Design and Development of Low Noise, High Speed High Electron Mobility Transistors

Final Report: R930030F

J. P. Kreskovsky, H. L. Grubin, B. J. Morrison and T. R. Govindan

May 1991

U. S. Army Reseach Office
Contract DAAL03-90-C-0021

Scientific Research Associates, Inc.
50 Nye Road, P.O. Box 1058
Glastonbury, CT 06033

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1991	Final 1 July 1990 - 31 December 1991

4. TITLE AND SUBTITLE
Design and Development of Low Noise, High Speed High Electron Mobility Transistors

5. FUNDING NUMBERS
DAAL03-90-C-0021

6. AUTHOR(S)
J. P. Kreskovsky, H. L. Grubin, B. J. Morrison and T. R. Govindan

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Scientific Research Associates, Inc.
50 Nye Road, P. O. Box 1058
Glastonbury, CT 06033

8. PERFORMING ORGANIZATION REPORT NUMBER
R930030F

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARO 27952.1-EL-SBI

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report summarizes work performed under contract DAAL03-90-C-0021. Under this study the moments of the Boltzmann Transport Equation were applied to investigate the performance of an AlInAs/GaInAs on InP HEMT. The results of the simulations are compared to experimental measurements for a related device structure. Numerical solutions of quantum Liouville equation were also introduced to investigate transport perpendicular to the conducting channel under the device contacts to provide insight as to how carriers enter the 2-D electron gas. The goals of this Phase I study were to demonstrate that the cost of fabricating low noise transistors could be reduced by using numerical simulation during the early stage of design. The study was successful.

14. SUBJECT TERMS
HEMT, Low Noise, MBTE, Liouville Equation, InP, InGaAs, Simulation, Fabrication

15. NUMBER OF PAGES
58

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

Introduction

AlInAs/GaInAs on InP HEMT technology is an extremely attractive means to realize high speed integrated circuits. In fact, this technology has provided the fastest three-terminal devices to date for room temperature operation with $f_T > 200$ GHz and $f_{max} > 300$GHz [1]. AlInAs/GaInAs HEMTs offer distinct advantages over their AlGaAs/GaAs counterparts: (1) The lower electron mass in GaInAs provides a mobility of about 11400 cm^2/vs in contrast to the value of 8000 cm^2/vs in GaAs. (2) The peak velocity also is about 1.2 times higher than in GaAs. (3) The conduction band discontinuity in the AlInAs/GaInAs system is 0.5 ev compared to 0.28 ev in the AlGaAs/GaAs system. This results in larger sheet carrier density, better carrier confinement and also larger current density. (4) This system also exhibits reduced DX centers and hence, is good for 77° k operation. (5) The low sheet resistance of the system results in low thermal noise and large intervalley separation results in small intervalley noise.

For the above reasons, AlInAs/GaInAs HEMTs have been investigated widely for low noise, high speed operation. This system, however, has some limitations and efforts are currently underway to overcome the limitations. For instance, the doped GaInAs layer has a very small Schottky barrier height. This can be disadvantageous in several ways: devices become highly temperature sensitive, high gate leakage currents, poor pinch-off. To reduce some of these effects, Schottky contacts are made on undoped (or non-intentionally doped) AlInAs. This layer thickness can be optimized. In addition, the high background doping of GaInAs is also not desirable. High output conductance, backgating, high gate leakage current and low breakdown voltage are the current technical problems in this technology [1]. Of these, the low break-down voltage is not a major issue for low noise applications.

As the technology and materials issues continue to progress, many of the above problems can be addressed and AlInAs/GaInAs on InP technology is certainly the most viable candidate for high speed circuits. With this in mind, **Scientific Research Associates, Inc. (SRA) proposed a SBIR program to both design and to fabricate AlInAs/GaInAs HEMTs. The fabrication was undertaken by GE Syracuse under subcontract to SRA.**

The goal of the SRA program was to demonstrate that the costs of fabricating low noise transistors could be significantly reduced through the introduction of numerical simulations during the early design stages. Accomplishing these goals requires several key tasks:

(1) implementation of a physics based simulation code to represent transport in the device;

(2) optimization studies based upon physics based simulations that incorporate variations in device dimensions, doping, barrier heights, etc.;

(3) incorporation of circuit contributions;

(4) fabrication and testing;

(5) comparison with the simulations.

Success in the program, a combined Phase I and Phase II program, would be demonstrated by the deliverable. The deliverable, **transistors fabricated on the basis of the algorithm.** Additionally, to provide the design tool to the transistor community, SRA would include as

a deliverable, **a user-friendly transportable algorithm for the design of low noise transistors.**

Phase III would involve the manufacturing of low noise transistors as well as the sale of the numerical procedures developed under this study.

The Phase I study, which this document summarizes, was limited in its goals. The objectives of the Phase I program were to demonstrate the feasibility of using simulation in the design and optimization of AlInAs/GaInAs HEMTs. Specifically SRA was to perform simulations of a preliminary device structure, supplied under subcontract by General Electric, to obtain the I-V and small signal characteristics of the device. General Electric was to provide experimental results for comparison. The ability to simulate the device performance would thus be established and the feasibility of the proposed development program thus demonstrated.

The Phase I study involved implementing its nonequilibrium transport model to examine AlInAs/GaInAs HEMTs, to compare the output of the results to data provided by GE Syracuse, and to modify the input to the algorithm to achieve agreement with experiment. Success in this study was achieved as judged by the degree to which the simulations replicated experiment, both dc-wise and small signal-wise.

Analysis

The initial device design aspect of the Phase I program was based on the numerical simulation of the device structure shown in Fig. 1. The details of the structure were supplied by General Electric and were taken from the actual device tested under the experimental portion of the program. The governing equations employed in the study are the first three moments of the Boltzmann transport equation (MBTE) expressing conservation of mass, momentum and energy. These are semiclassical equations, and have not been modified to deal with tunneling aspects of the structure. The incorporation of the quantum potential is one possible approach to treat tunneling. In the simulations transport across regions incorporating heterostructures requires carriers of sufficient energy to be thermionically transported over barriers.

The limitations imposed by use of transport equations that do not incorporate tunneling was assessed through a program whose development has advanced sufficiently far to be used in the present study. This program involved use of a quantum Liouville equation algorithm for examining transport perpendicular to the conducting channel. Indeed this study incorporates the **first** use of the Liouville equation in the study of transport in heterostructure FETs.

The implementation of the Liouville equation was undertaken to provide a foundation for several assertions made during the course of the study. The Liouville equation algorithm is not developed to the point where it can be used to study the full two dimensional transport within the FET. For example, as revealed by the semi-classical equations, it was found that when contacts were placed on heavily doped narrow band gap material, as indicated by the GE structure, only thermionic carriers would be transported over the wide band gap material into the quantum well. This led to numerically computed values of the source to drain current that were significantly below the measured values. To deal with these

differences two issues were raised: (1) will numerical simulations that included the metalization of the region underneath the source and drain contacts permit sufficient current to flow, if it assumed that the metalization does not destroy the heterostructure interface; and (2) is the integrity of the heterostructure intact, even after metalization. Clearly the former question can only be addressed quantum mechanically, while a negative answer to the second would indicate that the heterostructure is irrelevant under the source and drain regions.

The supplementary Liouville equations address the more complicated problem of transport through the heavily doped alloyed region and determined that **for sufficiently high doping in the narrow and wide band gap materials the tunnel barrier is reduced in height and becomes thin enough for a significant number of carriers to tunnel towards the quantum well. Thus a sufficient number of carriers can now get through the heterostructure and higher levels of current are possible.** (It is noted that virtually all previous simulations of heterostructure FETs ignore the narrow band gap region associated with the source and drain contacts and instead assert that contact is made directly to the wide band gap material. While this is satisfactory from a numerical point of view it is completely unsatisfactory from the point of view of developing an algorithm for designing HEMT structures.) Once it was established that the current could be transported from the narrow to the wide band gap material and then into the quantum well simulations using the moments of the Boltzmann transport equation were undertaken. Since these simulations do not permit the presence of tunneling, unless such concepts as the quantum potential are introduced, which was not done in this study, it was necessary to remove the conduction band discontinuity associated with the heterojunction in the vicinity of the source and drain contacts. This approximation is one that is within the spirit of the Phase I study, but will not be a limitation of the proposed Phase II study.

The discussion that follows is in two parts with respect to the simulations. First the principle discussion is connected with the simulations that were performed using the moments of the Boltzmann transport equation. The supplementary discussions involving solutions of the Liouville equation follow the discussion of the moment equations. A detailed discussion of the Liouville equations is not part of this text. It is incorporated as an appendix to this document. The appendix is a copy of a final report submitted to ARO during September 1990.

Under the present analysis two species of electrons, central (low effective mass) and satellite (high effective mass) valley carriers are considered. The governing equations are

Continuity

$$(1) \qquad \partial n_1/\partial t = -\nabla \cdot (n_1 v_1) - n_1 f_1 + n_2 f_2$$

$$(2) \qquad \partial n_2/\partial t = -\nabla \cdot (n_2 v_2) + n_1 f_1 - n_2 f_2$$

where n_1 and n_2 are the central valley and satellite valley carrier number densities respectively and V_1 and V_2 are the corresponding velocities. f_1 and f_2 are the scattering integrals for particle conservation.

Momentum

$$\partial(n_1\mathbf{P}_1)/\partial t = -\nabla\cdot(n_1\mathbf{V}_1\mathbf{P}_1) - \nabla p_1 - n_1 e\mathbf{F}_n - n_1\mathbf{P}_1 f_3 + (n_1\mathbf{V}_1\cdot\mathbf{V}_1\nabla m_1)/2 + (n_1 T_1\nabla m_1)/m_1 \quad (3)$$

where the momentum, $\mathbf{P}_1$, and the field, $\mathbf{F}_n$, are defined by

$$\mathbf{P}_1 = m_1\mathbf{V}_1 \quad (4)$$

$$\mathbf{F}_n = -(\nabla\psi + \nabla\chi/e) \quad (5)$$

m_1 is the mass of the central valley carrier, e is the electron charge, ψ is the electric potential and χ is the electron affinity. $\mathbf{F}$ is the field due to potential differences and conduction band discontinuity arising from material variations. The partial pressure, p_1, is related to the central valley carrier temperature, T_1, and number density by the perfect gas relationship, which results from the assumption of Boltzmann statistics,

$$p = n_1 k T_1 \quad (6)$$

where k is Boltzmann's constant. f_3 is the scattering integral for the central valley carrier momentum. Contributions to f_3 include impurity, acoustic phonon, polar phonon, nonpolar intervalley scattering. The effects of electron-hole scattering is accounted for through an enhancement of the impurity scattering. These quantities are computed from scattering integral expressions.

It is noted that the satellite valley electrons, due to their heavy mass, are extremely slow. Compared to the speed of the central valley electrons, the velocity of the satellite valley carriers is negligible for a wide range of bias conditions. For this reason, to first order, it is possible to ignore the contribution of the satellite valley carrier to the total current. Their velocity may be assumed to be zero and the satellite valley momentum equations need not be solved. This approximation is valid provided that relative local population of satellite valley carriers is, for example, less than an order of magnitude greater than that of the central valley carriers. If the population were approximate equal, the typical error in total current would only approach 5 to 10 percent.

On the basis of a broad range of calculations involving the solutions of the energy and momentum balance equation for the central and satellite valley carriers it has been our experience that for a wide range of conditions an accurate representation of device behavior can be obtained with this approach. Under this assumption, equation (2) for carrier continuity becomes simply a rate equation.

Energy

Next we write an energy equation for the central valley electrons. There are various forms in which the central and satellite valley carrier energy equations can be described. We choose to cast the energy equations in terms of the central and satellite valley temperature:

$$\partial(n_1T_1)/\partial t = -\nabla\cdot(n_1V_1T_1) - (2/3)\, n_1T_1\nabla\cdot V_1 + (2/3k)\,\nabla\cdot(\kappa\nabla T_1)$$

(7)

$$+3V_1\cdot V_1 m_1 [n_1(2f_3-f_1) + n_2f_2]$$

$$-n_1T_1f_5 + n_2T_2f_6 - (n_1T_1V_1\cdot\nabla m_1)/m_1$$

In equation (7) f_5 denotes energy relaxation within the central valley plus energy exchange with the satellite valley; f_6 denotes energy exchange between the satellite and central valley. Satellite valley carriers, due to their heavy mass, exchange very efficiently with the substrate, and hence their temperature is assumed to be equal to the background temperature. Again this has been verified in a wide number of simulations employing a satellite valley energy equation.

Finally, the potential is related to the total number density through Poisson's equation.

(8) $$\nabla\cdot\epsilon\nabla\psi = e(n_1+n_2-n_0)$$

In two dimensions, the complete problem description requires 2 continuity equations (1), and (2), two components of the vector momentum equation (3), an energy equation (7) and Poisson's equation (8), a total of six equations. The boundary conditions are as follows: The carrier densities at the Ohmic contacts are fixed at the value of local doping whereas for Schottky Contact, assumed to be zero. The temperature of the carriers at the contacts is assumed to be 300° k. For velocities, the normal gradient is assumed to be zero. The boundary condition for potential is the applied voltage plus an appropriate built-in potential.

The governing equations are solved numerically using a procedure similar to that developed by Kreskovsky and Grubin [2] for the drift and diffusion equations. The procedure is highly efficient and is based on the Linearized Block Implicit (LBI) method of Briley and McDonald [3].

Results

Preliminary Calculations: The results discussed below are presented in a manner that best emphasizes its didactic value. The discussion does not follow the chronological order in which they were obtained.

The preliminary device structure presented to SRA by GE is schematically represented in Fig. 1. The structure includes, successively:

(1) a narrow gap InGaAs cap layer 100Å thick, doped to $2\times10^{18}/cm^3$, providing a means of making ohmic source and drain contacts;

(2) a 200Å thick, nominally undoped AlInAs wide bandgap region;

(3) a Si planar-doped layer, 10Å thick, doped to yield an expected surface charge density of $5\times10^{12}/cm^2$,

(4) a 30Å AlInAs undoped spacer layer;

(5) an 800Å InGaAs undoped quantum well;

(6) an undoped 0.25μ AlInAs buffer layer;

(7) an InP substrate (not shown) which was not considered in the simulation.

The AlInAs barrier, including the planar-doped region and the 30Å barrier is 240Å thick. For the gate region, this AlInAs barrier is recessed, with the gate metalization resting 140Å from the AlInAs/InGaAs interface.

For the structure shown in figure 1, the principle issue is the one discussed above, namely, **how do which carriers reach the quantum well?** Typical simulations of HEMT structures ignore the narrow band gap cap regions associated with the source and drain contacts, and instead assume that contact is made directly to the wide band gap material. In that case the issue of carriers going from the source to the quantum well is not an issue-they are not required to climb the potential hill provided by the wide band gap material. For the structure considered here in which accurate device modeling is undertaken, climbing the potential hill is an issue. As discussed above the two approaches to dealing with this problem include ignoring the wideband gap material and assuming that the metalization destroys the integrity of the heterostructure region; or assuming that the metallization does not destroy the integrity of the heterostructure region, but that heavy doping results in a significant alteration of the self-consistent potential seen by the carriers, thereby permitting significant charge to enter the quantum well. The consequences of assuming that the metalization destroys the integrity of heterostructure are self-evident. The consequences of assuming that the metalization alters the potential is less clear; and to study this SRA's density matrix algorithm was invoked for solving the Liouville equation. A discussion of this algorithm and the relevant equations is contained in a report recently submitted to ARO and is incorporated here in the appendix.

The quantum mechanical algorithm was implemented to examine the equilibrium distribution of charge and potential along a line perpendicular to the quantum well. In examining the possibility of transport from the source to the quantum well two types of calculations were performed, as illustrated in figure 2. In this calculation, the InGaAs cap layer was somewhat thicker, 300Å, than that of figure 1. The structure labeled nominal, incorporated doping only within the InGaAs cap layer and in the silicon doped spike layer. The structure labeled metalization extended the indicated doping into the channel. **It is important to note that if the fabrication process results in doping to the quantum well, then in addition to transport arising from the two dimensional electron gas, there will also be a contribution from charge injected from the source region.** Thus both space charge injected transport as well as two dimensional electron transport occur. The space charge injection transport may be an issue in a structure of the type shown in figure 1, where the source to drain spacing is only 1.8 microns. This is an important design and device physics issue to be addressed during a Phase II portion of the study. These questions arise because results reported in the literature [4] indicate that both surface doped as well as surface undoped HEMTs have very similar characteristics. It was suggested in [4] that, since the cap layer on InGaAs was very thin, the metalization of the ohmic contacts could penetrate the 2 DEG region.

The distribution of charge in the nominal structure is shown in figure 3. There is an accumulation of charge on the InGaAs side of the barrier, as is expected at the narrow-band gap wide-band gap interface. There is very little spill over into the AlInAs region. The 2-DEG region in the channel to the right of the AlInAs/InGaAs interface is very evident as is an absence of charge in the region of the planar doped region. Thus, the planar doped region acts as expected, providing seeding for the 2-DEG. Note that the sheet charge density in the quantum well is approximately $3x10^{12}/cm^2$, which is close to the value expected. But, the most significant effect observed here is the absence of carriers in the AlInAs region. Thus this region, in the absence of metalization, is highly resistive and would not allow significant current flow. The potential energy distribution for this structure is shown in Fig. 4. Here the effects of accumulation on both sides of the AlInAs layer are apparent, the downward bending of the band structure, and the depletion of the planar doped layer is evident as the deep notch at -100Å. Of significance is the remaining wide barrier on the source side of the structure. While tunneling currents are anticipated the magnitude is expected to be small. The results of figures 3 and 4 suggest that the absence of any device processing to accommodate the wide barrier is likely to lead to low current levels in the structure.

The distribution of charge in the metallized structure is shown in figure 5. Several key features are apparent. (1) There is increased charge in the InGaAs cap region, particularly at the source-side interface region. (2) There is increased density of carriers within the wide band gap region, (3) The quantum well charge density has more than doubled. (4) Due to the presence of dopants in the AlInAs region, a secondary peak appears in the area of the planar doped layer and the carrier density remains within an order of magnitude of the doping in spite of the effects of the barriers. As can be seen, in comparison to the results of figure 3, the wide band-gap region is now, relatively, highly conductive. The self consistent potential energy distribution is shown in figure 6. The key point of this calculation is that the heavy doping reduces the energy of the barriers seen by those electrons entering from the source, and that the barriers are considerably reduced in thickness, permitting the presence of a larger tunneling contribution. **The broad conclusion of the results of figures 5 and 6, is that doping the heterostructure region sufficiently high is likely to permit all of the necessary current to flow from the source contact.**

The Liouville equation was also implemented to examine the structure of the 2-DEG in the region under the gate. In this simulation the gate is on the wide band gap AlInAs with the planar doped, $5x10^{18}/cm^3$, layer 200Å from the surface. The AlInAs/InGaAs quantum well interface is at 140Å. The density distribution in figure 7 shows that the peak concentration of the electron gas is reduced substantially below the nominal value under the source contact (with or without the metalization). The band structure for this result, shown in figure 8 reflects the depletion of the wide band gap region. The Schottky barrier height for this calculations was approximately 0.5 eV.

Transistor Calculations, Initial Structure:

The transistor calculations were preformed for the structure shown in figure 1. The parameters used in the calculations are shown in Tables I and II. The steady state, uniform field velocity field relation arising from these tabulated values is displayed in figure 9. Note that the steady state values were not used in the study but are included here for completeness in the discussion.

The initial transistor calculations included the presence of the undoped AlInAs barriers, and a planar-doped layer of $5x10^{18}/cm^3$, as in the quantum mechanical calculations. (Calculations with planar-doped layers of $5x10^{19}/cm^3$ did not converge). The grid structure used in these initial efforts is shown in figure 10. As shown, the grid is unequally spaced and consists of 94 grid points normal to the gate and 102 grid points in the source-drain direction. At a planar doped layer of $5x10^{18}/cm^3$ a surface charge density of $5x10^{11}/cm^2$ was obtained; this value is below that obtained with the quantum mechanical treatment. Two bias points were computed for this case, $V_{DS} = 1.0$ volts for $V_{gs} = 0$ and 0.5 volts. While converged solutions were obtained, the source-drain current level was very small and approached the limit of the accuracy of the simulation code. The explanation for this low current is consistent with the discussion associated with figures 3 and 4, and rests entirely with the lack of carriers tunneling through the wide band gap material. However, the simulation did show qualitatively the correct behavior, consistent with these low current levels. A high density 2-D electron gas was generated at the AlInAs/InGaAs interface in the channel, and a high degree of confinement was observed. This is shown in Fig. 11, where contours of carrier density are presented for $V_{DS} = 1$, $V_{gs} = 0$.

