm
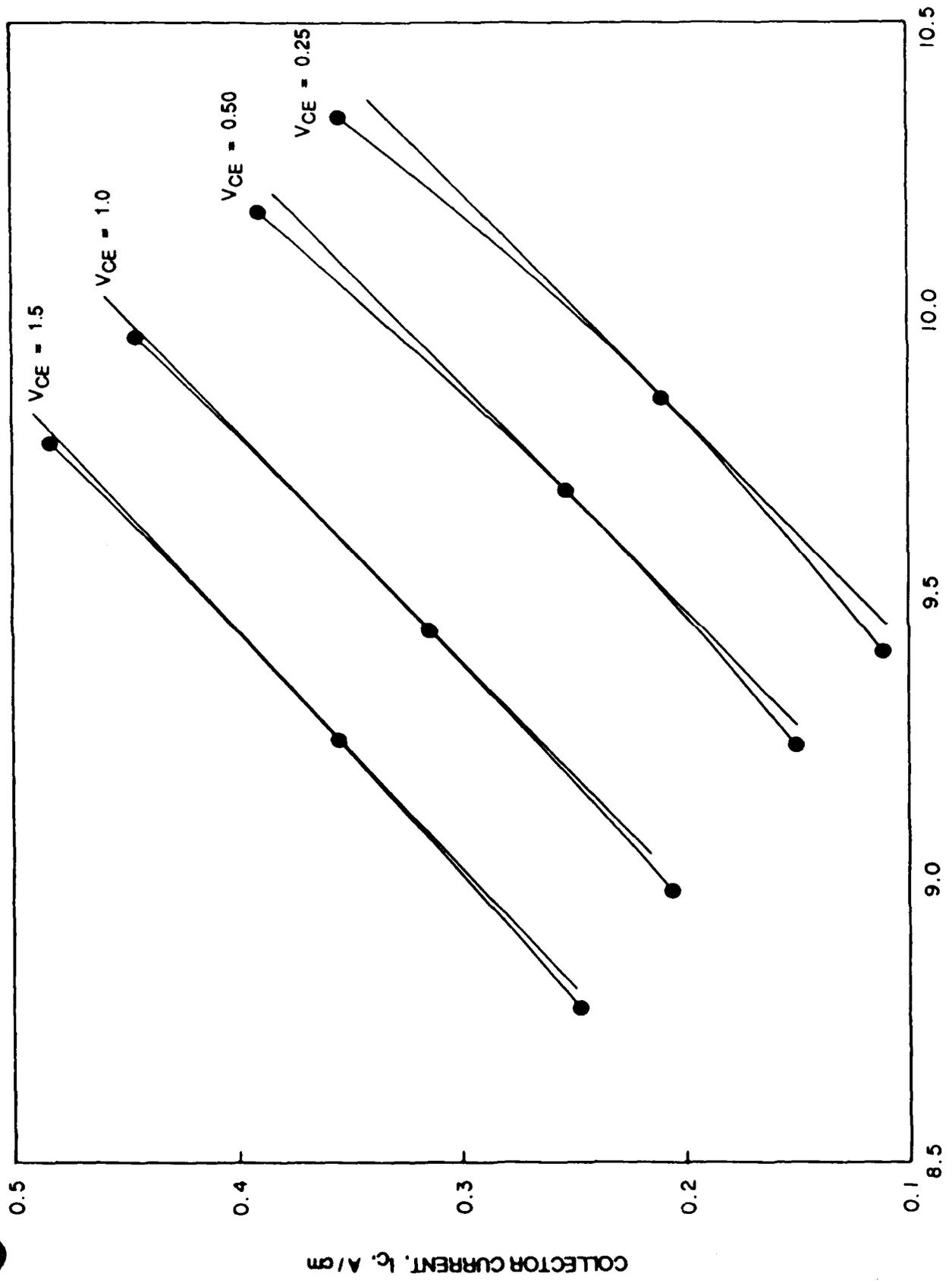
Fig. 26 – Collector current vs. charge.

COLLECTOR CURRENT, I$_C$, A/cm

V_{CE} = 0.25

V_{CE} = 0.50

V_{CE} = 1.0

V_{CE} = 1.5

d = 500 Å
h = 1000 Å
N_r = 5 x 10^{16}
V_{CE} = 1.0 volt
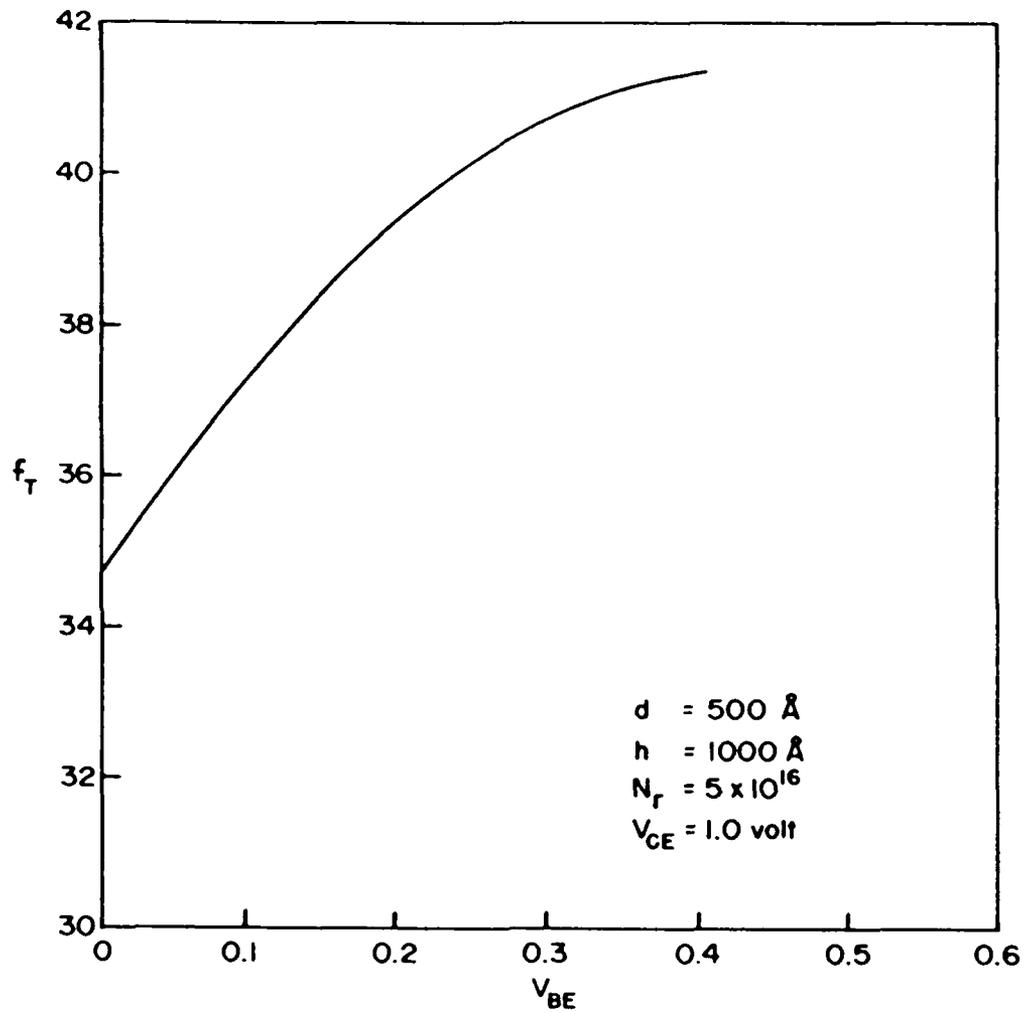
(27a)

Fig. 27a to 27c - Effect of base potential on f_T.

Graph with y-axis labeled f_T ranging from 24 to 36, and x-axis labeled V_{BE} ranging from 0 to 0.6.

$$d = 1000 \text{ Å}$$
$$h = 1000 \text{ Å}$$
$$N_r = 5 \times 10^{16}$$
$$V_{CE} = 1.0 \text{ volt}$$

(27b)

d = 1000 Å
h = 1000 Å
N_r = 1 x 10^{16}
V_{CE} = 1.0 volt

(27c)

62
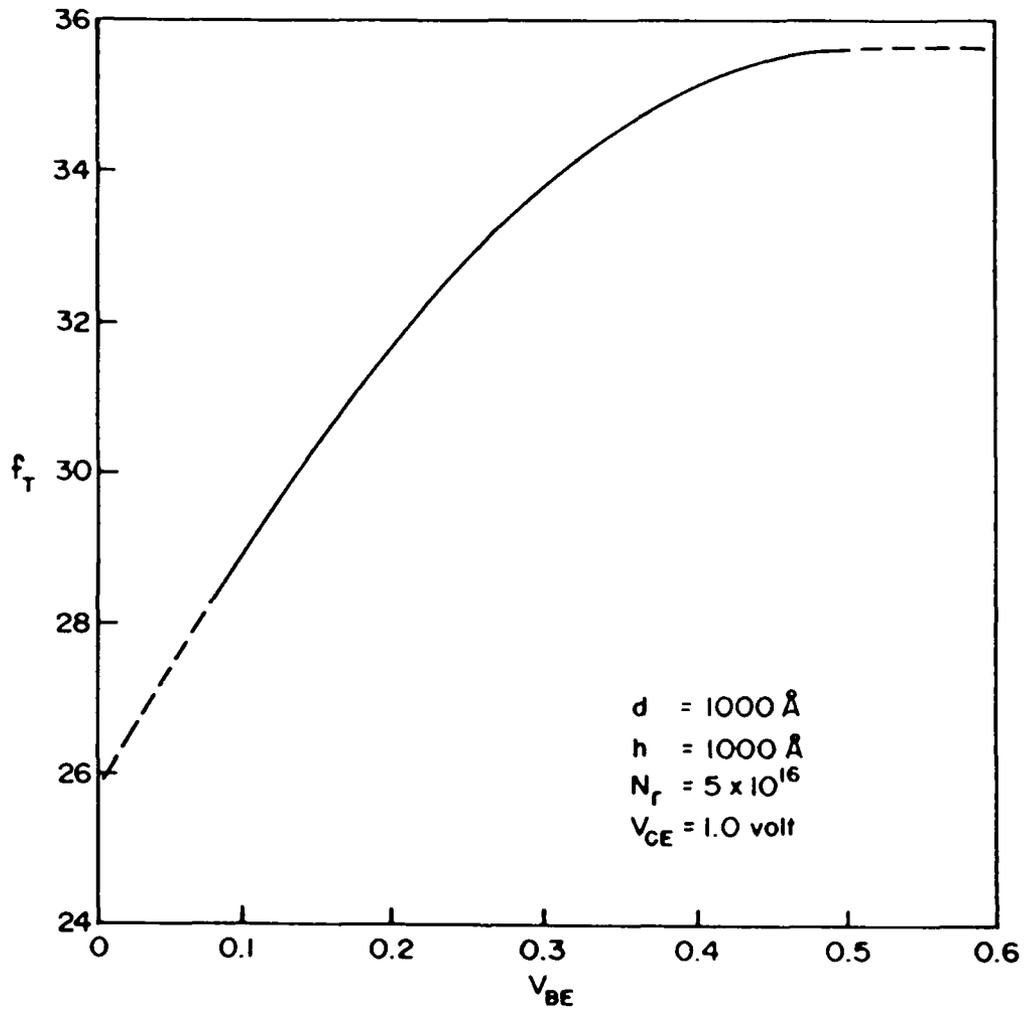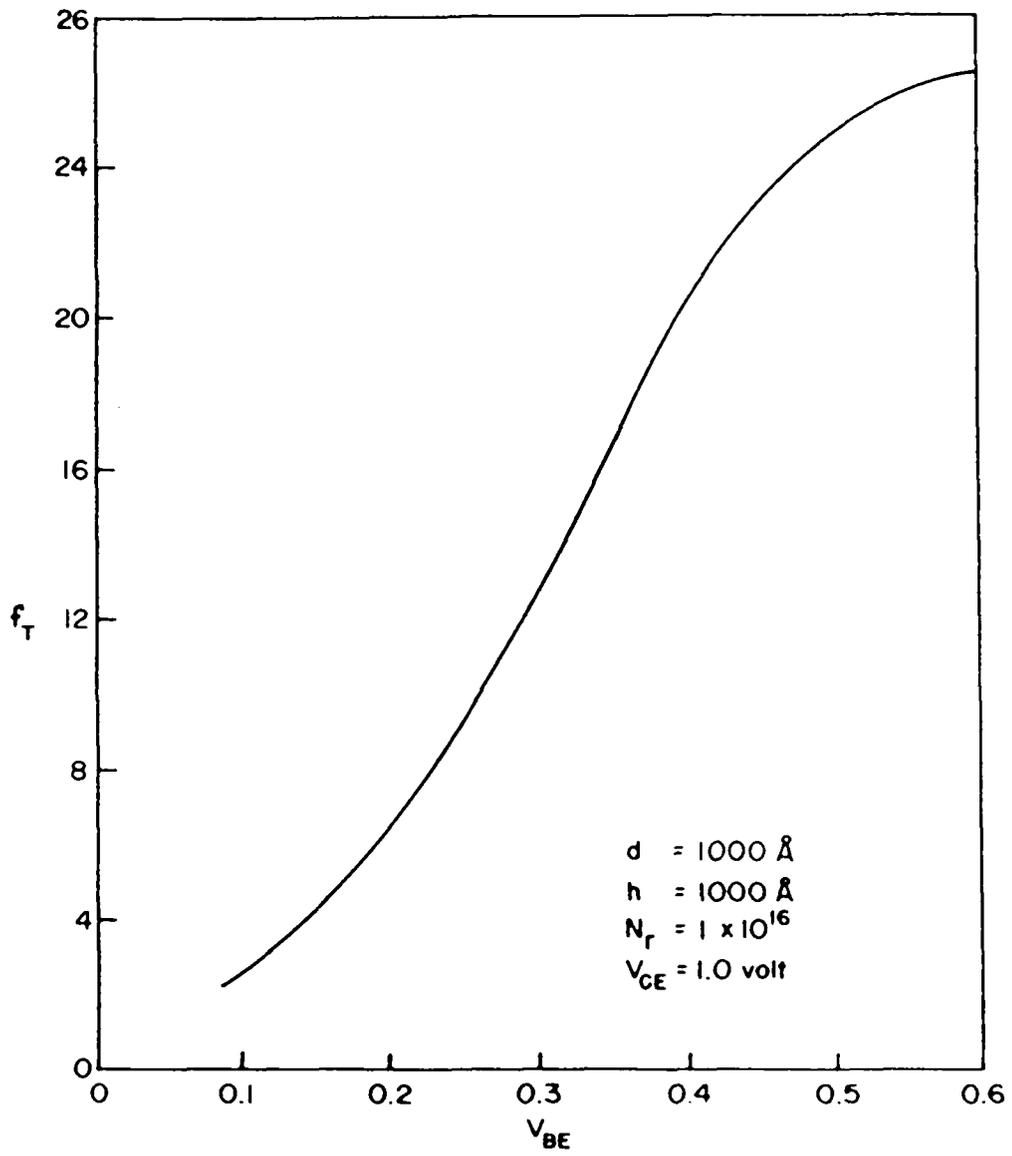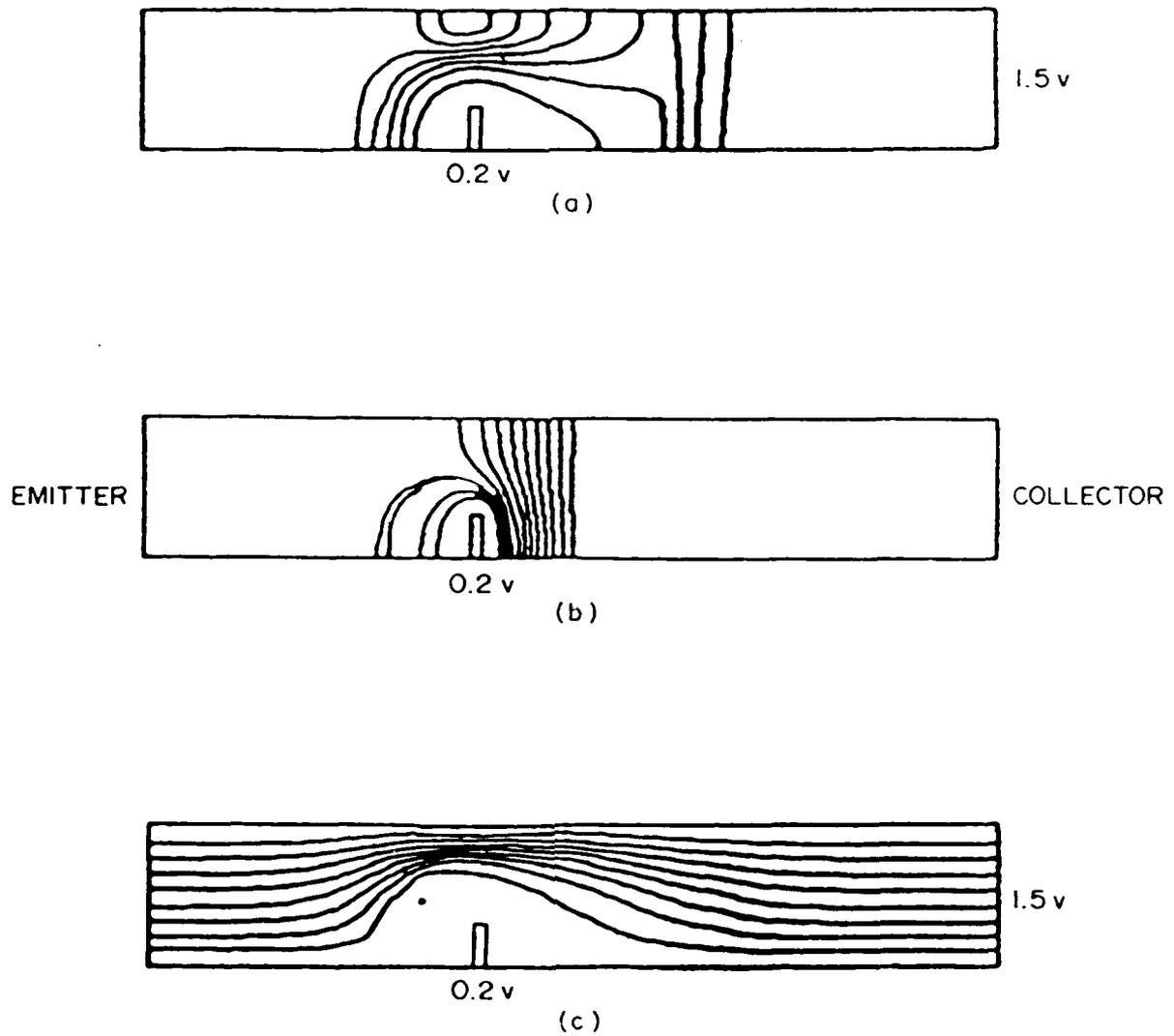
Fig. 28 – Electron density, potential and streamline contours
for the device corresponding to figure 23.
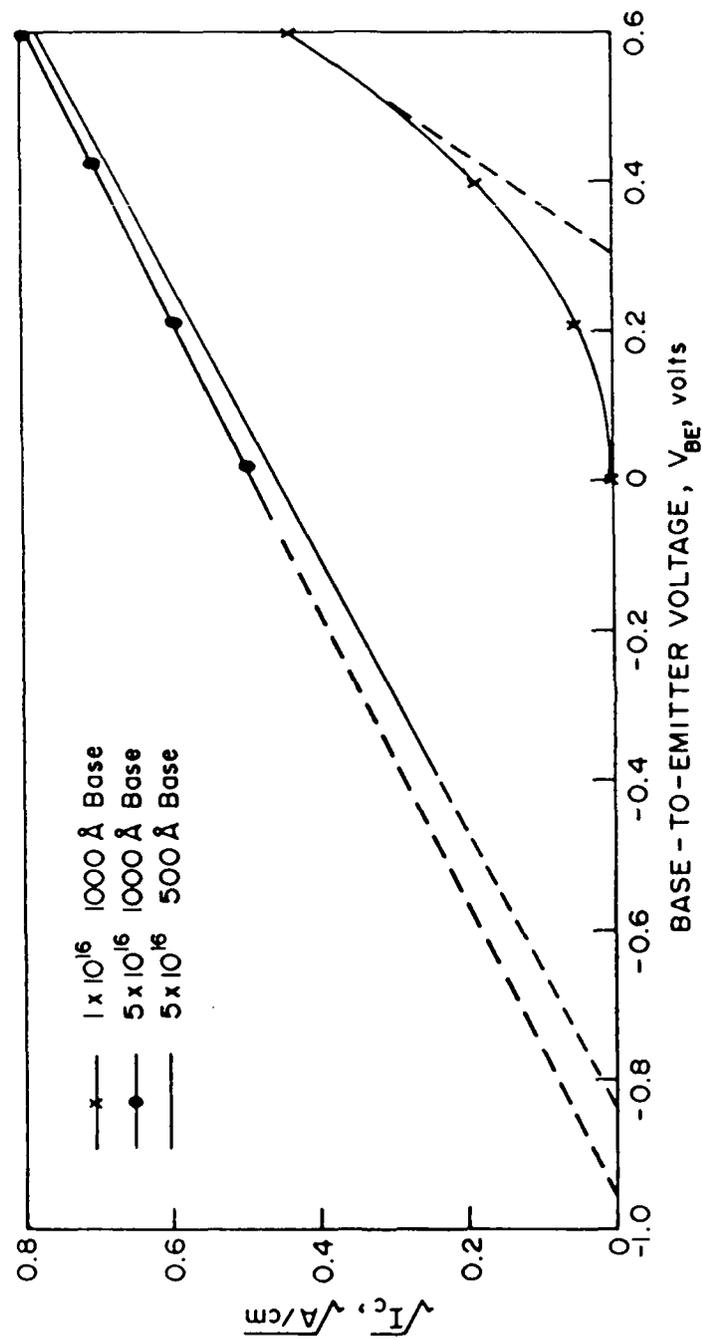
Fig. 29 – Threshold voltage dependence on doping and base penetration.
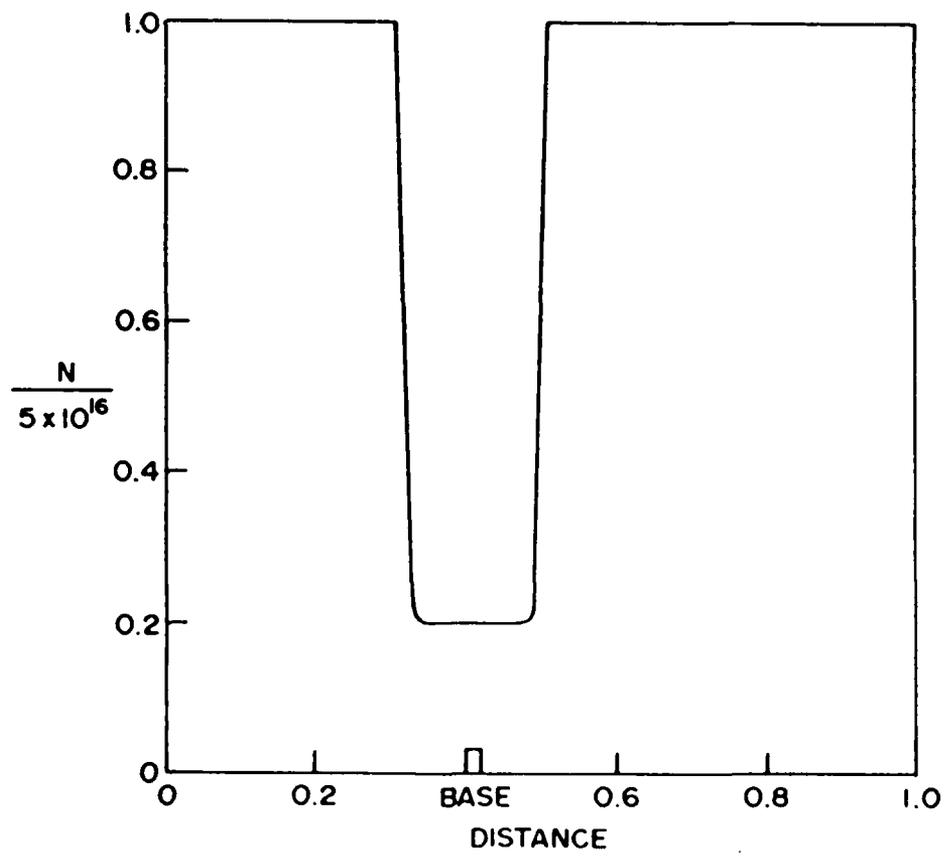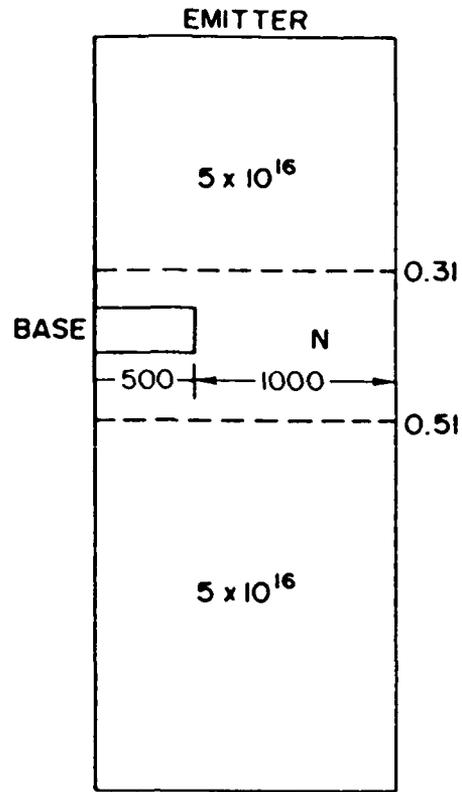
Fig. 30 - PBT with an N-N⁻-N doping profile.
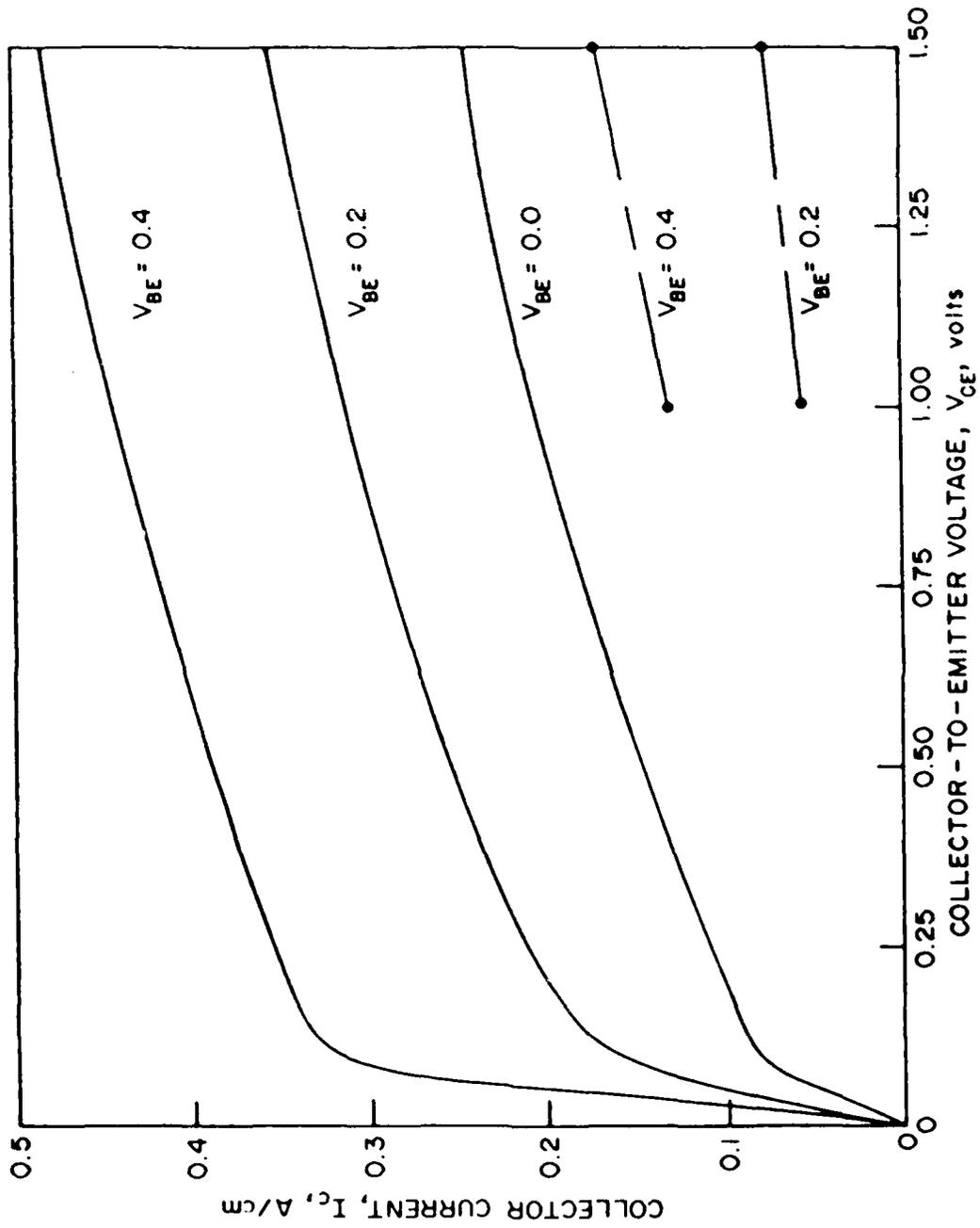
Fig. 31 - I-v characteristics for the device with high-low-high doping.

