AFCRL 71 0030

A CENTER OF COMPETENCE IN SOLID STATE MATERIALS AND DEVICES

by
Fred A. Lindholm, Arthur J. Brodersen, Eugene R. Chenette,
Stephen W. Director, Robert W. Gould, Larry L. Hench
John J. Hren, Sheng S. Li, James K. Watson, Aldert van der Ziel

Electrical Engineering Department
College of Engineering
University of Florida
Gainesville, Florida 32601

Contract No. F 19628-68-C-0058

Project No. 8687

Scientific Report No. 6

10 October 1970

Contract Monitor
Andrew C. Yong
Solid State Sciences Laboratory

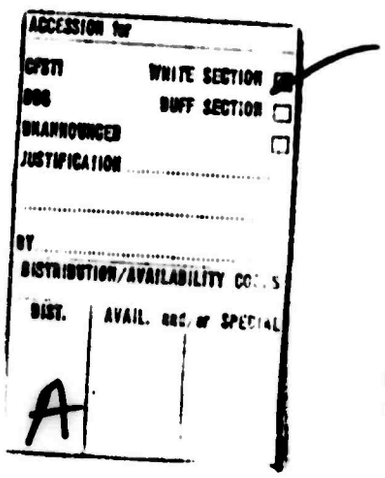
Sponsored by

Advanced Research Projects Agency
ARPA Order No. 1060

Monitored by

Air Force Cambridge Research Laboratories
AIR FORCE SYSTEMS COMMAND
United States Air Force
Bedford, Massachusetts 01730

Program Code No.	1D10
Effective Date of Contract	11 September 1967
Contract Expiration Dote	31 August 1971
Principal Investigotor and Phone No.	Dr. Fredrik A. Lindhalm, 904-392-0904
Project Scientist or Engineer and Phone No.	Dr. Andrew C. Yang, 617-861-2225



Qualified requestors may obtain additional copies from the Defense Documentotion Center. All others should apply to the Clearinghouse for Federal Scientific and Technical Infarmation.

BLANK PAGE

AFCRL – 71-0030

A CENTER OF COMPETENCE IN SOLID STATE MATERIALS AND DEVICES

by
Fred A. Lindholm, Arthur J. Brodersen, Eugene R. Chenette,
Stephen W. Director, Robert W. Gould, Lorry L. Hench
John J. Hren, Sheng S. Li, Jomes K. Wotson, Aldert von der Ziel

Electrical Engineering Deportment
College of Engineering
University of Florido
Goinesville, Florida 32601

Controct No. F 19628-68-C-0058

Project No. 8687

Scientific Report No. 6

10 October 1970

Controct Monitor
Andrew C. Yong
Solid Stote Sciences Loborotory

Sponsored by

Advonced Reseorch Projects Agency
ARPA Order No. 1060

Monitored by

Air Force Combridge Reseorch Loborotories
AIR FORCE SYSTEMS COMMAND
United States Air Force
Bedford, Massochusetts 01730

Program Code No.	1D10
Effective Date of Contract	11 September 1967
Cantract Expiration Date	31 August 1971
Principal Investigatar and Phane Na.	Dr. Fredrik A. Lindholm, 904-392-0904
Praject Scientist or Engineer and Phone No.	Dr. Andrew C. Yang, 617-861-2225

ABSTRACT

This report describes technical findings in three main subject areas: semiconductors and semiconductor devices, insulating and semiconducting glasses, and measurement techniques. In the research concerning semiconductors and semiconductor devices, a model is presented that identifies and fully characterizes the physical origins of burst noise in bipolar transistors. A new characterization of phototransistor noise performance is reported. The density of states and Fermi level for silicon is derived over the whole range of impurity concentrations of technological interest, taking account that the density of states depends on impurity concentration. Steady-state recombination and trapping processes in gold- and phosphorus-doped silicon are studied for impurity concentrations and temperatures such that the impurity density greatly exceeds the equilibrium carrier density. A unifying approach is described for deriving models for field-effect devices. Algorithms for the automated design of semiconductor integrated-circuit amplifiers are developed that grow transistors, in the computational sense, so as to attain specified design goals. In the research concerning insulating and semiconducting glasses, variations of thermal expansion coefficient, elastic modulus, and fracture strength of Li_2O-SiO_2 glass-ceramics are determined as functions of the nucleation treatment and volume fraction of crystals present. X-ray small-angle scattering results indicate that a metastable reaction precedes equilibrium crystallization in certain Li_2O-SiO_2 glasses. The results of our continuing research effort concerning the electrical properties of glass are summarized and related to results obtained elsewhere. In the research concerning measurement techniques the design for a more reliable tip-holder for field-ion microscopes is proposed. X-ray scattering, x-ray diffraction and transmission electron microscopy are employed to study the solid solution decomposition of aluminum-rich, aluminum-zinc-silver alloys.