Transistor Calculations, Initial Structure without AlInAs Buffer:

For the bias levels chosen the initial calculations demonstrated that the charge in the channel was confined to the interface between the spacer layer and the channel; very little charge was present at the substrate side of the quantum well. To eliminate the possibility that the small source to drain current levels did not arise from computational limitations, additional simulations were performed that excluded the AlInAs buffer region. Thus only two heterostructure contributions were considered. Quantum mechanical quasi-bound state contributions, which can only be treated within the framework of quantum transport, and were not treated in the initial calculation, are less of a feature in this semi-classical calculations. From a physics point of view, the present calculations are relevant only for situations in which the charge distribution is low enough throughout the quantum well. From a computational viewpoint eliminating the buffer layer permitted increased resolution in the active region of the device, to improve accuracy without increasing the total number of grid points. The new grid structure is shown on the same scale as that of figure 10 in figure 12. The same grid is shown with 3x magnification in the direction normal to the gate in figure 13. The $V_{DS} = 1.0$, $V_{gs} = 0$ solution was computed on this reduce domain with increased resolution and the results were found to be qualitative similar to those obtained for the complete device. Again, **the predicted current was extremely low. The conclusion: the low current levels are not grid dependent but were physics dependent.**

Transistor Calculations, Initial Structure without AlInAs Buffer and with Modification of the Source and Drain Regions:

The next stage of the approximation rested on the quantum mechanical calculations which suggested that the origin of the low current lies in the inability of the semi-classical equations to adequately represent the effects of tunneling through the barriers through, e.g., heavy doping. Thus, the first barrier between the InGaAs and AlInAs layer was removed to allow carriers to enter the quantum well freely.

The resulting simulations yield the calculated I-V characteristics shown in figure 14, along with the experimental data obtained by General Electric. As can be observed in the comparison some discrepancies exist, particularly at low voltages where the resistivity is higher in the calculated curves. Part of this discrepancy lies in the reduced density (in comparison with the quantum transport calculations) associated with the spike doping layer. However, **the overall level of agreement with experiment is excellent.** At $V_{gs} = 0$, the current level is slightly over predicted, at $V_{gs} = 0.2$ the agreement is better and at $V_{gs} = 0.4$ the current level is under predicted. There is also a discrepancy between the predicted and observed current levels for V_{DS} below about 0.5 volts. This would seem to be related to the effective low field mobility of the mobile electrons. Here we note that the scattering rates used in the present simulations were based on purely classical effects and no attempts were made to account for modified scattering in the region of the 2-D electron gas.

Contour plots of potential, carrier density, the magnitude of the central valley velocity and temperature are shown in figure 15 for $V_{DS} = 1.5$, $V_{gs} = 0$ and in Fig. 16 for $V_{DS} = 1.5$, $V_{gs} = 0.4$. The results indicate:

(1) The potential contours clearly show that a significantly higher field region exists downstream of the gate contact for $V_{gs} = 0.0$v, than for $V_{gs} = 0.4$v. In these two contour plots the voltage increment between contours is approximately equal. Thus, the greater the distance between contours the lower the field.

(2) The density contours display substantially more depletion under the gate for $V_{gs} = 0$ than for $V_{gs} = 0.4$. In both cases, the maximum carrier density occurs in the 2-DEG under the drain. Carriers are drawn to the drain end of the 2-DEG in the channel due to the field distribution in this region. For $V_{gs} = 0$, with its lower current level, the carriers are more uniformly distributed along the entire length of the 2-DEG.

(3) The contours of the magnitudes of the central valley velocity indicates, as would be expected from the carrier distribution, that the peak velocity is higher for $V_{gs} = 0$, due to the greater depletion under the gate and higher field in this region. This peak occurs in the area just downstream of the gate as does the temperature peak. However, we note that for $V_{gs} = 0.4$, the peak velocity occurs much closer to the interface between the spacer layer and the channel. For $V_{gs} = 0$, the peak velocity occurs at a position near the middle of the channel. This is due, in part, to the deeper penetration of the depletion layer under the gate for $V_{gs} = 0$.

(4) The central valley temperature contours indicate that for $V_{gs} = 0.4$ the temperature contours are more spread out than for $V_{gs} = 0.0$ due to the lower field under the drain side of the gate. The peak temperature is also lower.

An enlargement of the density contours are presented in Fig. 17. Here, the direction normal to the gate surface is magnified 3x. These plots clearly show the pinching off of the 2-DEG for $V_{gs} = 0$. The depletion region is deeper and much more heavily depleted. When these plots are viewed with the velocity magnitude plots in mind it becomes even more apparent why the peak velocity occurs deeper in the channel for $V_{gs} = 0$. Even though the current is reduced, the current path follows the carrier distribution and, therefore, extends deeper into the channel around the gate depletion region.

RF Device Characterization:

GE Syracuse provided small signal characterization of the structure shown in figure 1, in particular the transconductance, cutoff frequency, S parameters, etc. To begin to determine a correlation with these measurements as well as to begin to determine the noise characteristics of the structure, small signal admittance calculations were performed. These calculations involve using the algorithm for solving the moments of the Boltzmann transport equation.

For the admittance studies, at select drain and then gate bias points perturbations to the steady dc results were imposed, with the subsequent transient response computed. The current response was then Fourier analyzed to determine the 'Y' parameters which are defined as:

(9a) $$Y_{12}(\omega) = \delta I_g(\omega)/\delta \psi_d(\omega)$$

(9b) $$Y_{22}(\omega) = \delta I_d(\omega)/\delta \psi_d(\omega)$$

(9c) $$Y_{11}(\omega) = \delta I_g(\omega)/\delta \psi_g(\omega)$$

(9d) $$Y_{21}(\omega) = \delta I_d(\omega)/\delta \psi_g(\omega)$$

Here I and ψ are current and potential; the subscript g and d stand for gate and drain respectively. The Y parameters are dependent on the space charge and potential distribution, and hence are bias dependent. Usually it is assumed that about a given bias condition, the small signal currents add linearly, permitting an equivalent circuit representation. While the calculations arising from equations (9) admit an equivalent circuit interpretation, during the Phase I study only the simplest zeroth order equivalent circuit model was used for interpretation of the results. As the discussion below indicates the equivalent circuit model requires generalization.

The zeroth order equivalent circuit model indicates:

(10a) $$Y_{110}(\omega) = -j\omega(C_{gs} + C_{gd})$$

(10b) $$Y_{120}(\omega) = j\omega C_{gd}$$

(10c) $$Y_{210}(\omega) = g_m + j\omega C_{gd}$$

(10d) $$Y_{220}(\omega) = g_{ds} - j\omega C_{gd}$$

where C_{gs} denotes that gate-source capacitance, C_{gd}, the gate-drain capacitance, g_m the transconductance, and g_{ds} the forward conductance. The zeroth order model does not permit the introduction of a realistic gate impedance or the contributions from the source metalization. Rather, it emphasizes the intrinsic circuit elements.

The admittance calculations are displayed in figures 18, for a gate perturbation and in figure 19 for a drain perturbation. The dc points for this calculation are $V_d = 1.0$v, $V_g = 0.2$v. The dc points for the measurements are $V_d = 0.9$v, $V_g = 0.25$v. The results

shown in figure 14 suggest that the transconductance comparisons at these drain bias levels would show good agreement at values of gate voltage closer toward pinchoff. Under forward gate bias the calculations underpredict the current levels and the transconduction values. Indeed, the calculated transconductance at this point is 777mS/mm; the measured transconductance is 1273mS/mm.

The gate-source capacitance is a measure of the ability of the gate region to store charge. Measurements yield C_{gs} = 0.0642pf, and the calculations yield an order of magnitude smaller capacitance, 0.006pf. However, the gate-drain capacitance measurements and calculations reveal order of magnitude similarities: The measurements yield 0.007pf, the calculations 0.005pf. The discrepancies in the source-gate capacitance are attributed to the model dependent distribution of charge that arises from eliminating the heterostructure interface in the source region, and the lack of charge confinement under the source. This region needs to be modeled more carefully during the Phase II aspects of the study. Along the drain side where there is a broad Schottky barrier depletion region, there is less of an issue.

The cutoff frequency is obtained from the admittance parameters. The calculated values are anticipated to be much higher than experiment because of the underestimation of the gate-source capacitance. The measured cutoff frequency was 196GHz, that calculated was approximately 700GHz.

During the small signal studies, consideration was given to calculating the noise figure for the structure. While the noise figure in principle can be calculated from a generalized admittance parameter calculation, realistic calculations from this procedure would require a more accurate representation of the properties of, e.g., the gate region. For example, one of the representations of the noise figure is that given below:

$$F_O = 1 + K_\ell L f (g_m (R_g + R_s)^{1/2} \qquad (11)$$

where K_ℓ is a fitting parameter, R_g is the gate resistance and R_s is the source resistance. In the present model the gate is represented as a boundary condition, and no resistance is attributed to it, while the inadequacy of the modeling of the source contact region, led to an underprediction of the current flow at the low values of drain bias. Thus it was not deemed useful at this point to examine the details of the noise figure from the admittance parameters at this point; but instead to pursue this during the Phase II portion of the study.

Conclusions and Recommendations

The Phase I study demonstrates that a realistic numerical model based upon nonequilibrium transport equations coupled to concepts obtained from quantum mechanical transport equations can provide order of magnitude agreement with the dc electrical characteristics of the high frequency low noise pseudomorphic HEMT structures.

The Phase I study exposed the requirements for obtaining a meaningful device model for examining transport in these structure:

(1) Quantum transport must be incorporated into the study, either in the complete simulation, or as a guide to the necessary approximations within the framework of the

moment equations. In this regard, incorporating the quantum potential into the framework of the algorithm should be given very serious consideration.

(2) The properties of the source and drain metalization are an issue that must be addressed if predictions of the noise properties are to be made. In the case of the source metalization, the procedures for examining transport perpendicular to the heterointerfaces are likely to provide the key frequency dependence of the effects of the source and drain resistance on the admittance parameters. Recently, a similar procedure involving the quantum Liouville equation has been undertaken at SRA to examine the details of transport through the Schottky barrier. This should also be pursue during the Phase II part of the program.

(3) One of the goals of the SRA program is to provide a tool that will be used by design engineers to fabricate devices. Presently design engineers are accustomed to using such codes as Libra for examining the small signal properties of devices. The code developed at SRA is a physics based code that while capable of providing the full admittance parameters, will not pass the 'familiarity' test. Thus a means must be provided for interfacing the SRA code with those commercially available. One proposed task will be to find a means of interfacing the SRA code with the commercial codes.

(4) The test of the efficacy of the program is the ability to design a successful transistor. Thus, it is recommended that any future program involving the design of HEMTs provide a demonstration that it can successfully design devices. The proposed Phase II program will provide approximately one-third of the anticipated funding to the fabrication and testing of HEMT structures.

References

1. A. S. Brown et al, GaAs IC Symposium, Tech. Digest, p. 143, 1989.
2. J. P. Kreskovsky and H. L. Grubin, J. Comp. Phys., Vol. 68, No. 2, Feb. 1987.
3. W. R. Briley and H. McDonald, J. Comp. Phys., Vol. 34, No. 1, Jan. 1980.
4. Y. C. Pao, et al, IEEE Trans. Elec. Dev., Vol. 37, No. 10, Oct. 1990.

TABLE 1. $In_{(.53)}Ga_{(.47)}As$ Parameters Used in Calculations

Parameter	Γ	L	Common
Number of equivalent valleys	1	4	
Effective mass (m_e)	.041	.0291	
Γ-L separation (eV)			.55
Polar optical scattering			
Static dielectric constant			13.91
High-frequency dielectric constant			11.32
LO phonon (eV)			.0327
Γ-L scattering			
Coupling constant (eV/cm)			$1.0x10^9$
Phonon energy (eV)			.0278
L-L scattering			
Coupling constant (eV/cm)		$5x10^9$	
Phonon energy (eV)		.0327	
Acoustic scattering			
Deformation potential (eV)	7.0	9.29	
Nonpolar scattering (L)			
Coupling constant (eV/cm)		$3x10^8$	
Phonon energy (eV)		.0327	
Acoustic velocity (cm/sec)			$4.101x10^5$
Density (cm/cm^{xx3})			5.48

TABLE 2. $Al_{(.48)}In_{(.52)}As$ Parameters Used in Calculations

Parameter	Γ	X	Common
Number of equivalent valleys	1	3	
Effective mass (m_e)	.084	.47	
Γ-X separation (eV)			.55
Polar optical scattering			
Static dielectric constant			11.25
High-frequency dielectric constant			10.25
LO phonon (eV)			.038
Γ-X scattering			
Coupling constant (eV/cm)			$8x10^8$
Phonon energy (eV)			.03
X-X scattering			
Coupling constant (eV/cm)		$5x10^9$	
Phonon energy (eV)		.038	
Acoustic scattering			
Deformation potential (eV)	7.0	9.27	
Acoustic velocity (cm/sec)			$4.46x10^5$
Density (cm/cm^{8x3})			4.774

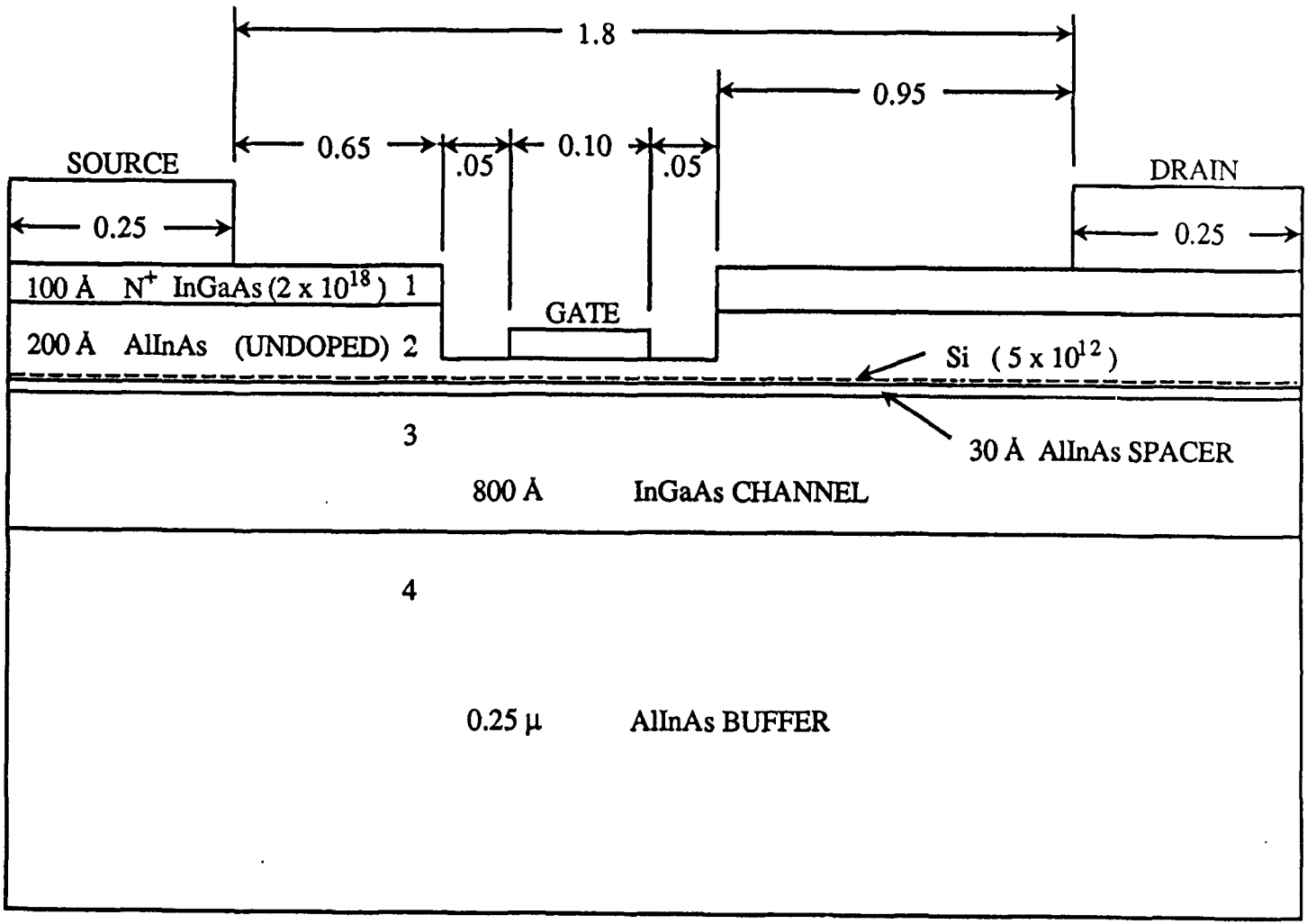


Figure 1. HEMT Structure used in Phase I Study.

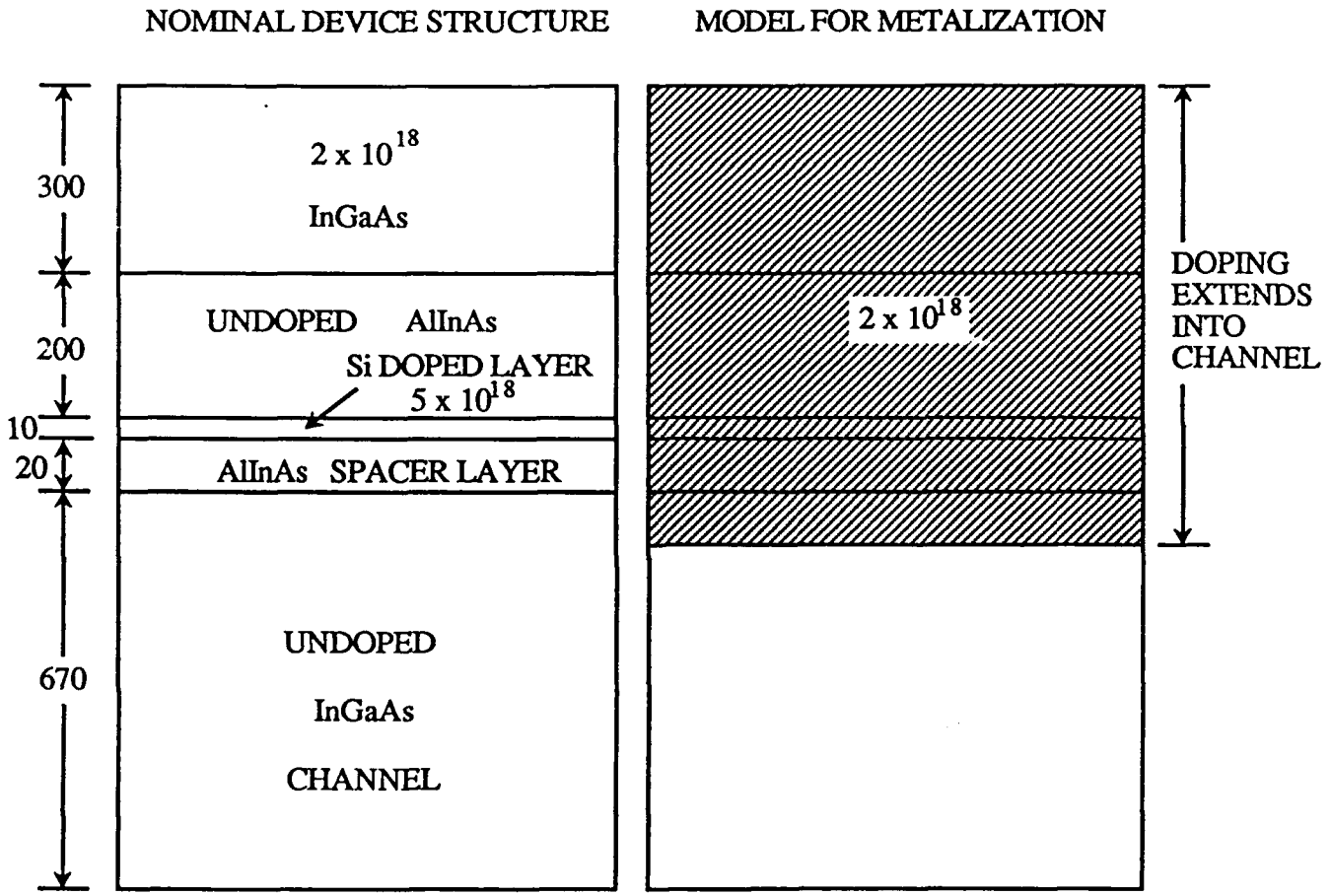


Figure 2. Schematic of Nominal Device Structure and Structure used to Model Metalization in Density Matrix Calculations.