66

(a)

EMITTER   COLLECTOR
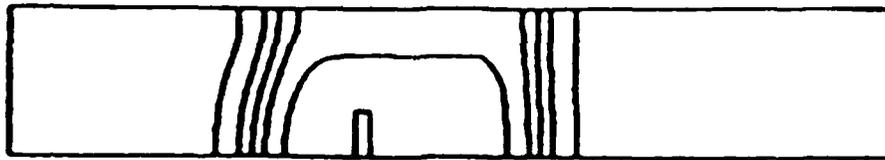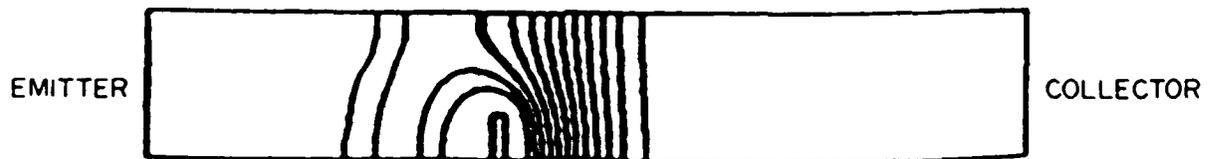
(b)
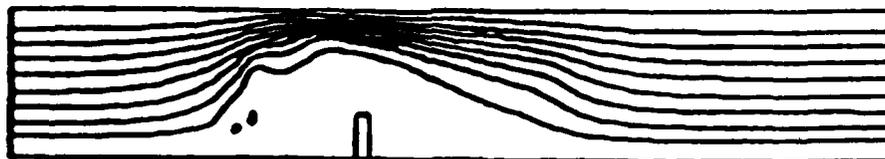
(c)

$V_{BE} = 0.2$ v

$V_{CE} = 1.0$ v

(32A)

Fig. 32A - Electron density, potential and streamline contours
to 32D   for the device with high-low-high doping.

(a)

EMITTER   COLLECTOR

(b)

(c)

$V_{BE} = 0.2$ v

$V_{CE} = 1.5$ v

(32B)
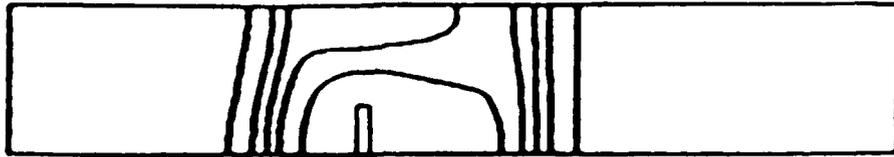
(a)
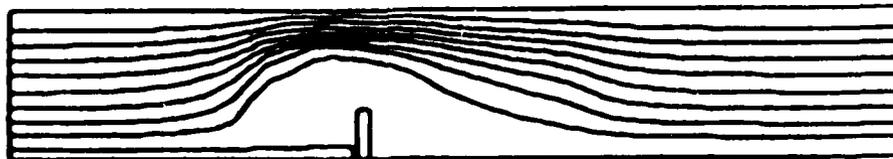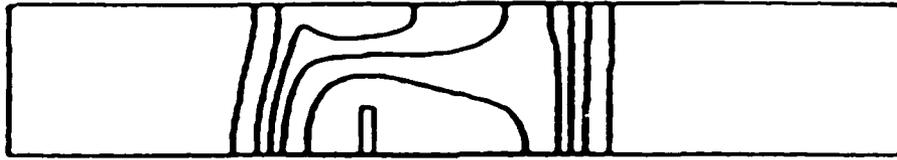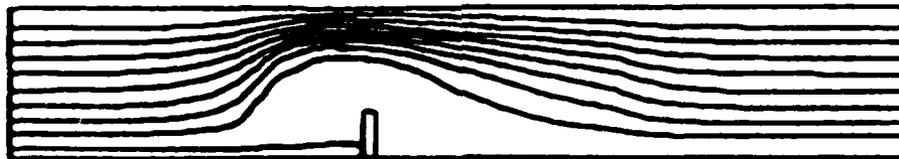
EMITTER   COLLECTOR

(b)

(c)

$V_{BE} = 0.4$ v

$V_{CE} = 1.0$ v

(32C)
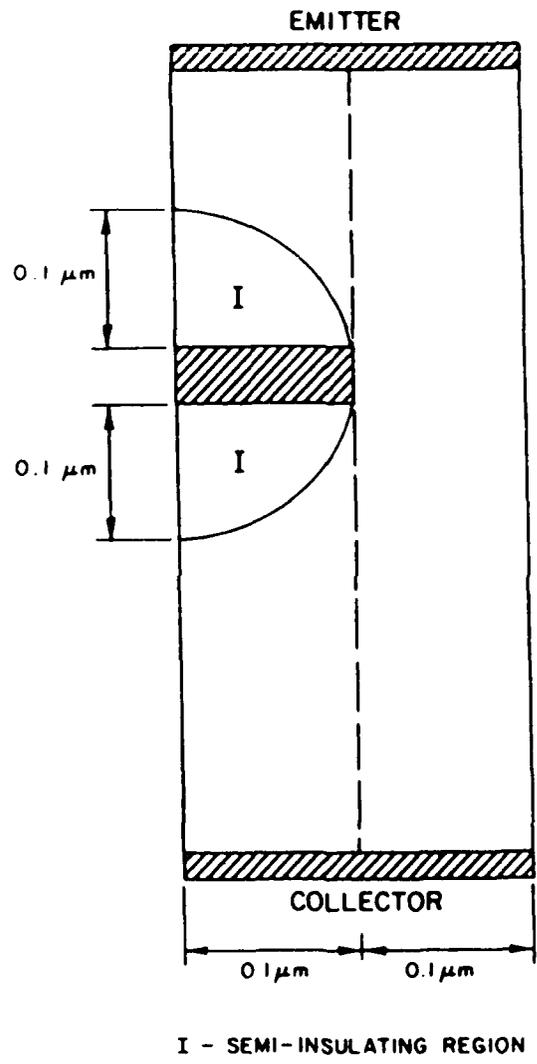
69

(a)

EMITTER   COLLECTOR

(b)

(c)

V_{BE} = 0.4 v

V_{CE} = 1.5 v

(32D)

70

EMITTER

0.1 μm

I

0.1 μm

I

f_T = 49.6 GHz

COLLECTOR

0.1 μm   0.1 μm

I — SEMI-INSULATING REGION

Fig. 33 — Structure with a semi-insulating region around the base.

71

EMITTER

0.4 μm   27%

0.1 μm   I

0.02 μm

0.1 μm   I   66%
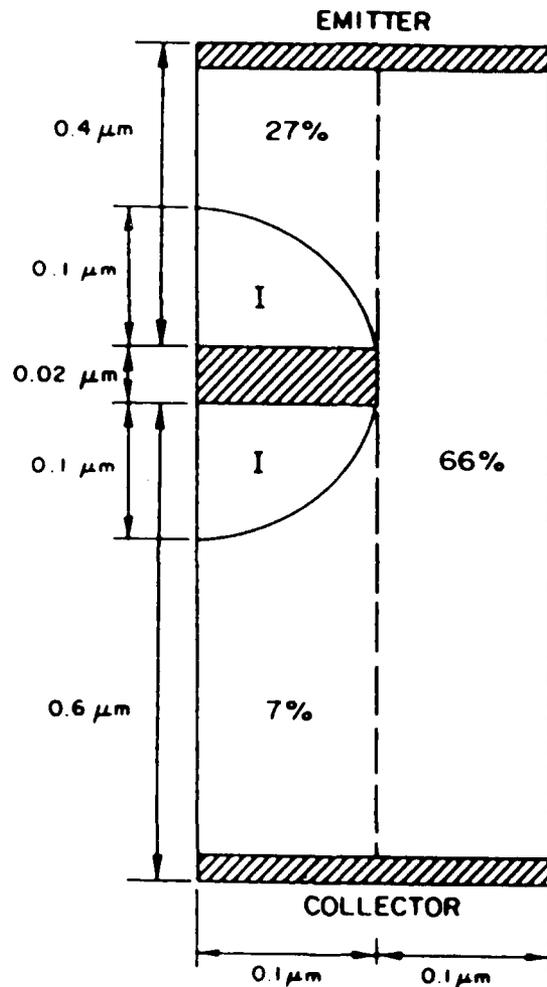
N = 1.25 x 10^{17}

f_T = 49.6 GHz

V_{CE} = 1.0 volt

V_{BE} = 0.363 volt

0.6 μm   7%

I_D = 1.2 A/cm

COLLECTOR

0.1 μm   0.1 μm

I - SEMI-INSULATING REGION

Fig. 34 – Capacitance distribution in the structure with
a semi-insulating region around the base.
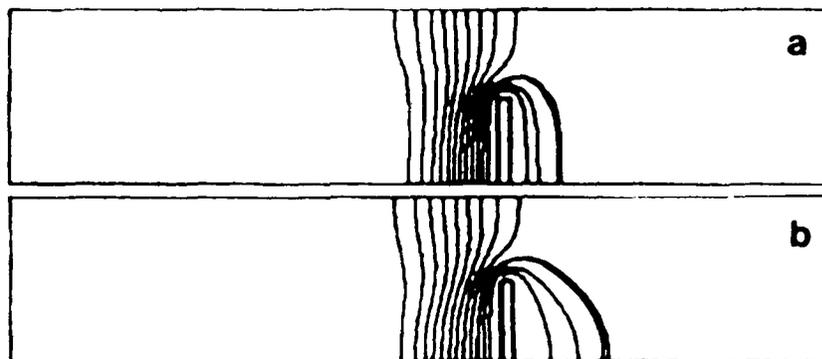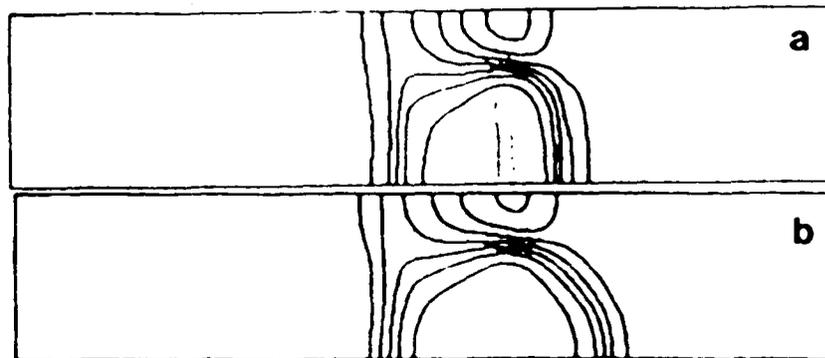
Fig. 35 – Electron density and potential contours: Comparison
of cases (a) without and (b) with semi-insulating region.
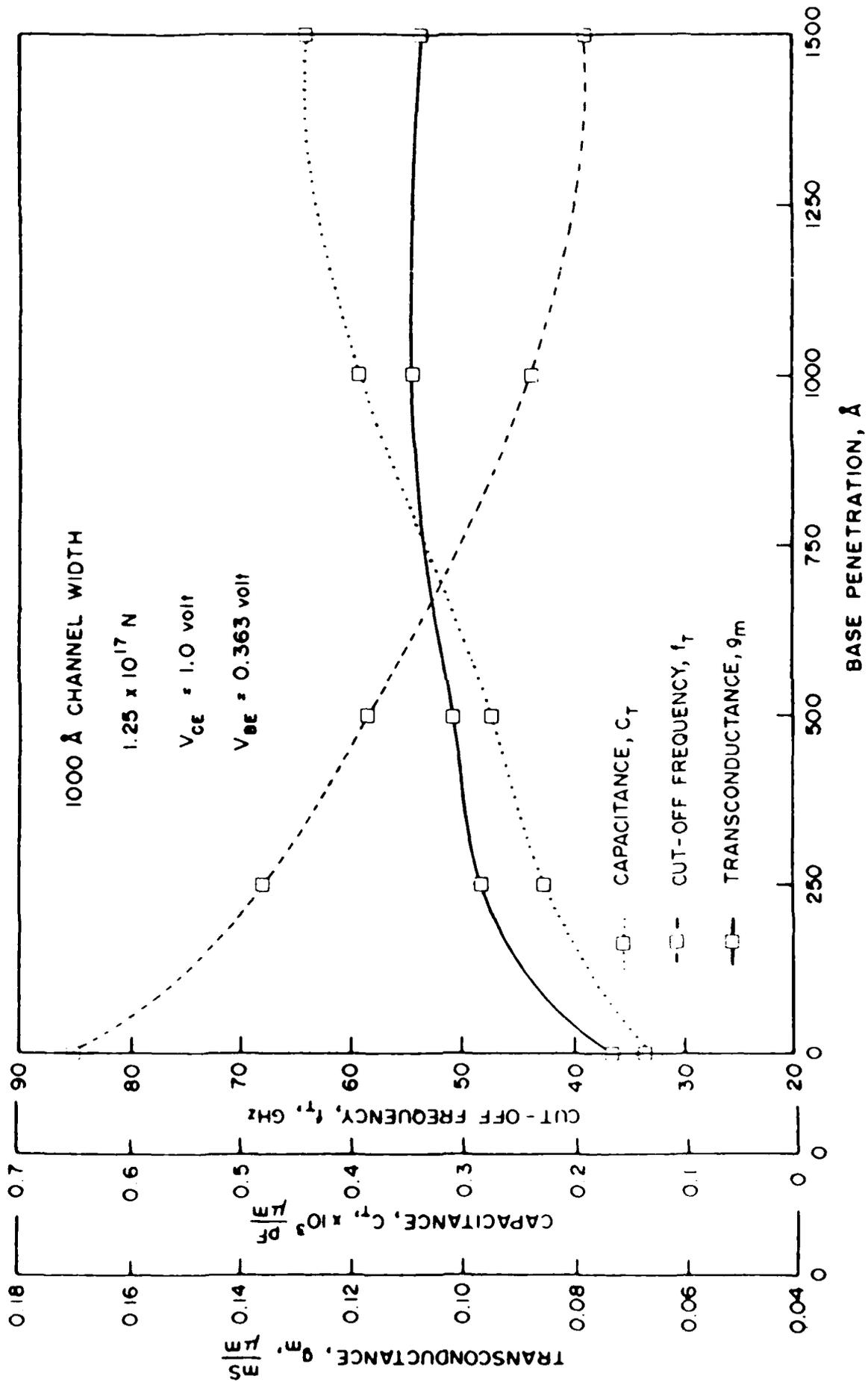
Fig. 36 - Small signal parameters for the device with a semi-insulating region around the base: variation with base penetration.

EMITTER

23%

500 Å

I

500 Å

100 Å

58%

0.501 μm   500 Å   I   19%

N = 2.0 × 10^17

f_T = 67.2 GHz

I_D = 0.797 A/cm

0.05 μm   0.05 μm

COLLECTOR
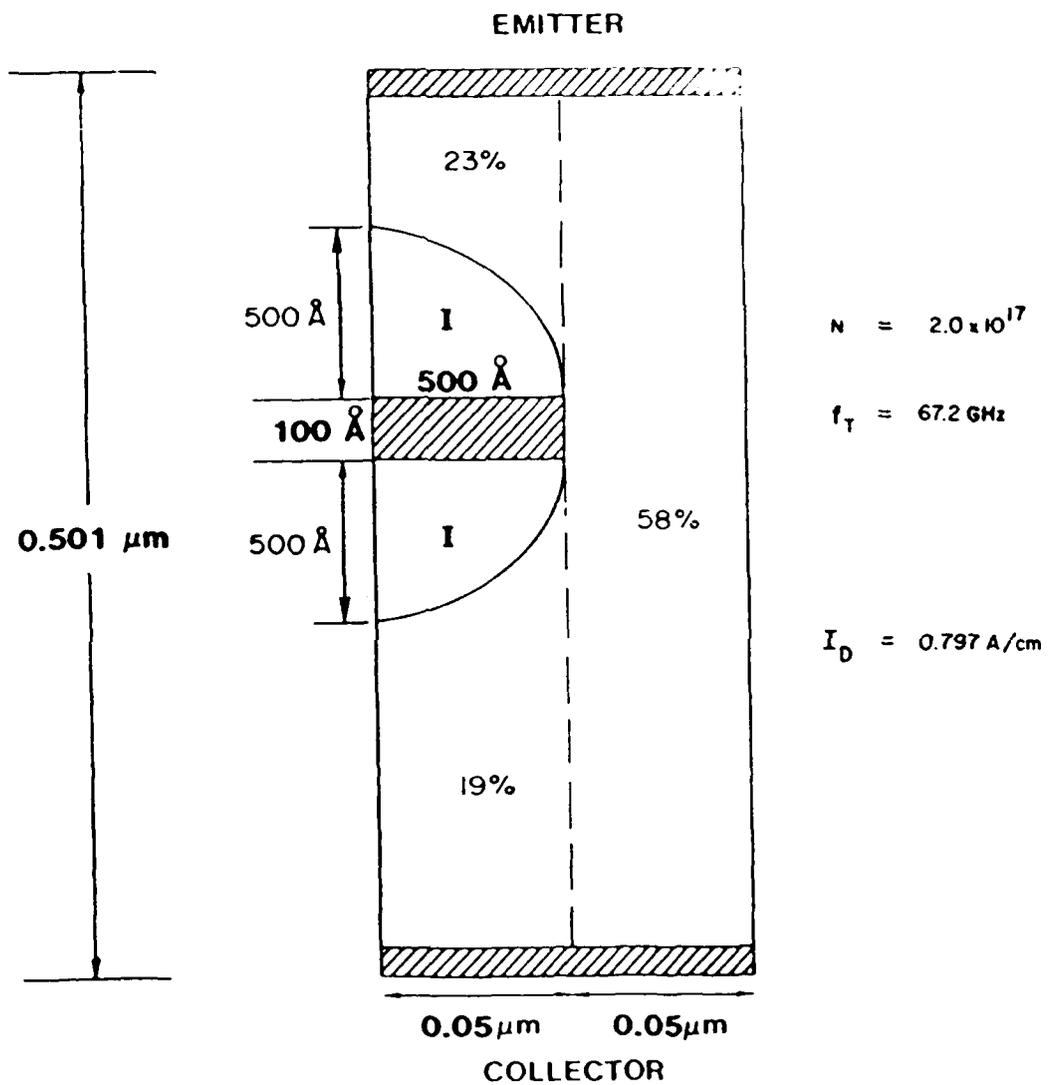
I ⇒ SEMI−INSULATING REGION

(37a)

Fig. 37a and 37b − Shape of the semi-insulating region: circular vs. elliptic

75

EMITTER

20%

0.1 μm

I

500 Å

100 Å

I

0.501 μm

0.1 μm

72%
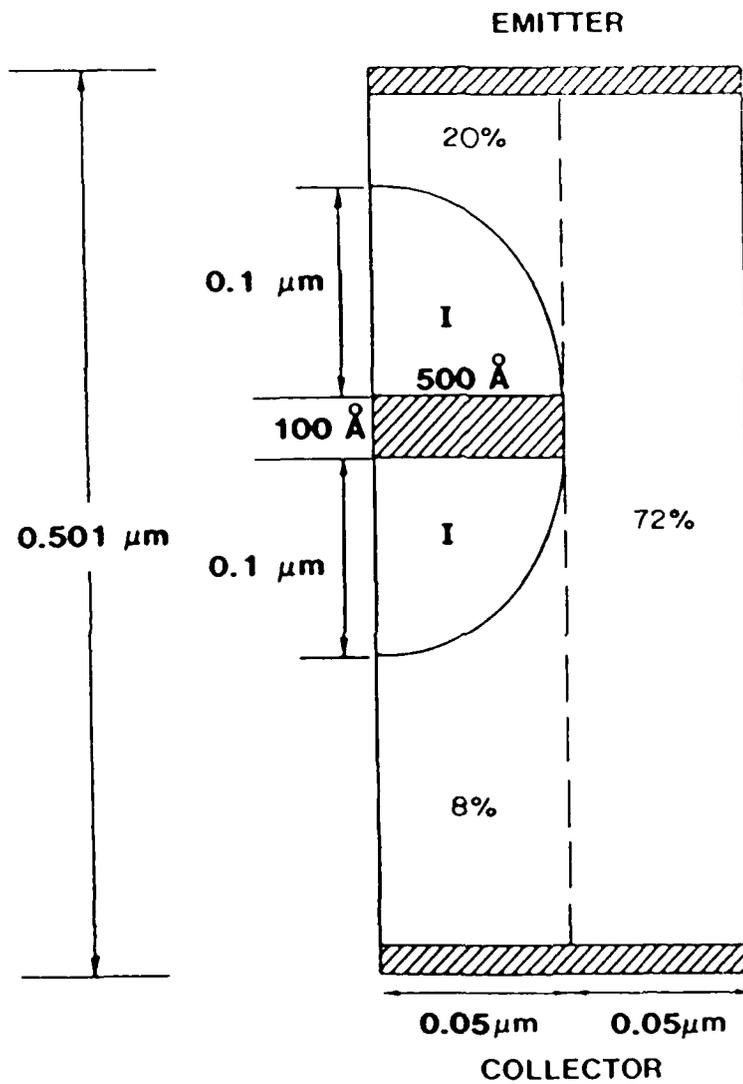
N = 2.0 × 10^{17}

f$_T$ = 80.5 GHz

I$_D$ = 0.737 A/cm

8%

0.05 μm   0.05 μm

COLLECTOR

I ⇒ SEMI-INSULATING REGION

(37b)

76

C. MICROWAVE PROPERTIES

The calculations discussed in Section B for the drift and diffusion equations, as well as the calculations using the moments of the Boltzmann transport equation included calculations of the cutoff frequencies as obtained from the dc ratio.

$$f_T = \frac{g_m}{2\pi C_T} \tag{1}$$

These type of calculations while providing a figure of merit for the device are of limited value when high frequency performance characterization is required. Of interest, particularly for microwave and millimenter wave application are such quantities as fmax, the frequency at which the maximum available gain is unity. To obtain high frequency characteristics for the PBT transient calculations were performed. These calculations to obtain the high frequency characteristics were typical switching-type calculations, with the constraint that either the base or collector contact was held at a fixed value. For the common emitter configuration, "Y" parameter calculations were performed. The "Y" parameter evaluation required two transient accurate calculations. In one calculation, the base potential was fixed and the collector potential was perturbed by a prespecified value. In the second calculation, the collector potential was fixed and the base potential was perturbed by a prespecified value. The resulting current transients are then accurately calculated. After the calculations are completed the voltage perturbations and respective current transients are approximated by a Fourier series. The Fourier expansion coefficients are then used in calculating the admittance parameters, as discussed below.

The microwave calculations were performed for the $5 \times 10^{16}/cm^3$ PBT with the ratio of base penetration to channel opening of 1:2 (see Section A-V). The calculations were performed at the quiescent point of $V_{CE} = 1.0$ volts and $V_{BE} = 0.4$ volts, where the current level was 0.45A/cm. A small step change in potential equal to 0.1 volts was imposed first on the collector and then a change of 0.05 volts was imposed on the base contact. The resulting change in emitter, base and collector currents were then computed subject to the constraint

$$\delta I_E(t) = \delta I_C(t) + \delta I_G(t) \tag{2}$$

The base and collector current, and the collector potential are then Fourier analyzed. From the Fourier coefficients, the following ratios, for a fixed base potential are identified as the admittance parameters:

$$Y_{12}(\omega) = \delta I_B(\omega)/\delta \Psi_C(\omega) \tag{3}$$

$$Y_{22}(\omega) = \delta I_C(\omega)/\delta \Psi_C(\omega) \tag{4}$$

A similar exercise is performed for a perturbation on the base contact, with

the identification of two additional admittance parameters.

$$Y_{11}(\omega) = \delta I_B(\omega)/\delta \Psi_B(\omega) \tag{5}$$

$$Y_{21}(\omega) = \delta I_C(\omega)/\delta \Psi_B(\omega) \tag{6}$$

The "Y" or admittance parameters are dependent on the space charge and potential profiles and are thus bias dependent. It is noted that an assumption usually made is that about a given bias point, the small signal currents add in a linear way, permitting an equivalent circuit representation for the PBT. A simplified circuit is shown in figure 1. The relationship of the admittance parameters to those of equations 3 through 6 is discussed below. To begin we point out that the following breakdown is usually made:

$$\delta I_B = (Y_{11} + Y_{12}) \, \delta \Psi_B - Y_{12} \, (\delta \Psi_B - \delta \Psi_C) \tag{7}$$

$$\delta I_C = (Y_{21} - Y_{12}) \, \delta \Psi_B + Y_{12} \, (\delta \Psi_B - \delta \Psi_C) \, Y_{22} + Y_{12}) \, \delta \Psi_C \tag{8}$$

The term $Y_{11} + Y_{12}$ is identified as the emitter base admittance parameter, while $-Y_{12}$ is referred to as the base-collector admittance. The advantage of this type of representation is that a heurisitic experimental connection can be made to the circuit elements.

The frequency dependent Y parameters (equations (3) and (6)) are displayed in figures 2 through 5 as conductance and capacitance. In figure 2

$$G_{12} \equiv Re \, Y_{12} \tag{9a}$$

$$\omega C_{12} \equiv Im \, Y_{12} \tag{9b}$$

in figure 3

$$G_{22} \equiv Re \, Y_{22} \tag{10a}$$

$$\omega C_{22} \equiv Im \, Y_{22} \tag{10b}$$

in figure 4

$$G_{11} \equiv Re \, Y_{11} \tag{11a}$$

$$\omega C_{11} \equiv Im \, Y_{11} \tag{11b}$$

in figure 5

$$G_{21} \equiv Re \, Y_{21} \tag{12a}$$

$$\omega C_{21} \equiv Im \, Y_{21} \tag{12b}$$

The admittance parameters in figures 2 through 5 are expressed in multiples of

$$G_o = \frac{eN_r \, X_r V_r}{\Psi_r} \qquad (13)$$

and are plotted as function of the Fournier index n. Thus figure 2a is a plot of ReY_{12}/G_o. Figure 2b is a plot of

$$ImY_{12}/G_o = \omega C_{12}/\omega_o C_o \qquad (14)$$

where

$$C_o = \frac{eN_r X_r^2}{\Psi_r} \qquad (15a)$$

$$\omega_o = \frac{V_r}{X_r} \qquad (15b)$$

where $N_r = 5 \times 10^{16}/cm^3$, $X_r = 1.0\mu m$, $V_r = 1.0 \times 10^7 cm/sec$ and $\Psi_r = 1.0$ volt. Thus, $G_o = 80S/mm$, $C_o = 8.0pF/mm$, and $\omega_o = 10^{11}/sec$. We note, $\omega = 2\pi nf$, where f is expressed as a multiple of $10^{11}/sec$. Here $f = \omega_o/10$.

In examining these results we refer to the zeroth order equivalent circuit (within the bracket) of figure 1. For the zeroth order circuit

$$Y_{120} = j\omega C_{12} = j\omega C_{BC} \qquad (16)$$

$$Y_{220} = G_{22} - j\omega C_{22} = G_{CE} - j\omega C_{BC} \qquad (17)$$

For earlier calculation within HEMT and FET structures, to within 10% and for frequencies below 200GHz, $Y_{12} \approx - Y_{22} \, G_{11} \approx 0$, $G_{22} \approx 0$. This is not the case with the PBT. Also, for the zeroth order equivalent circuit associated with the gate perturbation, Y_{11} is separated into

$$Y_{110} = -j\omega C_{BE} - j\omega C_{BC} \qquad (18)$$

$$Y_{210} = G_m + j\omega C_{BC} \qquad (19)$$

For the zeroth order element, the imaginary part of Y_{21} is the same as that of Y_{12} and Y_{22}. The computed results indicate otherwise. In particular, $|Y_{21}|^2 > |Y_{12}|^2$. Thus the zeroth order circuit is inadequate.

The cutoff frequency was computed from several parameters. From the dc parameters $f_T = 43GHz$. The cutoff frequency was also obtained from the Y

parameters using

$$f_T = \frac{\omega \, \text{Re} \, Y_{21}}{2\pi |\text{Im}(Y_{11} + Y_{12})|}$$ (20)

and evaluated in the dc limit where $f_T = 50.2$GHz.

The microwave performance of the structure was also evaluated using several other expressions. The current gain was computed from the magnitude of A_i, where

$$A_i = \frac{Y_{21} \, Y_l}{Y_{11}(Y_{21} + Y_l) - Y_{12}Y_{21}}$$ (21)

and Y_l is the load admittance. A_i is dependent upon the load admittance, and an extrema occurs at

$$\frac{\partial A_i}{\partial Y_l} = 0$$ (22)

requiring

$$(Y_{11} - Y_{12})Y_{21} = 0$$ (23)

While the above condition is generally not satisfied, it yields

$$A_{i,max} = \frac{Y_{21}}{Y_{11}}$$ (24)

and is independent of load. The magnitude of the current gain is plotted in figure 6. The cutoff frequency occurs when the log of the maximum current gain is zero, and is near 50 GHz. We note: this value is somehwat higher than that obtained using equation 20.

The current gain is sensitive to loading, and for purely resistive load becomes negligibly small as Y_l approaches zero. This is displayed in figures 7 through 9.

The power gain was computed from

$$G = \frac{Y_{21}^2 \, Y_l}{[Y_{11}(Y_{22}+Y_l) - Y_{12}Y_{21}][Y_{22}+Y_l]}$$ (25)

80

The condition where G is at an extreme occurs when

$$\frac{\partial G}{\partial Y_1} = 0 \qquad (26)$$

This is load-dependent and occurs when

$$Y_1^2 = Y_{22}[Y_{11}Y_{22} - Y_{12}Y_{21}]/Y_{11} \qquad (27)$$

A plot of G at its extreme is displayed in figure 10. The power gain is zero near 50 GHz. The dependence on load (real) is displayed in figures 11 through 13. At both low and high values of Y_1, the power gain is reduced, and the frequency at which the power gain is zero is reduced.

The more common evaluation of the microwave performance of a device is through the maximum available power gain, MAG,

$$MAG = \frac{|Y_{21}|}{|Y_{21}|} \times [k - \sqrt{k^2-1}] \qquad (28)$$

where

$$k = \frac{2Re\ Y_{11}\ Re\ Y_{22} - Re(Y_{11}\ Y_{21})}{|Y_{12}\ Y_{21}|} \qquad (29)$$

and occurs when the input and output ports of a two-port network are simultaneously matched to their conjugate admittances. Note that k is the stability factor. k must be $\geq$ unity for stability.