SUMMARY

This report, for the sixth semiannual period of contract support, describes technical findings in three main subject areas: semiconductors and semiconductor devices, insulating and semiconducting glasses, and measurement techniques.

SEMICONDUCTORS AND SEMICONDUCTOR DEVICES:

A model based on the random occupancy of flaw centers associated with crystallographic defects is developed to fully explain the characteristics of the burst noise seen in bipolar transistors. In contrast to previous explanations, this new model takes account of the influence of defects at the surface as well as those in the bulk. For many transistors the defects at the surface dominate in determining the observed burst noise. Experimental results on specially fabricated transistors, some of which were gold-doped, compare well with the predicted behavior. From the experimental data one can infer values of some of the material parameters that describe the defects.

A full characterization of phototransistor noise performance is reported. As a by-product of this characterization, we demonstrate that noise measurements made on a phototransistor can be used to characterize its performance by determining the common-emitter curreng gain and cut-off frequency. For devices in which the base lead is unavailable, this technique yields information important to device and current design that was not previously attainable.

Taking account that the density of states in a semiconductor depends not only on energy but also on impurity concentration, we calculate the density of states and the Fermi level for silicon over the whole range of impurity concentrations of technological interest. The calculated density of states is shown to agree with the implications of pertinent experimental results.

Steady-state recombination and trapping processes of injected carriers in gold- and phosphorus-doped silicon are studied for impurity concentrations and temperatures such that the impurity density greatly exceeds the equilibrium carrier density. The fluctuation of the charge in the impurity centers resulting from injection is taken into account by assuming the condition of charge neutrality. From the empirical data we formulate a power-law expression that describes the relationship between injected electrons and injected holes. The study provides the dependence of lifetime on injected carrier concentration for silicon both over-compensated or under-compensated by the

gold impurities.

A unifying approach is described for deriving models for field-effect devices. The most general model yie ded by this approach contains, as special cases, models for the small- and large-signal behavior of the two major types of field-effect transistor and of related devices such as the semiconductor current limiter and the pinch resistor. From this most general model emerge, by application of particular approximations and constraints, several models proposed earlier for use in computer-aided circuit analysis. The unifying approach thus demonstrates the relationship borne among various of the existing models for field-effect devices. Additionally, it provides a wide diversity of new models of different complexities and degrees of accuracy useful in the simulation and computer-aided design of semiconductor integrated circuits.

Algorithms for the automated design of semiconductor integrated-circuit amplifiers are developed. These algorithms grow transistors, in the computational sense, so as to attain specified design goals. Under the reasonable assumption that certain fabrication parameters remain fixed, suitable design parameters for transistors are the dc collector current and the area and the perimeter of the emitter. The algorithms developed allow transistors to grow to meet ac gain specifications; dc design requirements arise naturally. The design of a semiconductor integrated-circuit differential amplifier illustrates the feasibility of the approach.

INSULATING AND SEMICONDUCTING GLASSES:

Variations of thermal expansion coefficient, elastic modulus, and fracture strength of Li_2-SiO_2 glass-ceramics are determined as functions of nucleation treatment and volume fraction of crystals present. Strength enhancement is attributed to an increase in fracture energy, which is proportional to the mean-free path between crystals. Strength reduction is attributed to the development of localized cracks at the crystal-glass interface arising from volumetric changes occurring during crystallization.

X-ray small-angle scattering results indicate that a metastable reaction precedes equilibrium crystallization in 30 mole % Li_2O-SiO_2 and 33 mole % Li_2O-SiO_2 glasses. Heterogeneous dielectric relaxation behavior provides additional evidence for the reaction sequence in the non-phase-separating 33 mole % glass.

The results of our continuing research effort concerning the electrical

properties of glass are summarized and related to results obtained elsewhere. We present first the theoretical basis for understanding the electrical behavior of glass. Second, we describe the role played by composition and structure. Third, important problems in the relationships borne among structure, composition, and electrical properties are discussed. Following a description of a general theory of conductivity, the theoretical discussion then treats ionic conductivity and electronic conductivity in glasses. Dielectric properties are considered; the discussion includes theory, polarization mechanisms, and the current status of understanding of composition and structural effects. Switching phenomena receive attention.