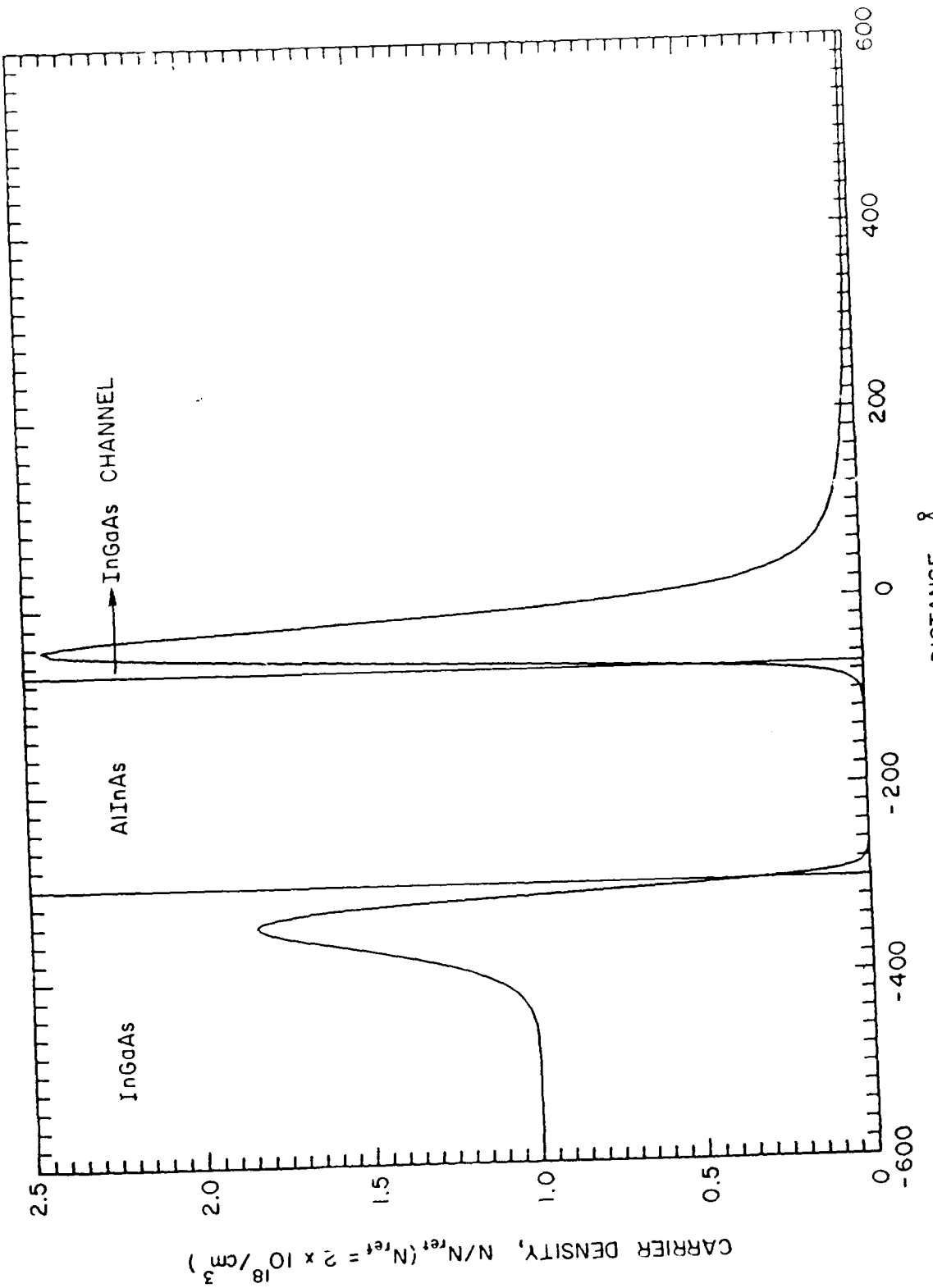


Figure 3. Density Distribution Predicted for Nominal Device Structure using Density Matrix Equation.

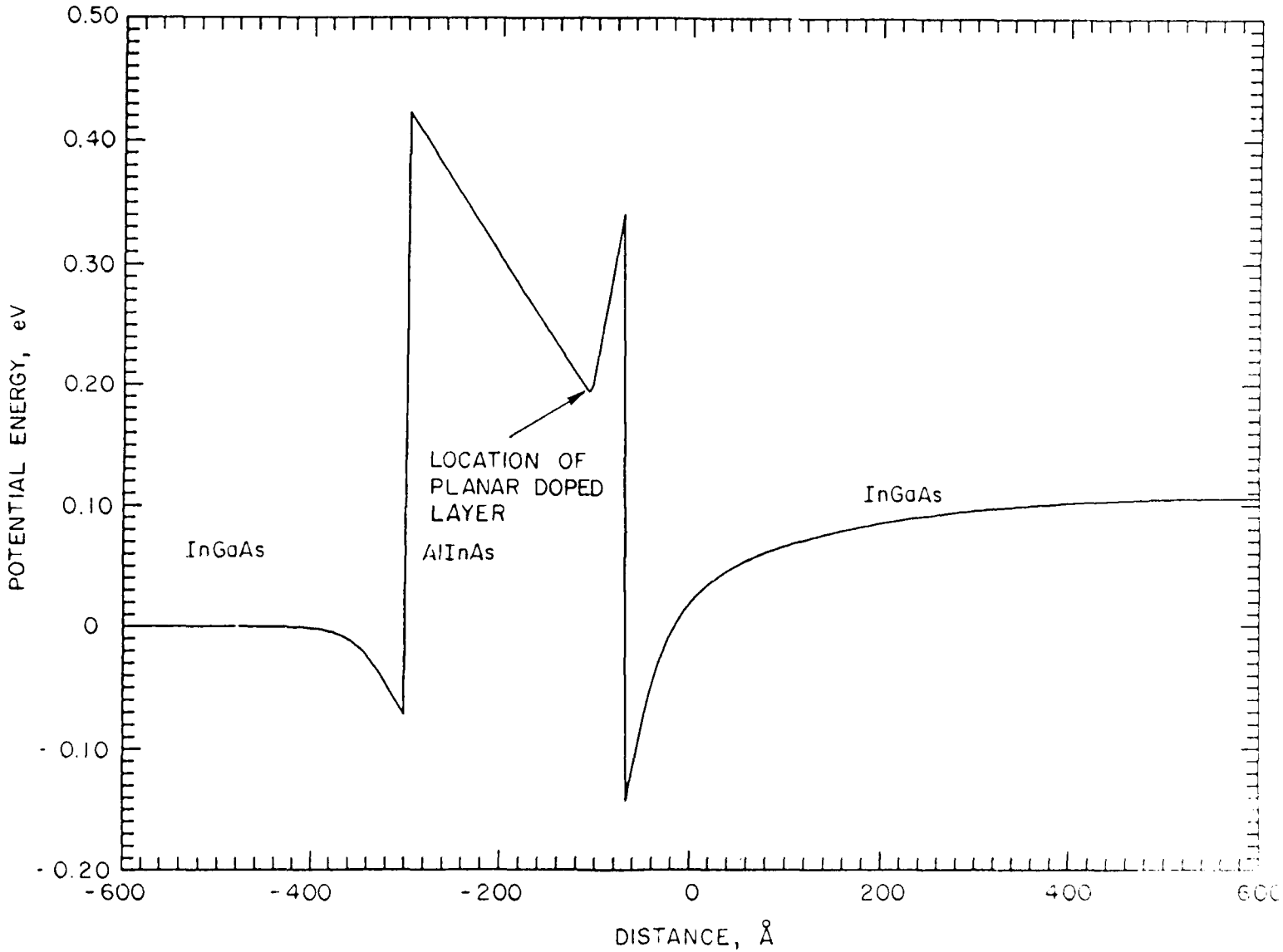


Figure 4. Self-consistent Energy Distribution from Density Matrix Solution for Nominal Device Structure.

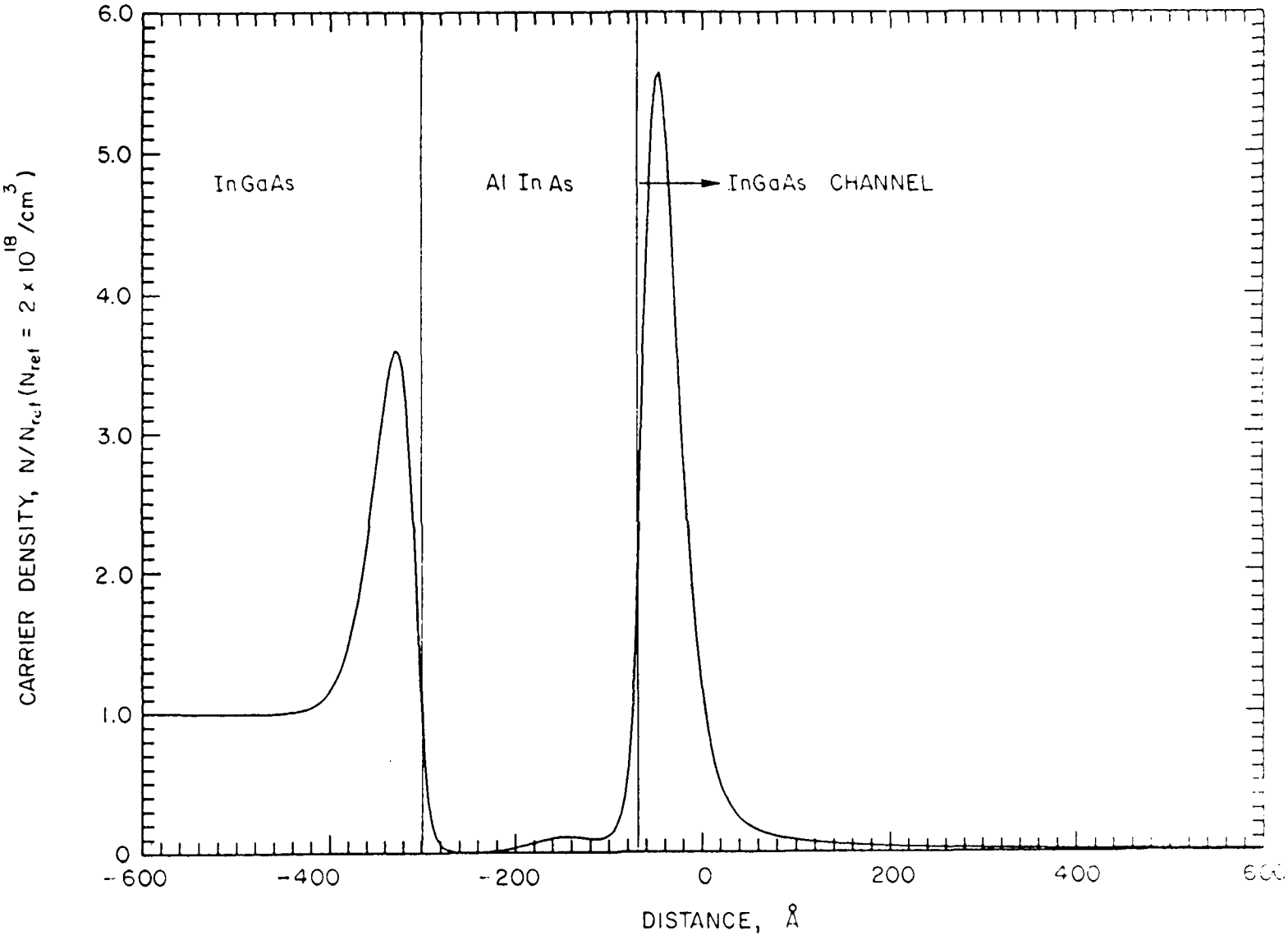


Figure 5. Density Distribution Predicted for Nominal Device Structure with Metalization using Density Matrix Equation.

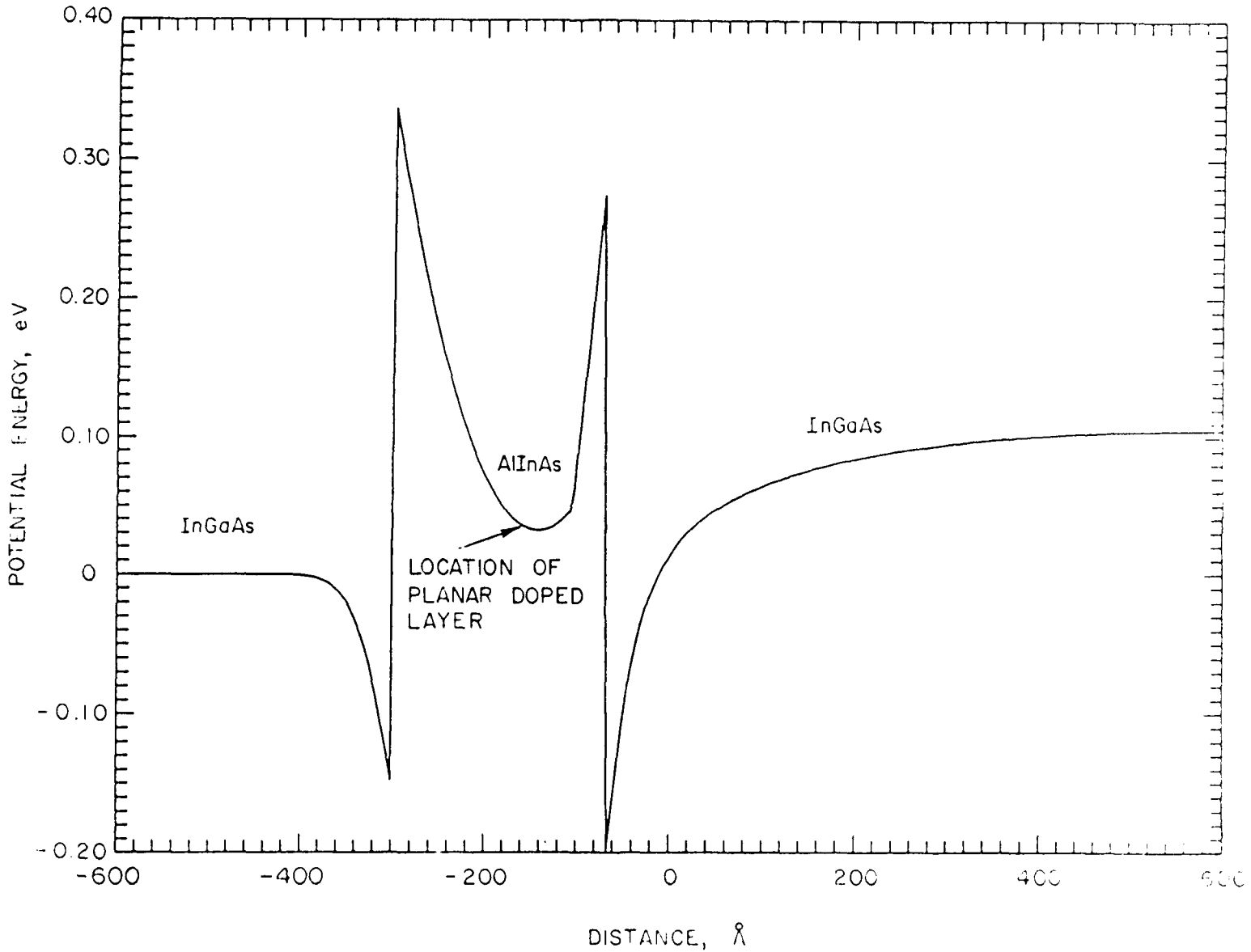


Figure 6. Self-consistent Energy Distribution from Density Matrix Solution for Nominal Device Structure with Metalization.

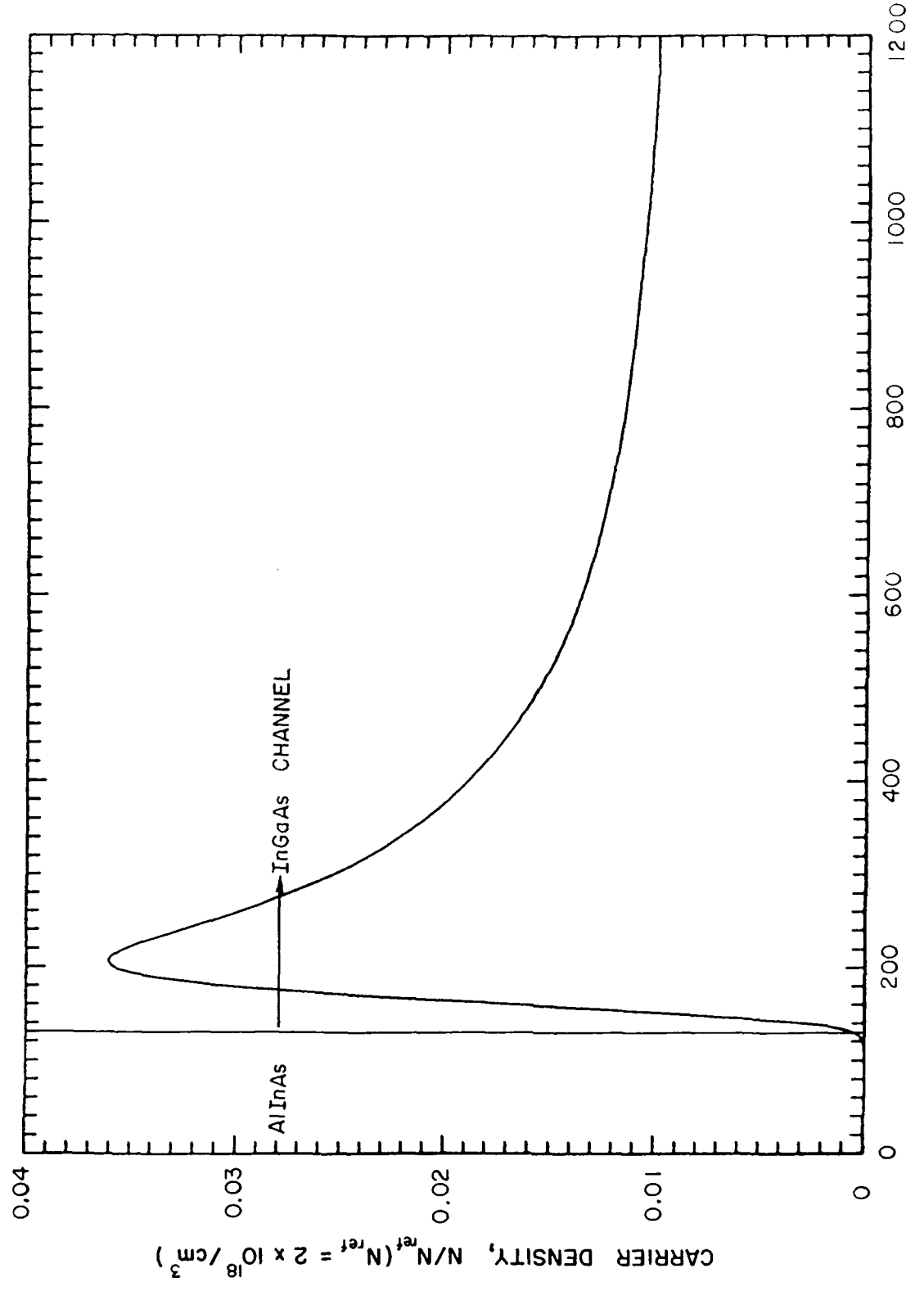


Figure 7. Carrier Density Distribution Under the Gate as Determined from Density Matrix Equation.

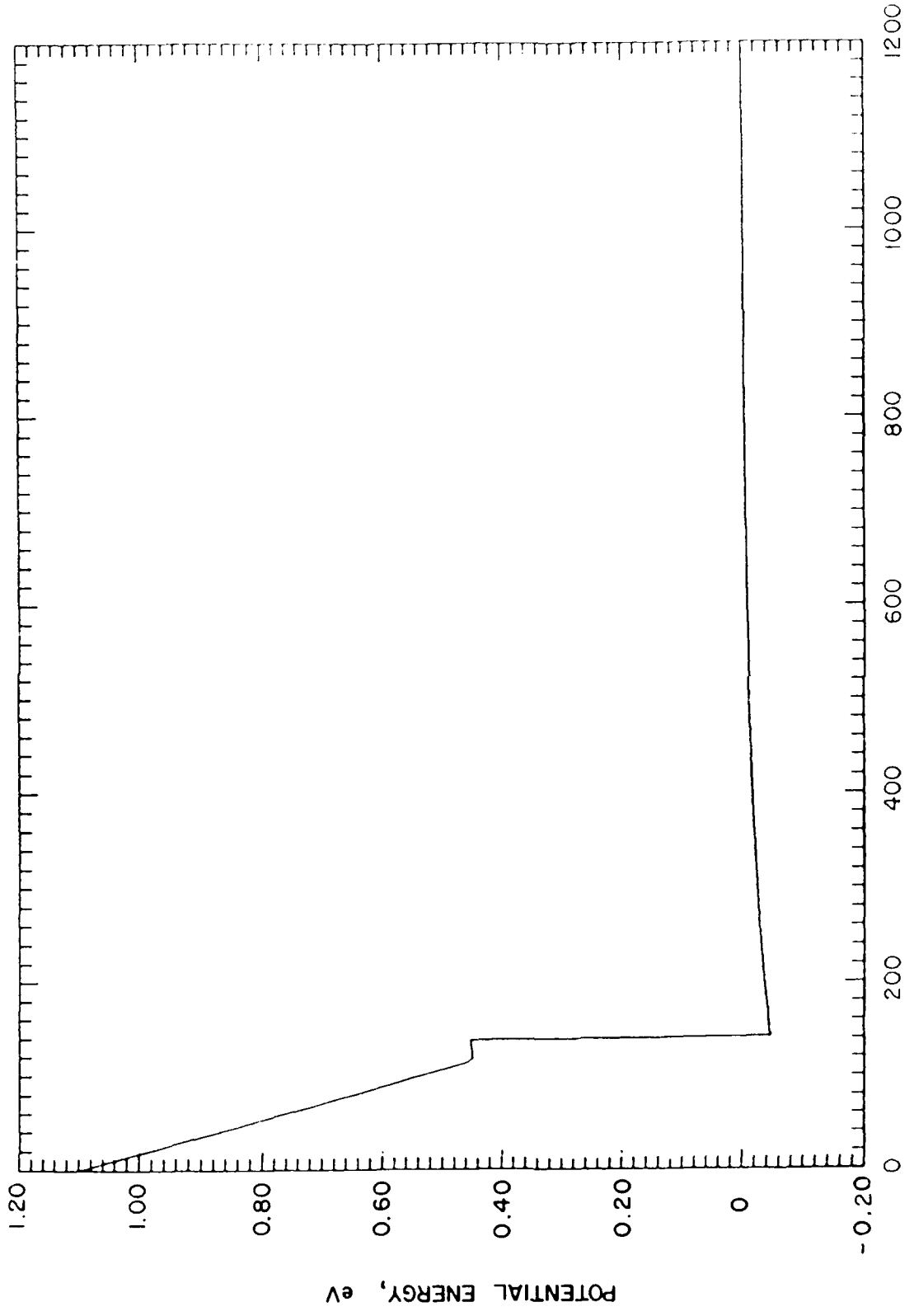


Figure 8. Energy Distribution Normal to Gate as Determined from Density Matrix Equation.