If k <1 then lossy admittances can be placed at the ports to make the network stable. The maximum power gain then tends toward the maximum stable power gain

$$MSG = \frac{|Y_{21}|}{|Y_{12}|} \qquad (30)$$

MAG, K and MSG are plotted in figures 14 through 16. We note that there is approximately 12.5 dB of MSG at 20 GHz, 7.7 dB at 60 GHz and 5.3 at 100 GHz. The system is stable at 110 GHz and yields 4.1 db of MAG. We note that f_{max}, the frequency at which 10log MAG = 0, is approximately 200 GHz.

The detailed "Y" parameter calculations were performed at only one pair of bias points, but show the possibilities of satisfactory gain at high frequency. Additional calculations under Phase III, show a bias dependence of gain.

REFERENCES

1. Rollet, J.M.   IRE Trans. Circuit Theory. CT 9 29 (1962).

2. IRE Standards, Proc. IRE 45, 983 (1957).
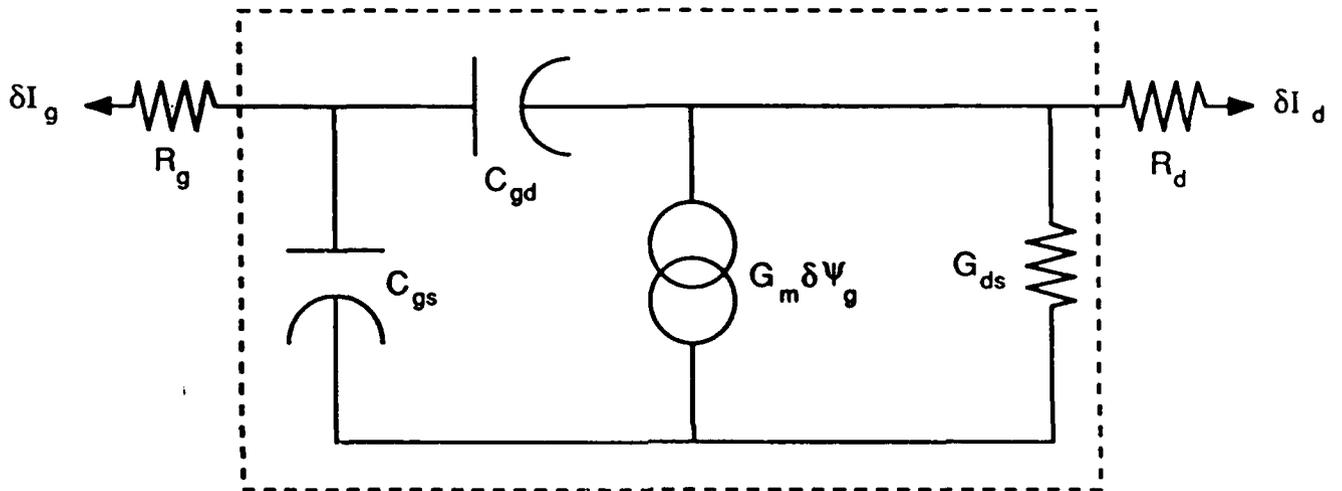
Fig. 1. Simplified small signal equivalent circuit.
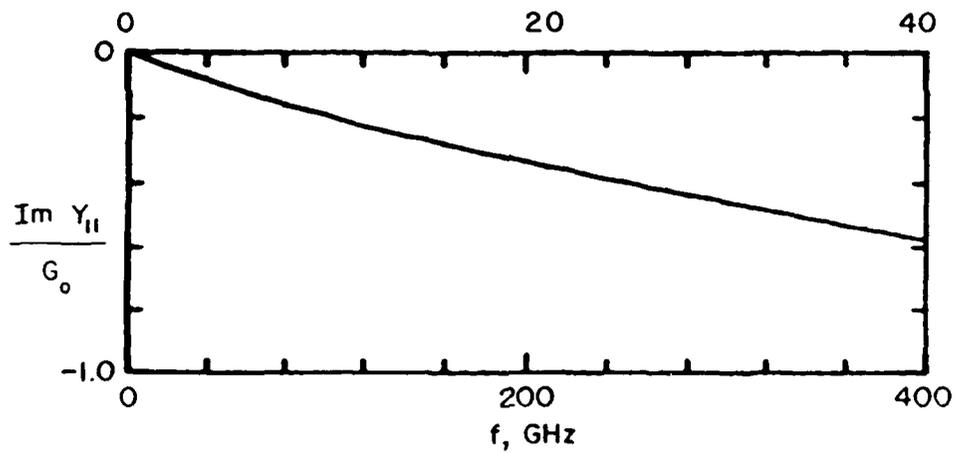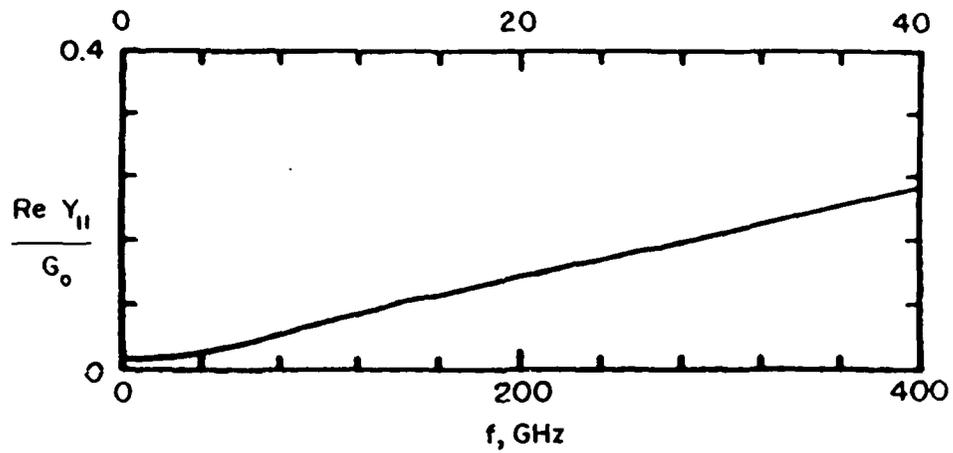
Fig. 2. Real and imaginary part of Y_{11}.
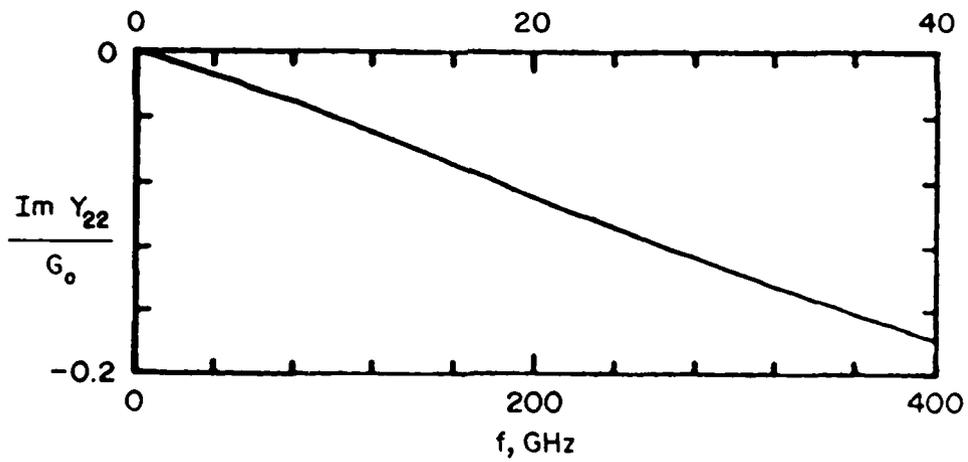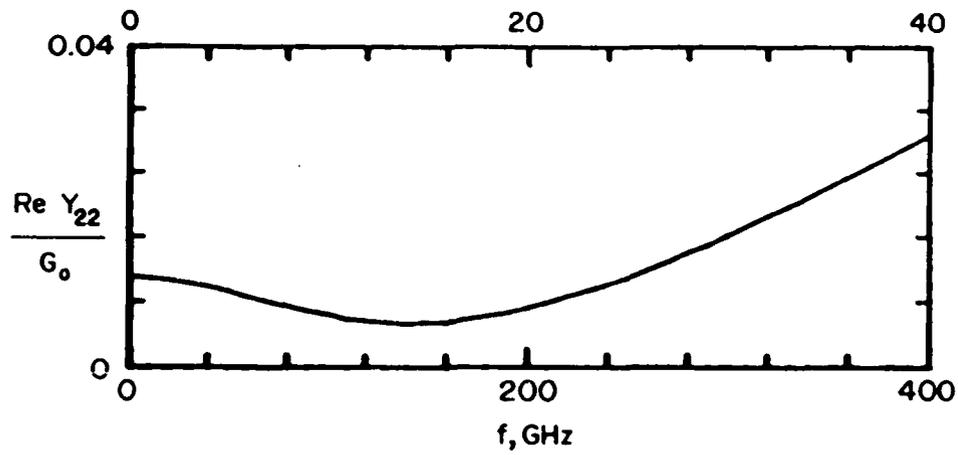
Fig. 3. Real and imaginary part of Y_{22}.
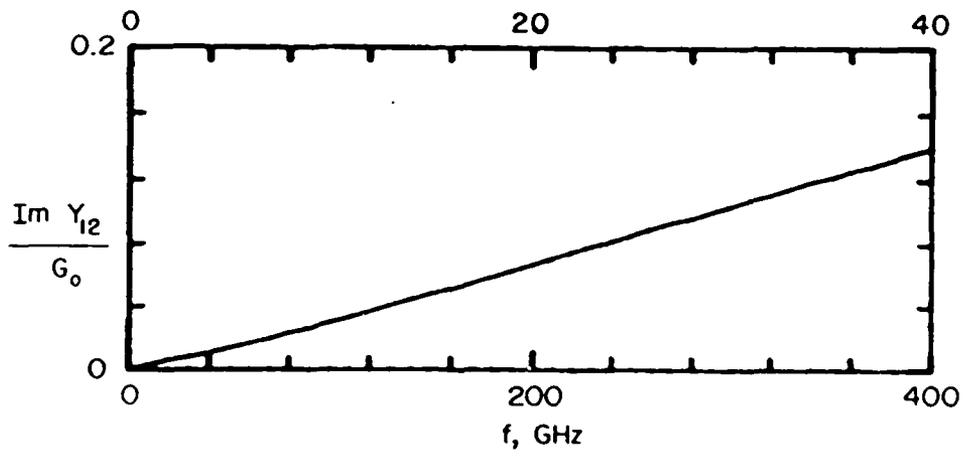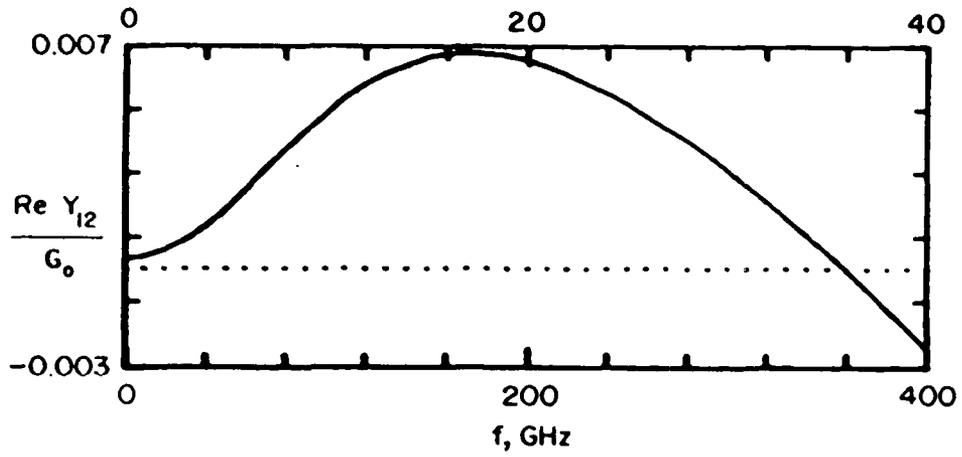
Fig. 4. Real and imaginary part of Y_{12}.
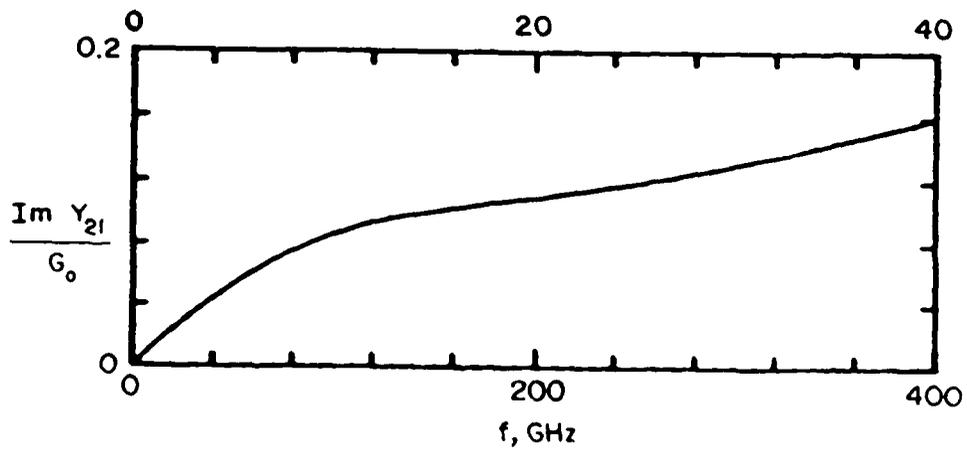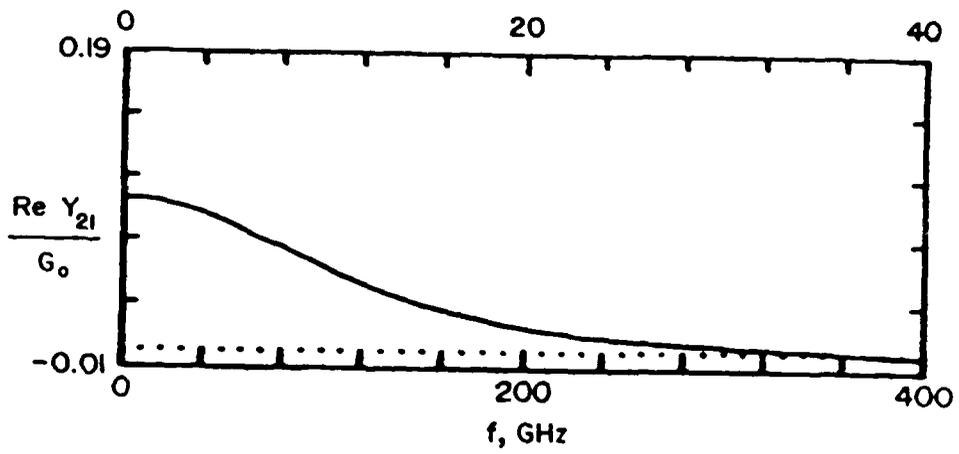
Fig. 5. Real and imaginary part of Y_{21}.
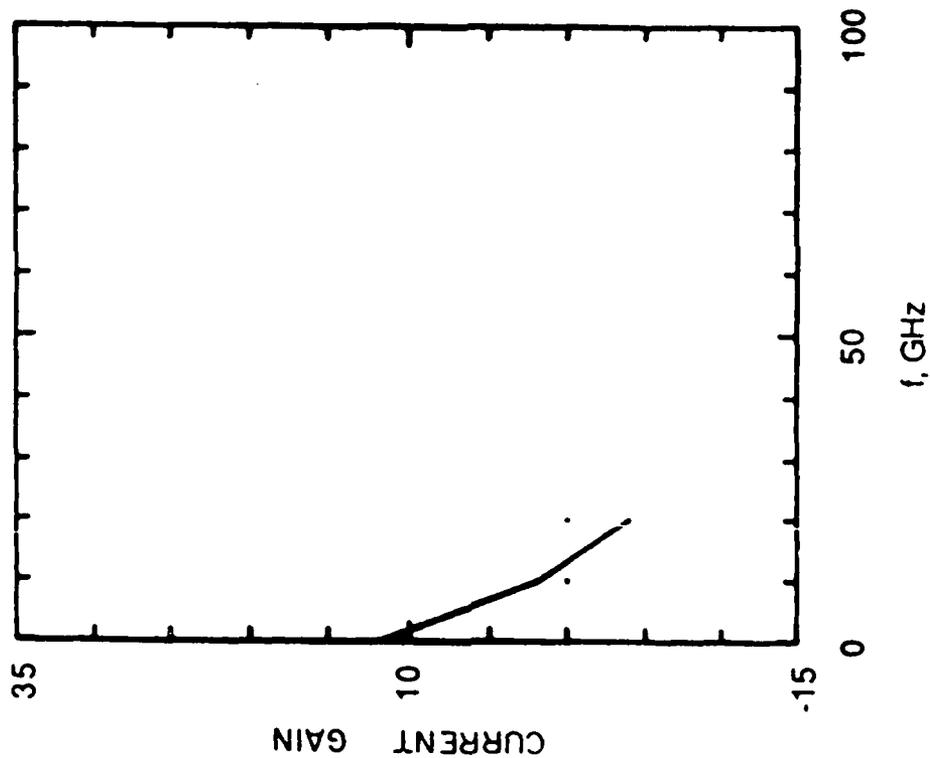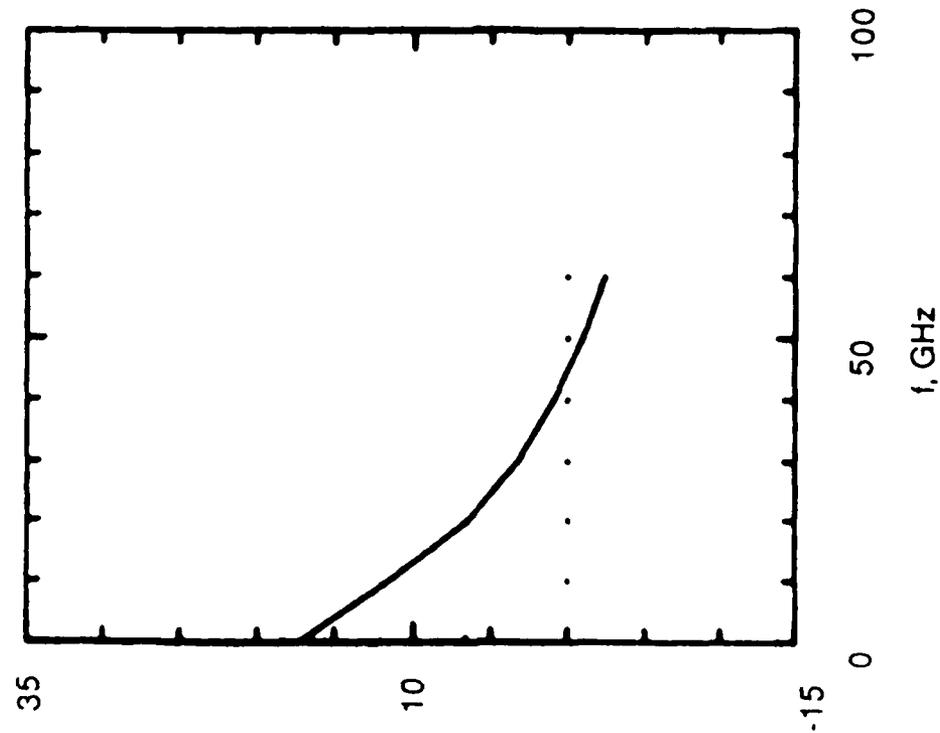
Fig. 6. Maximum current gain versus frequency for the Y parameters of Figures 2 through 5.

Fig. 7. Current gain from Equation (21) for the Y parameters of Figures 2 through 5 with $Y_L = 0.01$.

Fig. 9. As in Figure 8, but for $Y_1 = 1.0$.

Fig. 8. As in Figure 7, but for $Y_1 = 0.1$.
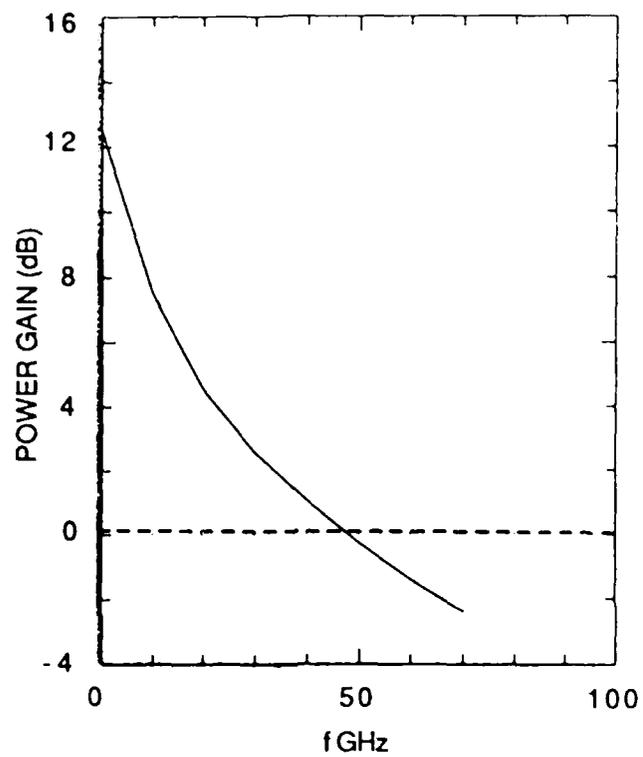
Fig. 10.   The power gain as computed from Equation 25, with Y_1 given by
            Equation 27.   Y parameters are as in Figures 2 through 5.
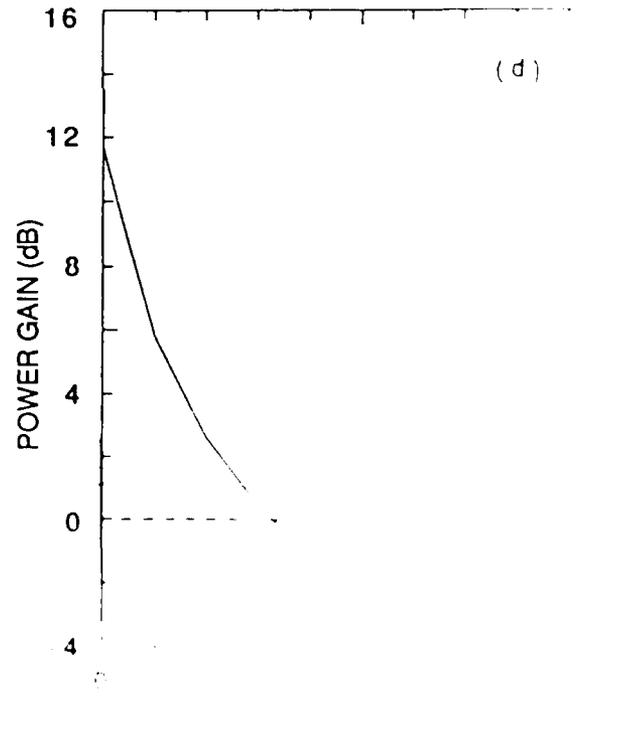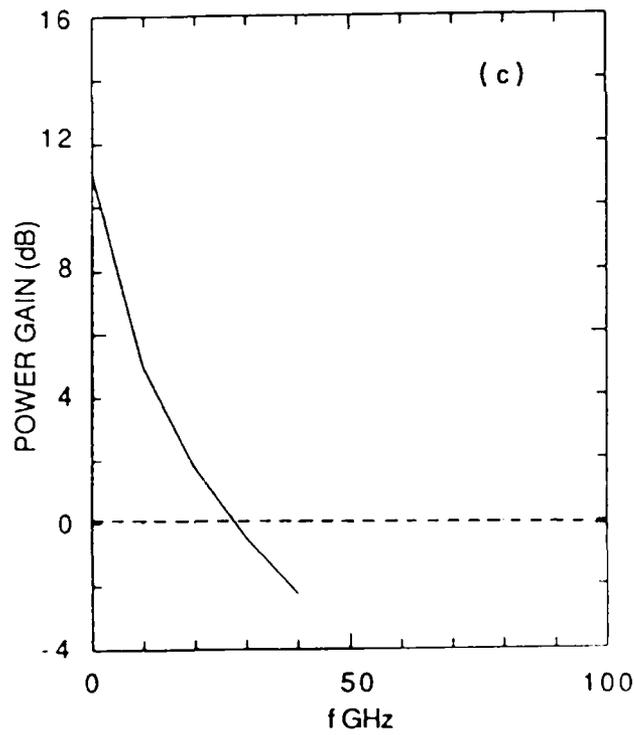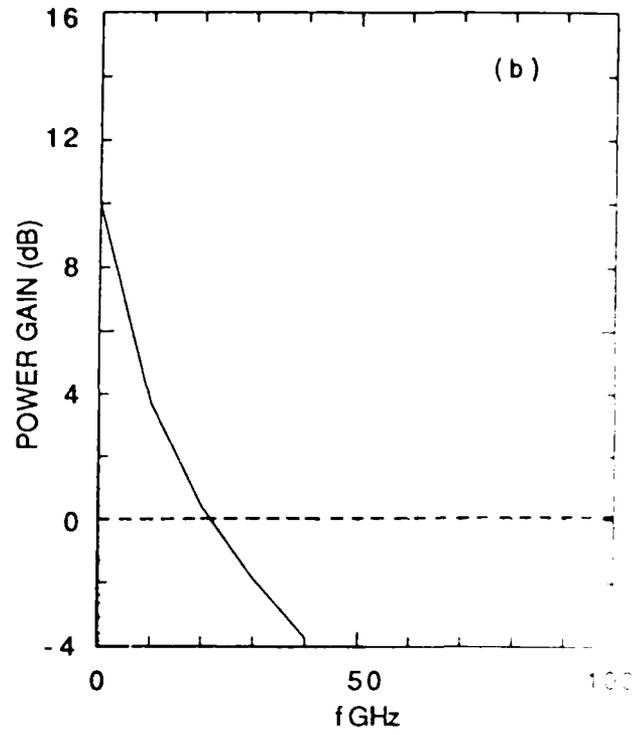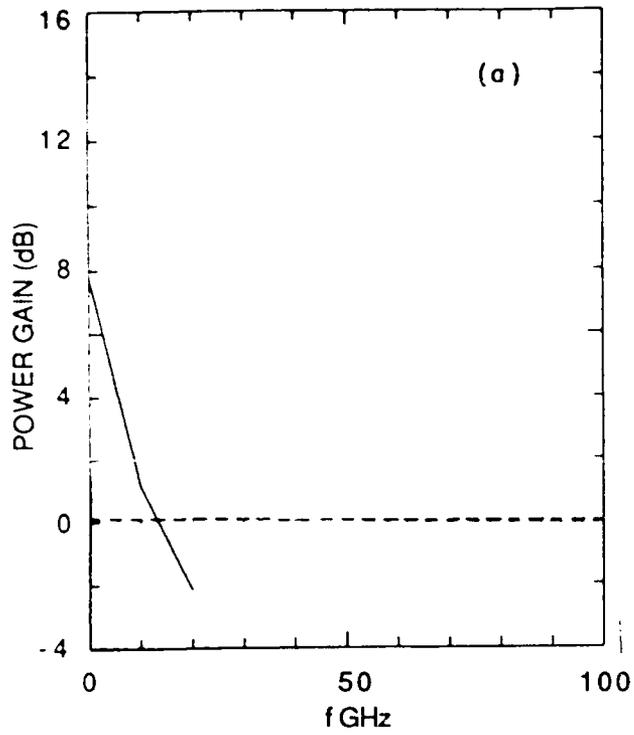
Fig. 11. As in Figure 10, but with
c) $Y_1 = 0.003$, d) $Y_1 = $
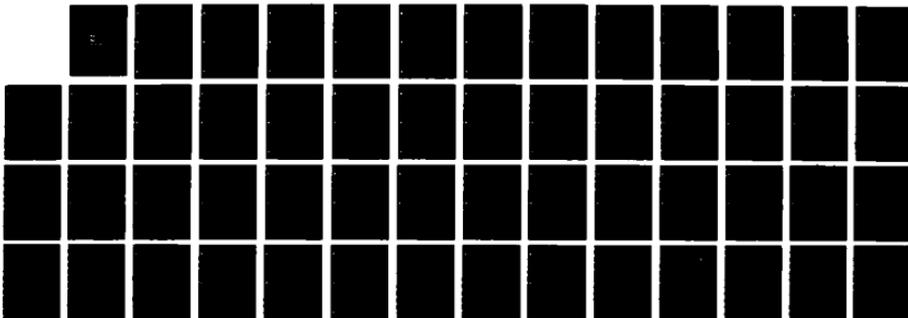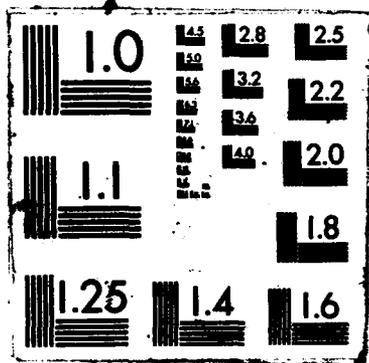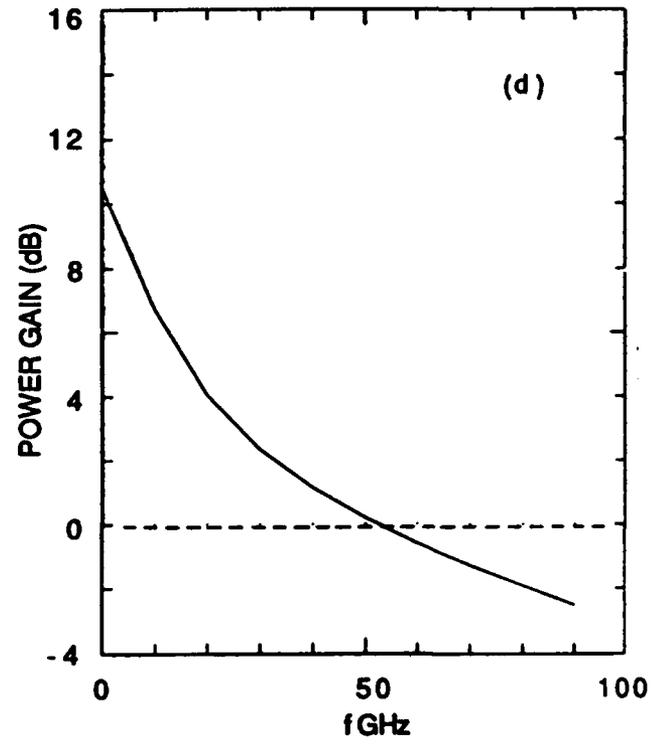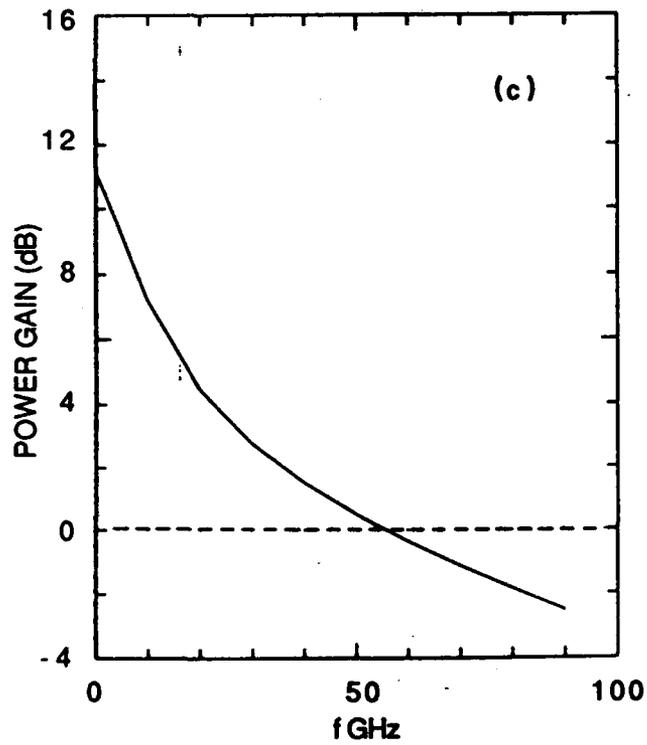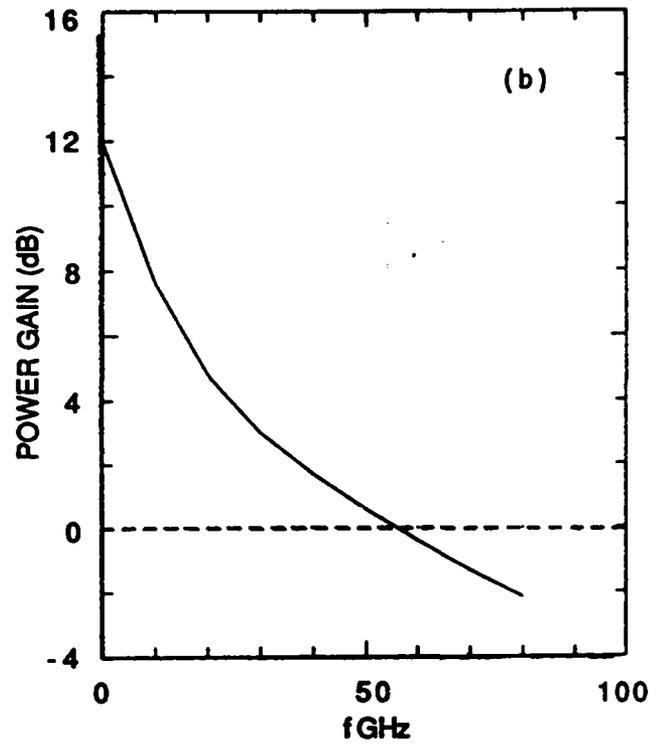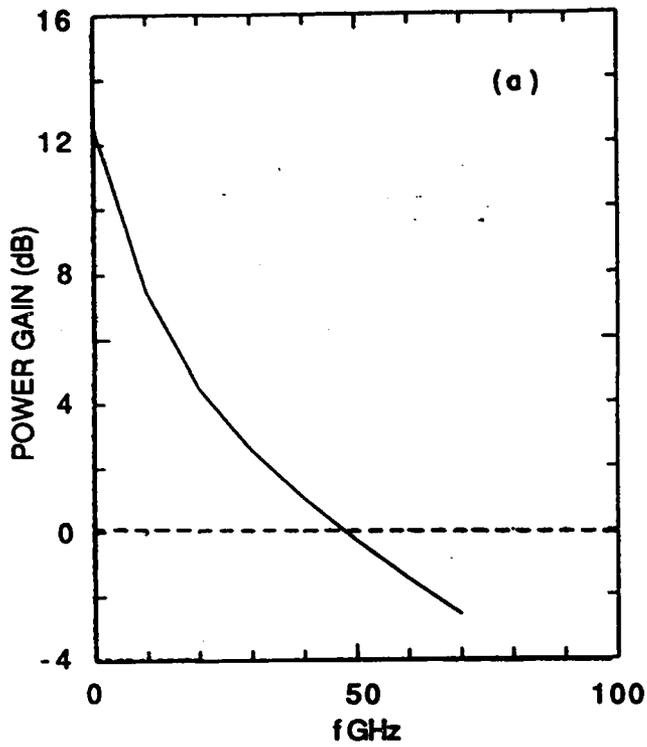
Fig. 12. As in Figure 10, but with: a) $Y_1 = 0.01$, b) $Y_1 = 0.02$, c) $Y_1 = 0.03$, d) $Y_1 = 0.04$.