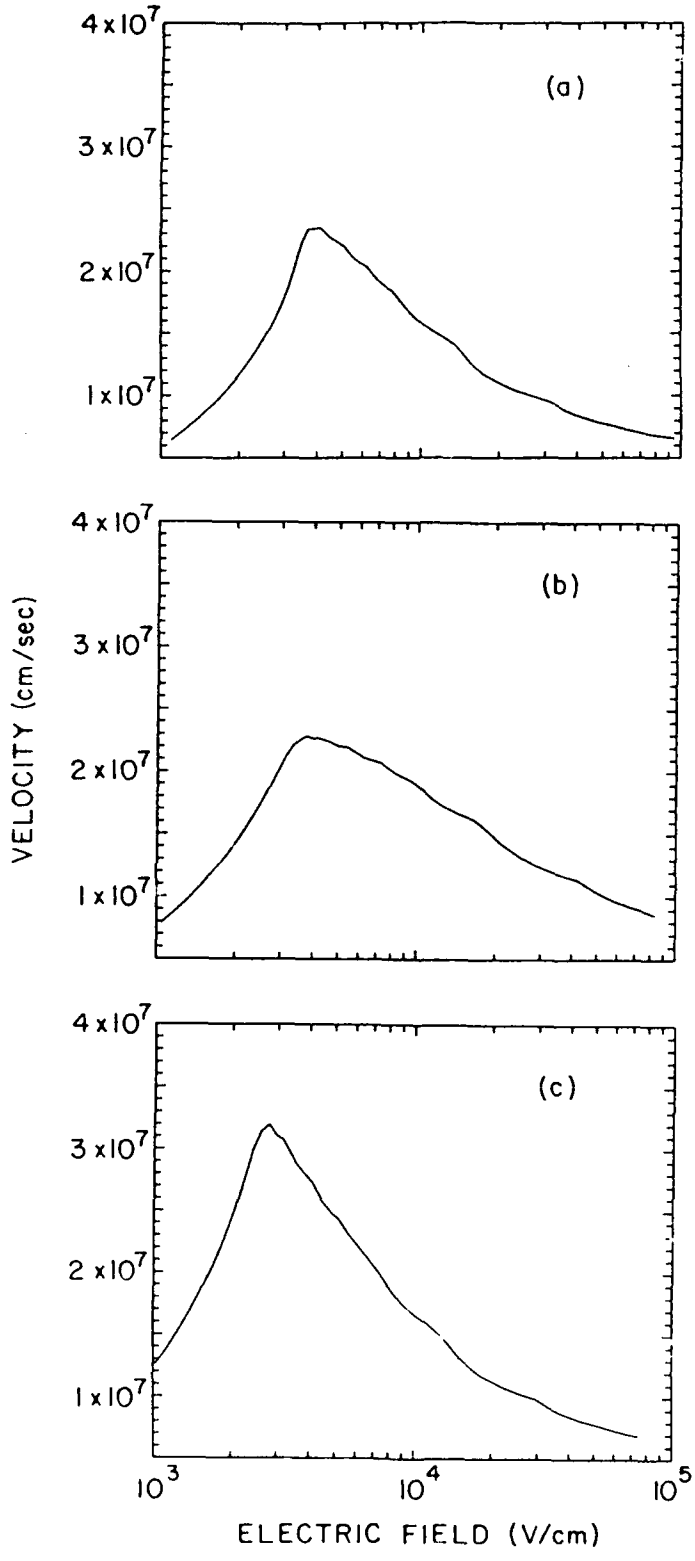


Figure 9. Velocity Field Curve Obtained using Data of Table 1, a) $In_{53}Ga_{47}As$ 1 x $10^{18}/cm^3$ Doping, b) $Al_{48}In_{52}As$ Undoped, c) $In_{53}Ga_{47}As$ Undoped.

1 = N^+ InGaAs CAP LAYER
2 = UNDOPED AlInAs SCHOTTKY ENHANCEMENT LAYER
3 = Si DOPED PLANAR & SPACER LAYERS
4 = InGaAs CHANNEL
5 = AlInAs BUFFER LAYER

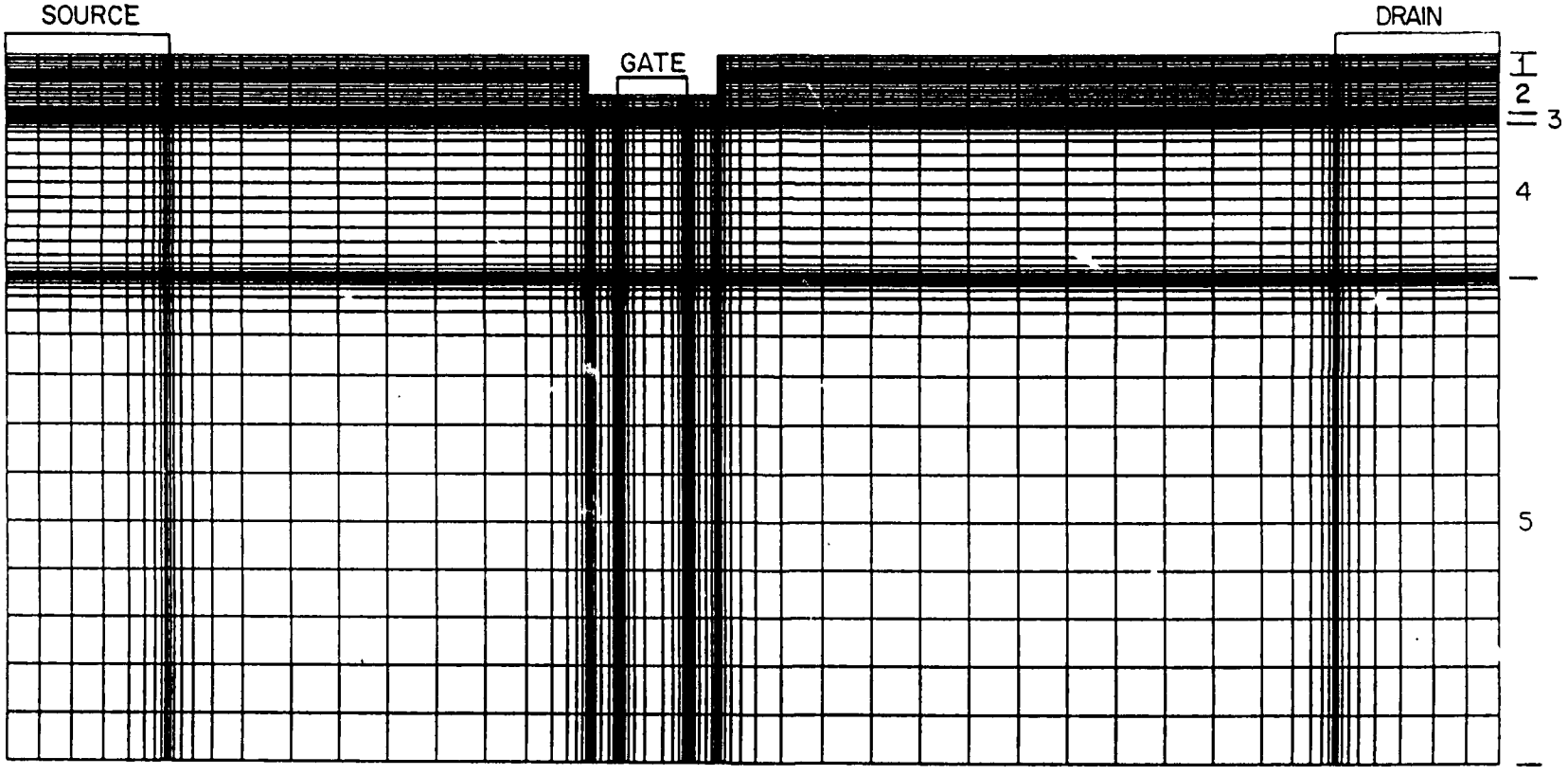


Figure 10. Grid Structure Used in Initial Simulations of Complete Device.

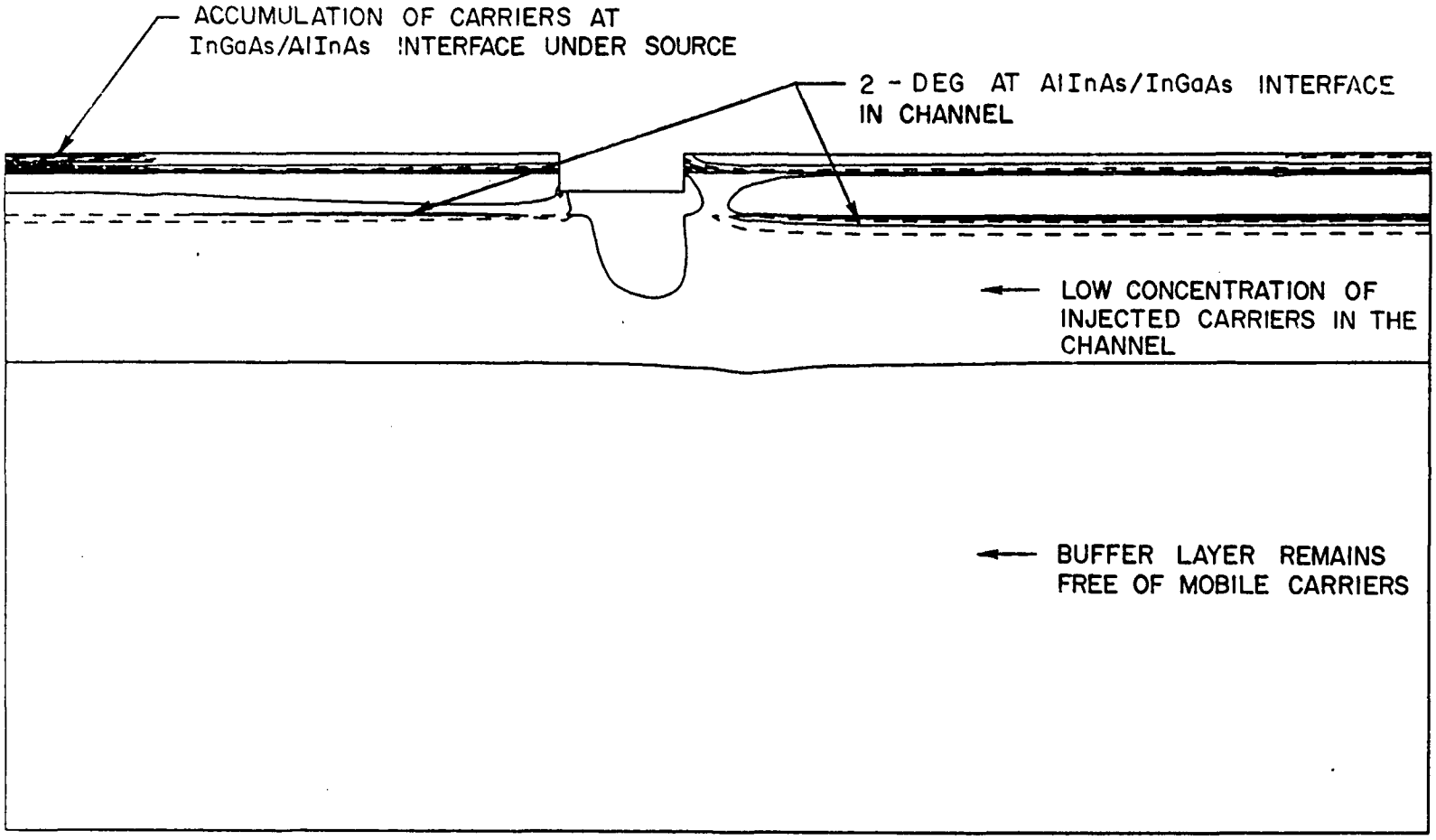


Figure 11. Contours of Carrier Density in the Complete Device for $V_{gs} = 0$, $V_{ds} = 1.0$.

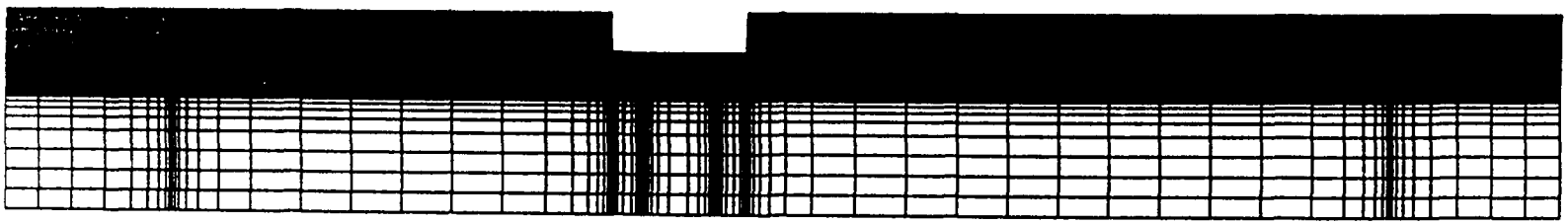


Figure 12. Grid used for Simulations of Reduced Device Domain shown on Same Scale as Fig. 10.

1 = N^+ InGaAs CAP LAYER

2 = UNDOPED AlInAs SCHOTTKY ENHANCEMENT LAYER

3 = Si DOPED PLANAR & SPACER LAYERS

4 = InGaAs CHANNEL

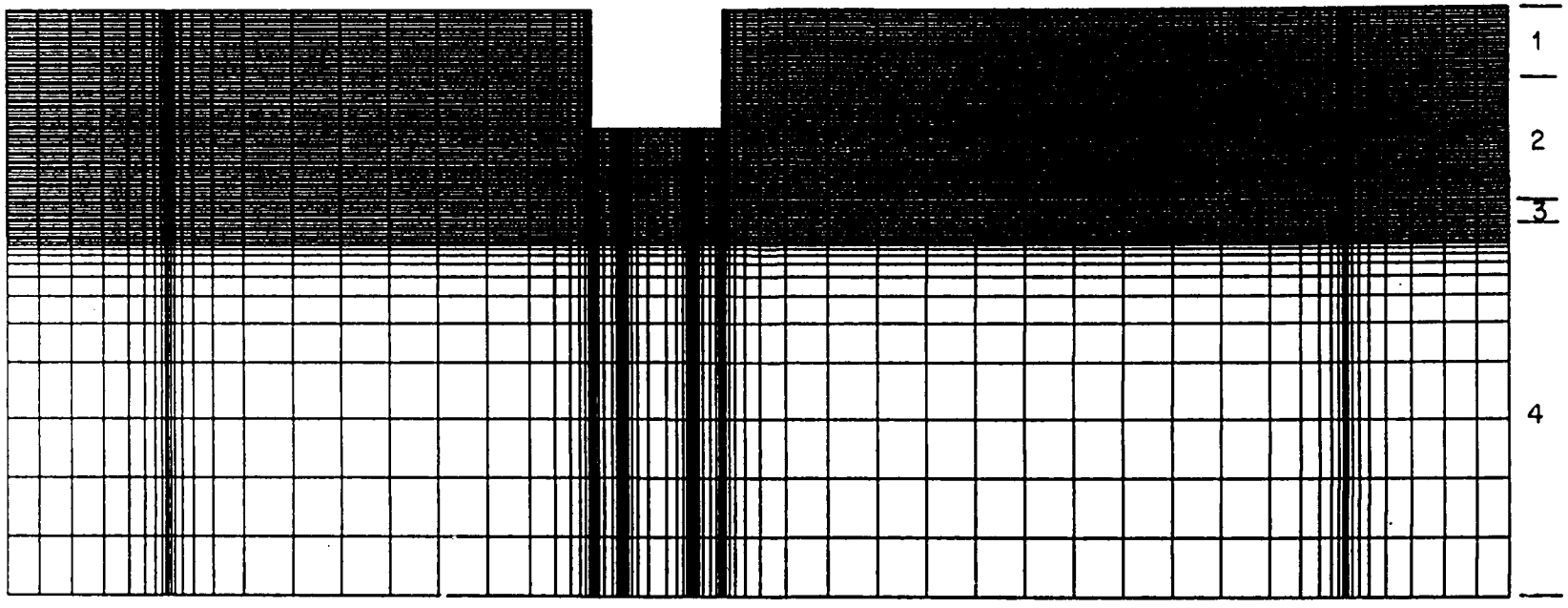


Figure 13. Same as Fig. 12, but with Vertical Scale Magnified 3X.

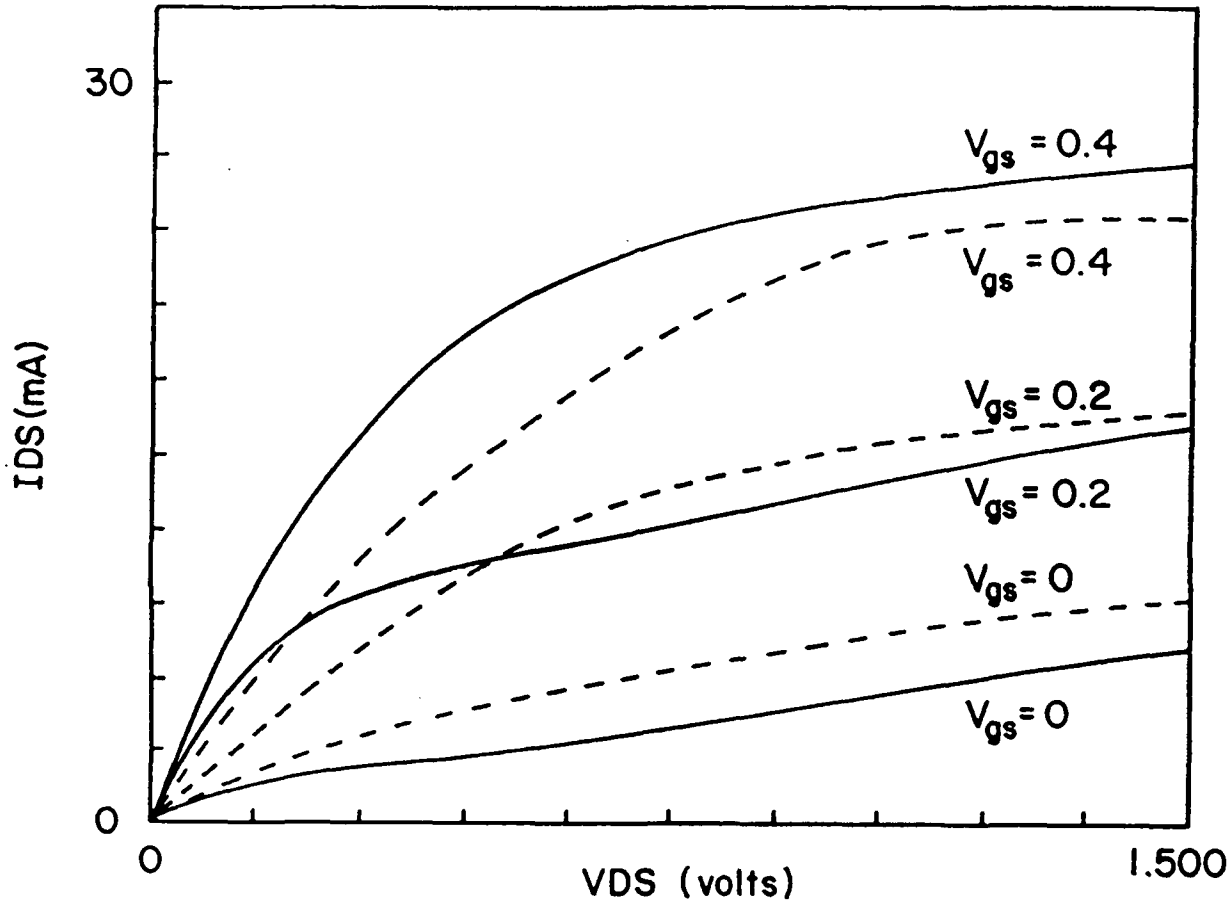


Figure 14. Predicted and Measured I-V Characteristics.
——— Measured, - - - - - Predicted.

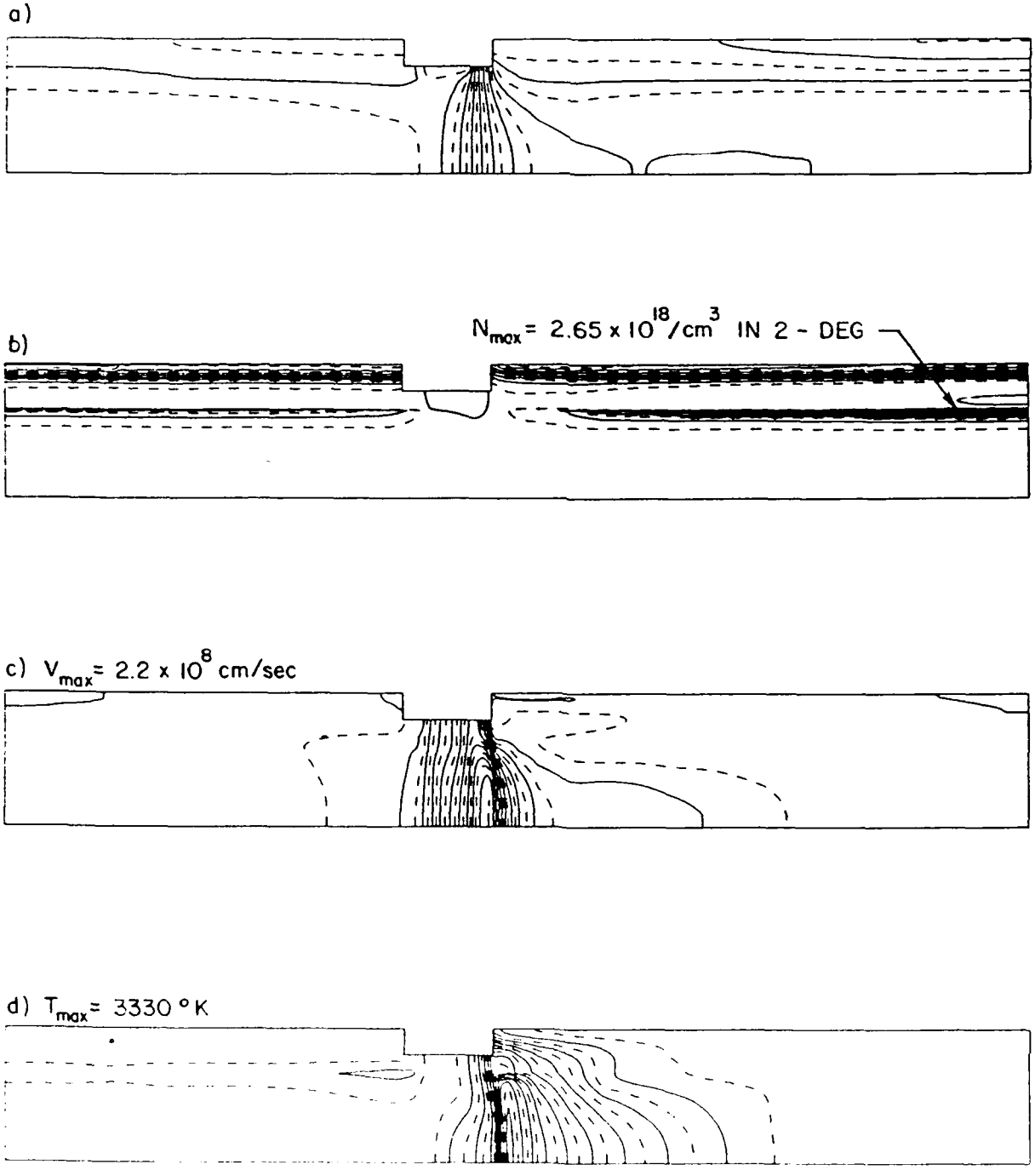


Figure 15. Contours of a) Potential, b) Carrier Density, c) Central Valley Velocity Magnitude, and d) Central Valley Temperature for V_{gs} = 0.0, V_{ds} = 1.5.

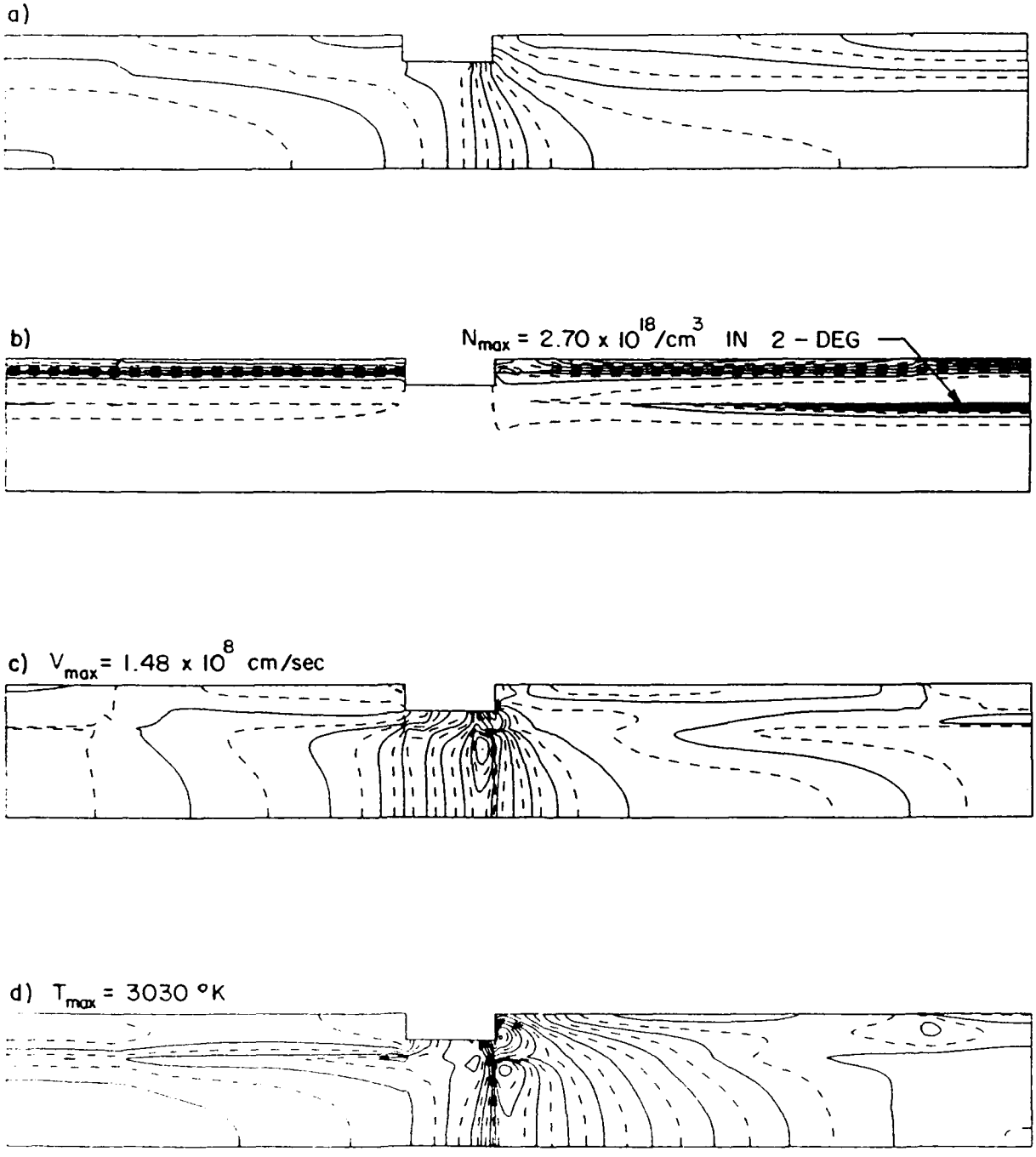
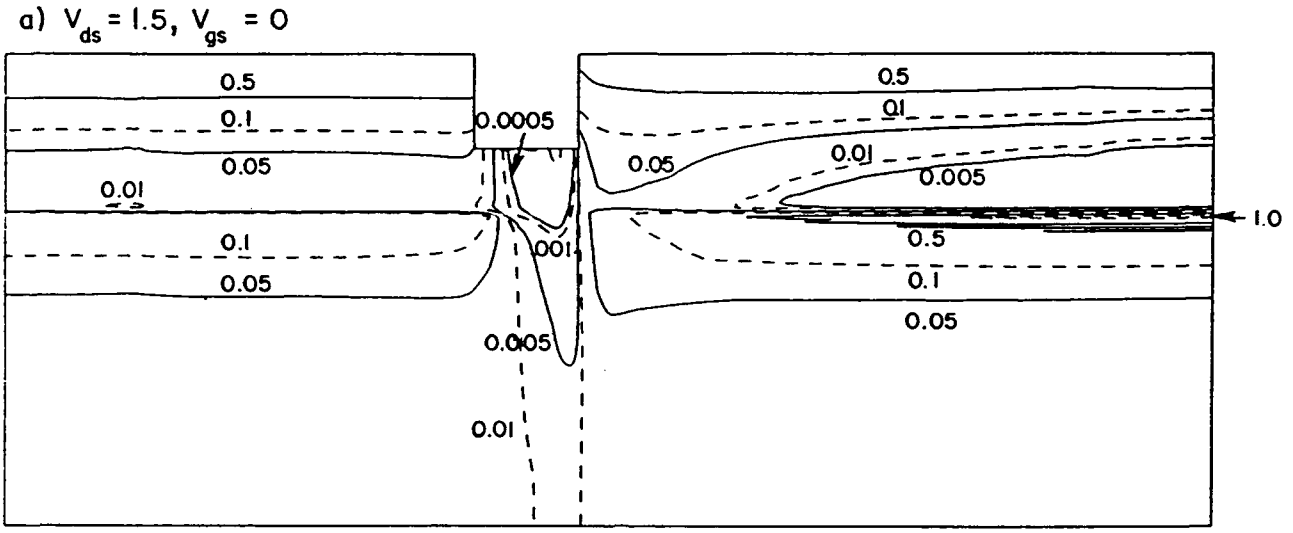


Figure 16. Contours of a) Potential, b) Carrier Density, c) Central Valley Velocity Magnitude, and d) Central Valley Temperature for $V_{gs} = 0.4$, $V_{ds} = 1.5$.



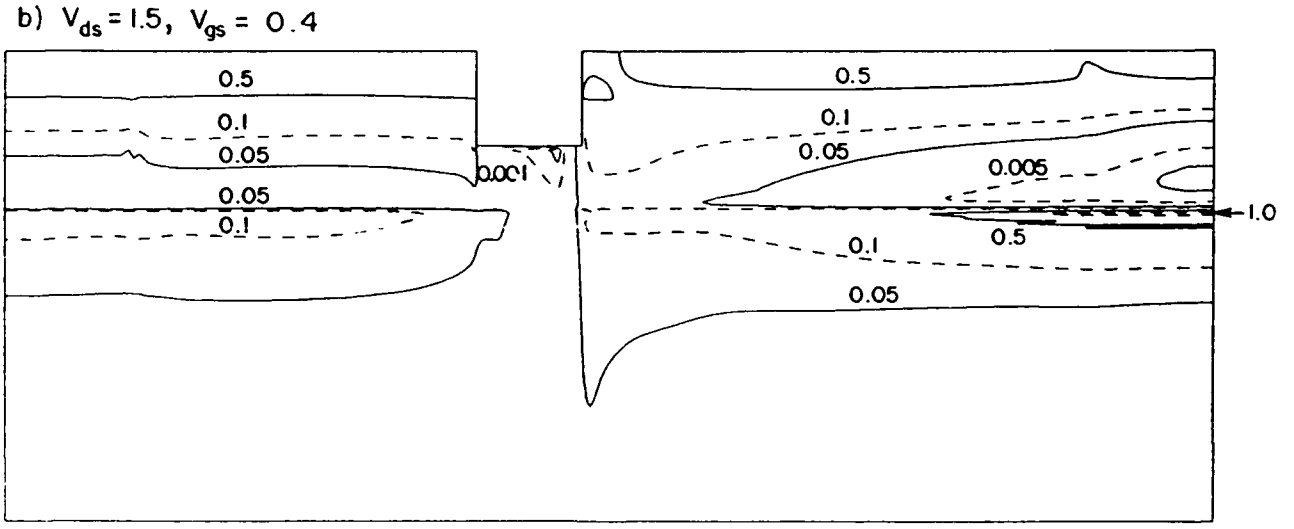


Figure 17. Contours of log N/N_{ref} for a) $V_{ds} = 1.5$, $V_{gs} = 0.0$ and b) $V_{ds} = 1.5$, $V_{gs} = 0.4$, $N_{ref} = 2.0 \times 10^{18}/cm^3$.

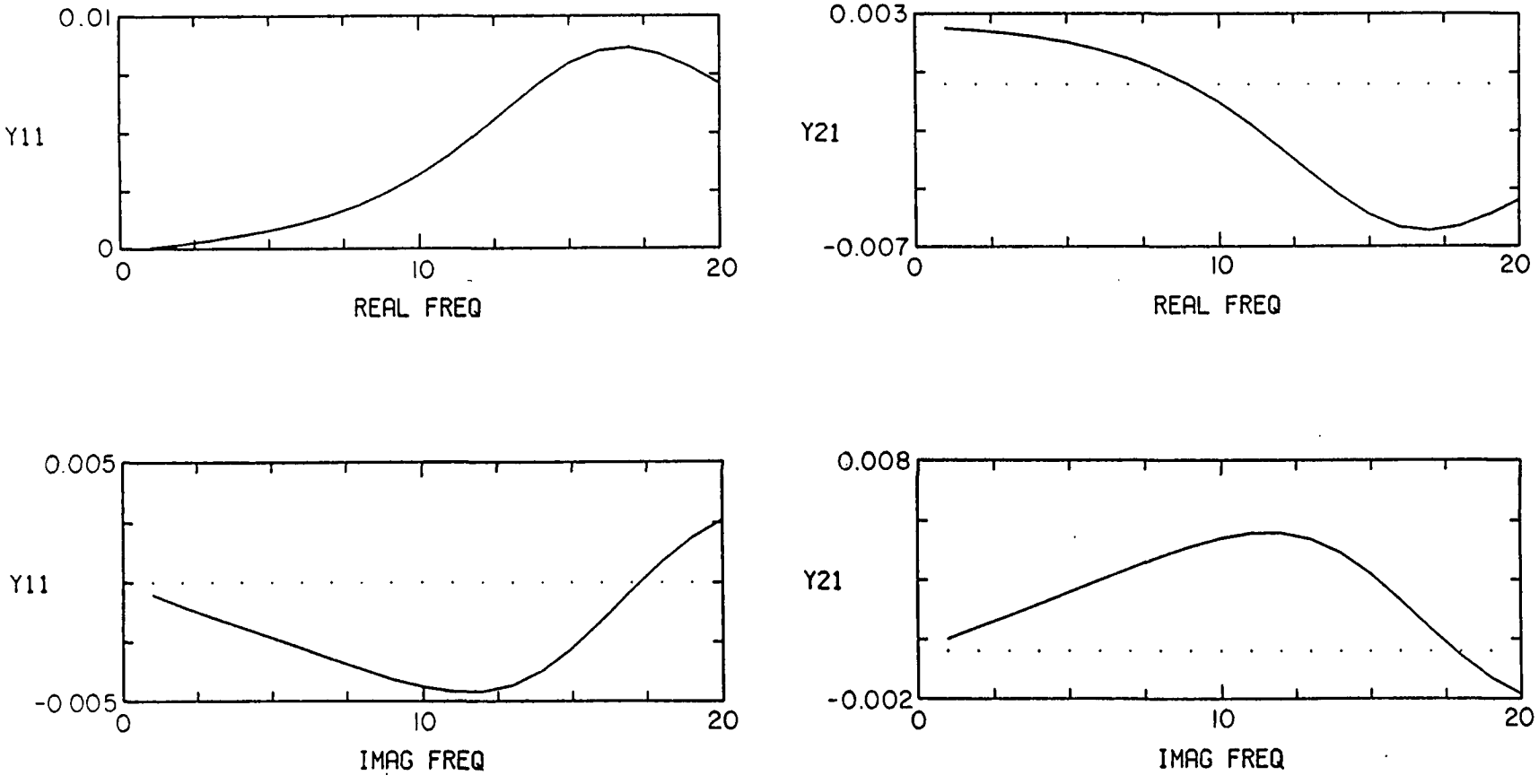


Figure 18. Small Signal Analysis, $V_g = 0.2$, $V_d = 1.0$, Gate-Gate and Drain-Gate Perturbed.

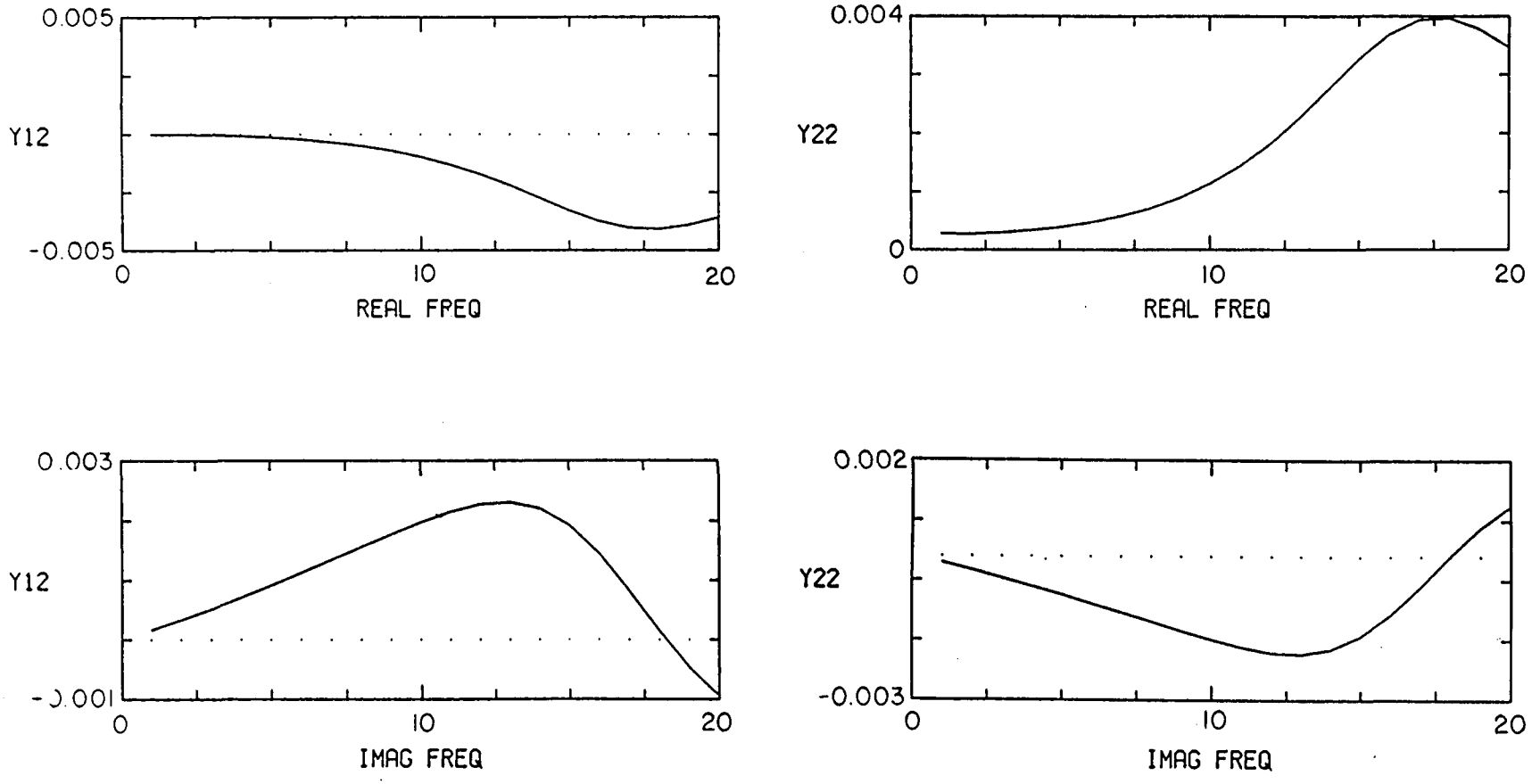


Figure 19. Small Signal Analysis, $V_g = 0.2$, $V_d = 1.0$, Gate-Drain and Drain-Drain Perturbed.

Appendix

TRANSPORT VIA MOMENTS OF QUANTUM DISTRIBUTION FUNCTIONS

Final Report: R930026F

H. L. Grubin and T. R. Govindan

December 10, 1990

U. S. Army Research Office

Contract DAAL03-90-C-0005

Scientific Research Associates, Inc.
50 Nye Road, P. O. Box 1058
Glastonbury, CT 06033

SCIENTIFIC RESEARCH ASSOCIATES, INC.