92
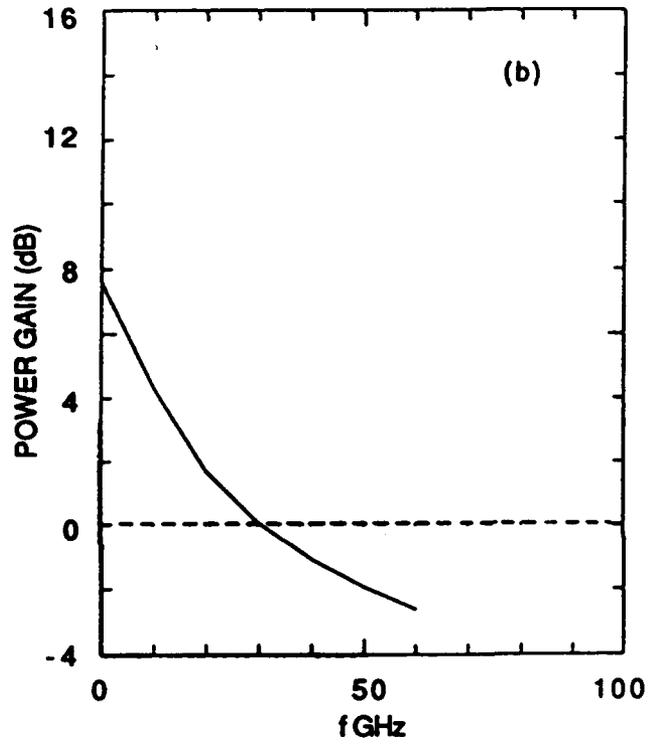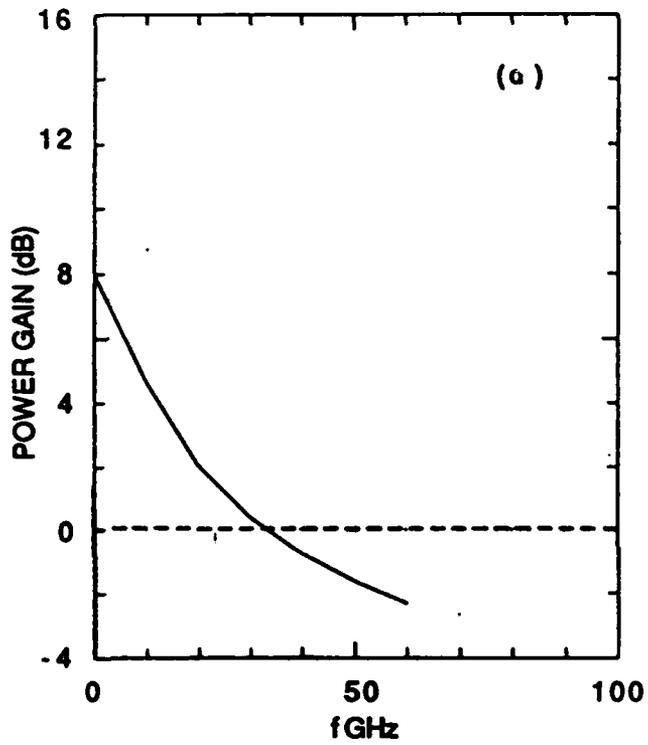
Fig. 13. As in Figure 10, but with: a) $Y_1 = 0.09$, b) $Y_1 = 0.1$.
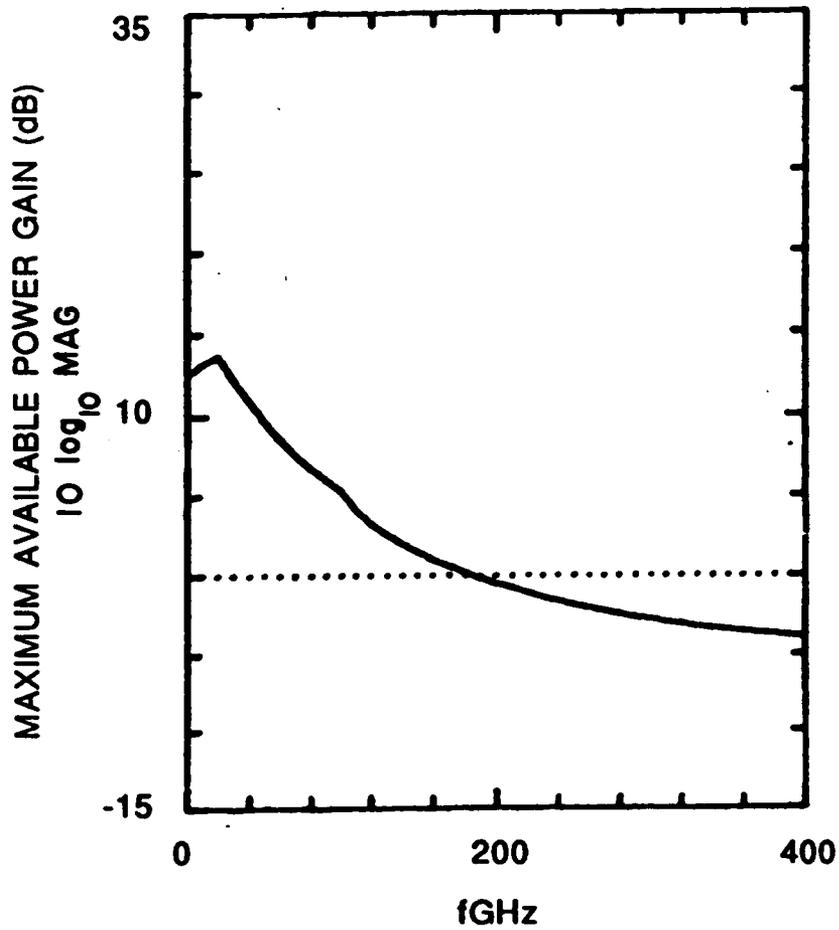
Fig. 14. Maximum available gain, from Equation (28) with the Y parameters of Figures 2 through 5.
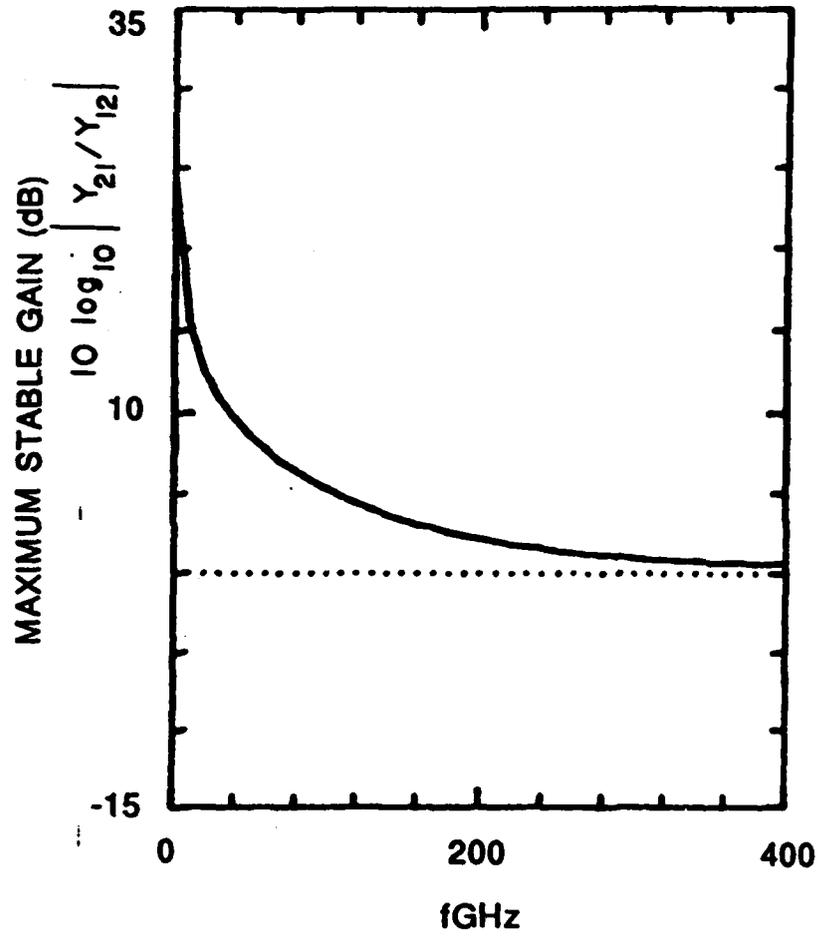
Fig. 15. Maximum stable gain, from Equation (30) with the Y parameters of Figures 2 through 5.
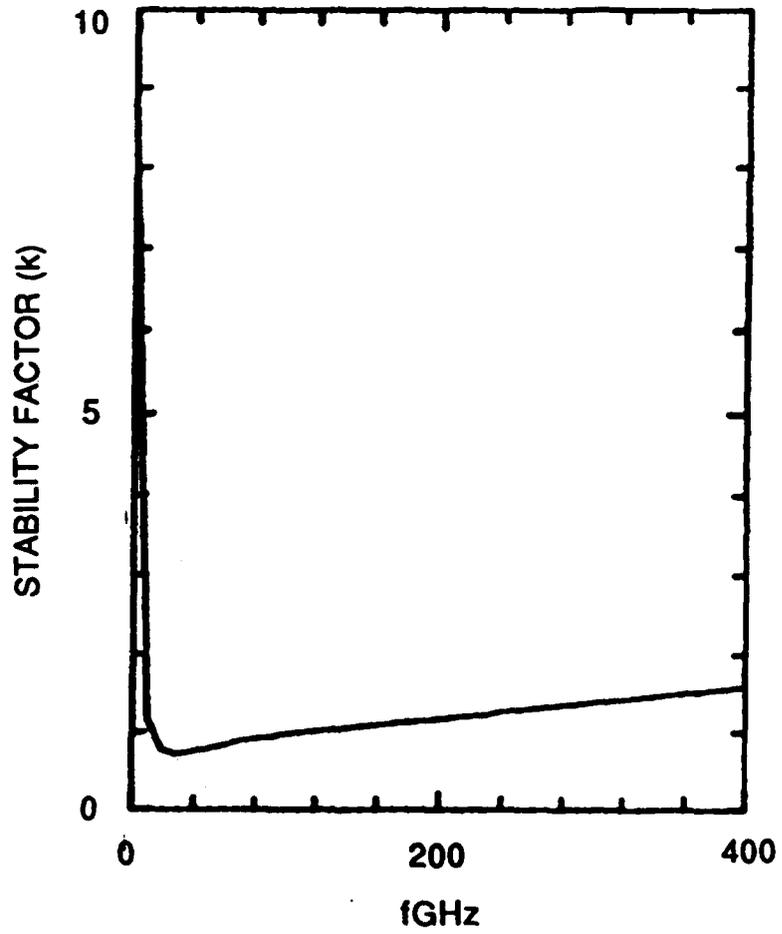
Fig. 16. Stability parameter from Equation (25) with
the Y parameters of Figures 2 through 5.

D. BREAKDOWN CHARACTERISTICS OF THE PBT

Because of the unusual structure of the PBT, and the extended depletion layer properties of the base contact, the breakdown characteristics of the PBT represent a major uncertainty in its evaluation. Also, the breakdown characteristics of the PBT relative to coplanar configured GaAs FET are also unknown. The first published study of the high bias characteristics of the PBT was performed by Alley [1] whose calculations were for $4 \times 10^{16}/cm^3$ and showed generation in excess of 18 volts. These calculations were significantly deficient in one key area, they did not include hole transport. Rather, the calculations were carried to the point where it was clear that the electrons were multiplying. In the Phase II study both electron and hole transport were included in the breakdown investigation. But here too there are limitations in the approach. The most serious limitations in the breakdown study is that generation occurs at a prespecified value of electric field, rather than energy. Additionally, avalanche coefficients are not obtained from fundamental principles, and must be treated cautiously. Of importance then are the breakdown characteristics of the PBT relative to another device. In an ONR-sponsored study, the breakdown characteristics of a coplanar GaAs FET with a 3 micron source-to-drain separation was studied. The remarkable result is that breakdown for the FET occurred at lower voltages than breakdown for the PBT. This will be discussed below.

The breakdown characteristics of the PBT were examined through incorporation of holes into the calculation. The PBT equations were generalized as follows:

$$\frac{\partial N}{\partial t} = \frac{1}{e} \nabla \cdot Jn + G - R \tag{1}$$

$$\frac{\partial P}{\partial t} = -\frac{1}{e} \nabla \cdot Jp + G - R \tag{2}$$

$$Jn = e\left[N\mu n \ \nabla(\psi) + Dn\nabla N\right] \tag{3}$$

$$Jp = -e\left[P\mu p \ \nabla(\psi) + Dp\nabla P\right] \tag{4}$$

with Poisson's equation becoming

$$\nabla \cdot \epsilon \nabla \psi = -\rho = e(N\text{-}P\text{-}C) \tag{5}$$

where

$$C(x,z) = N_D - N_A \tag{6}$$

97

Avalanche generation is through

$$G_i = \frac{1}{e} \left(\alpha_n |J'_n| + \alpha_p |J'_p| \right) \tag{7}$$

where J'_k incorporates only the drift component of particle current (not diffusion). Thus, generation is proportional to the numbers of electrons and holes within the structure. The coefficient α_k, in equation 7 is given by

$$\alpha_k = A_k \exp \left[-\left(\frac{b_k}{|F|}\right)^{m_k} \right] \tag{8}$$

with values: $A_n = 1.899 \times 10^5$, $A_p = 2.215 \times 10^5$, $b_n = 5.75 \times 10^5 v/cm$, $b_p = 6.57 \times 10^5 v/cm$, $m_n = 1.82$, $m_p = 1.75$.

The carrier velocity for electrons is as discussed earlier. For holes, the mobility is field dependent and equal to

$$\mu = \frac{\mu'}{\left[1 + \left[\frac{\mu'|F|}{V_{sat}} \right]^\beta \right]^{1/\beta}} \tag{9}$$

Here $\mu' = 200$, $\beta = 2$, and $V_{sat} = 1 \times 10^7$ cm/sec. Hole diffusivity is given by the Einstein relations

$$D = \frac{kT}{e} \mu \tag{10}$$

Recombination is included in this calculation, and is given by

$$R = \frac{NP - N_i^2}{\tau_p(N+N_i) + \tau_n(P+N_i)} + r(N+P)(NP-N_i^2) \tag{11}$$

where $\tau_p = \tau_n = 1.0ns$, $N_i = 1.79 \times 10^6/cm^3$, and $r = 2 \times 10^{-31}$ cm^6/sec. The second item represents Auger recombination.

The structure for which the breakdown study was performed is that discussed in Section B-V. The current voltage characteristics for this structure at a base potential of 0.4 volts is shown in Figure 7. It is seen that the effects of avalanche generation are initiated at 15 volts, but that breakdown is not apparent until 17 volts.

In earlier sections, particularly Section B, contours of the potential
distribution showed fairly recognizable one-dimensional potential
distributions, both upstream and downstream of the base contact. Very similar
features occur at high bias levels. Now in both the PBT and the FET, high
fields occur both upstream and downstream of the base and drain contacts, and
these fields are high enough to cause avalanching. However, at a given value
of bias the fields at the base contact corners of the PBT are lower than those
at the gate contact of the FET.

The distribution of potential and charge along the central line of symmetry is
shown in Figure 2. First we note the hole distribution, and point out that
the holes drift in the direction of decreasing potential. Thus they drift
toward the emitter contact, and tend to pile up both upstream and downstream
of the base contact. We note that the largest concentration of holes occurs
upstream from the base contact. These holes (as well as electrons) are
generated at the sides upstream and downstream from the base contact, travel
laterally around the base where the fields are reduced below the contact
corner fields. Holes are also generated in the channel opening. The increase
in electron density along the line of symmetry is a consequence both of
avalanching and current continuity. Downstream from the base contact there is
precipitous drop in excess holes. This is not to be construed as a reduction
in hole generation. Rather, the holes generated near the base corners are
swept around to the emitter contact. Similar arguments apply to the generated
electrons. Qualitatively similar behavior is seen at higher bias levels.

Breakdown calculations under an ONR study were performed for the FET shown in
Figure 3, with the indicated dimensions. The current voltage relation for
this structure is shown in Figure 4. The lower breakdown voltage is apparent
and supports the idea that the PBT breakdown characteristics, while obtained
at only one base voltage, show very promising expectations vis-a-vis coplanar
FETs.

In earlier sections, particularly Section B, contours of the potential distribution showed fairly recognizable one-dimensional potential distributions, both upstream and downstream of the base contact. Very similar features occur at high bias levels. Now in both the PBT and the FET, high fields occur both upstream and downstream of the base and drain contacts, and these fields are high enough to cause avalanching. However, at a given value of bias the fields at the base contact corners of the PBT are lower than those at the gate contact of the FET.

The distribution of potential and charge along the central line of symmetry is shown in Figure 2. First we note the hole distribution, and point out that the holes drift in the direction of decreasing potential. Thus they drift toward the emitter contact, and tend to pile up both upstream and downstream of the base contact. We note that the largest concentration of holes occurs upstream from the base contact. These holes (as well as electrons) are generated at the sides upstream and downstream from the base contact, travel laterally around the base where the fields are reduced below the contact corner fields. Holes are also generated in the channel opening. The increase in electron density along the line of symmetry is a consequence both of avalanching and current continuity. Downstream from the base contact there is precipitous drop in excess holes. This is not to be construed as a reduction in hole generation. Rather, the holes generated near the base corners are swept around to the emitter contact. Similar arguments apply to the generated electrons. Qualitatively similar behavior is seen at higher bias levels.

Breakdown calculations under an ONR study were performed for the FET shown in Figure 3, with the indicated dimensions. The current voltage relation for this structure is shown in Figure 4. The lower breakdown voltage is apparent and supports the idea that the PBT breakdown characteristics, while obtained at only one base voltage, show very promising expectations vis-a-vis coplanar FETs.