FINAL REPORT: R930026F

CONTRACT DAAL03-90-C-0005

TRANSPORT VIA MOMENTS OF QUANTUM DISTRIBUTION FUNCTIONS

H. L. GRUBIN and T. R. GOVINDAN

ABSTRACT

This report summarizes work performed under Contract DAAL03-90-C-0005. The study represents the first comprehensive effort to assess through (1) numerical simulations of the equation of motion of the density matrix, and (2) approximate analytical procedures involving an expansion of the Liouville potential, the multiplicative constant associated with the quantum potential as used in quantum hydrodynamic transport. It was found that when quantum effects are 'corrections' to classical calculations, as in $N^+N^-N^+$ structures the factor '1/3' is a satisfactory multiplicative constant. When the quantum potential is no longer a correction, as in the case of double barrier diodes, the multiplicative constant is no longer 'constant' but is position dependent and less than unity.

SCIENTIFIC RESEARCH ASSOCIATES, INC.

FINAL REPORT: R930026F

CONTRACT DAAL03-90-C-0005

TRANSPORT VIA MOMENTS OF QUANTUM DISTRIBUTION FUNCTIONS

H. L. GRUBIN and T. R. GOVINDAN

INTRODUCTION

The development of crystal growth techniques has made it possible to fabricate devices with very small dimensions and sharp interfaces. Consequently, a variety of device concepts based upon small dimensions with sharp interfaces has emerged. These include such high speed devices as the HEMT, HBT, quantum resonant tunneling devices, quantum wire structures and quantum dots. The structure and operation of these devices is based upon a perceived physical picture, often based upon results of quantum transport theory.

While all devices are governed and sometimes limited by quantum mechanical properties, many devices do not require quantum mechanics for an understanding of their basic operation, but do need quantum mechanics to provide key electrical properties. For example, heterostructure bipolar transistors sustain low levels of current at low bias levels; these currents are dominantly tunneling currents. Thermionic contributions to current occur at high bias levels. Until recently, the drift and diffusion equations as well as the moments of the Boltzmann transport equation did not include a description of tunneling currents in any way other than through ad hoc arguments. Tunneling requires quantum mechanical contributions.

The recent advance in the description of transport in devices with quantum contributions is due to the discussion of Ancona and Iafrate (1989), who demonstrated that the quantum corrections to multi-particle transport, first discussed by Wigner (1932), could be incorporated in the drift and diffusion equations by adding to the classical potential a term proportional to a quantum mechanical based potential:

$$(1) \qquad Q \equiv -(\hbar^2/2m)[(\rho)^{1/2}_{xx}]/(\rho)^{1/2}$$

where the subscripts denote derivatives. Within the framework of the drift and diffusion equation the 'Q' term modifies the particle current as follows:

$$(2) \qquad j(x,t) = n\mu k_b T[(V + aQ)/k_b T + \ell n(n)]_x$$

The proportionality constant 'a' was evaluated [Ancona and Iafrate (1989)] in the high temperature limit from the equation of motion of the Wigner function and was found to be equal to 1/3. The factor '1/3' is a serious issue, in that intuitively, as well as through arguments associated with the single particle Schrodinger equation, it would be anticipated that the constant 'a' would be unity.

The problem addressed in the last paragraph, namely the origin of the factor a = 1/3, as well as the **procedures under which quantum contributions to the balance equations are obtained** formed the basis of the study under Contract DAAL03-90-C-0005. This was

complemented by numerical simulations, using 'characteristic' algorithms to examine the extent to which the quantum corrected results are relevant.

The means by which the problem is addressed is through an expansion of the equation of motion of the Wigner function, and an expansion of the equation of motion of the density matrix. The discussion of the Wigner function is in the form of a review, in that a comprehensive treatment appears in Grubin and Kreskovsky (1989). The work on the density matrix is new.

QUANTUM CORRECTED EQUATIONS OF MOTION FOR THE WIGNER FUNCTION AND THE DENSITY MATRIX

The Wigner Equation of Motion: The analysis under this study was restricted to one space and one momentum direction and includes Fokker-Planck scattering. The equation of motion of the Wigner function with Fokker-Planck dissipation as discussed by Stroscio (1986) is:

$$(3)\qquad f_t + (p/m)f_x + (1/i\hbar)(1/2\pi\hbar)\int_{-\infty}^{+\infty} dp' \int_{-\infty}^{+\infty} dx' f(p',x)[V(x,x',t)]\exp[i(p-p')x'/\hbar]$$
$$= 2\gamma[pf]_p + Df_{pp}$$

where:

$$(4)\qquad V(x,x',t) = [V(x+x'/2,t) - V(x-x'/2,t)]$$

The potential energy V(x,x',t) is referred to below as the Liouville potential. As discussed by Frensley (1990) the first term of the Fokker-Planck dissipation corresponds to a frictional damping term, the second corresponds to thermal fluctuations. This will be discussed in more detail below.

It is direct, but nontrivial to demonstrate that the integral in equation (3) reduces in the classical case to $V_x f_p$; thus equation (3) reduces to the Boltzmann transport equation with Fokker-Planck dissipation:

$$(5)\qquad f_t + (p/m)f_x - V_x f_p = 2\gamma[pf]_p + Df_{pp}$$

To second order in $\hbar$, the Wigner equation of motion, or equivalently, the quantum corrected Boltzmann equation is:

$$(6)\qquad f_t + (p/m)f_x - V_x f_p + (\hbar^2/24)V_{xxx} f_{ppp} = 2\gamma[pf]_p + Df_{pp}$$

The left hand side of equation (6) has been discussed in detail by Ancona and Iafrate (1989) and Grubin and Kreskovsky (1989).

The Equation of Motion of the Density Matrix: This equation has been discussed in the context of quantum structures by Frensley (1990). While the equation of motion is a fundamental equation, it is equivalent to and can be obtained directly from the Wigner equation of motion. In the latter case multiply equation (3) by $dp[\exp[i2\zeta p/\hbar]]$ and integrate over p. For this integration p varies from $-\infty$ to $+\infty$, and it is assumed that the Wigner function and all necessary derivatives with respect to momentum vanish as $p \Rightarrow \pm\infty$. Identifying the density matrix through the Weyl transformation:

$$(7)\qquad \rho(x+\zeta, x-\zeta) = [1/(2\pi\hbar)]\int_{-\infty}^{\infty} dp f_W(p,x)\exp[2ip\zeta/\hbar]$$

the transformed equation becomes the density matrix in the coordinate representation:

$$(8) \qquad \rho_t + (\hbar/2mi)\rho_{x\zeta} - (1/i\hbar)[V(x+\zeta,t) - V(x-\zeta,t)]\rho + 2\gamma\zeta\rho_\zeta + [4D\zeta^2/\hbar^2]\rho = 0$$

An expansion of the Liouville potential about the point x yields:

$$(9) \qquad \rho_t + (\hbar/2mi)\rho_{x\zeta} - (1/i\hbar)[2\zeta V_x + (\zeta^3/3)V_{xxx}]\rho + 2\gamma\zeta\rho_\zeta + [4D\zeta^2/\hbar^2]\rho = 0$$

which is the transformed equivalent of equation (6).

Retaining only the terms that are linear in ζ yields the transformed equivalent to the classical Boltzmann equation. As will be seen the presence of the factor '1/3' in equation (9) is the key ingredient in obtaining the multiplicative factor in the quantum potential. It is emphasized that quantum corrections to $O(\hbar^2)$ are contained in the term $[(\zeta^3/3)V_{xxx}]\rho$.

Equations (6) and (9) are the relevant equations for the analytical part of this study. Equations (3) and (8) are, however, the ones to deal with in a fundamental approach to quantum transport. Results with equation (8) are introduced later into the discussion.

THE APPROXIMATE EQUILIBRIUM DISTRIBUTION FUNCTION

In the absence of dissipation the approximate Wigner distribution function to second order in $\hbar$ is Wigner (1932):

$$(10) \qquad f_W = \exp-\beta[p^2/2m + U(x)]\{1-(\lambda^2\beta/4)[(V_{xx}-\beta V_x^2/3)-\beta(p^2/3m)V_{xx}]\}$$

Where $\lambda^2 = \hbar^2\beta/2m$, and $\beta = 1/(k_bT)$. The equivalent approximate density matrix solution to order $O(\hbar^2)$ is:

$$(11) \qquad \rho(x+\zeta,x-\zeta) = N\exp-[\zeta^2/\lambda^2 + \beta V]\{1-(\lambda^2\beta/6)(1+\zeta^2/\lambda^2)V_{xx} + \lambda^2\beta^2 V_x^2/12\}$$

where,

$$(12) \qquad N = 2(m/2\pi\beta\hbar^2)^{1/2}.$$

Note: First, the classical density matrix equivalent to the Boltzmann distribution is:

$$(13) \qquad \rho(x+\zeta,x-\zeta) = N\exp-[\zeta^2/\lambda^2 + \beta V];$$

Second, the brackets $\{\bullet\bullet\bullet\}$ contains quantum corrections to the density matrix. Third, the approximate density matrix is real, indicating that the current is zero. Fourth, the diagonal components of the approximate density matrix, which are obtained for $\zeta = 0$, are equal to:

$$(14) \qquad \rho(x,x) = N[\exp-\beta V][1-(\lambda^2\beta/6)(V_{xx}-\beta V_x^2/2)]$$

This result also emerges from an integration of the Wigner function over all values of momentum (Ancona and Iafrate (1989).

The significance of equation (14) for device applications is that it describes the way in which quantum corrections alter the built-in potential. For example, when it is recalled that, classically, the potential and carrier density are related through the expression:

$$(15) \qquad \rho = N\exp-\beta V$$

and therefore that $\beta V_x = -\rho_x/\rho$; then through incorporation of the quantum potential (equation (1)), it is straightforward to show that $Q = (\hbar^2/4m)\beta[V_{xx} - \beta V_x^2/2]$, and that:

(16) $\rho = N[\exp -\beta V][1-(\beta Q/3)]$

Several points are clear: First there may be an alteration of the built-in potential arising from gradients in the carrier density; second, **the factor of '3' that appears in equation (16) is seen to be a consequence of the truncation of the Taylor expansion of the potential in both the density matrix and Wigner equations.**

Quantum Corrected Energy Density Matrix: There is a characteristic energy associated with the equilibrium system obtained from both the Wigner function and the density matrix. To obtain it from the density matrix we define an energy density matrix through a Weyl-like transformation:

(17) $E(x+\zeta, x-\zeta) = [1/(2\pi\hbar)]\int_{-\infty}^{\infty} dp (p^2/2m) f_w(p,x) \exp[2ip\zeta/\hbar]$

Note: for $\zeta = 0$, the above reduces to mean kinetic energy of the system. It is direct to demonstrate that

(18) $E(x+\zeta, x-\zeta) = -(\hbar^2/8m)\rho_{\zeta\zeta}$

The expectation value of the energy is the diagonal component of the energy density matrix, and yields:

(19) $E(x,x) = [\rho k_b T/2[[1-(\lambda^2/6)(\ell n \rho)_{xx}]$

$= [k_b T/2 + w_q]\rho$

The term w_q was first introduced by Wollard et al (1990). Equation (19) is valid only to second order in $\hbar$. Thus the density multiplying the term '$(\ell n \rho)_{xx}$' is strictly only the classical density. **The significance of equation (11) is that there is a change in energy due to the gradients of carrier density**, as first predicted by Wigner (1932).

THE APPROXIMATE NONEQUILIBRIUM DENSITY MATRIX

Classical moment balance equations are often obtained through representing the nonequilibrium state by a displaced Maxwellian, $\exp-[\beta\{(p-p_d)^2/2m + V\}]$ where, e.g., p_d, the density and a particle temperature, are to be determined. The Weyl transformation indicates that a displaced Maxwellian yields a density matrix with the following modification:

(20) $\rho(x+\zeta, x-\zeta) \Rightarrow \rho(x+\zeta, x-\zeta)\exp-[2ip_d\zeta/\hbar]$

Thus the classical density matrix corresponding to the displaced Maxwellian is (21)

$\rho(x+\zeta, x-\zeta) = N\exp-[\zeta^2/\lambda^2 + \beta V + 2ip_d\zeta/\hbar]$

Note that the density matrix contains an imaginary part, as required for a finite current. The nonequilibrium quantum-corrected density matrix is given by a modification of equation (11):

(22) $\rho(x+\zeta,x-\zeta)=N\exp-[\zeta^2/\lambda^2+\beta V+2ip_d\zeta/\hbar]$

$\{1-(\lambda^2\beta/6)(1+\zeta^2/\lambda^2)V_{xx}+\lambda^2\beta^2V_x^2/12\}$

Quantum Corrected Current (velocity flux) Density Matrix: Since the density matrix contains imaginary components we are interested in the expectation values of the current flux density. To obtain this we define a current density matrix as:

(23) $j(x+\zeta,x-\zeta)=[1/(2\pi\hbar)]_{-\infty}\int^{\infty}dp(p/m)f_W(p,x)\exp[2ip\zeta/\hbar]$

Note: for $\zeta=0$, the above reduces to mean velocity flux of the system. It is direct to show that:

(24) $j(x+\zeta,x-\zeta)=(\hbar/(2mi))\rho_\zeta$

The expectation value of current is the diagonal component:

(25) $j(x,x)=(p_d/m)\rho$

The quantum corrections to the current are obtained from the quantum corrections to the density. Note that the form of the current density along the diagonal is the same for classical transport as it is for quantum corrected transport. Thus of relevance are the equations defining p_d, density and temperature. To establish these we reconsider the equation of motion of the density matrix taking successive derivatives.

THE APPROXIMATE NONEQUILIBRIUM BALANCE EQUATIONS AS OBTAINED FROM THE EQUATION OF MOTION OF THE DENSITY MATRIX

Particle Balance: The first balance equation is obtained by rewriting the equation of motion of the density matrix and then dealing only with the diagonal component. Thus there is reduced content. Using the definition of velocity flux, equation (9) becomes:

(26) $\rho(x+\zeta,x-\zeta)]_t+[j(x+\zeta,x-\zeta)]_x-(1/i\hbar)[2\zeta V_x+\zeta^3V_{xxx}/3]\rho(x+\zeta,x-\zeta)$

$+2\gamma\zeta[\rho(x+\zeta,x-\zeta)]_\zeta+[4D\zeta^2/\hbar^2]\rho(x+\zeta,x-\zeta)=0$

The first balance equation is obtained from the diagonal component of equation (24):

(27) $[\rho]_t+[\rho p_d/m]_x=0$

which is the equation of continuity.

Momentum Balance: To determine the next governing equation to second order in $\hbar$ we take derivatives of equation (2) with respect to 'ζ', multiply by $\hbar/2i$, and obtain:

(28) $[p(x+\zeta,x-\zeta)]_t+2[E(x+\zeta,x-\zeta)]_x+(1/2)[2V_x+\zeta^2V_{xxx}]\rho(x+\zeta,x-\zeta)$

$-(1/i\hbar)[2\zeta V_x+\zeta^3V_{xxx}/3]p(x+\zeta,x-\zeta)$

$+2\gamma p(x+\zeta,x-\zeta)-i[4D\zeta/\hbar]\rho(x+\zeta,x-\zeta)$

$-i\hbar\gamma\zeta[\rho(x+\zeta,x-\zeta)]_{\zeta\zeta}+[4D\zeta^2/\hbar^2][p(x+\zeta,x-\zeta)]=0$

where $p(x+\zeta,x-\zeta)=j(x+\zeta,x-\zeta)m$. The diagonal component of energy under finite current conditions is required for the second balance equation. It is:

$$(29) \qquad E(x,x) = [p_d^2/2m + (k_bT/2)(1-(\lambda^2/6)(\ell n\rho)_{xx}]\rho$$

With the diagonal components of energy given by equation (29), identifying a relaxation time $\tau = 2/\gamma$, the second balance equation is obtained from the diagonal components of equation (29):

$$(30) \qquad (\rho p_d)_t + (\rho p_d^2/m)_x + (\rho kT)_x + \rho(Q/3)_x + \rho V_x + \rho p/\tau = 0$$

where we recognize that $[(\ell n\rho)_{xx}\rho]_x = -4m\rho Q_x/\hbar^2$. The above equation differs from its classical analog through the presence of the quantum potential. The form of the scattering term in the above equation also confirms Frensley's (1990) statement of the first part of the Fokker-Planck scattering as a frictional term.

Energy Balance: If we regard the quantum corrected density matrix as a distribution function with three undetermined parameters, the third being the temperature, then a third equation is needed to complete the system of equations. This third equation is obtained in a manner similar that of equation (28), namely by taking a second derivative of the density matrix equation of motion, with respect to ζ. We are only interested in the diagonal components of this equation, which with the energy given by equation (18), are:

$$(31) \qquad E_t + (1/(2m^2))P^3{}_x + jV_x = (E_t)_{coll}$$

where P^3 is the diagonal component of the third moment of the off-diagonal element:

$$(32) \qquad [1/(2\pi\hbar)]_{-\infty}\int^{\infty} dp(p^3)f_w(p,x)\exp[2ip\zeta/\hbar] = (\hbar/2i)^3\rho_{\zeta\zeta\zeta}$$

Using equation (21) for the density matrix, equation (32) becomes:

$$(33) \qquad E_t + \{(p_d/m)[E + (\rho/\beta)(1-[\lambda^2/6](\ell n\rho)_{xx})]\}_x + (\rho p_d/m)U_x + 2E/\tau + 8D\rho = 0$$

where the energy in equation (33) is given by equation (29) (Note: the equilibrium value of P^3 is zero). The above equation can be rearranged to read:

$$(34) \qquad E_t + (p_dE/m)_x + (p_d\rho k_bT/m)_x + (\rho p_d/m)[Q/3 + V]_x$$
$$-\rho(\lambda^2 k_bT/6)[(\ell n\rho)_{xx}](p_d/m)_x + 2E/\tau - 8D\rho = 0$$

If we assume equilibrium values for D, namely $D = mk_bT/\tau$, equation (34) becomes:

$$(35) \qquad E_t + (p_dE/m)_x + (p_d\rho k_bT/m)_x + (\rho p_d/m)[Q/3 + V]_x$$
$$-\rho(\lambda^2 k_bT/6)[(\ell n\rho)_{xx}](p_d/m)_x + (2/\tau)[E - k_bT/2]\rho = 0$$

And the second part of the Fokker-Planck dissipation involves a relaxation to a non-zero thermal energy.

The equation of motion of the density matrix involves two NxN matrices (real and imaginary parts) where N represents the numbers of grid points along a linear dimension. The density matrix equation is the equation of interest. An approximate representation of the equation of motion of the density matrix is provided by the first three moments of the density matrix. These moments include the continuity equation, momentum balance, and

energy balance.

As in all moment equation representations, the form of the moment equation depends upon the primary equation, in this case equation (9), and the generic form of the distribution function, in this case equation (22). The most dramatic consequence of the approach, as expressed in the momentum balance equations, is the appearance of the quantum potential. The quantum potential with its associated factor of '1/3', the latter arising solely as a consequence of the expansion of the Liouville potential, permits a quantum mechanical "correction" description of classical transport and allows tunneling to be incorporated into the latter. This result was also the initial conclusion of Ancona and Iafrate (1990).

It is important to reiterate that one key result of the present study is that through the expansion of Liouville potential in the equation of motion of the density matrix the quantum corrections are the same as those obtained from the equation of motion of the Wigner distribution function. The origin of the factor of '1/3' is due solely to the expansion of the Liouville potential, and is not limited the high temperature limit discussed by Ancona and Iafrate (1989).

The energy balance equation is also driven by the quantum potential, but additionally there appears to be a contribution that may behave as a quantum correction to the pressure. This is the terms $\rho(\lambda^2 k_b T/6)[(\ell n \rho)_{xx}]$.

In addition to the quantum mechanical contributions, the moment equations include an incorporation of Fokker-Planck dissipation in the moment formulation, and twenty of the relative contribution of each of the two terms. A more general treatment is provided by Stroscio (1986).

The next question concerns the significance of the quantum potential.

SELF CONSISTENT NUMERICAL SOLUTIONS

This section contains a discussion of the **interpretation** of the distribution of charge within the quantum well and quantum barrier device and the values obtained thereof. The calculations involve solutions to the density matrix coupled to Poisson's equation:

$$\partial/\partial x[\epsilon(x)\partial V/\partial x] = -e^2[\rho(x,x)-\rho_0(x)] \quad (36)$$

Classical $N^+N^-N^+$ Structures: Since most resonant tunnel structures are designed with the heterostructures placed within the interior of a low doped region, the first self-consistent problem discussed involves transport through a classic $N^+N^-N^+$ structure. The structure is 1200Å long with a nominal doping of $10^{18}/cm^3$ and a centrally placed 250Å, $10^{15}/cm^3$ region. The variation in background doping was over one grid point or 4Å.

There are a variety of questions to be addressed here; among them is the issue of whether there are any quantum contributions associated with an ostensibly classical structure. It is relevant, in this matter to recall one of the conclusions of a paper by Iafrate, Grubin and Ferry (1981), where it was argued that quantum corrections are important if the density sustained an approximate Gaussian distribution with a width at half maximum of approximately 80Å of less.

The equilibrium charge density and potential distribution as obtained from the complete equation of motion of the density matrix, equation (8) without dissipation, for the $N^+N^-N^+$ structure are shown in figures (1a) and (1b), respectively. The results appear classical. We also show the diagonal component of energy density of the system, figure (1c),

as computed from equation (18) and the quantum potential, figure (1d), as computed from equation (1). For completeness we also show the two dimensional plot of the density matrix in figure 2.

Figures (1a) and (1b) display charge density and self-consistent potential distributions that are ostensibly classical. There is an increase in potential across the N^- region, which accompanies a decrease in charge density across this same region. From the point of view of quantum transport, we are also dealing with a tunneling problem, particularly with those carriers whose energy is below the potential barrier, which in this case has a height of approximately 45 mev. We note that the mean energy of the entering carriers is $k_bT/2 < 45$ mev.

Quantum transport permits tunneling, thereby increasing the numbers of particles in the N^- region over the classical value. Continuity of the wavefunction and its derivative through the barrier region (at the metallurgical interface the energy density is approximately $k_bT\rho/2$, and decreases into the N^- region) prevents the density from approaching it's classical value, and instead assumes a smaller value. This result which is a consequence of wave function continuity has been referred to by others as quantum "repulsion" (Kluksdahl et al (1989)). The "increased" value of density in the barrier region and "decreased" value of density in the classically accessible region are represented by equation (16) if the quantum is negative within the barrier region and positive within the region of the 'turning points' at the boundary of the metallurgical junction. The quantum potential as evaluated from equation (1) and depicted in figure 1d displays these qualitative features. Note that the quantum potential at the center of the structure is approximately 25% of the energy at the boundaries. For a very wide N^- region the density at the center region of the structure will be relatively constant and $Q \lesssim 0$, at the center.

For the calculation of figures 1 and 2, the density, potential and quantum potential are respectively

(37) $\rho_{cal}(v=0) = 1.847 \times 10^{17}/cm^3$

(38a) $V_{cal}(x=0) = 44.725$ mev

(38b) $Q_{cal}(x=0) = -3.0228$ mev

Calculations $\rho(x=0)$ from equation (16) with the potential energy and quantum potential given by equation 38 yield

(39) $\rho_{Eq.(38)}(x=0) = 1.832 \times 10^{17}/cm^3$

The energy density at the center of the structure as computed from the density matrix is

(40) $E_{cal}(x=0) = 2.3015$ mev/cm^3

The energy as computed from equation (19) is

(41) $E_{Eq.(19)}(x=0) = 2.286$ mev/cm^3

The hydrodynamic moment equations appear in the case of the simple $N^+N^-N^+$ to yield the same result as the density matrix for the charge distribution.

Double Barrier Diodes: The situation for double barrier diodes is quantitatively different than that of the $N^+N^-N^+$ structure. Here the quantum mechanics is not a correction to the classical solution. However, as in the $N^+N^-N^+$ structure, it is expected that outside of the double barrier that continuity of the wave function would again yield a density below that of the classical value. Thus the quantum potential would be positive in this region. Within the barrier quantum mechanical tunneling permits a greater number of carriers than those associated with the classical distribution; thus the quantum potential is negative. Within the quantum well continuity of the wavefunction results in a charge distribution that is below that obtained classically; and Q is negative. The variation of density, potential, energy and quantum potential are displayed in figures 3 and 4.

Figure 3a displays the density distribution which shows a small residual change in the quantum well and an insignificant amount of charge within the barriers. Potential energy is displayed in figure 3b, where we note that the low density in the quantum well region contributes to an elevation of the potential relative to the end points. The mean energy density is shown in figure 3c, and displays negligible values at the center of the structure. The quantum potential is displayed in figure 3d. For the calculations of figure 3, the density potential energy and quantum potential are respectively

(42) $\rho_{cal}(x=0) = 6.156 \times 10^{16}/cm^3$

(43a) $V_{cal}(x=0) = 15.479$ mev

(43b) $Q_{cal}(x=0) = 92.149$ mev

Calculating $\rho(x=0)$ from equation (16) within the potential energy and quantum potential given by equation (43) requires that the quantum potential be multiplied by a constant other than 1/3. For the present situation the constant is closer to 2/3. This difference is value is not surprising in light of the fact that within the quantum well the value of the quantum potential is not longer a correction. Indeed Q is approximately equal to the quasi-bound state energy.

The value of $\rho(x=0)$ from equation 16, with Q/3 replaced by 2Q/3 is:

(44) $\rho_{(Eq.(38)}(x=0) = 5.043 \times 10^{16}/cm^3$

We note that the density upstream of the emitter is satisfactorily represented by equation (16). It would appear that our results are consistent with the discussion of Ancona and Iafrate (1989) where the quantum potential appearing in the moment equations should be

(45) Phenomenological quantum potential: $aQ(x)$

where a is a position dependent positive constant less than unity.

The energy calculated from equation (18) with the center of the quantum well is

(46) $E_{cal}(x=0) = 2.864$ mev/cm^3

The energy calculated from equation (19) with Q/3 replaced by 2Q/3 is:

(47) $E_{Eq.(19)}(x=0) = 2.684$ mev/cm^3

The low value is present, even though the quantum potential tends to push up the mean carrier energy, and arises from the reduced charge in the quantum well.

CONCLUSION

This study represents the first comprehensive effort to assess through (1) numerical simulations of the equation of motion of the density matrix, and (2) approximate analytical procedures involving an expansion of the Liouville potential, the multiplicative constants associated with the quantum potential, as used in quantum hydrodynamic transport. It was found that when quantum effects are 'corrections' to classical calculations, as in $N^+N^-N^+$ structures the factor '1/3' is a satisfactory multiplicative constant. When the quantum potential is no longer a correction, as in the case of double barrier diodes, the multiplicative constant is no longer 'constant' but is position dependent and less than unity. Additional work is required to narrow the range of variation of this constant before effectively incorporating the quantum potential in detailed quantum hydrodynamic simulations of ultrasmall devices.

REFERENCES

M.A.Ancona and G.J.Iafrate, Physical Review B 39, 9536 (1989)

W.A.Frensley, Reviews of Modern Physics, 62, 745 (1990)

H.L.Grubin and J.P.Kreskovsky,Solid State Electronics 32, 12 (1989)

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

N.C.Kluksdahl, A.M.Kriman, D.K.Ferry and C Ringhofer, Physical Review B, 39, 7720 (1989)

M.A.Stroscio, Superlattices and Microstructures, 2, 83 (1986)

E. Wigner, Physical Review, 40 749 (1932)

D.L.Wollard, M.A.Stroscio, M.A.Littlejohn, R.J.Trew and H.L.Grubin, Proceedings of the Workshop on Computational Electronics (To be published, 1991)

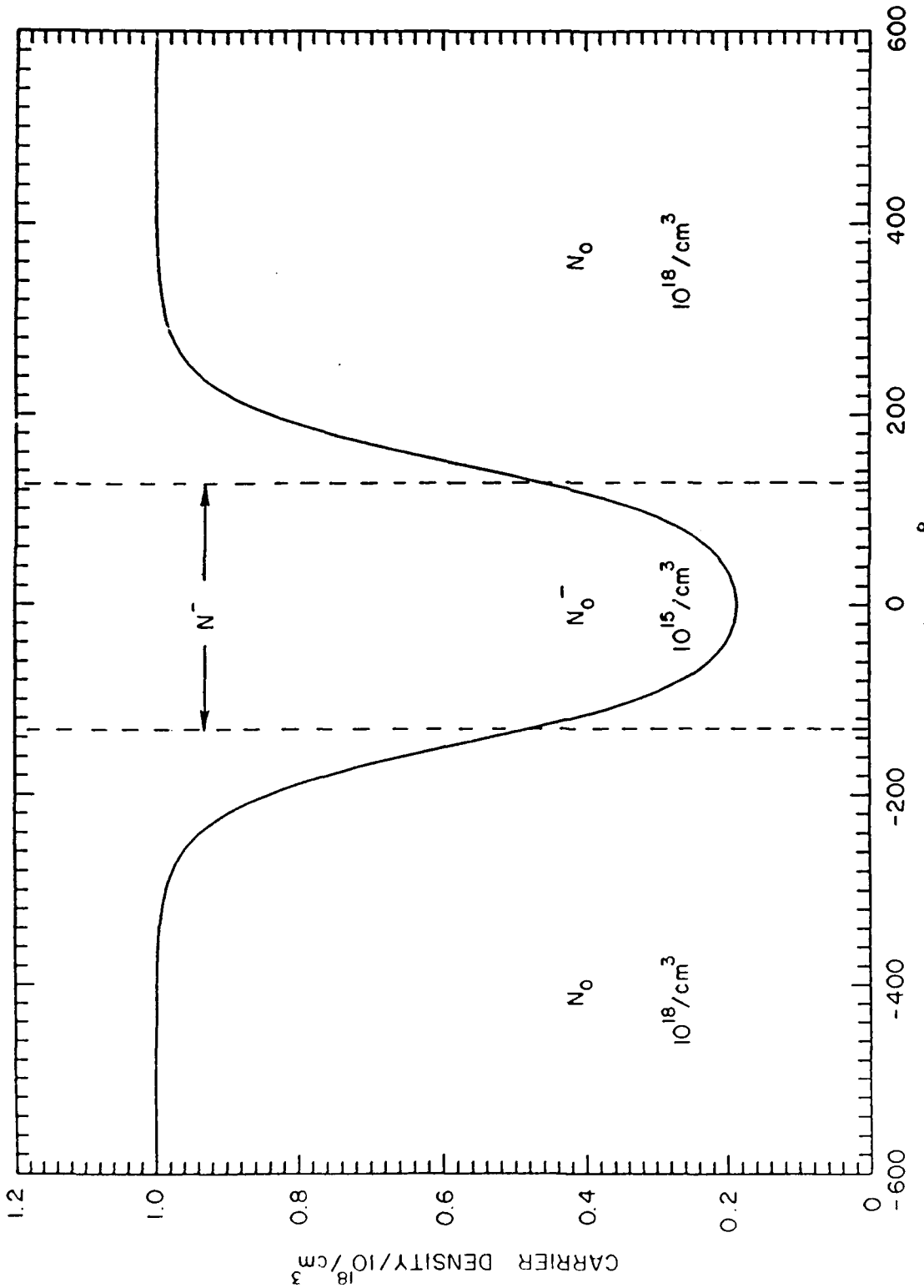


Fig. 1a. Carrier and background distribution for a NIN structure with free particle boundary conditions, as obtained from a solution of Equation (8) with zero dissipation and Poisson's Equation.

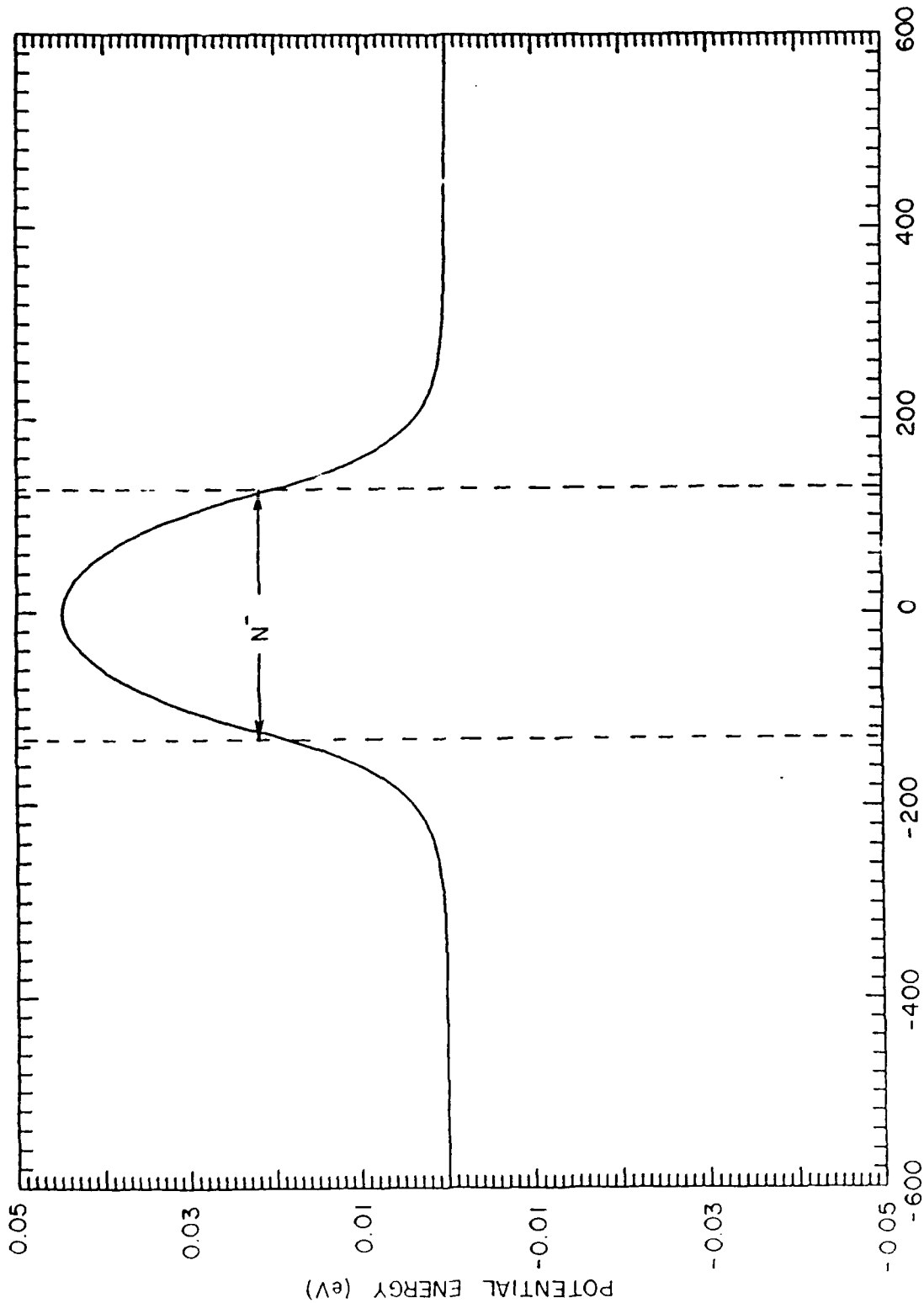


Fig. 1b. Self consistent potential associates with the density distribution of Figure 1a. The N^+N^- and N^-N^+ regions are indentified by the dashed lines.

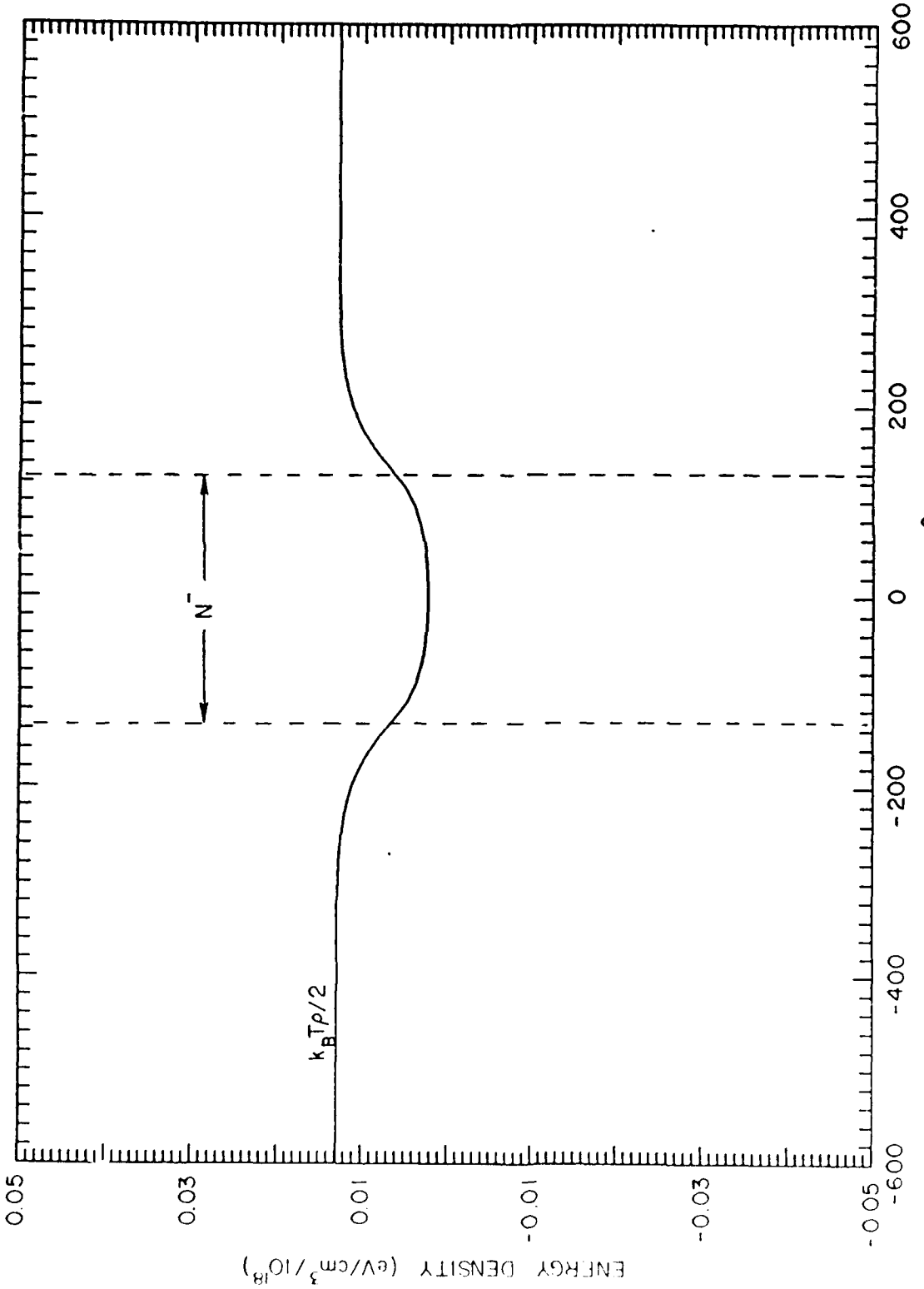


Fig. 1c. Energy density as obtained from Equation (18). Density at boundary region is entirely thermal.

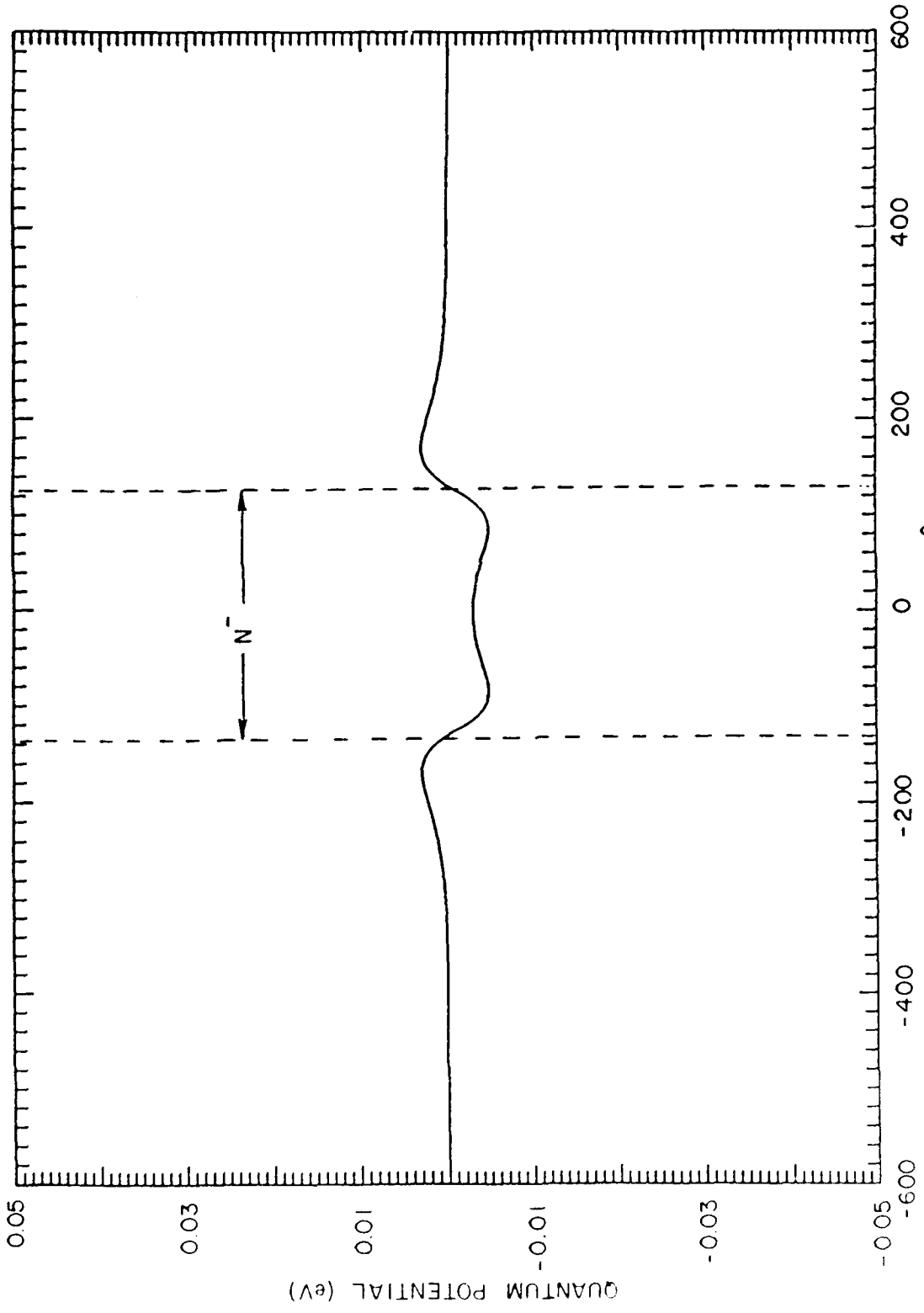


Fig. 1d. Quantum potential for NIN structure, as obtained from Equation (1).