REFERENCES

1.   G.D. Alley, IEEE Trans Electron Devices <u>ED-30</u>, 52 (1983)
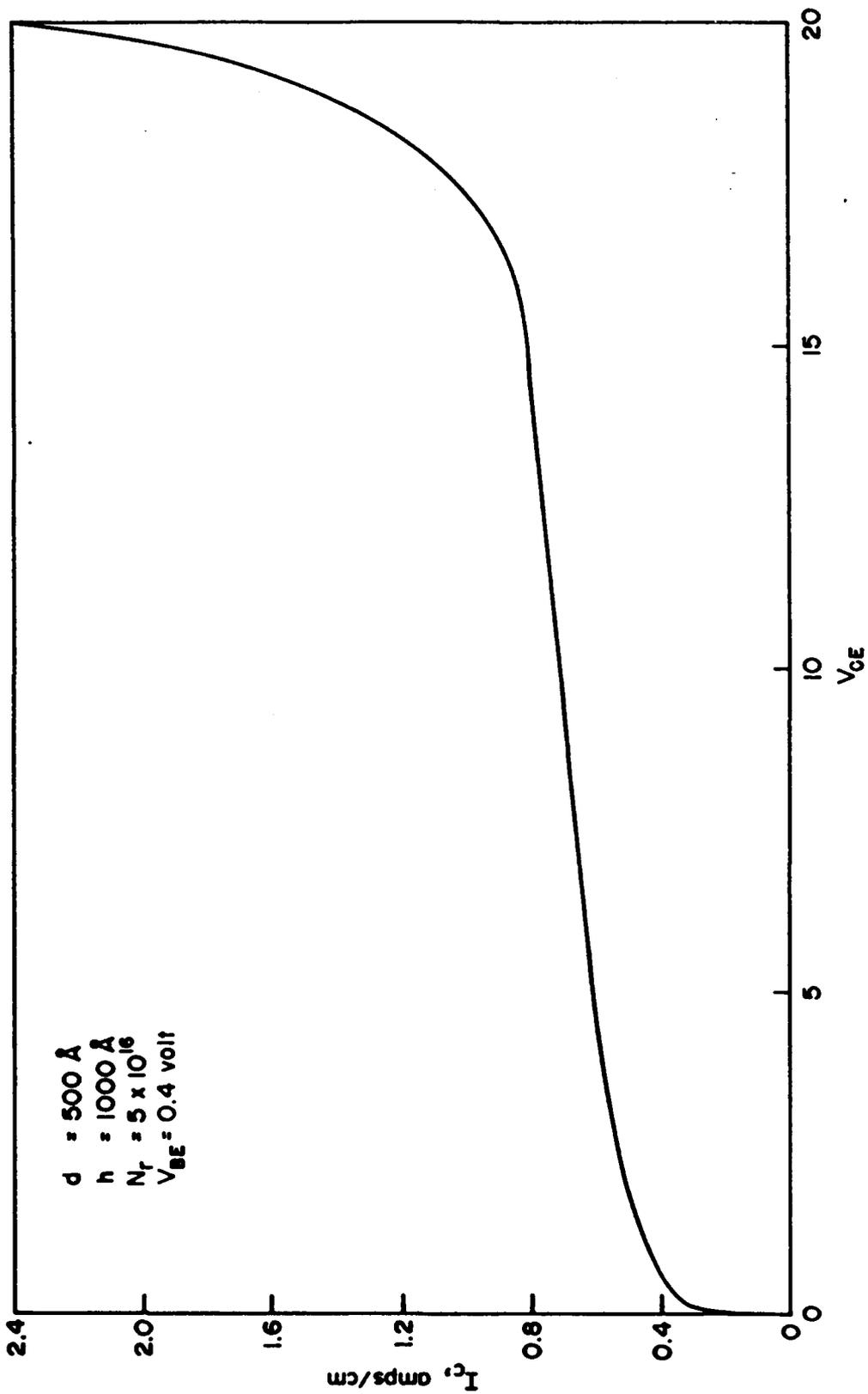
Figure 1. Breakdown voltage calculations for the narrow base penetration-high doping density BPT.
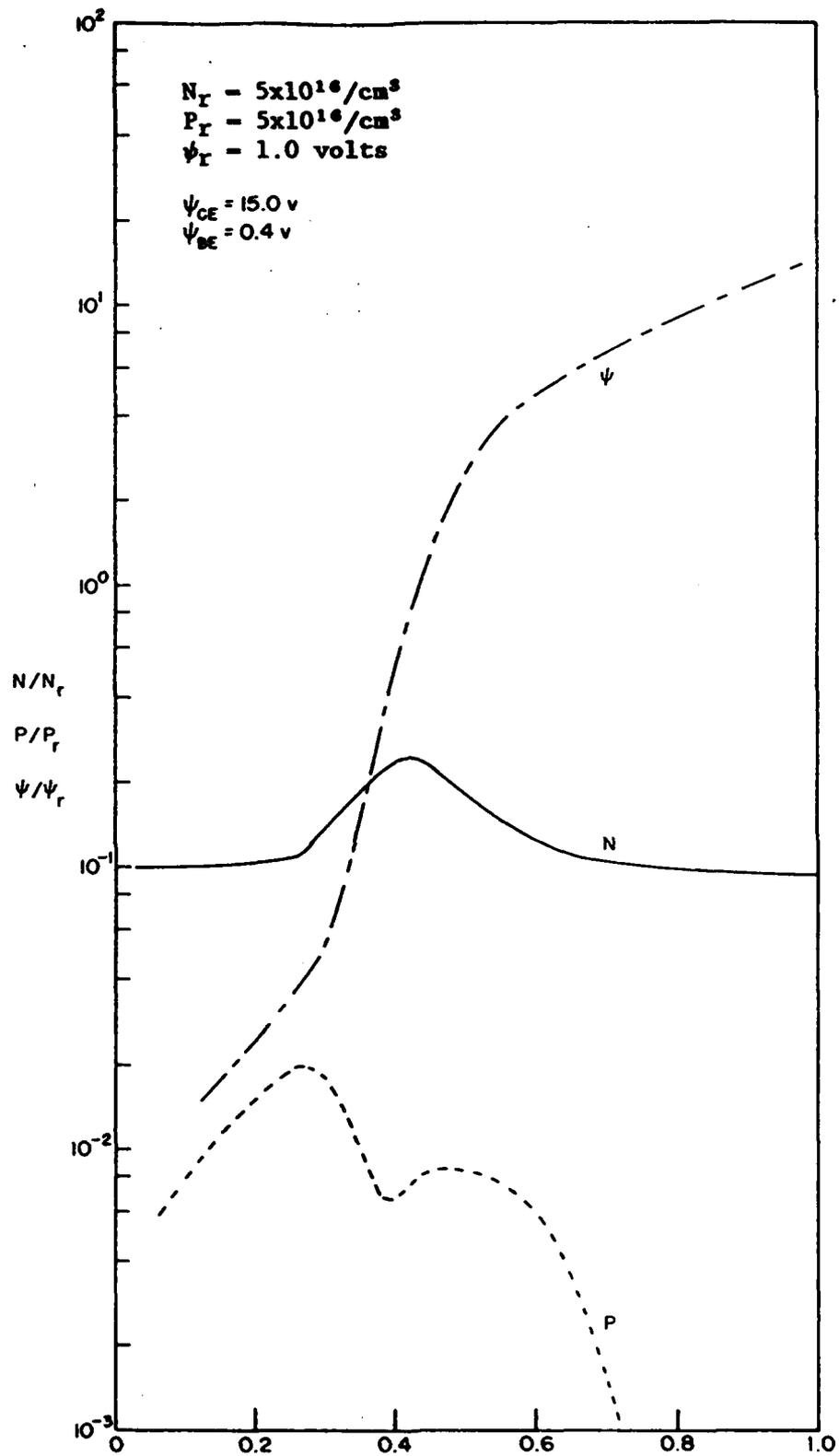
Figure 2. Distribution potential and carrier density along a line of symmetry.
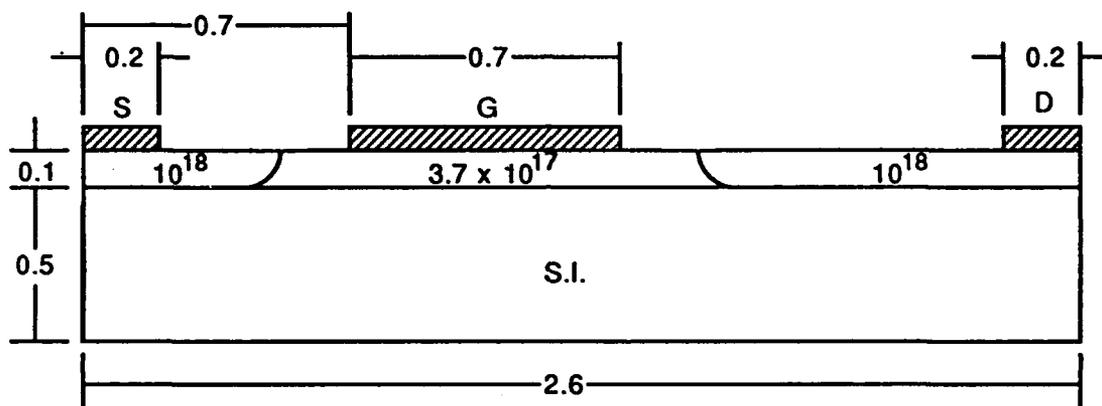
Figure 3. GaAs FET for which breakdown studies were performed. Dimensions are in microns. Note: the active region is approximately 1.2 microns long.
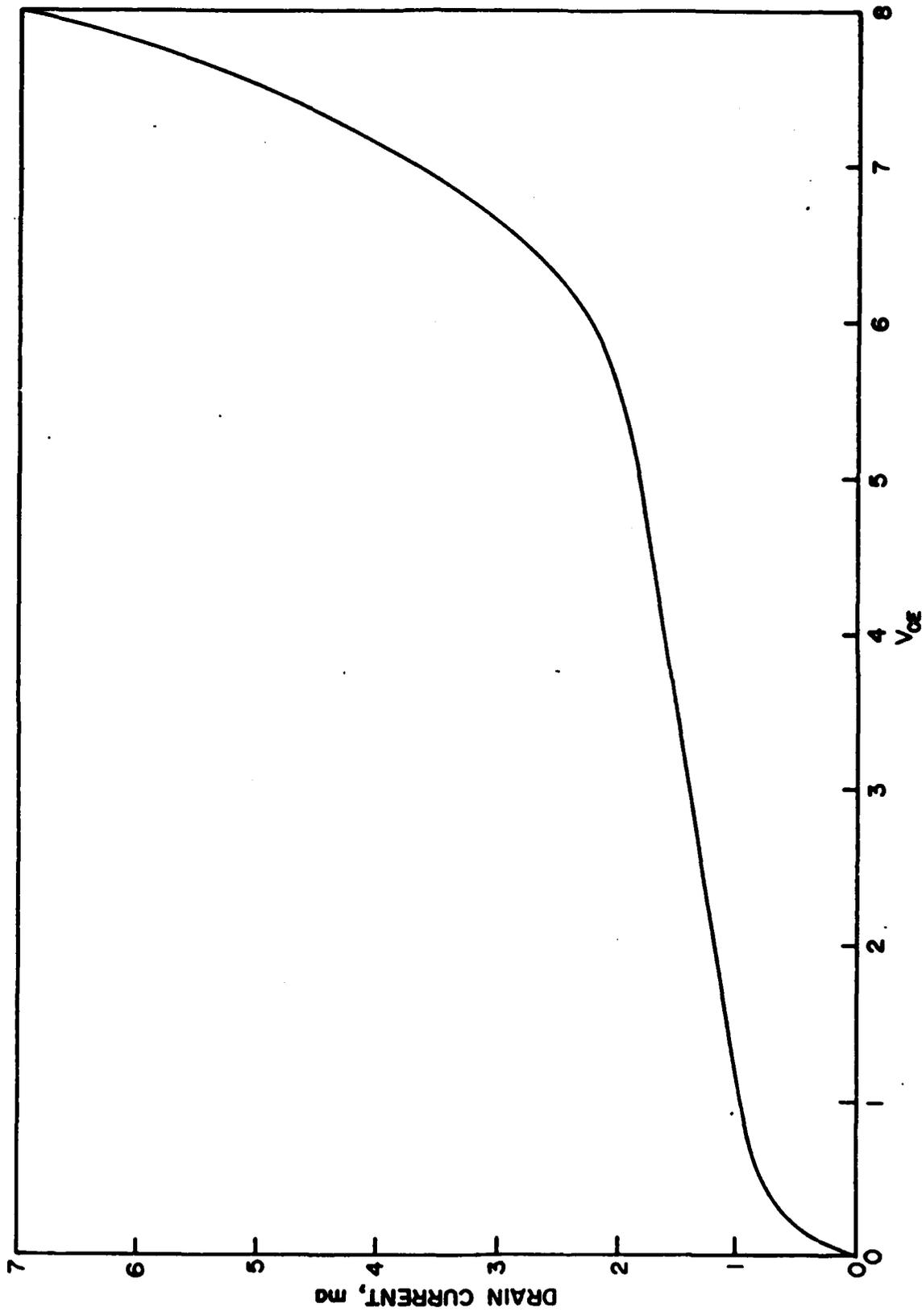
Figure 4   Current voltage characteristics for the structure of Figure 3, with $V_g = .4v.$

E. PBT SIMULATIONS USING THE MOMENTS
OF THE BOLTZMANN TRANSPORT EQUATION

Under the Phase I effort of the present study it was demonstrated that
numerical solutions to the first three moments of the Boltzmann transport
equation, for two level transfer, could be obtained [1]. However, due to
problems with convergence in the region of the tip of the base contact, it was
necessary to utilize a very small time step in the pseudo time marching
algorithm of the solution to steady state. As a result, a large number of
time steps had to be taken to reach a steady state. This made solutions
computationally intensive and, as a consequence excessively expensive.
Additionally, since the time steps were so small, the solution changed very
little over each time step, thereby making it extremely difficult to judge
true convergence.

As a result of these combined problems of slow convergence and excessive
costs, it was proposed that under the Phase II Contract an attempt would be
made to speed the convergence rate of MBTE code. Since the convergence
problems seemed to be isolated to the region of the tip of the base contact,
i.e., re-entrant corner points, one proposed solution to the problem was to
introduce a boundary fitted, nonorthogonal coordinate system. In such a
coordinate system the plane of symmetry which passes through the base and the
surface of the base contact are treated as a single coordinate line. The
sharp corners at the tip of the base contact are replaced by a smooth, but
very small radius continuous curve. Since the sharp corners of the base
contact would be removed, and since the convergence problems exhibited under
the Phase I effort were a result of the corners, it was argued that
convergence could improve. However, after careful consideration this approach
was abandoned for several reasons. First, an accurate representation of the
base contact could not be obtained if the radius of curvature of the
coordinate line representing the boundary surface was to remain within
acceptable computational limits. Second, the skewing of the coordinate lines
could introduce more severe problems than those which were present
originally. Third, and perhaps most significant, the additional number of
grid points required to resolve the shape curvature in the region of the base
tip and resulting increased computational effort per iteration would likely
offset any increase in convergence rate thus obtained.

An alternate approach to speeding the convergence of the MBTE solutions is to
spatially scale the time step; i.e., the time step varies at each grid point
in the solution domain in such a manner that a nearly optimum time step
(measured against convergence rate) is used everywhere. Such a procedure has
been routinely used in obtaining solution to the Navier-Stokes equations by
researchers at SRA [B] and the procedure has been successfully applied to FET
calculation performed using the MBTE by the present authors as well. However,
this procedure is based on the mathematical characteristics of the governing
equations and, while applicable to the PBT problem, its implementation did not
improve the convergence rate of the PBT solutions.

This indicated that the convergence problems experienced under the Phase I

effort were directly related to the PBT structure and, in particular, the protruding base contact. It should be noted that similar problems were not encountered with the drift and diffusion PBT simulations reported here, thus the problem was thought to lie in the interaction of the discretization technique and/or boundary conditions applied to the MBTE and the re-entrant corner points on the base contact. As will be discussed subsequently, the convergence problems in the corner region of the base contact were eliminated through a reformulation of the difference approximations and a treatment of boundary conditions that was consistent with this reformulation.

THE GOVERNING EQUATIONS

The governing utilized in this study are the first three moments of the Boltzmann transport equation. These equations are derived for two different species of carriers, in this case, satellite valley and central valley carriers in gallium arsenide with the parameter established in Table 1. The governing equations are obtained by taking the collisional invarient moments [3] of the Boltzmann transport equation, via., the moments with respect to the mass, momentum and energy of the two carriers. This yields a set of governing equations which are similar in form to the equations utilized for two-phase fluid dynamic flow. The governing equations reflect the conservation laws of mass, momentum and energy for the two carriers and are often referred to as the moments of the Boltzmann transport equation (MBTE).

Using vector notation, the two particle conservation equations can be expressed as

$$\frac{\partial n_1}{\partial t} = -\nabla \bullet (n_1 V_1) - n_1 f_1 + n_2 f_2 \qquad (1)$$

and

$$\frac{\partial n_2}{\partial t} = -\nabla \bullet (n_2 V_2) + n_1 f_1 - n_2 f_2 \qquad (2)$$

where n_1 and n_2 are the satellite valley and central valley carrier number densities respectively while V_1 V_2 are the corresponding velocities. f_1 and f_2 are the corresponding scattering integrals for particle conservation and in general are functions of the corresponding carrier temperature and momenta [3]. For the purposes of this study, the dependence of all scattering integals on momenta is neglected. Defining a total or global number density, n, by

$$n = n_1 + n_2 \qquad (3)$$

and adding equation (1) to equation (2) yields the equivalent relationship, i.e., the global continuity equation

$$\frac{\partial n}{\partial t} = -\nabla \bullet [n_1 V_1 + (n - n_1) V_2] \qquad (4)$$

107

Conservation of momentum for the satellite valley carrier can be expressed as

$$\frac{\partial}{\partial t}(n_1 P_1) = -\nabla \bullet (n_1 V_1 P_1) - \nabla \wp_1 - \nabla \bullet \sigma_1 - n_1 e F - n_1 P_1 f_3 \tag{5}$$

where the momentum, P_1, and the electric field, F, are defined by

$$P_1 = m_1 V_1 \tag{6}$$

$$F = -\nabla \psi \tag{7}$$

where m_1 is the mass of the satellite valley carrier, e is the electronic charge and ψ is the electric potential. The partial pressure, $\wp_1$ (7), is related to the satellite valley carrier temperature, T_1, and number density by the perfect gas relationship

$$\wp = n_1 k_1 T_1 \tag{8}$$

where k is Boltzmann's constant. f_3 is the scattering integral for the satellite valley carrier momentum. The term $\nabla : r_1$ represents the stress forces. In this study, the stress tensor, σ_1, is approximated by the Stokes relationship

$$\sigma_1 = -\mu_1 \left[\nabla V_1 + \nabla V_1^T - \frac{2}{3} \nabla \bullet V_1 \right] \tag{9}$$

where μ_1 is the viscosity associated with the satellite valley carriers, and V_1^T is the transpose of the velocity vector. Substitution of equation (6) and equation (7) into equation (5) and dividing by m_1 yields the final form of the momentum equation for the satellite valley carriers:

$$\frac{\partial}{\partial t}(n_1 V_1) = -\nabla \bullet (n_1 V_1 V_1) - \frac{\nabla \wp_1}{m_1} - \frac{\nabla \bullet \sigma_1}{m_1} + \frac{n_1}{m_1} e \nabla \psi - n_1 V_1 f_3 \tag{10}$$

It is observed that equation (10) is a vector equation and hence can in general be considered as ℓ independent equations where ℓ is the number of relevant physical dimensions.

An analogous derivation for the central valley carriers plus the identity of equation (3) yields the momentum equation for the species Γ valley species:

$$\frac{\partial}{\partial t}\left[(n-n_1)V_2\right] = \nabla \bullet \left[(n-n_1)V_2 V_2\right] - \frac{\nabla \wp_2}{m_2} - \frac{\nabla \bullet \sigma_2}{m_2} + (n-n_1)\frac{e}{m_2}\nabla \psi - (n-n_1)V_2 f_4 \tag{11}$$

where

$$\mathcal{P}_2 = (n-n_1)kT_2 \tag{12}$$

and

$$\sigma_2 = -\mu_2(\nabla V_2 + \nabla V_2^T - \tfrac{2}{3}\nabla \bullet V_2) \tag{13}$$

m_2 is the mass of a central valley carrier and μ_2 is the corresponding viscosity coefficient. f_4 is the scattering integral for central valley carrier temperature, T_2 only.

There are various forms in which the satellite valley and central valley carrier energy equations can be described. This study chooses to cast the energy equations in term of the satellite and central valley temperatures, T_1 and T_2. As with the momentum equation, we start with the satellite valley carrier energy equation. This equation can be expressed as a balance between the time rate of change of the total energy (intregal plus kinetic energy), the convection of that energy, the pressure, stress and electric field work, the heat conduction and the production and depletion of energy due to the transfer of satellite valley carriers to the central valley carriers and vice versa. Mathematically, the above physical statement can be expressed as

$$\frac{\partial}{\partial t}\left[n_1\left(m_1\frac{V_1 \bullet V_1}{2} + u_1\right)\right] = -\nabla\bullet\left[n_1 V_1\left(m_1\frac{V_1 \bullet V_1}{2} + u_1\right)\right] -\nabla\bullet(\mathcal{P}_1 V_1) -\nabla\bullet(\sigma_1 \bullet V_1) + n_1 e\nabla\psi\bullet V_1$$

$$+\nabla\bullet(\tilde{\kappa}_1\nabla T_1) - \frac{3}{2}k\left[n_1 T_1 f_6 -(n-n_1)T_2 f_6\right] \tag{14}$$

where U_1 is the specific internal energy of the satellite valley carriers, κ_1 is the corresponding thermal conductivity and f_6 and f_6 are the scattering integrals for the satellite valley energy equation (which are again assumed to be functions of temperature only). To obtain the static temperature version of equation (14) requires the elimination of the kinetic terms $V_1.V_1/2$ of equation (14) through the use of the mechanical energy equation. The mechanical energy equation is obtained by dotting the satellite valley carrier velocity, V_1, with the satellite valley carrier momentum equation, applying the vector identities

$$V_1 \bullet \nabla(n_1 V_1 V_1) = \nabla\bullet\left[\left(n_1\frac{V_1 \bullet V_1}{2}\right)V_1\right] + \frac{V_1 \bullet V_1}{2}\nabla\bullet(n_1 V_1) \tag{15}$$

$$V_1\bullet\frac{\partial}{\partial t}(n_1 V_1) = \frac{\partial}{\partial t}\left[n_1\frac{V_1 \bullet V_1}{2}\right] + \frac{V_1 \bullet V_1}{2}\frac{\partial n_1}{\partial t} \tag{16}$$

109

and the satellite valley particle conservation equation, equation (1). This yields upon multiplying by m_1

$$\frac{\partial}{\partial t}\left[m_1 n_1 \frac{V_1 \bullet V_1}{2}\right] = -\nabla \bullet \left[(m_1 n_1 \frac{V_1 \bullet V_1}{2})V_1\right] - V_1 \bullet \nabla p_1 - V_1 \bullet (\nabla \bullet \sigma_1) + n_1 e V_1 \bullet \nabla \psi$$

$$+ \frac{V_1 \bullet V_1}{2} m_1 \left[n_1(2f_3 - f_1) - (n - n_1)f_2\right] \tag{17}$$

Equation (17) is next subtracted from equation (14) and with the use of the vector identities

$$\nabla \bullet (\sigma \bullet V_1) = V_1 \bullet (\nabla \bullet \sigma_1) + \sigma_1 : \nabla V_1 \tag{18}$$

and

$$\nabla \bullet (\mathcal{P}_1 V_1) = V_1 \bullet \nabla \mathcal{P}_1 + \mathcal{P}_1 \nabla \bullet V_1 \tag{19}$$

the internal energy version of the energy is obtained:

$$\frac{\partial}{\partial t}(n_1 u_1) = \nabla \bullet (n_1 V_1 u_1) - \mathcal{P}_1 \nabla \bullet V_1 - \sigma_1 : \nabla V_1 + \nabla \bullet (\tilde{\kappa}_1 \nabla T_1)$$

$$+ \frac{V_1 \bullet V_1}{2} m_1 \left[n_1(2f_3 - f_1) - (n - n_1)f_2\right] - \frac{3}{2}k\left[n_1 T_1 f_3 - (n - n_1)T_2 f_6\right] \tag{20}$$

It is to be noted that the electric field term has disappeared in this form of the equation; in addition, the term $\sigma_1 : \nabla V_1$ is often called the dissipation term. Finally, to obtain the static temperature version of the satellite valley carrier energy equation, the relationship between internal energy and static temperature

$$u_1 = \frac{3}{2}kT_1 \tag{21}$$

and the equation of state, equation (8) must be utilized. This yields upon multiplying by 2/3k the static temperature energy equation.

$$\frac{\partial}{\partial t}(n_1 T_1) = -\nabla \bullet (n_1 V_1 T_1) - \frac{2}{3}n_1 T_1 \nabla \bullet V_1 - \frac{2}{3k}\sigma_1 : \nabla V_1 + \frac{2}{3k}\nabla \bullet (\tilde{\kappa}_1 \nabla T_1)$$

$$+ 3V_1 \bullet V_1 m_1 \left[n_1(2f_3 - f_1) - (n - n_1)f_2\right] - n_1 T_1 f_5 - (n - n_1)T_2 f_6 \tag{22}$$

110

In an analogous manner the energy equation can be derived for the central valley carrier. The results are

$$\frac{\partial}{\partial t}\left[(n-n_1)T_2\right] = -\nabla\bullet\left[(n-n_1)V_2T_2\right] - \frac{2}{3}(n-n_1)T_2\nabla\bullet V_2 - \frac{2}{3k}\sigma_2:\nabla V_2 + \frac{2}{3k}\nabla\bullet(\tilde{\kappa}_2\nabla T_2)$$

$$+ 3V_2\bullet V_2 m_2\left[(n-n_1)(2f_4-f_2) + n_1 f_1\right] + n_1 T_1 f_8 - (n-n_1)T_2 f_7 \qquad (23)$$

where f_2 and f_8 are the scattering integrals for the central valley energy equation and κ_2 is the thermal conductivity.

The partial differential equations, equations (1), (4), (10), (11), (22) and (23) with equations of state, equations (8) and (12), and the constitutive relations for the stress tensors, equations (9) and (13), constitute the governing equations utilized in this study. Equations (8), (9), (12) and (13) can be substituted into the six partial differential equations, thus relating the dependent variables utilized in this study, viz. n_1, n, V_1, V_2, T_1, T_2. In addition, a governing equation must be supplied for the electric potential ψ as the term occurs in both moment equations.

The electric potential is related to the total number density through a Poisson's equation

$$\nabla^2\psi = \frac{e}{\epsilon}(n-N_0) \qquad (24)$$

where N_0 is the doping number density (given), ϵ is the permitivity and e the electron charge. It is to be noted that the particle conservation equations, the momenta equations and the energy equations all have time (rate) terms while the Poisson's equation for electric potential has no such term. In addition, if we restrict ourselves to two dimensional problems, it is seen that the governing partial differential equations, equation (1), (4), (10), (11), (22), (23) and (24) relate nine independent variables for nine equations (the momentum equations are vector equations and in two dimension are each two separate, independent equations). This governing system of equations is coupled and many of the terms are nonlinear. As a result, no general closed form solution is know. Thus, numerical solution techniques must be implemented to obtain solutions under arbitrary conditions.

The numerical method used in the solution of the governing equations, which were derived in the previous section, is based on an application of consistently split, linearized, block implicit (LBI) methods as developed by Briley and McDonald [4,5]. LBI methods have been highly successful in the field of computational fluid dynamics (CFD) where they have been applied in obtaining solutions to a closely related system of governing equations, the Navier-Stokes equations, [c.f. Ref. 4]. Thus, application of such methods to solutions of the moments of the Boltzmann transport equation can draw on a vast amount of related experience generated using LBI techniques.

LBI techniques center about the use of a formal linearization procedure in which systems of coupled nonlinear PDE's in one space dimension are reduced to a system of linear equations, which upon application of spatial differencing, may be expressed as a block coupled matrix system. The resulting system may then be solved efficiently, without iteration, to advance the solution in time. .Steady solutions are obtained as the long-term asympotic solution. The benefits of the procedure are retained for multidimensional problems through application of ADI schemes in their natural extension to block coupled systems. The ADI procedures reduce the multidimensional system of equations, having broad-banded matrix structures to systems of one-dimensional equations with narrow block-banded structures which are solved efficiently using fundamental block-elimination methods.

Bri'.ey and McDonald [4] considered the coupled system of nonlinear, time-dependent, multidimensional equations given by

$$\frac{\partial H(\phi)}{\partial t} = D(\phi) + S(\phi) \tag{25}$$

In equation (25), ϕ represent the vector of dependent variables $\phi = (n, n_1, V_1, B_2, T_1, T_2, \psi)^T$, $H(\phi)$ and $S(\phi)$ are nonlinear functions of ϕ, and $D(\phi)$ is a general, nonlinear, multidimensional, partial differential operator. Equation (25) is first time differenced about $t^n + \beta \Delta t$

$$\frac{H^{n+1} - H^n}{\Delta t} = \beta(D^{n+1} + S^{n+1}) + (1-\beta)(D^n + S^n) \tag{26}$$

where $\Delta t = t^{n+1} - t^n$. The parameter $\beta = 1$ for a fully implicit scheme or $\beta = 0.5$ for the Crank-Nicolson formulation. The implicit level nonlinear operators H, D and S are then formally linearized using a Taylor series expansion about the explicit time level. For example

$$H^{n+1} = H^n + \left[\frac{\partial H}{\partial \phi}\right]^n (\phi^{n+1} - \phi^n) + O(\Delta t^2) \tag{27}$$

Equation (26) may then be expressed at each grid point in the solution domain

112

as a matrix equation of the form

$$(A - \beta\Delta t L^n)(\phi^{n+1} - \phi^n) = \Delta t(D^n + S^n) \tag{28}$$

where

$$A = \left(\frac{\partial H}{\partial \phi}\right)^n - \beta\Delta t\left(\frac{\partial S}{\partial \phi}\right)^n \tag{29}$$

and

$$L^n = \left(\frac{\partial D}{\partial \phi}\right)^n \tag{30}$$

As a result, the nonlinear, coupled system of PDE's given by equation (25) has been reduced to a block coupled, linear system of temporal difference equations, equation (28), which upon spatial differencing, need only be solved once per time step to obtain a solution. Additionally, since the linearization error is at worst of the same order as the temporal discretization error, the linearization is not expected to introduce significant inaccuracies.

Application of equation (28) to second order PDE's in one space dimension, using standard three-point spatial difference approximations requires the solution of one block tridiagonal system per time step. Such a system can be solved efficiently using standard block tridiagonal elimination procedures. However, application of the LBI algorithm given by equations (28) through (30) to multidimensional problems results in the loss of this narrow, block banded matrix structure. The discretization of the multidimensional spatial operator results in a broad-banded matrix structure, which, if solved by direct or iterative methods, can be extremely inefficient. Such observations led Briley and McDonald [4,5] to develop consistently split LBI algorithms for multidimensional problems. The splitting is accomplished by dividing the multidimensional spatial operator, L, into one dimensional operators associated with each coordinate direction.

$$L = L_1 + L_2 + L_3 \tag{31}$$

Equation (28) is then split following the scalar ADI development of Douglas and Gunn [6].

$$(A - \beta\Delta t L_1^n)(\phi^* - \phi^n) = \Delta t(D^n + S^n) \tag{32a}$$

$$(A - \beta\Delta t L_2^n)(\phi^{**} - \phi^n) = A(\phi^* - \phi^n) \tag{32b}$$

$$(A - \beta\Delta t L_3^n)(\phi^{***} - \phi^n) = A(\phi^{**} - \phi^n) \tag{32c}$$

Here $\Delta\phi^*$, $\Delta\phi^{**}$, $\Delta\phi^{***}$, are intermediate solutions of equations (32a) through (32c). Again, if three-point operators are used to approximate

113

the spatial operators, L_i, each of equations (32a) through (32c) will be block-tridiagonal and can be efficiently solved. The block size and band width are independent of the number of grid points, hence the computational effort required to solve the sequence varies linearly with the total number of grid points regardless of the number of space dimensions considered. For two dimensional problems, equation (32c) is omitted. Elimination of the intermediate steps in equations (32a) through (32c) yields

$$(A - \beta\Delta t L_1^n)A^{-1}(A - \beta\Delta t L_2^n)A^{-1}(A - \beta\Delta t L_3^n)(\phi^{n+1} - \phi^n) = \Delta t(D^n + S^n) \tag{33}$$

thus

$$\phi^{n+1} = \phi^{***} + O(\Delta t^*)^2 \tag{34}$$

The development given above presents a brief outline of the LBI method used in the present investigation. A more detailed development, as well as in-depth discussion of LBI methods, the linearization procedure and related topics may be found in the article by Briley and McDonald [4].

Application of LBI procedures to equations (1), (4), (10), (11), (22), (23) and (24), in two space dimensions give rise to an A matrix of the form

$$
A = \left|
\begin{array}{cccccccc|c}
X & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & X & X & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & X & 0 & X & 0 & 0 & 0 & 0 & 0 \\
X & X & 0 & 0 & X & 0 & 0 & 0 & 0 \\
X & X & 0 & 0 & 0 & X & 0 & 0 & 0 \\
0 & X & 0 & 0 & 0 & 0 & X & 0 & 0 \\
\underline{X} & \underline{X} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{X} & \underline{0} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{array}
\right| \tag{35}
$$

This matrix, due to the absence of a time derivative in Poisson's equation, is singular, thus equation (28) cannot be split following equation (32). However, the partitioning of the A matrix indicated by the dotted lines in equation (35) suggests that if the L matrix (the linearized D-operator) could be similarly partitioned, Poisson's equation could be decoupled from the rest of the system. The remaining coupled equations could then be solved by a direct application of the LBI method outlined, followed by the solution of Poisson's equation, completing a time step. To accomplish the decoupling of Poisson's equation, it is only necessary to lag the electric field terms (the potential gradient) which appear in the momentum equations. While this formally reduces the accuracy of the temporal integration to $O(\Delta t)$, it does not adversely effect the stability of the solution algorithm. A similar procedure has been employed by Kreskovsky and Grubin [7] in solving the

114

semiconductor drift and diffusion equations. Solution of Poisson's equation is performed using a scalar ADI procedure with cycled acceleration parameters [4,6]. The overall solution algorithm proceeds as follows:

1. Initial and boundary conditions are specified for all variables throughout the computational domain.
2. The continuity, momentum and energy equation are solved using an LBI scheme to advance the carrier densities, velocities and temperatures form time t^n to time $t^{n+\Delta t}$.
3. Poisson's equation is solved using the carrier densities at time $t^{n+\Delta t}$ to obtain the advanced time potential distribution.
4. Steps 2 and 3 are repeated until a steady solution is reached or until the calculation is terminated.

SPATIAL DIFFERENCE APPROXIMATIONS

In the original MBTE code, as used under the Phase I program, the spatial difference approximations were obtained by a straight forward application of three point central difference approximations to the conservative form of the governing equations. Numerical or artificial diffusion was then introduced as required to stabilize the scheme in regions of extremely steep gradients which are under-resolved. Such a procedure has been routinely and successfully applied to the Navier-Stokes equations, and also worked well when applied to the MBTE simulation of FET's.

To examine why such a procedure did not work well for the PBT, consider the one dimensional, scalar equation

$$\frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial X} = 0 \tag{36}$$

On an equally spaced mesh, the central difference approximation used for the spatial derivative is given as

$$\frac{\phi_{i+1} - \phi_{i-1}}{2\Delta X} = \frac{\partial \phi}{\partial X} + \frac{\Delta X^2}{6} \frac{\partial^3 \phi}{\partial X^3} + \cdots \tag{37}$$

On an unequally spaced mesh, the approximation is generalized to

$$\frac{\Delta X_- \phi_{i+1} - (\Delta X_- - \Delta X_+)\phi_i - \Delta X_+ \phi_{i-1}}{2\Delta X_+ \Delta X_-} = \frac{\partial \phi}{\partial X} + \frac{(\Delta X_+ - \Delta X_-)}{4} \frac{\partial^2 \phi}{\partial X^2}$$

$$+ \frac{(\Delta X_+^2 + \Delta X_-^2)}{12} \frac{\partial^3 \phi}{\partial X^2} + \cdots \tag{38}$$

where $\Delta X_+ = X_{i+1} - X_i$, and $\Delta X_- = X_i - X_{i-1}$. However, if the difference approximations to the spatial derivative is obtained from a

115

control volume approach, then a strictly conservative approximation is given by

$$\frac{2(\phi_{i+1/2} - \phi_{i-1/2})}{(\Delta X_+ + \Delta X_-)} - \frac{\partial \phi}{\partial X} \tag{39}$$

If $\phi_{i\pm1/2}$ is computed from simple averages of ϕ_i and $\phi_{i\pm1}$, then this approximation reduces to

$$\frac{\phi_{i+1} - \phi_{i-1}}{(\Delta X_+ + \Delta X_-)} - \frac{\partial \phi}{\partial X} \tag{40}$$

It is easily shown that the accuracy of this approximation is given by

$$\frac{\phi_{i+1} - \phi_{i-1}}{(\Delta X_+ + \Delta X_-)} - \frac{\partial \phi}{\partial X} + \frac{(\Delta X_+ - \Delta X_-)}{2} \frac{\partial^2 \phi}{\partial X^2} + \frac{(\Delta X_+^3 + \Delta X_-^3)}{6(\Delta X_+ + \Delta X_-)} \frac{\partial^3 \phi}{\partial X^3} + \cdots \tag{41}$$

Several observations relative to equations (37), (38) and (41) are in order. First, equations (38) and (41) have the same order truncation error. Second, equations (38) and (41) reduced to equation (37) on an equally spaced mesh, indicating that equation (37) is a strictly conservative approximation in itself. Third, the generalization of equation (37) to equation (38) does not retain this strictly conservative property. As a result, when applied to the continuity equations, for example, local sources and sinks may be erroneously introduced in regions of unequal mesh spacing. This explains, in part, the poor convergence behavior observed under the Phase I effort. Approximations consistent with those given by equation (41) were implemented under the Phase II program.

The implementation of difference approximations consistent with equation (41) also requires a consistent treatment of boundary conditions at the contacts, as well as on free surfaces. The control volume boundaries lie halfway between the grid points. Thus, if the device surfaces are taken to be coincident with the mesh lines, special approximation must be developed at the boundaries to account for half cells. Alternately, the device surface can be taken to be coincident with the control volume boundaries and the boundary conditions specified in such a manner to assure that appropriate fluxes at the boundaries are specified. For example, a zero flux condition at $i-1/2$ requires that $\phi_{i-1/2} = 0.0$. Since $\phi_{i-1/2}$ is computed using an average of ϕ_i and ϕ_{i-1} the imposed boundary conditions require that $\phi_{i-1} = -\phi_i$.

This approach to boundary conditions was used in the present simulations along with the difference technique outlined previously. As a result, the convergence problems experienced under the Phase I effort were eliminated. While the cost of the MBTE simulations are significantly higher that those using the drift and diffusion equation, considerably more insight into PBT

116

operation is obtained from such simulations. Presently, MBTE .simulations of the PBT using the procedure described above require an effort of the same order as that of MBTE simulations of simpler FETs with a similar mesh.

COMPUTATIONAL RESULTS USING THE MOMENTS OF THE BOLTZMANN TRANSPORT EQUATION

The moments of the Boltzmann transport equation were implemented for GaAs with two level transfer, with parameters listed in Table 1. For undoped GaAs, and uniform fields, the resulting field dependent velocity, carrier density in the gamma and L valleys, mean velocity in the gamma and L valleys, and electron temperature in the gamma and L valleys are shown in figure 1. In the calculation ionized impurity scattering was introduced. This yielded a low field mobility of approximately 4500cm2/V-s, and a peak velocity of 1.88x10⁷cm/sec. The results of the calculation were presented at several conferences. Copies of these papers are included in the appendices, and are briefly discussed below.

The nonequilibrium calculations were performed for the structure shown in figure 2. It is the structure for which most of the detailed calculations with the drift and diffusion equations were performed.

The most significant result for the nonequilibrium calculation is the current-voltage curves. These are displayed in figure 3, which also includes, for comparison calculations using the drift and diffusion equations. The calculation using the moments of the Boltzmann transport equation are denoted by the closed circles. We point out that the a constraint of zero current on the base contact was imposed. This constraint was not imposed on the drift and diffusion equation calculations. Further, the results for the MBTE were extrapolated to the origin, as indicated by the long broken lines.

The results of figure 3 show the presence of current levels that are approximately three to four times greater than that associated with he drift and diffusion equation calculations. This as may be anticipated would arise from nonequilibrium velocity overshoot. In addition, for the high current levels, the presence of negative forward conductance is also observed. This negative forward conductance arises solely from electron transfer, and is not thermal in origin. It is important to note that such structure is not predicted in the drift and diffusion equation calculations which includes electron transfer, but only for uniform field values. It is important to note that negative forward conductance is observed experimentally.

The details of the calculation reveal some very interesting features of the device physics arising from the MBTE calculations. These features are discussed through comparison with the drift and diffusion equation solutions. To begin, figure 4 displays the potential distribution along the line of symmetry indicated in bold in the inset. For a one volt potential difference between the emitter and collector, the field changes more gradually in the MBTE calculation than in the DDE calculation. For example in the DDE calculation the one volt potential difference falls across 2000A for an average field of approximately 50kv/cm. The average field in the MBTE

117

calculation is low, and approximately 20kv/cm.

The distribution in charge arising from these field differences is shown in figure 5. For the DDE calculation we note the appearance of a dipole layer, with charge accumulation beginning very near the upstream base region. In the MBTE calculation there is charge depletion, where there is charge accumulation from the DDE calculation. There is additional structure in the MBTE calculation showing a small amount of charge accumulation followed by a local region of charge depletion. The details of the structural differences are important, for we note that along the line of symmetry the gradient of potential normal to the line of symmetry is zero.

The details of the potential and charge distribution indicate the following: Initially for both the DDE and the MBTE the potential is undergoing an increase at a rate that is approximately linear. The potential increase in the vicinity of the base contact shows a positive curvature for both the DDE and the MBTE. This positive curvature indicates charge accumulation for the DDE and the MBTE. As the details of the calculation reveal, there is a small amount.of charge accumulation in the MBTE. Depletion is accompanied by a change in curvature of the potential. This change in curvature occurs earlier for the MBTE than for the DDE, occuring at approximately 35 microns in the former, and at approximately 0.45 microns for the DDE. There are subsequent curvature changes for the MBTE but not for the DDE, as indicated in figures 4 and 5.

It is important to note that initially the carrier distribution is all gamma valley carriers. Electron transfer is not an issue until half way downstream of the emitter contact. Indeed, for approximately 20% of the line of symmetry only 20% of the total number of carriers are gamma valley carriers.

Now the constraints of current continuity require that a decrease in concentration be accompanied by an increase in velocity. In this regard it is important to note that the L valley carriers sustain very low mobilities and as a result carry very little current. Indeed at a point 0.6 microns downstream from the emitter contact the gamma valley carriers sustain approximately 70% of the total current. Thus the quantity of interest is the gamma valley velocity, as displayed in figure 6. Note that carrier velocities near $7x10^7$cm/sec are possible. In this regard the electron temperature for the gamma valley is also shown (see figure 7). It is noted that the carriers, when reaching the collector contact, are not in equilibrium with the lattice. Rather it is necessary that the carriers travel a further distance in the PBT structure before being collected for them to be in equilibrium with the lattice.

In addition to the above calculations the cutoff frequency was obtained from the ratio of transconductance and capacitance. These results are displayed in figure 8, as a function of base emitter voltage. We note that the bias dependence of the cutoff frequency for a collector-emitter voltage of 0.5v shows an increase in cutoff frequency with increasing voltage. This is qualitatively similar to that obtained using the DDE equation. The situation at higher values of base voltage show a cutoff frequency that decreases with

118

increasing bias. At least part of this decrease must be attributed to compression in the current-voltage relation as displayed in figure 3, which also lead to a negative forward conductance. The remarkable feature of these results is, however, the range of cut off frequencies in excess of 120GHz.

REFERENCES

1. Buggeln, R.C.; Kreskovsky, J.P. and Grubin, H.L., SRA Report R84-91005-F, May 1984.

2. Briley, W.R., Buggeln, R.C. and McDonald, H., AIAA paper No. 85-1520-CP, 1985.

3. Grubin, H.L., Ferry, D.K., Iafrate, G.J., and Baker, J.R., <u>VSLI Electronic</u> (ed. N. Einsprush), <u>3</u>, Academic Press, N.Y. (1982).

4. Briley, W.R. and McDonald, H., J. Comp. Phys., Vol. 24, 372, (1977).

5. Douglas, J. and Gunn, J.E., Numerische Mathematik, Vol. 6, 428, (1964).

6. Kreskovsky, J.P. and Grubin, H.L., J. Comp. Phys., Vol. 68, 420, (1987).

TABLE I			
GALLIUM ARSENIDE			
PARAMETERS AND BOUNDARY CONDITIONS USED IN CALCULATION			

PARAMETERS	Γ	L	COMMON
NUMBER OF EQUIVALENT VALLEYS	I	4	
EFFECTIVE MASS (m_e)	0.067	0.222	
Γ-L SEPARATION (ev)			0.33

POLAR OPTICAL SCATTERING

STATIC DIEL. CONST.	12.90
HIGH FREQ. DIEL. CONST.	10.92
LO PHONON (ev)	0.0354

Γ-L SCATTERING

COUPL. CONS. (ev/cm)	0.800×10^9
PHONON ENERGY (ev)	0.0278

L-L SCATTERING

COUPL. CONST. (ev/cm)	2.0×10^9
PHONON ENERGY (ev)	0.0354

ACOUSTIC SCATTERING

	Γ	L
DEFORM. POT. (ev)	7.0	9.2

NONPOLAR SCATTERING (L)

COUPLING CONSTANT (ev/cm)	0.300×10^9
PHONON ENERGY (ev)	0.0343

Table 1. Scattering parameters used in the Gallium Arsenide calculation.
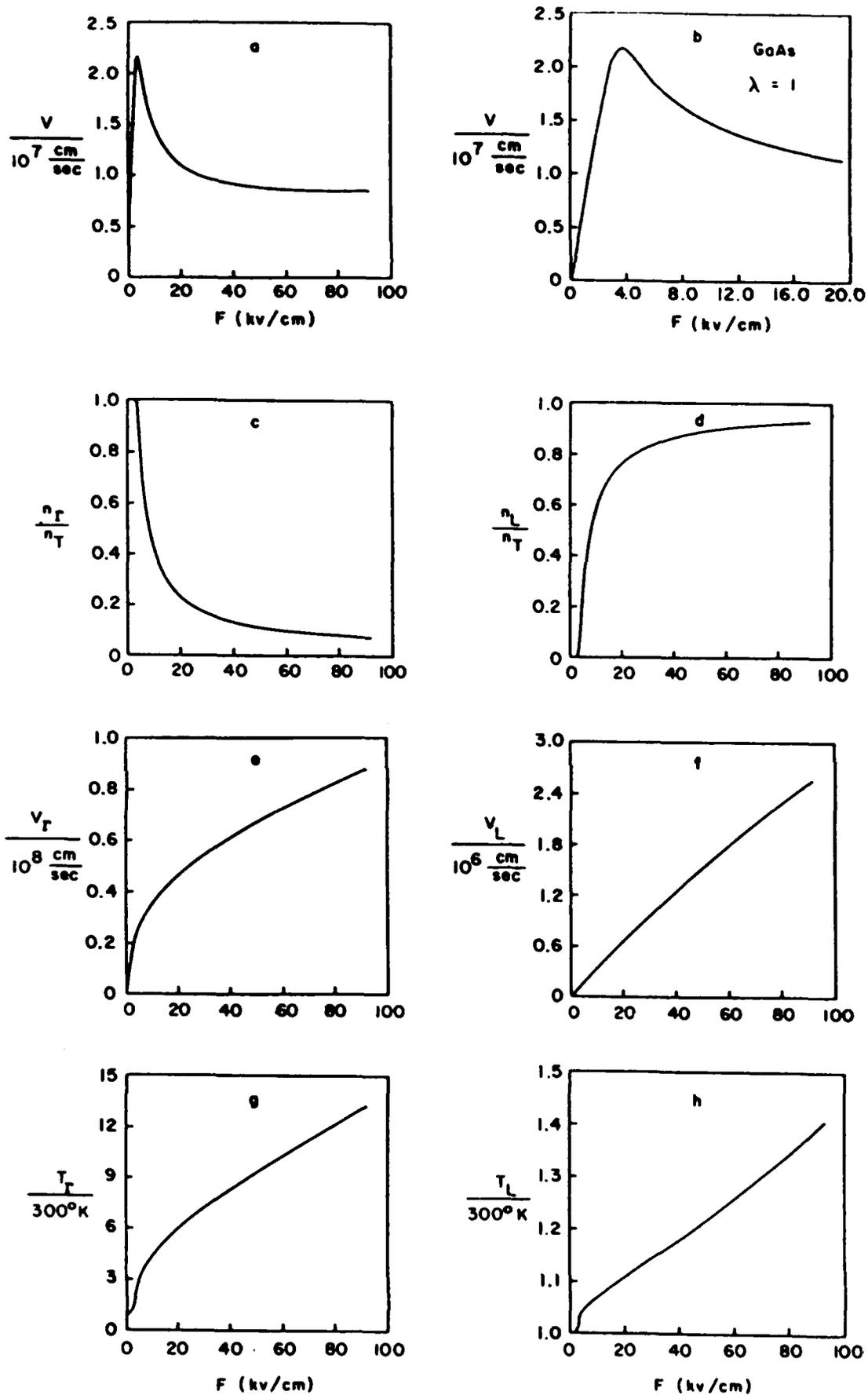
Figure 1.   Uniform field characteristics of GaAs, for the parameters of Table
1. a) and b) field dependent mean velocity, c) and d) population
of central and satellite valleys, e) and f) velocity of central
and satellite valleys, g) and h) electron temperature of central
and satellite valleys.

122
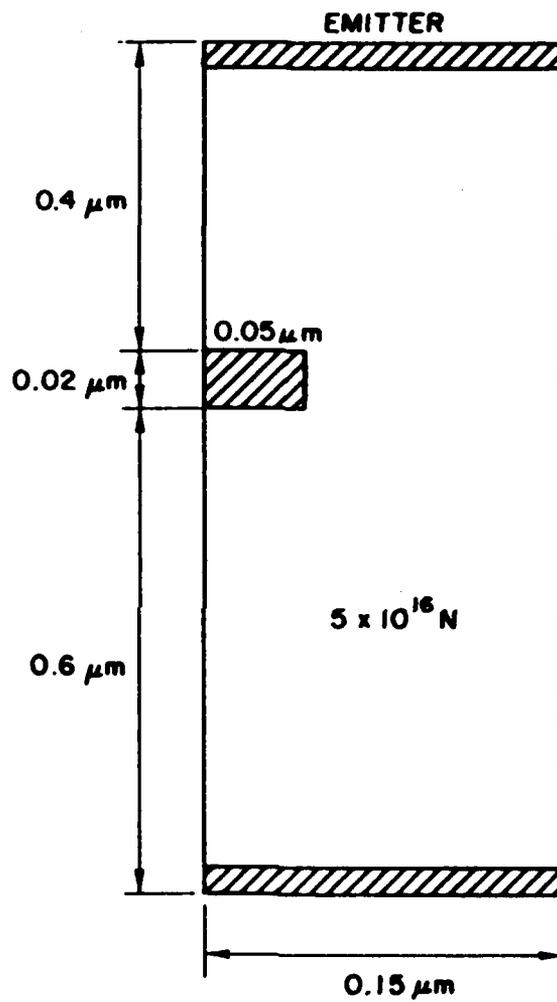
Figure 2. Structure for the MBTE calculations.
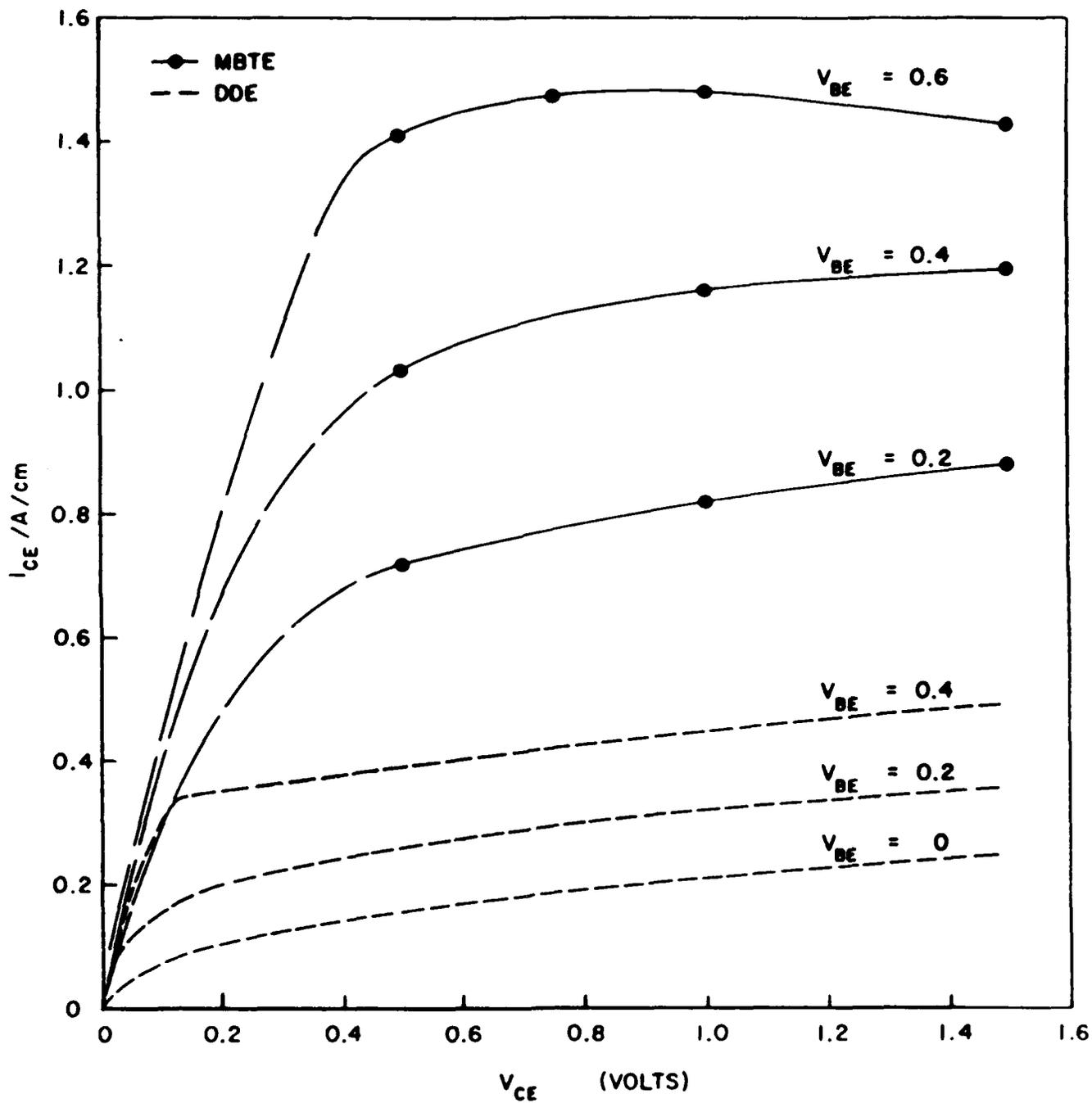
Figure 3. Current-voltage characteristics for the PBT.
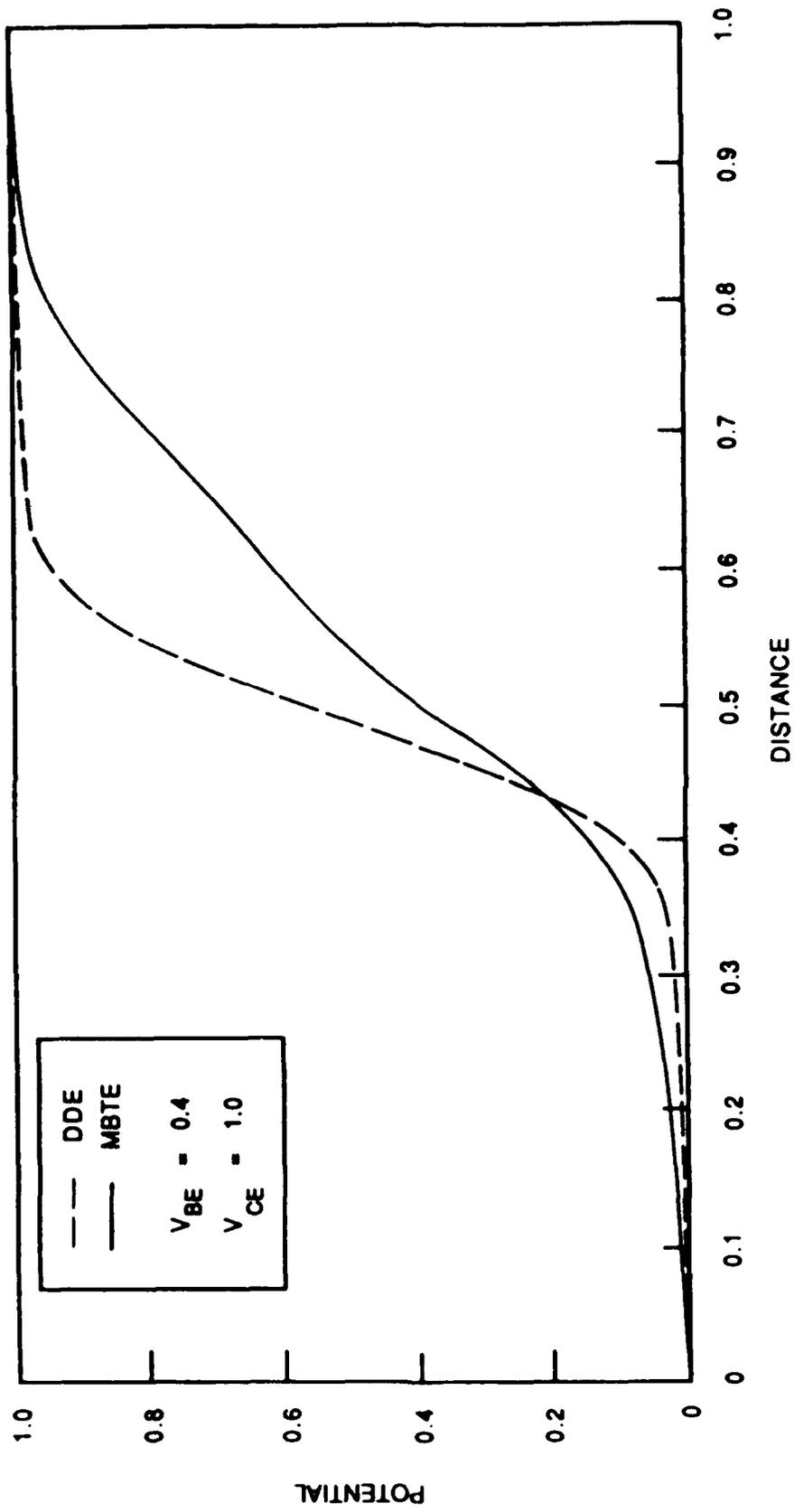
Figure 4. Potential vs. distance along the center of the channel for the PBT.
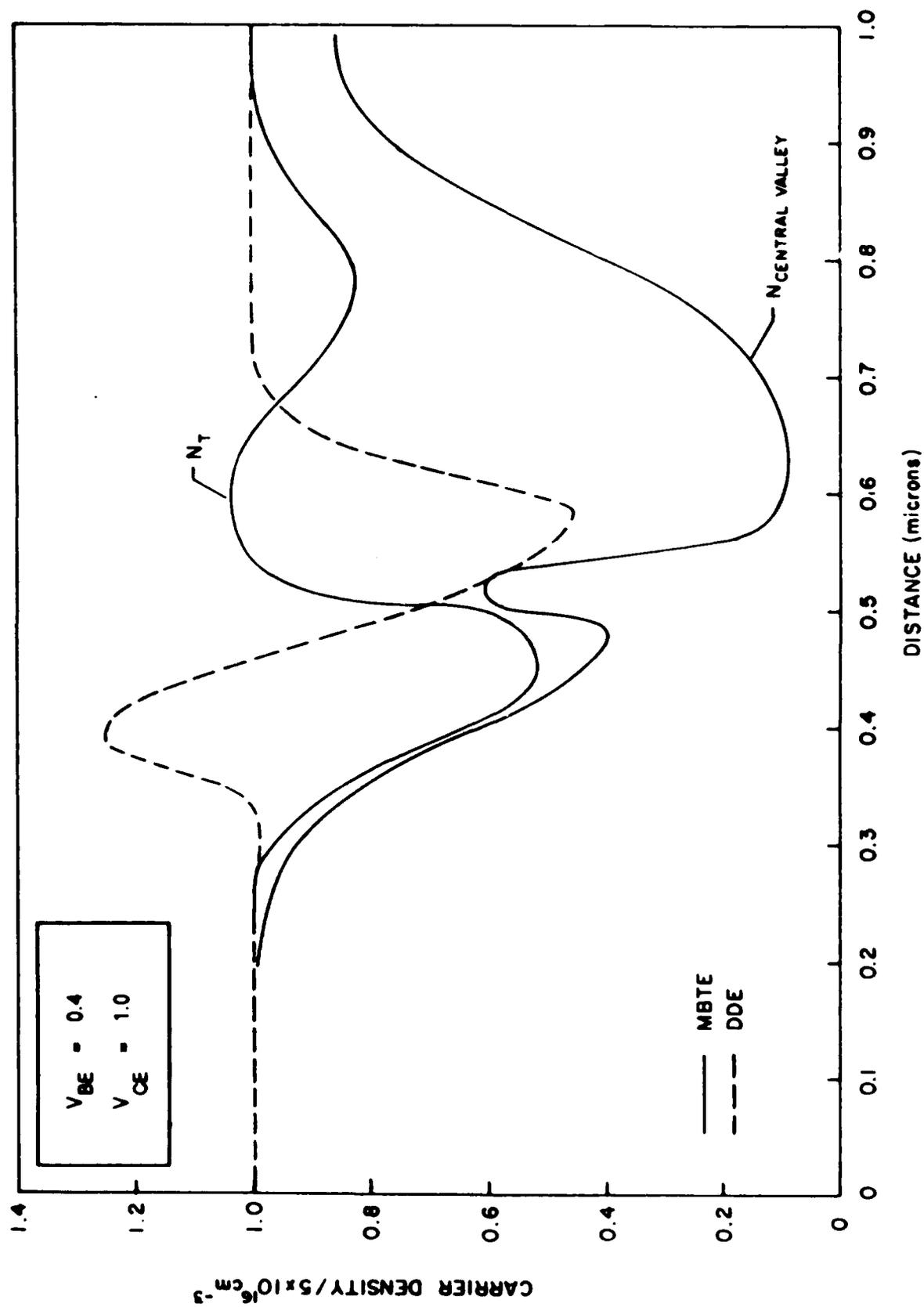
Figure 5. Carrier density vs. distance along the center of the channel for the PBT.

126
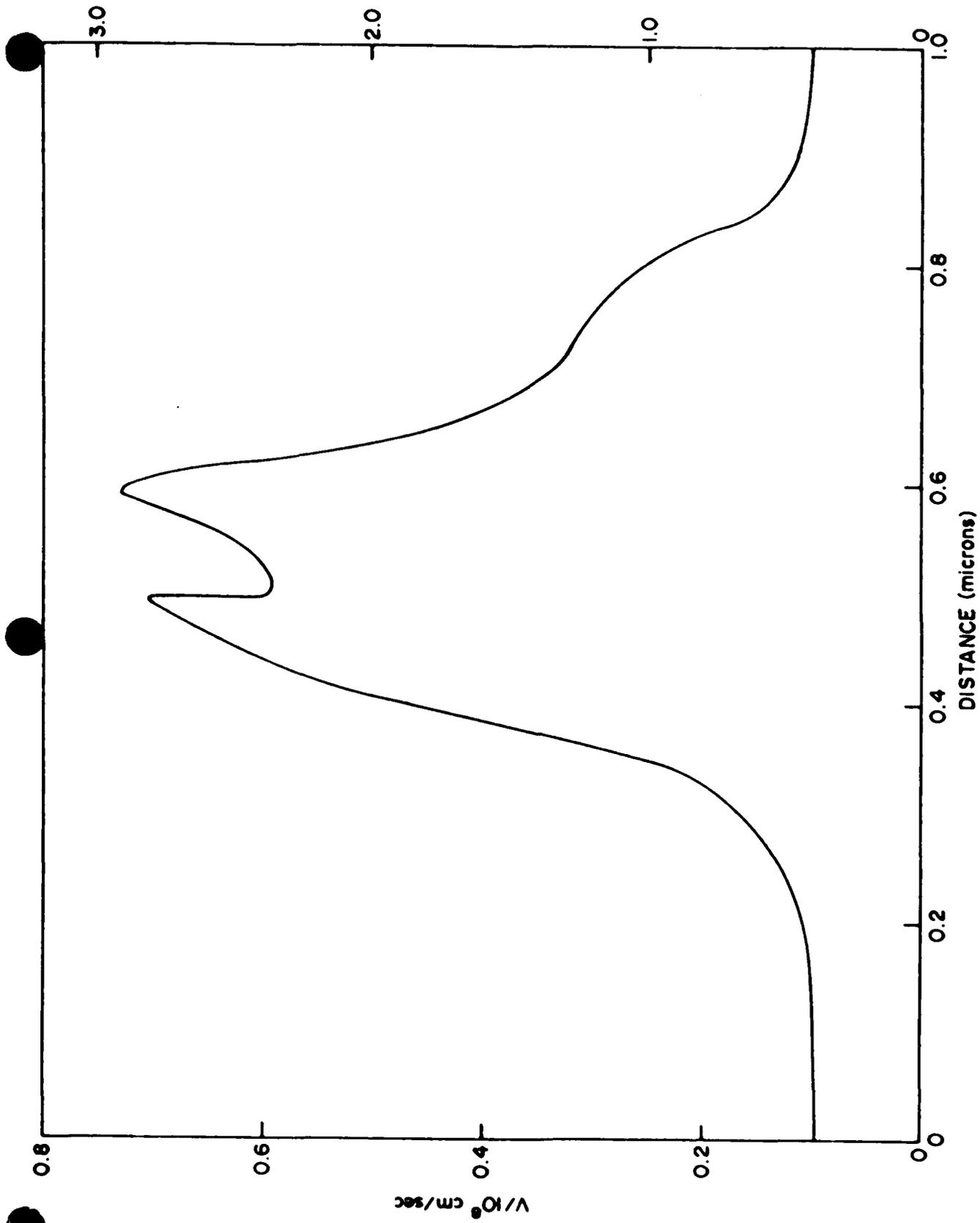
Figure 6. Mean carrier velocity vs. distance along the center of the channel for the PBT.