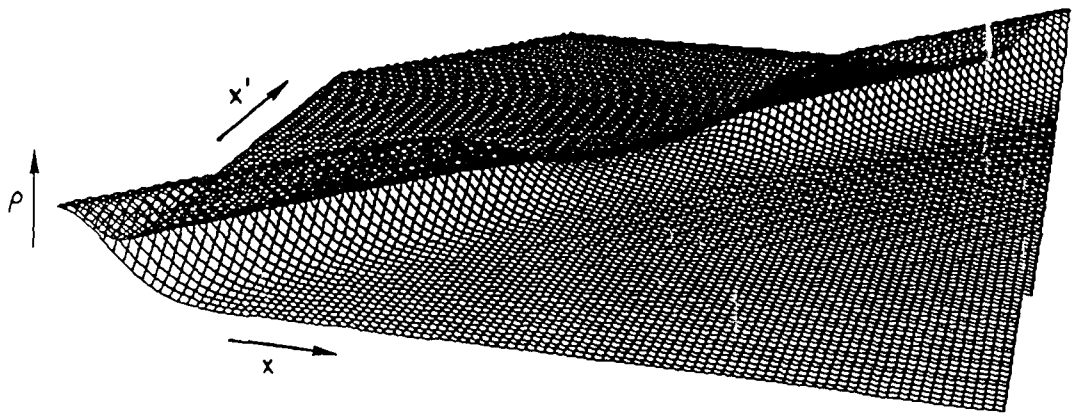


Fig. 2. Two dimensional representation of the density matrix, $\rho(x,x')$, for the NIN structure whose diagonal components are shown in Figure 1.

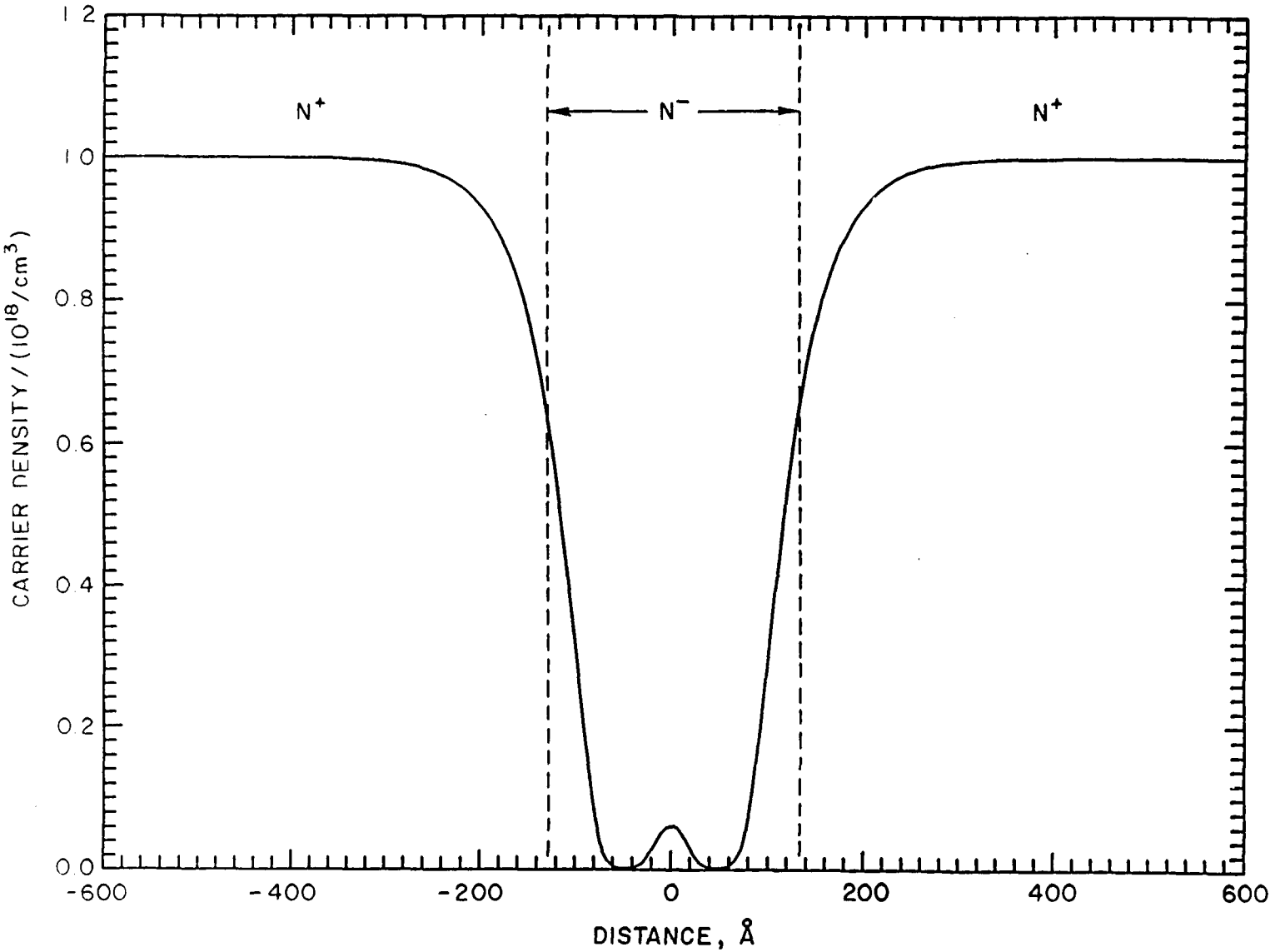


Fig. 3a. Carrier and background distribution for a $N^+N^-N^+$ structure with double barrier potentials in the N^- region. The barriers are each 30 meV high, 50 Å wide separated by a well that is 50 Å wide. The effective mass for this calculation is constant and equal to that of gallium arsenide. Free particle boundary conditions are assumed, and the equation of motion of the density matrix, Equation (3), is solved with zero dissipation, along with Poisson's equation.

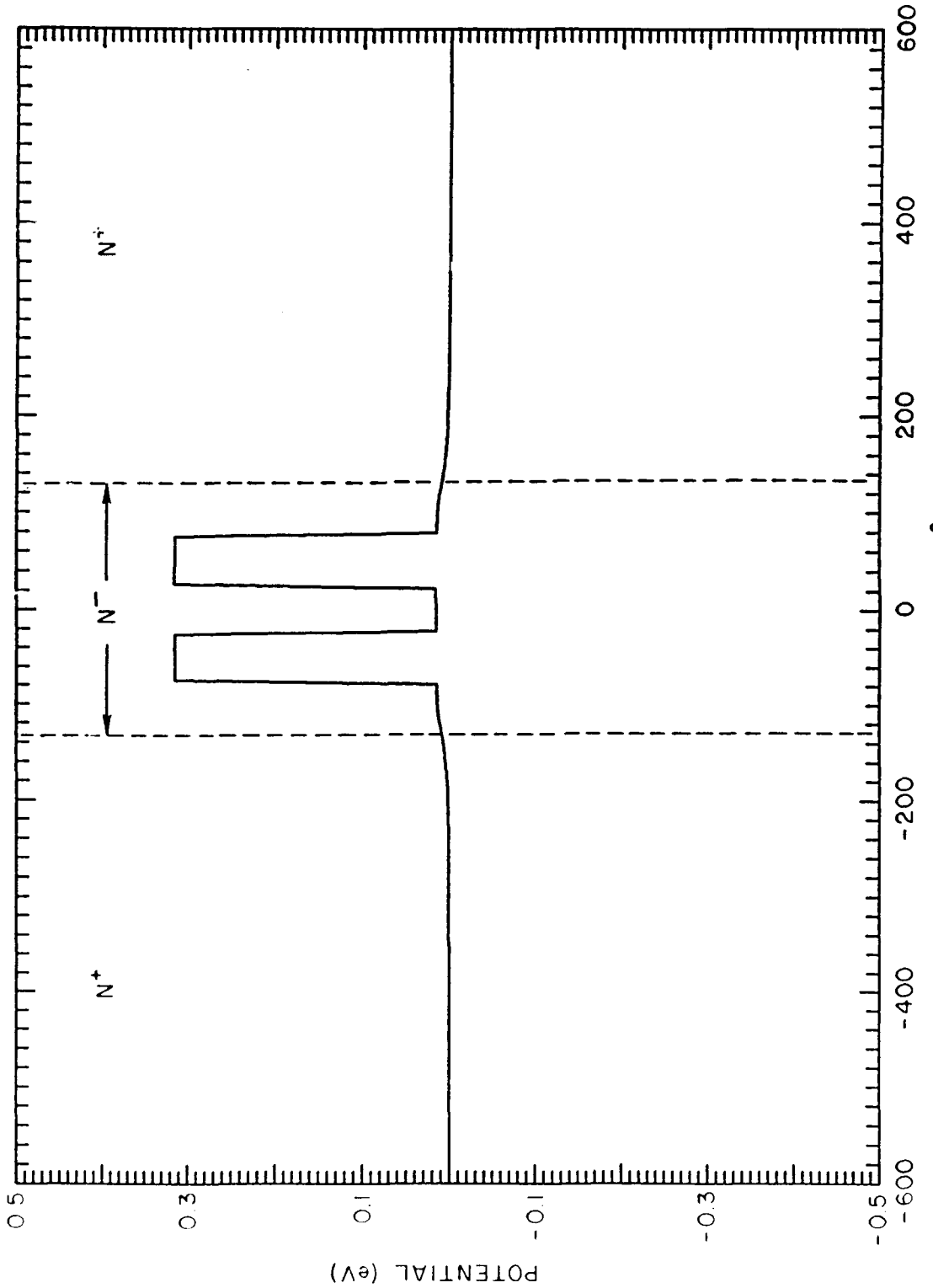


Fig. 3b. Self consistent potential plus heterostructure potential associated with the charge distribution of Figure 3a. The N^+N^- and N^-N^+ regions are identified by the dashed lines.

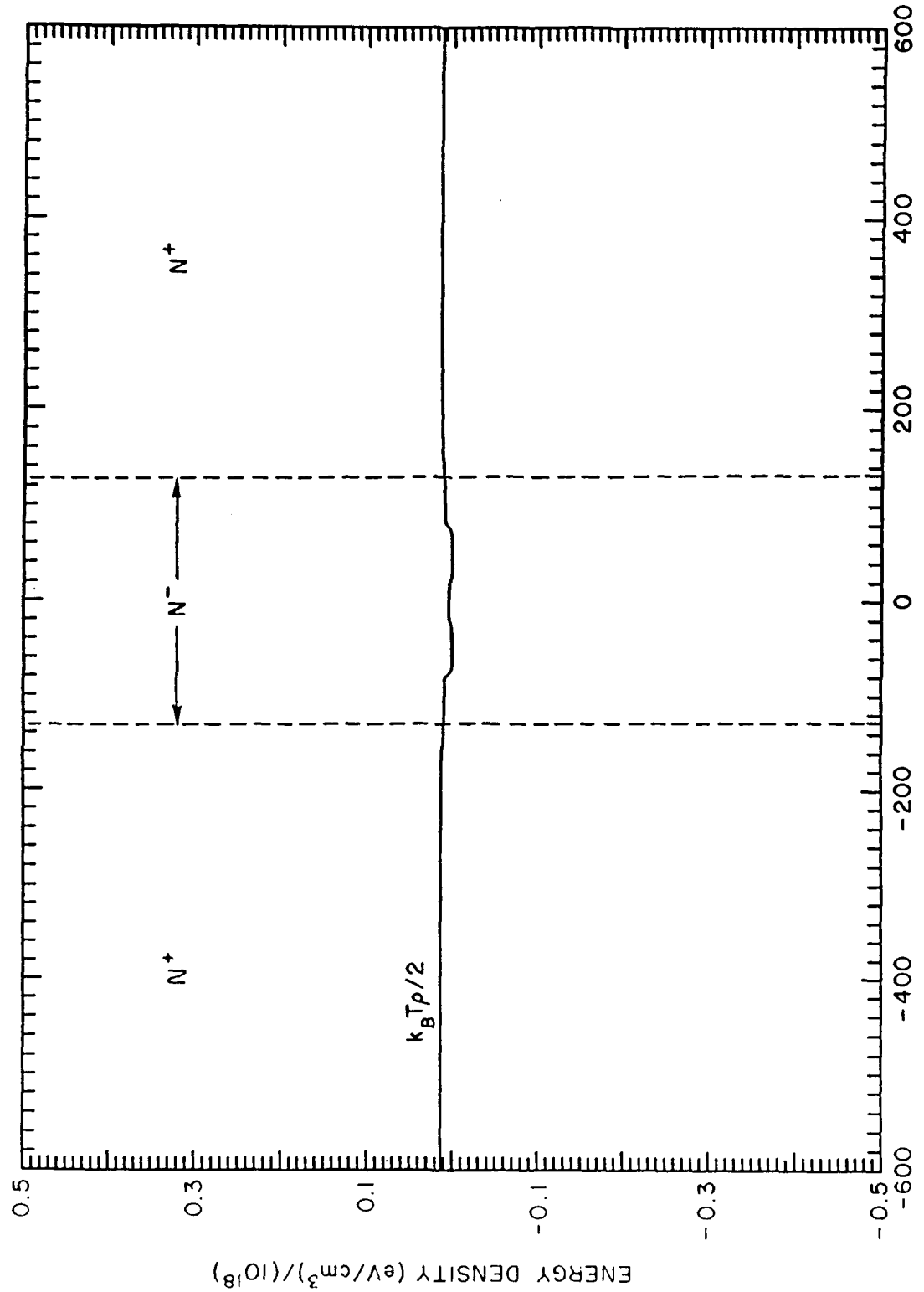


Fig. 3c. Energy density as obtained from Eq.(18). Energy density at boundary region is entirely thermal.

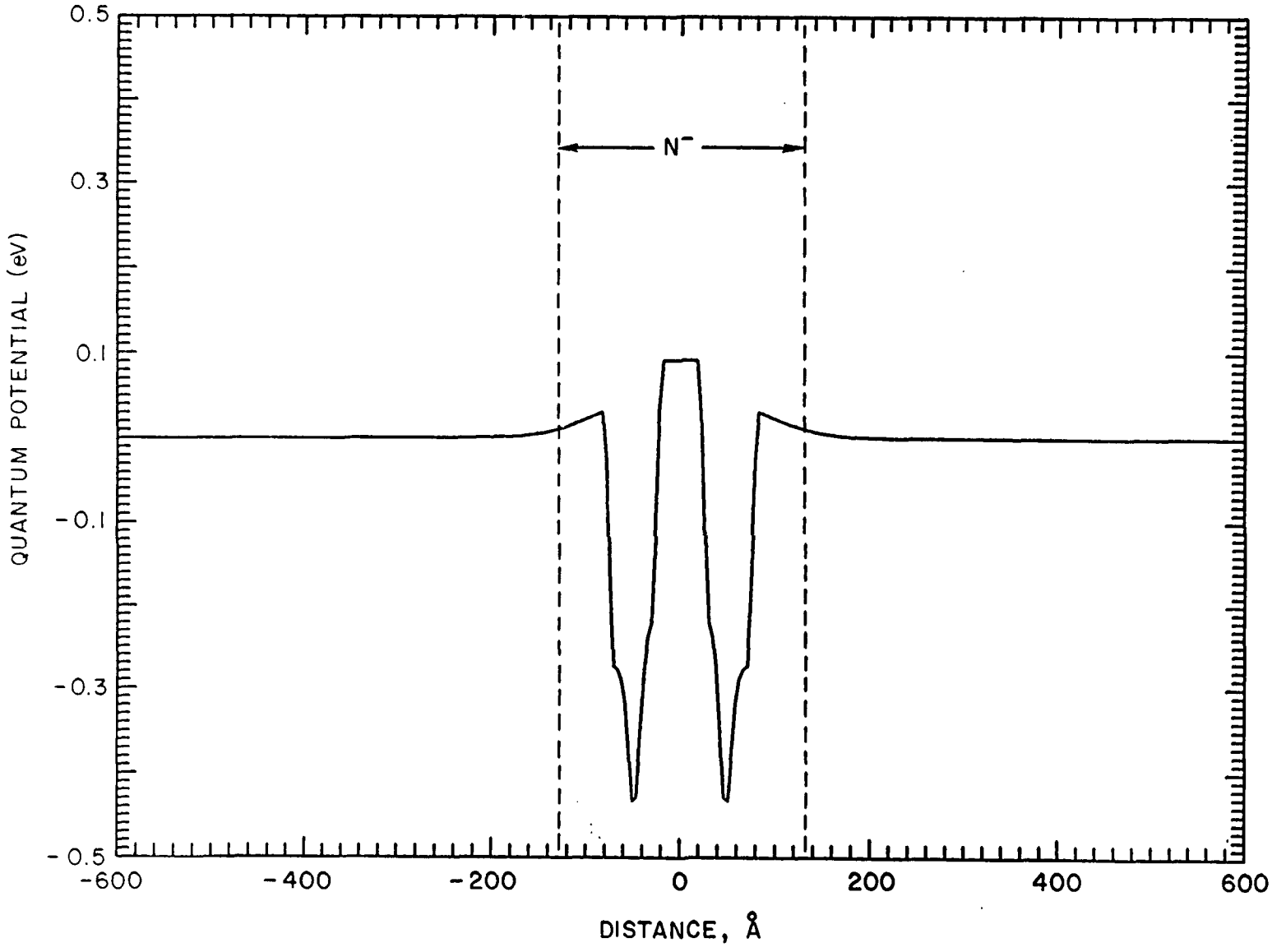


Fig. 3d. Quantum potential for the charge distribution of Figure 3a as obtained from Equation (1).

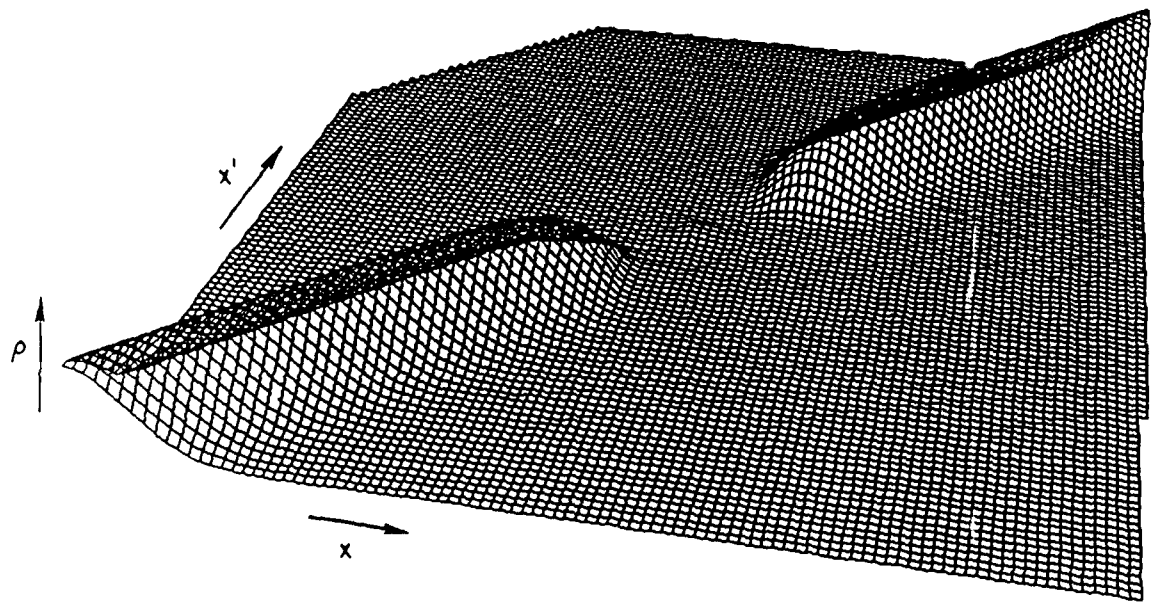


Fig. 4. Two dimensional representation of the density matrix, $\rho(x,x')$, for the double barrier structure whose diagonal components are shown in Figure 3.