127
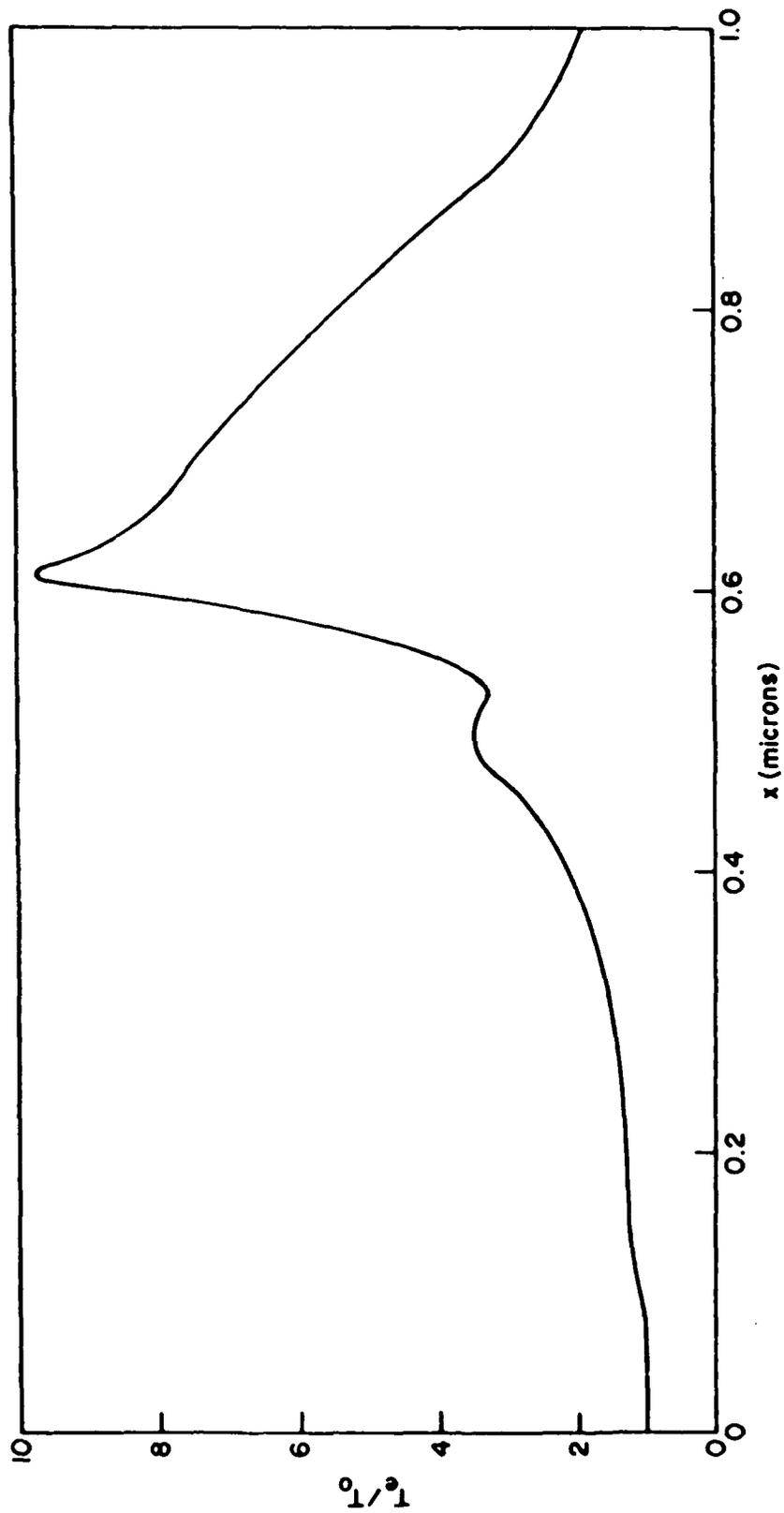
Figure 7. Electron temperature of central valley vs. distance along the center of the channel for the PBT.
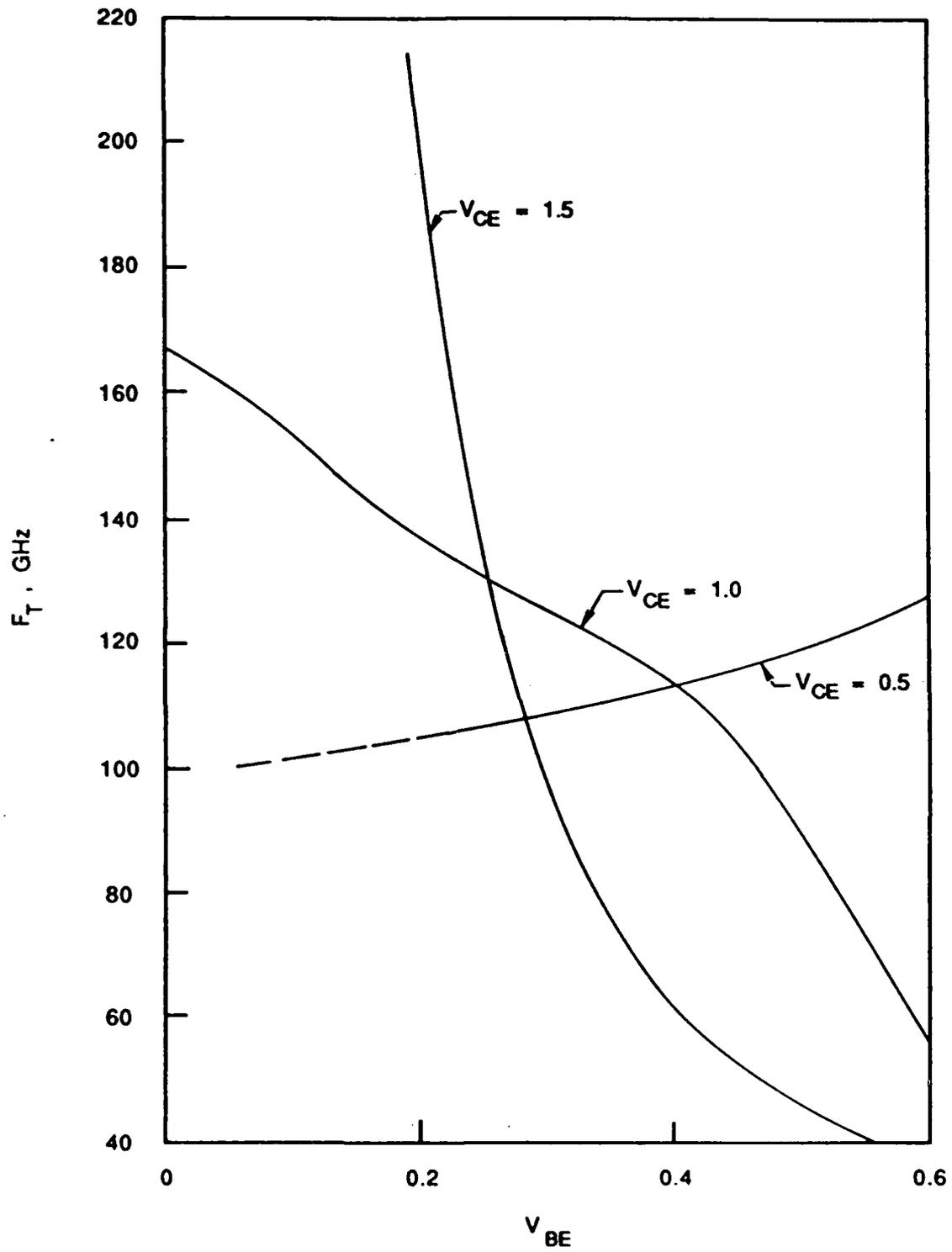
Figure 8.   Cutoff frequency vs. base emitter voltage.

129

APPENDIX A

THE MOMENTS OF THE BOLTZMAN TRANSPORT EQUATION AS APPLIED
TO THE GALLIUM ARSENIDE PERMEABLE BASE TRANSISTOR

J.P. KRESKOVSKY, M. MEYYAPPAN, H.L. GRUBIN
SCIENTIFIC RESEARCH ASSOCIATES, INC.
GLASTONBURY, CT 06033, USA

ABSTRACT

Solutions to the first three moments of the Boltzmann transport equation and Poisson's equation are obtained for a permeable base transistor (PBT) using linearized, block implicit (LBI) and ADI techniques. Two level electron transfer is considered. The results of the simulations are compared to results obtained from the drift and diffusion equations. The comparison indicates that nonequilibrium transport and velocity overshoot are important in the PBT. The predicted I-V characteristics of the device show substantially higher current levels and a higher cutoff frequency are obtained with the moment equations.

INTRODUCTION

Numerical and experimental studies of the PBT were first performed by Bozler and Alley [1]. Their results suggested that a PBT could obtain cutoff frequencies as high as 300GHz. This result has not yet been obtained. In some of their drift and diffusion computations Bozler and Alley [1] used a two piece velocity-field curve in an attempt to account for nonequilibrium effects. The velocity varied linearly with field up to a value of 4KV/cm after which it was held constant at 2×10^7cm/sec. This approach will yield higher current levels, but it offers no improvements in either the qualitative or quantitative consequences of nonequilibrium transport. Nonequilibrium effects, in the presence of nonuniform fields, result in carrier and velocity distributions that are significantly different than those offered by excess saturation velocity coupled to the drift and diffusion equations.

Nonequilibrium effects in a PBT have been investigated using Monte Carlo methods by Hwang, Navon and Tang [2]. Their results showed significant differences compared to the drift and diffusion results; the peak velocity along the center of the channel exceeded 4×10^7cm/sec and the density distribution was also significantly different. The Monte Carlo results also indicated a cutoff frequency approximately 60% higher than the corresponding drift and diffusion prediction.

In the present note, the performance of the PBT structure shown in Fig. 1 is analyzed using both the drift and diffusion and the moments of the Boltzmann transport equation, (MBTE). The base penetration of the present device structure is only one half of the channel width. This in itself results in a higher f_T,

compared to devices with equal base penetration and channel width [3]. However, further and substantial increases in f_T are observed from the MBTE predictions as well as higher current levels.

ANALYSIS AND NUMERICAL METHOD

The equations to be solved are obtained by taking the first three moments of the Boltzman transport equation. The resulting equations represent conservation of mass, momentum, and energy, and may be expressed for the central valley carriers as

$$\frac{\partial n_1}{\partial t} = -\frac{\partial n_1 v_1^i}{\partial x_i} - n_1 \Gamma_{11} + n_2 \Gamma_{12} \tag{1}$$

$$\frac{\partial n_1 v_1^i}{\partial t} = -\frac{\partial n_1 v_1^j v_1^i}{\partial x_j} - \frac{k_0}{m_1}\frac{\partial n_1 T_1}{\partial x_i} - \frac{1}{m_1}\frac{\partial \sigma_1^{ij}}{\partial x_j} + \frac{en_1}{m_1}\frac{\partial \psi}{\partial x_i} - n_1 v_1^i \Gamma_{13} \tag{2}$$

$$\frac{\partial n_1 T_1}{\partial t} = -\frac{\partial n_1 v_1^j T_1}{\partial x_j} - \frac{5}{3} n_1 T_1 \frac{\partial v_1^j}{\partial x_j} - \frac{2}{3}\frac{\sigma^{ij}}{k_0}\frac{\partial v^i}{\partial x_j} + \frac{2}{3k_0}\frac{\partial \kappa \frac{\partial T_1}{\partial x_j}}{\partial x_j} \tag{3}$$

$$- n_1 T_1 \Gamma_{14} + n_2 T_2 \Gamma_{15} + \frac{m_1}{3k_0}(2n_1 \Gamma_{13} - n_1 \Gamma_{11} + n_2 \Gamma_{12})$$

A similar system is obtained for the satellite valley carriers. These equations are coupled to Poisson's equation

$$\epsilon \nabla^2 \psi = e(n_1 + n_2 - N_D) \tag{4}$$

In eqs. (1-4) n is carrier density, v^i the velocity vector, T the carrier temperature, ψ the electrostatic potential, k_0 is Boltzman's constant, m is the effective electron mass, Γ_{ij} are scattering rate constants, σ^{ij} is a stress tensor arising from the nonspherical nature of the distribution function, and κ is a thermal conductivity. The scattering rates are determined *a priori* from evaluation of the collision integrals [4].

Eqs. (1-3), their satellite valley counterparts, and Poisson's equation form a coupled system of nine nonlinear PDE's in two space dimensions of the form

$$\frac{\partial H(\phi)}{\partial t} = D(\phi) + S(\phi) \tag{5}$$

where ϕ represents the vector of unknowns, H and S represent nonlinear functions of ϕ, and $D(\phi)$ represents a nonlinear, partial differential operator. To solve this system Poisson's equation is decoupled from the remaining eight equations by differencing the electric field appearing in the momentum equation at the explicit time level. The remaining equations are solved by direct application of linearized block implicit (LBI) techniques [5]. The equations are first time differenced

$$\frac{\Delta H(\phi)}{\Delta t} = \beta\left[D(\phi)^{n+1} + S(\phi)^{n+1}\right] + (1-\beta)\left[D(\phi)^n + S(\phi)^n\right] \tag{6}$$

The nonlinear operators are then linearized using Taylor series, as for example:

$$D(\phi^{n+1}) = D(\phi^n) + \frac{\partial D}{\partial \phi}\bigg|^n \Delta\phi^{n+1} + O(\Delta t^2) \tag{7}$$

Eq. (6) may then be expressed as

$$(A + \Delta t L)\Delta\phi^{n+1} = \Delta t\left[D(\phi^n) + S(\phi^n)\right] \tag{8}$$

where

$$A \equiv \left(I - \beta\Delta t \frac{\partial S}{\partial \phi}\bigg|^n\right) \tag{9a}$$

$$L \equiv -\beta\frac{\partial D}{\partial \phi}\bigg|^n \tag{9b}$$

Eq. (8) is then split following the ADI procedure of Douglas and Gunn [6].

$$(A + \Delta t L_x)\Delta\phi^* = \Delta t\left[D(\phi^n) + S(\phi^n)\right] \tag{10a}$$

$$(A + \Delta t L_y)\Delta\phi^{**} = A\Delta\phi^* \tag{10b}$$

and

$$\Delta\phi^{n+1} = \Delta\phi^{**} + O(\Delta t') \tag{11}$$

As a result of the ADI splitting, the number of operations required to solve the remaining coupled equations varies linearly with the total number of grid points, and due to the linearization process, the solution can be advanced in time without introducing nonlinear iteration. After solution of the continuity, momentum, and energy equations, Poisson's equation is solved using a scalar ADI procedure to complete a time step. Steady solutions, such as those to be presented here, are obtained from the long time asymptotic transient solution. Since transient accuracy is not of interest under such conditions, the time step can be spatially scaled to speed convergence to steady state.

The solutions to the drift and diffusion equations reported here were obtained using a related procedure, described in detail in [7].

COMPUTED RESULTS

The computed I-V characteristics of the device shown in Fig. 1 are presented in Fig. 2. The results for the MBTE calculations were extrapolated to the origin, as indicated by the long broken lines. The shorter dashed curves show the results for the DDE [3]. The comparison shows that the predicted current levels are significantly higher for the MBTE solutions, a result consistent with FET calculations

performed by the present authors, as well as by Cook and Frey [8] who used a highly simplified momentum-energy transport model. The present MBTE results also indicate a region of negative differential forward conductivity at $V_{BE} - 0.6$ volts. The origin of this phenomena is believed to be a consequence of electron transfer. The presence of a dc negative forward conductance is also a feature of PBT measurements [9], but is clearly absent from DDE simulations.

A comparison between the total carrier density distribution along the center of the channel for drift and diffusion and MBTE solutions is shown in Fig. 3 for $V_{CE} - 1.0$ volt and $V_{BE} - 0.4$ volts. The MBTE prediction for the gamma valley carrier density is also shown. The drift and diffusion result was obtained using an equilibrium velocity-field curve, and yields results qualitatively similar to that of Bozler and Alley.

As seen in Fig. 3 for the DDE simulations, the carrier density reaches a maximum between base contacts. Here, with the velocity in saturation and the cross sectional area at a minimum, the carrier density must increase to maintain current continuity. In the MBTE simulation the constraints of current continuity are more complex. First a decrease in the cross sectional areas is, as in the DDE, accompanied by an increase in field along the channel. The field increase under both equilibrium and nonequilibrium conditions is qualitatively similar, as may be observed from Fig. 4 which shows the potential distribution along the center of the PBT channel. However, consequent changes in electron temperature, both increasing and decreasing, lag behind the equilibrium state. This leads to velocity overshoot and a delay in electron transfer. As a result, for nearly the first half of the device transport is almost exclusively gamma valley transport. The implication is that if the gamma valley carrier velocity increases with increasing field, then the product of density and cross sectional area normal to current flow must decrease to maintain current continuity. Since the velocity increases faster than the area decrease the carrier density decreases.

At moderate bias levels typical FET calculations show a decreasing field as the gate region is passed. This also occurs in the PBT. Now, as the cross sectional area increases the gamma valley carriers exhibit a decrease in velocity. It must be noted, however, that for the parameters of the calculations the L valley carriers make a negligible contribution to current. Thus a decrease in carrier velocity results in a net increase in carrier concentration. However initially, the decrease in field is not accompanied by a corresponding temperature decrease (as experienced in the uniform field calculations). Thus, the high gamma valley temperature results in transfer to the L valley giving rise to the second minimum in the gamma valley carrier density shown in Fig. 3. Further toward the drain, the field decreases. However, relaxation is incomplete and the field at the collector is not equal to the field at the emitter. Also note, the MBTE . potential distribution gives rise to a slightly higher field upstream of the base,

and a lower field, over a longer distance, downstream of the base compared to the DDE result. More significantly the electron temperature at the collector exceeds that at the emitter. It is noted that that as the field relaxes, the electrons transfer back to the central valley.

All of the results presented thus far are qualitatively similar to the Monte Carlo results of Hwang, et. al. [2]. However, the results for cutoff frequency, shown in Fig. 5, are qualitatively different. At $V_{CE} = 0.5$ volts, f_T is seen to increase with V_{CE} from approximately 100GHz to 130GHz, value 2.5 to 3 times greater than that obtained using the DDE. However, for higher values of V_{CE}, f_T is shown to decrease with increasing V_{BE}. For $V_{CE} = 1.5$ volts, an f_T in excess of 200GHz is predicted for $V_{BE} = 0.2$ volts; while at $V_{BE} = 0.6$ volts, f_T has dropped below 40GHz, a result that is qualitatively consistent with the presence of the negative forward conductance, and one that would find no explanation in the DDE calculations.

ACKNOWLEDGEMENTS

REFERENCES

1. Bozler, C.O. and Alley, G.D., IEEE Trans. Elec. Dev., ED-27, No. 6, 1980.

2. Hwang, C.-G., Navon, D.H., and Tang, T.-W., IEEE Elec. Dev. Lets., EDL-6, No. 3, 1985.

3. Meyyappan, M., Kreskovsky, J.P., and Grubin, H.L., SRA report in preparation.

4. Grubin, H.L., Ferry, D.K., Iafrate, G.J., and Barker, J.R., VLSI Electronics, Vol. 3, 1982.

5. Briley, W.R., and McDonald, H., J. Comp. Phys., 1980.

6. Douglas, J. and Gunn, J.E., Num. Math., Vol. 6, 1964.

7. Kreskovsky, J.P. and Grubin, H.L., J. Comp. Phys., 1987.

8. Cook, R.F. and Frey, J., Compel, Vol. 1, No. 2, 1982.

9. See, e.g., Bozler, C.O., Hollis, M.A., Nichols, K.B., Rabe, S., Vera, A., and Chen, C.L., IEEE Electron Device Letters, Vol. EDL-6, No. 9., September 1985.
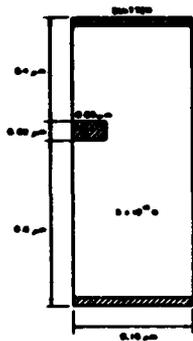
Figure 1. Dimensions and Doping Level of the Simulated PBT.
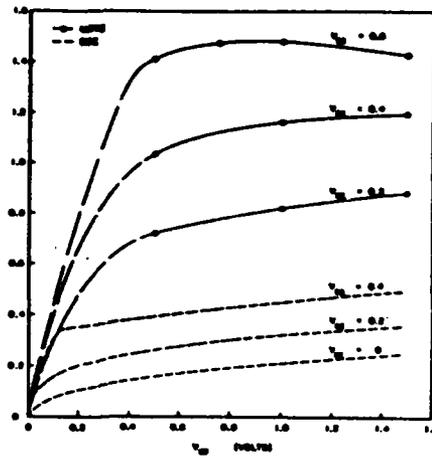
Figure 2. Collector Current vs. Collector Emitter Voltage for Different Values of Base Emitter Voltage.
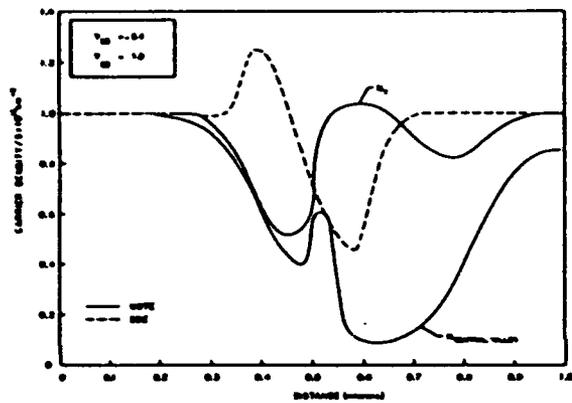
Figure 3. Carrier Density vs. Distance along Center of Channel for the PBT.
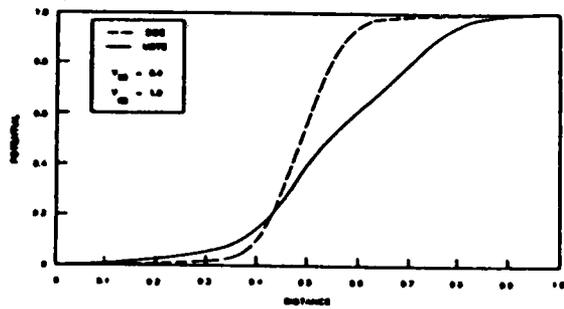
Figure 4. Potential vs. Distance along Center of Channel for the PBT.
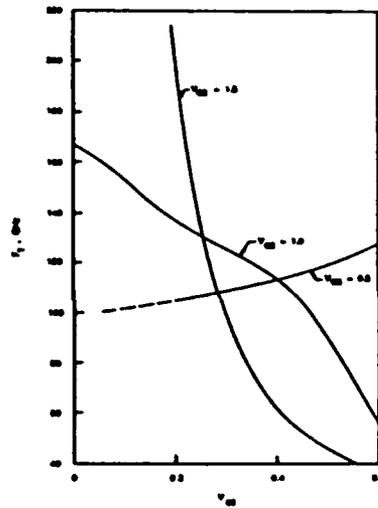
Figure 5. Cutoff Frequency vs. Base Emitter Voltage.

APPENDIX B

137

Ultrashort transients in semiconductors

H.L. Grubin, M.A. Osman and J.P.Kreskovsky
Scientific Research Associates, Inc.
P.O. Box 1058, Glastonbury, Connecticut 06033-6058

Abstract

Over the past decade a variety of physical phenomena have emerged which are exercising profound constraints on the speed of devices whose primary transient is electronic in origin. For example, the broad high frequency extrapolations of high speed arising from velocity overshoot were found to be limited by the effects of space charge and the imposition of fields that take a finite time to reach their specified level. Much of the latter problem may be minimized through optical processes, but in its place rests the constraints of electron-hole interaction and other processes that may require quantum coherence for implementation. In the case of electron-hole interaction it has been found that it provides an important channel through which electron energy is transferred to the lattice, and that the electron cooling rates at high hole concentrations are higher when the electron-hole interaction is included, than when it is ignored. At low hole concentrations the reverse is true. The effect of electron-hole interaction also influences velocity overshoot. In the case of quantum coherence, of interest is the effect, e.g., of temporal scatterers (phonons) on the time dependent wave functions and modifications of interference effects. The influence of space charge on transient velocity overshoot, the effects of electron-hole interaction on the relaxation of an electron-hole plasma, and the effects of time dependent and time independent scattering on quantum interference will be discussed.

Introduction

The early suggestions of Ruch[1] and subsequent calculations of others, have suggested that high mobility semiconductors were capable of extraordinary high speeds (10^8cm/sec) provided they were sampled for sufficiently short time intervals. For submicron and ultrasubmicron devices, this translates into time scales of 0.1 ps, and cutoff frequencies of 10^3GHz. While these predictions are clearly impressive, of serious concern is an inability of realizing the modest goal of room temperature operation at 100GHz, and speeds of less than 5ps, within the framework of a three terminal device. The difficulty of achieving anything close to that offered by the Ruch calculations has been discussed by a number of workers over the past several years. Perhaps the simplest statement summarizing these studies is that Ruch's assumptions are not realized experimentally. These assumptions include: 1) uniform fields within the device, 2) zero rise time of the electric field, and 3) neglect of displacement currents. Because of the importance of these assumptions, each will be briefly reviewed, along with the influence of nonequilibrium holes and phonons. Additionally the influence of quantum mechanical constraints on transport will be discussed.

Nonuniform fields

Electric fields in devices are nonuniform, as illustrated for the three terminal permeable base transistor (PBT). The PBT as currently configured, is a micron-length structure with ultra-submicron features. A section of the PBT is shown in figure 1, along with its an equivalent planar FET. As shown in figure 2[2], for a bias of one volt between the emitter and the collector the potential distribution indicates the expected field non-uniformity. For example, near the cathode, the field is low, increasing to a value near 20 kv/cm and finally approaching a low value near the collector contact. In uniform field studies the predicted peak velocity approaches 10^8cm/sec, followed by a decay to steady state. The peak velocity is achieved with a carrier concentration that is close to its equilibrium value. Decay is dominated by electron transfer, with a weaker component due to momentum relaxation. In steady state the gamma valley carrier velocity is high, while the concentration of gamma valley carriers is low. In the PBT the initial high peak in current associated with velocity overshoot is not achieved. Rather the high velocity is often achieved at the expense of carrier concentration. This is displayed in figures 3 and 4 which show the distributon of carrier density and velocity for the PBT. It is worthwhile noting that while the distribution of charge and velocity are not as encouraging as that from the uniform field calculations, the current voltage levels are between a factor of 3 and 4 above that obtained from the drift and diffusion equation calculations.

Finite rise time contributions

A second issue is the zero rise time. In virtually all transient transport calculations, overshoot phenomena is obtained by examining the response of carriers to an electric field that is suddenly turned on. Under conditions of finite rise time, relaxation effects are coupled to the increasing field and the peak velocity is diminished, as shown in figure 5[3]. We note that there is a qualitative similarity between the finite rise time and the response of carriers to a gradual spatially-varying electric field.

Displacement current effects

To carry the discussion of non-uniform fields even further, calculations have been performed for the paradigm two-terminal $N^+N^-N^+$ submicron structure shown in figure 6[4]. The structure results in non-uniform fields and when transients are computed it is a current transient that is involved rather than a velocity transient. Indeed, in virtually all structures it is a current transient rather than a velocity transient that is measured. As shown in figure 7, there is a peak value to the current that occurs within the time scale over which velocity overshoot may be expected. However, the details of the transient show that the initial time dependent behavior is associated entirely with field rearrangement, and that the transient is entirely a displacement current transient[4].

Nonequilibrium electron hole phonon interaction

In overshoot studies most inelastic collisions are treated assuming the emission of a single phonon and that the emitted phonon is in equilibrium with the lattice. For energetic electrons, a sequence of as many as ten to fifteen phonons maybe emitted as the electron is relaxing, and as a result an excess number of phonons may result in a lack of phonon equilibrium with the lattice. This phenomena is illustrated for an unbiased slab of gallium arsenide subject to a uniform photoexcited laser pulse of energy 1.8 ev. Under this situation holes as well as electrons are excited. The energy of the excited electrons is 250 Mev, while that of the holes is 40 Mev. The time evolution of the resulting electron, hole and phonon distribution is calculated using Monte Carlo techniques. Figure 8 shows the relaxation of the electrons to equilibrium with and without hot phonons. As seen, the inclusion of the hot phonons results in a delay of the relaxation to steady state.

The situation in the presence of finite fields and equilibrium phonons is discussed next and shown in figures 9 and 10[5]. These calculations were performed assuming nonequilibrium interactions between the gamma valley electrons and holes. L-valley hole interaction was not examined. The first point of note is that the electron hole interaction is an inelastic collision. In addition, since both electrons and holes are responding to the presence of the applied fields, momentum scattering of the electrons and holes can result in a reduction of the electron velocity, and in some cases turn the electron in the "wrong" direction. The net result is a decrease in the peak electron velocity. It is important to note that while the peak velocity is lower for the calculation with the electron-hole interaction, there is a cross-over where, for a certain interval of time, the carrier velocity in the presence of holes exceeds that obtained in the absence of holes. This arises because the electron hole interaction is an efficient method of energy exchange and as a result, the electrons are not energetic enough to transfer into the L-valley. In other words, there is a greater fraction of electrons that are retained in the gamma valley when holes are included than when they are not included. The possibility exists that the velocity of injected electrons in p-type material may, for a range of values of electric field, be higher than that of thermally generated carriers in N-type material.

Quantum contributions

It will be noticed that all of our discussion has been confined for electric fields and scalar potentials. It has been presumed that we are dealing with a gauge invariant system, and that the choice of gauge is not relevant. As discussed by Aharanov and Bohm and[6], the vector potential can lead to observables such as interference patterns for electrons that traverse ostensible field-free regions. In other words, interference patterns are accounted for in terms of differences in the vector potentials encountered by electrons on two opposite sides of a flux line. To see how these effects enter into the problem it is useful to re-examine an approach to quantum transport introduced earlier[7].

In reference 7, which introduced the Wigner function as an approach to quantum transport in electron devices, a set of hydrodynamic balance equations was introduced. While the hydrodynamic balance equations were developed for a system of particles, the equations may be reduced to examine transport for carriers not subject to scattering. Thus, the single particle wave functions are of interest.

Determining the single particle wave functions requires solutions to the single particle Schrodinger's equation. Rewriting Schrodinger's equation in terms of a hydrodynamic model requires solutions to the following two equations:

$$\frac{\partial \varrho}{\partial t} + \text{div } J = 0 \qquad (1)$$

$$m \frac{d\upsilon}{dt} = -\nabla (Q + V) \qquad (2)$$

where $V(x)$ is a classical scaler potential and Q is the quantum potential

$$Q = -\frac{h^2}{2m} \frac{\nabla^2 \sqrt{\varrho}}{\sqrt{\varrho}} \qquad (3)$$

The reference 7 study suggests that both $\varrho = \Psi^*\Psi$ and $J = \varrho\upsilon$ be regarded as the basic physical quantities representing properties of the continuous fluid. It is noted that the vector potential has dropped out of the basic equations, which have a completely classical form.

Philippidis, Bohm and Kaye (PBK)[8] argued that the above equations were incomplete. The most important objection arose because certain subsidiary conditions were not satisfied. To see this, the wave function for the system is written as $\Psi = R\exp[iS/h]$, with the carrier velocity expressed as

$$\upsilon = \frac{\nabla S}{m} - \frac{e}{mc} A \qquad (4)$$

If it is now required that the wave function be single valued, then we are forced to admit the condition

$$\oint \nabla S \cdot dx = \oint \left(m\upsilon + \frac{e}{c} A \right) \cdot dx = nh \qquad (5)$$

This means that the observable movements of electron transport in a device must necessarily be restricted in such a way that the integral of its velocity around a circuit containing a flux line depends upon the flux within the circuit. AB argue that there is no intuitively clear justification for the equation (5) constraint. Rather, they propose to assume that the vector potential has a physical meaning.

Figure 11 is a plot of the quantum potential for an electron passing through a dual slit in the absence of a magnetic field. The smooth parabolic hills are in the immediate vicinity of the slit. Particles subject to such a potential accelerate smoothly in the forward and sideward direction. This can be seen from the family of velocity trajectories displayed in figure 12. Note, near the slits where the wave functions are localized packets, the quantum potential exhibits two broad peaks. As the wave packets recede and spread, they begin to overlap significantly and their interference properties come into play. The shape of the quantum potential, a long way from the slits, is determined by the boundary conditions, the shape and size of the slits. It is through the force exerted by the quantum potential on the particle that the boundary conditions are made physically significant far from the slits. For the case when a flux line is place between the slits, the quantum potential is altered, as shown in figure 13. The velocity trajectories are altered as shown in figure 14. The important point to notice is that the trajectories are displaced.

The above discussion reached initial device implementation by Datta, et al.[9], who conceptually altered the path length of a beam of particles as shown in figure 15, where the slit in the GaAs between the AlGaAs is exaggerated. Additionally they altered the phase through application of a magnetic field. These alterations lead to interference in the downstream carrier wave functions. This result, as discussed in reference 9, is the appearance of oscillations in the conductance of the system.

An important point not addressed in reference 9 is the temporal duration of the conductance oscillations. It must be considered that the sporadic appearance of phonons will result in a loss of coherence and limit constructive interference to time scales shorter than 100 femtoseconds. For time scales longer than 100 femtoseconds, the constructive interference effects will certainly be diminished. Whether they will disappear is a matter to be determined. Here we note that coherence is essential for quantum well devices. There are other issues that may tend to mitigate the appearance of the AB effect such as elastic scattering arising from surface roughness.

Conclusions

The driving force toward high speed devices is limited by the presence of carrier-carrier scattering, phonons and submicron feature sizes that tend to reduce the coherence of scattered waves. While these effects do not lessen the importance of the nonequilibrium transients, unless these effects are incorporated into the design of the device, applications may be limited.

Acknowledgements

References

1. J.G. Ruch, IEEE Trans. Electron Devices ED-19, 652 (1982).

2. J.P. Kreskovsky and H.L. Grubin, Proceedings of NUMOS I (1987).

3. H.L. Grubin, J.P. Kreskovsky, Surface Sci. 132, 594 (1983).

4. H.L. Grubin, J.P. Kreskovsky, VLSI Electronics Microstructure Sci. 10, 237 (1985).

5. M.A. Osman, PhD Thesis, Arizona State University (1986).

6. Y. Aharonov and D. Bohm, Phys. Rev., 115, 485 (1985).

7. G.J. Iafrate, H.L. Grubin, and D.K. Ferry, J. de Physique, C7, 307 (1981).

8. C. Philippidis, D. Bohm, and R.D. Kaye, Il Nuovo Cimento, 71B, 75 (1982).

9. S. Datta, M.R. Mellcoh, S. Bandyopadhyay, and M.S. Lundstrom, Appl. Phys. Letts. 48, 487 (1986).
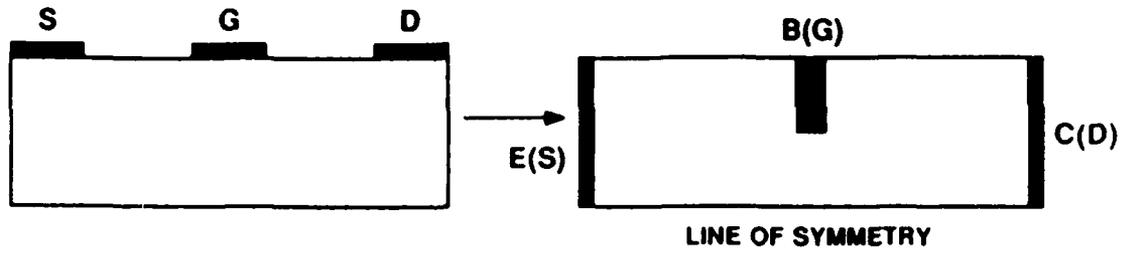
Figure 1. Schematic of FET and PBT.
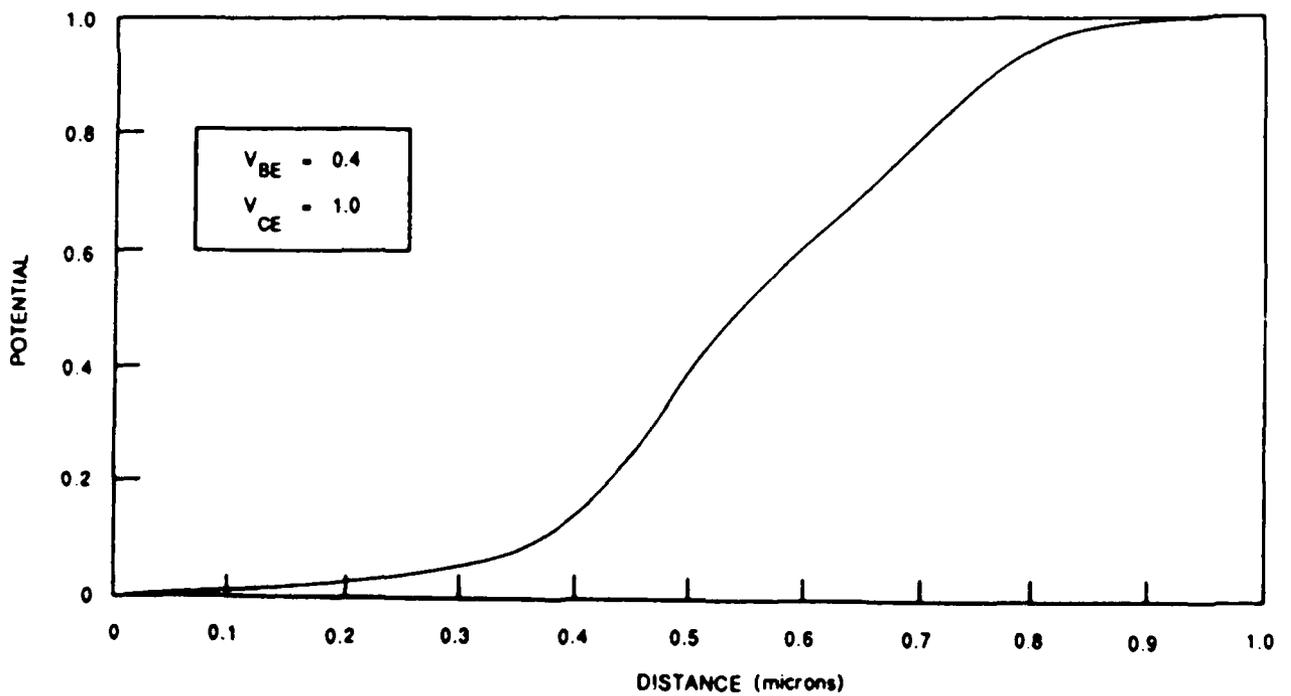
Figure 2. PBT potential distribution along the central line of symmetry.

142
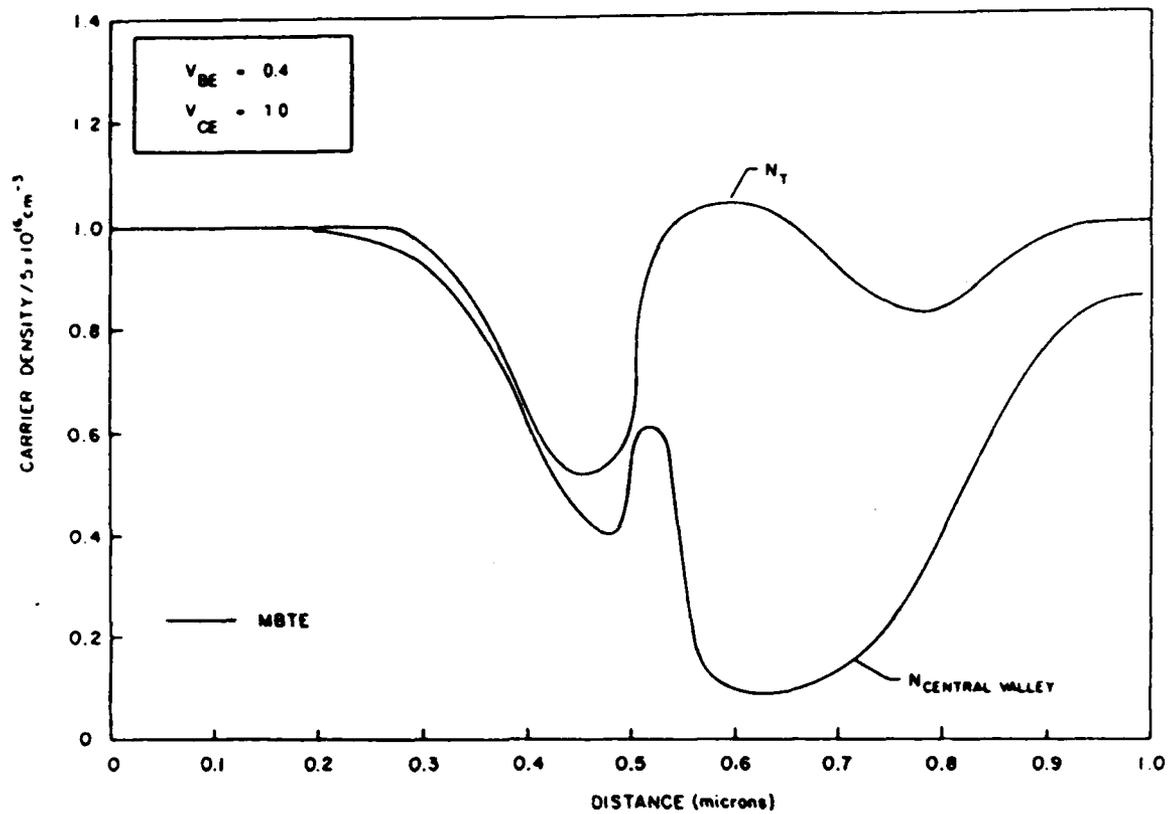
Figure 3. PBT carrier distribution along line of symmetry.
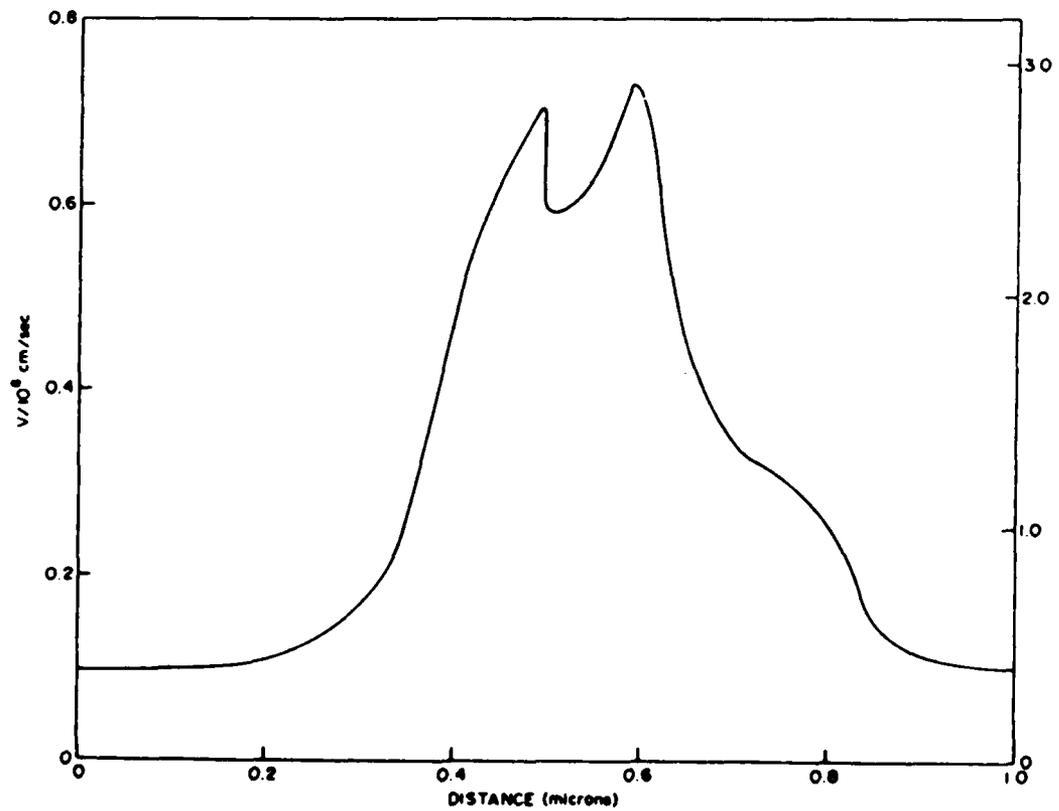
Figure 4. PBT gamma valley velocity along line of symmetry.

Figure 5. Dependence of peak overshoot
velocity on rise time. (From [3]).
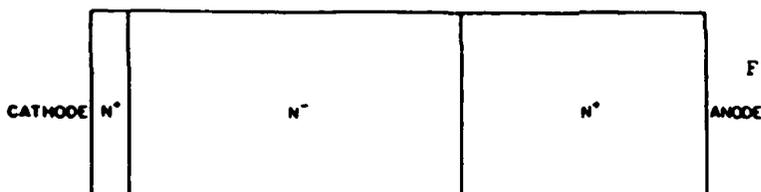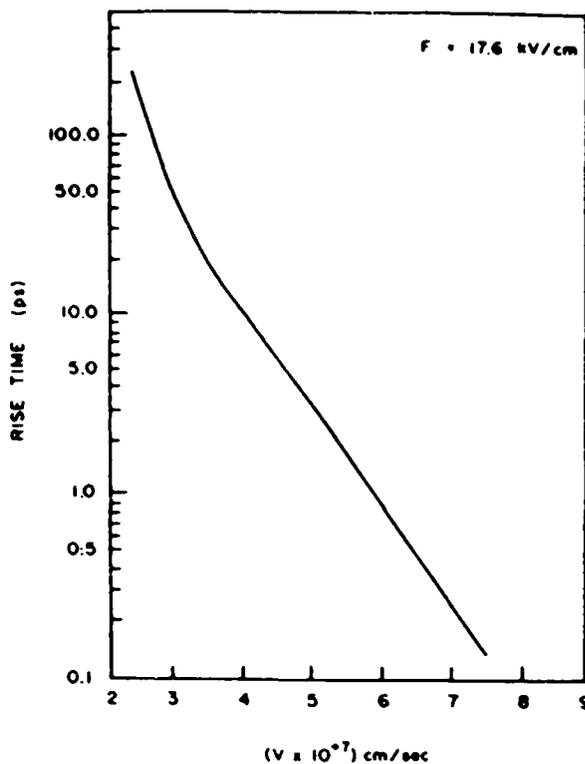
Figure 6. Schematic of structure for
examining displacement current
effects. (From [4]).

Figure 7. Initial current transient for
the structure of figure 6.

144
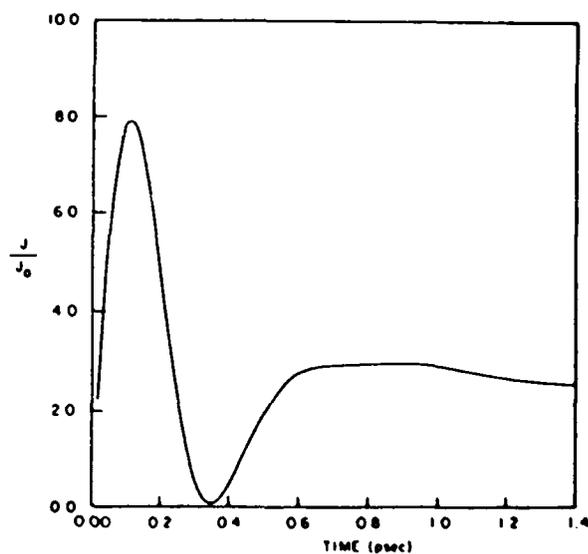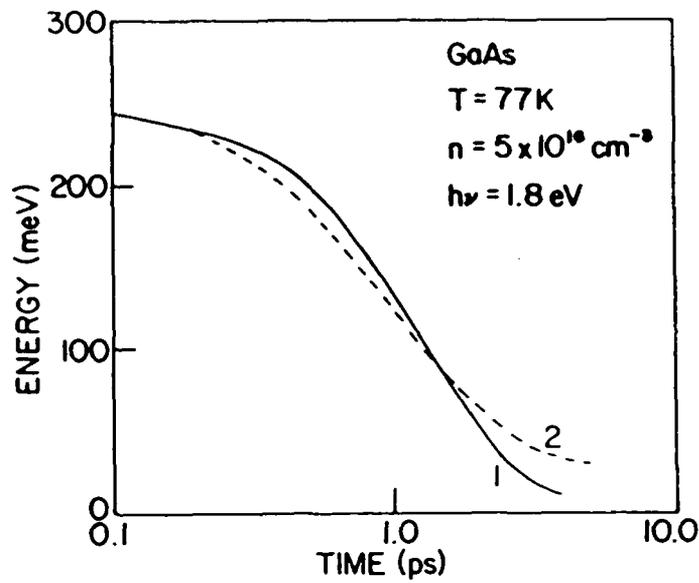
Figure 8. Energy relaxation with equilibrium
phonons (1) and nonequilibrium
phonons (2).
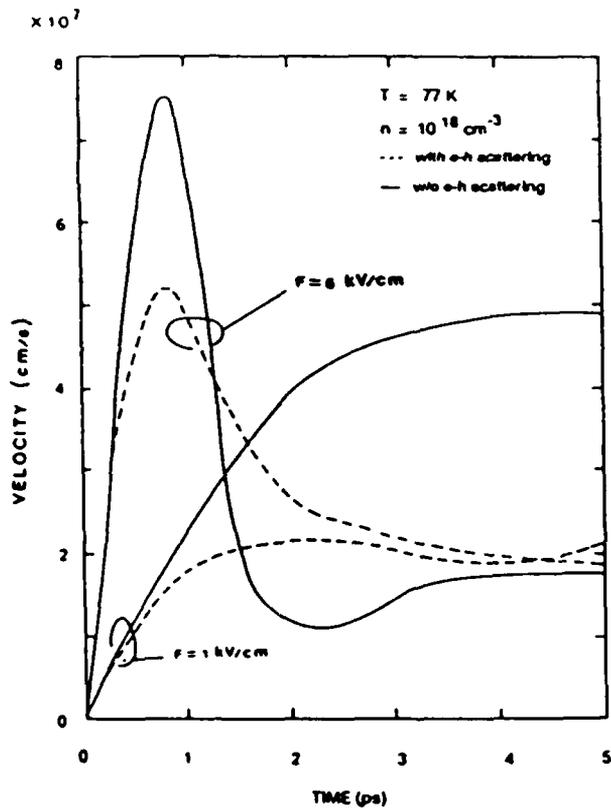
Figure 9. Overshoot with and without
e-h scattering. (From [5]).
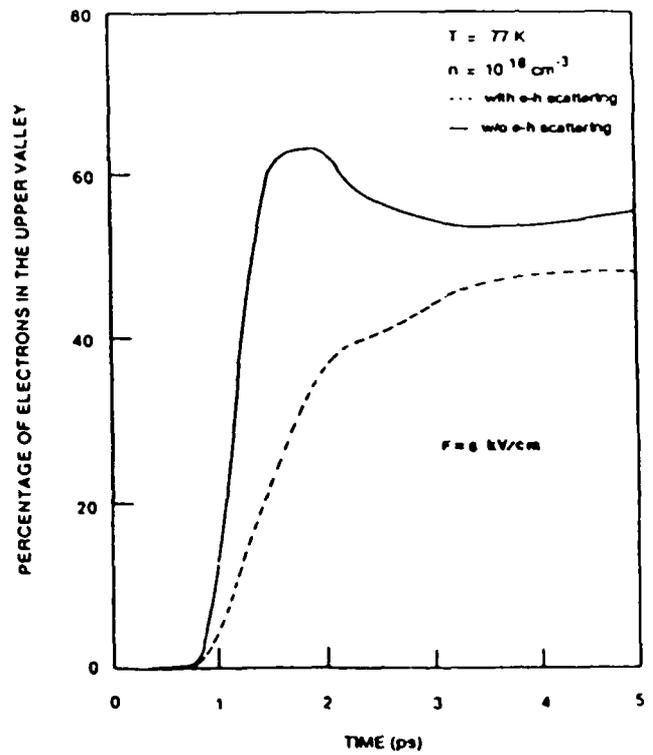
Figure 10. Carrier distribution with and without
e-h scattering. (From [5]).

145
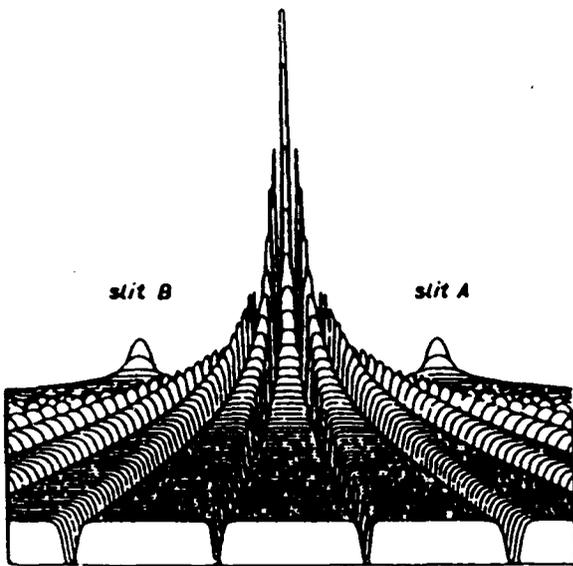
Figure 11. Quantum potential for the
two-slit system. (From [8]).
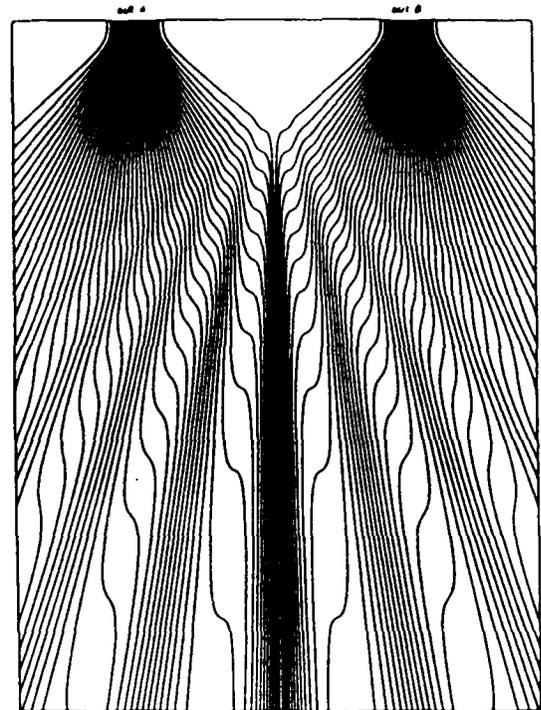
Figure 12. Particle trajectories for the
two-slit system. (From [8]).
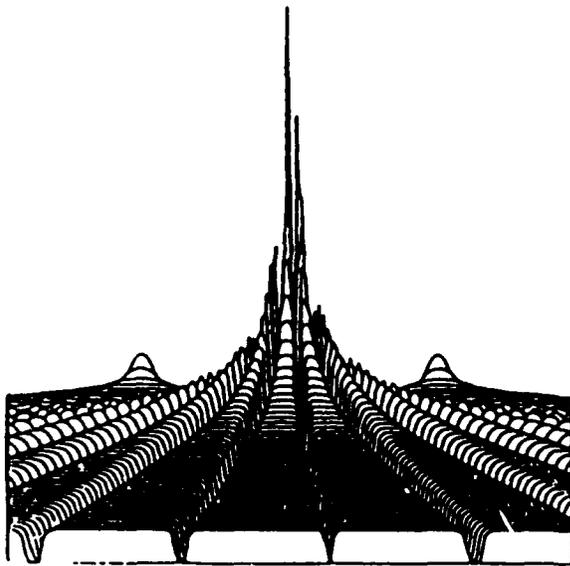
Figure 13. Quantum potential for the AB effect
(note the asymmetry). (From [8]).
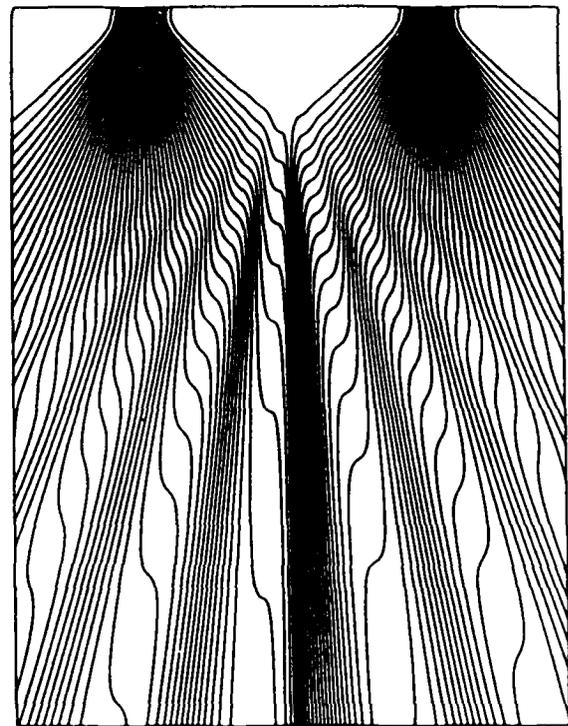
Figure 14. Trajectories for the AB effect
(note the shift of the overall
pattern). (From [8]).
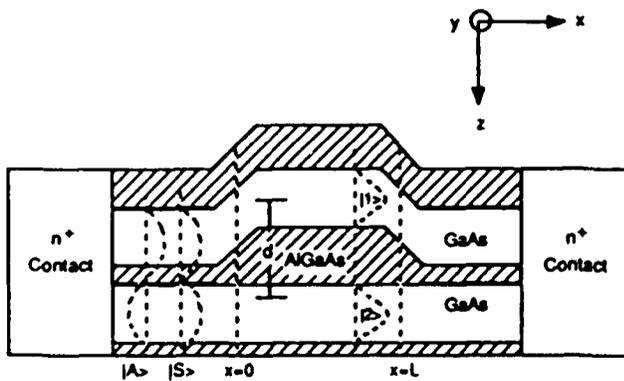
Figure 15. Proposed structure for quantum
interference effects. (From [9]).

146

END

7-81

DTIC