MEASUREMENT TECHNIQUES:

X-ray scattering, x-ray diffraction and transmission electron microscopy are employed to study the solid solution decomposition of several aluminum-rich, aluminum-zinc-silver alloys. Hot-stage electron microscopy provides a direct determination of the decomposition sequence. G. P. zones are observed to exist in all of the ternary aluminum-zinc-silver alloys used in this study. Only one type of G. P. zone is detected. Evidence is given supporting the contention that G. P. zone formation is governed by the existence of a metastable miscibility gap. Experimental results indicate that the presence of the G. P. zone distorts the lattice of the matrix.

Tip-holders for field-ion microscopes are traditionally made using glass-to-metal seals in various configurations. Such assemblies suffer occasional catastrophic failures. Thus a need exists for a durable tip-holder with electrical and thermal properties comparable to those of the glass-metal types. We describe the design and properties of such a holder which uses no glass in its construction.

TABLE OF CONTENTS

I. Introduction

The original general objective of this research program was to establish at the University of Florida a "Center of Competence in Solid-State Materials and Devices." From the efforts expended in developing this center of competence have evolved technical findings: technical findings concerning such materials as glass-ceramics, semiconducting glasses, magnetic films, and degenerate semiconductors; concerning devices made from these materials; concerning measurement techniques; and concerning methods of fabrication. Previous findings in these various areas are described in five previous scientific reports[1,2,3,4,5]. In Scientific Report No. 1 the reader will find a more detailed statement of research objectives given here and a discussion of the means to be used in achieving these objectives.

The present report sets forth major findings of the sixth semiannual period of support. In the presentation to follow, Section II describes the results of research concerning semiconductors and semiconductor devices. Section III reports findings concerned with insulating and semiconducting glasses and Section IV treats advances in measurement techniques.

REFERENCES

1. F. A. Lindholm et al, Scientific Report No. 1, Contract #F 19628-68-C-0058, College of Engineering, University of Florida, 10 April 1968.
2. F. A. Lindholm et al, Scientific Report No. 2, Contract #F 19628-68-C-0058, College of Engineering, University of Florida, 10 October 1968.
3. F. A. Lindholm et al, Scientific Report No. 3, Contract #F 19628-68-C-0058, College of Engineering, University of Florida, 10 April 1969.
4. F. A. Lindholm et al, Scientific Report No. 4, Contract #F 19628-68-C-0058, College of Engineering, University of Florida, 10 October 1969.
5. F. A. Lindholm et al, Scientific Report No. 5, Contract #F 19628-68-C-0058, College of Engineering, University of Florida, 10 April 1970.

II. Semiconductors and Semiconductor Devices (A. J. Brodersen, E. R. Chenette, S. W. Director, R. W. Gould, L. L. Hench, J. J. Hren, S. S. Li, F. A. Lindholm, C. T. Sah, A. van der Ziel)

A. PHYSICAL ORIGINS OF BURST NOISE IN TRANSISTORS (K. B. Cook, Jr and A. J. Brodersen)

Introduction

Burst noise is an important source of low frequency noise in bipolar transistors[1-6] Burst noise is characterized by random, but discrete, shifts in the dc currents of semiconductor devices. One of the first investigations into the physical mechanisms of burst noise was done by Card and Chaudhari[5,7]. Their work led to the conclusion that burst noise was due to the modulation of current flowing over a potential barrier. This current could be described by the conventional Schottky diode equation. However, their theory could not predict the observed temperature dependence of the amplitude of the discrete shifts of the dc current.

Recently, the presence of burst noise has been correlated with the presence of "excess currents" in transistors and pn junctions[8,9] The excess current is generally attributed to tunneling near an inverted base surface or leakage through crystallographic defects either near the surface or in the bulk. The excess current can also be described by current flow over a potential barrier such as a Schottky barrier.

Hsu, Whittier and Mead[10] recently proposed that crystallographic defects, located in the emitter-base depletion layer in contact with the semiconductor, form the Schottky barriers responsible for both excess current and burst noise. If a minor flaw, such as a generation-recombination center, is located in proximity to a Schottky barrier, the random occupancy of the g-r center by an electron or hold modulates the height of the potential barrier causing the discrete fluctuations of dc current.

Fig. 1 shows several regions in bipolar junction transistors which are possible locations of the crystallographic defects: the bulk emitter-base space-charge region; the base surface; and the area of the emitter-base space-charge region at the surface. Region One accounts for bulk defects which exist in the emitter-base space-charge region. Impurities which precipitate into such a defect during device fabrication form a low resistance current path. If the resistance is small enough, a substantial portion of the total emitter-base junction current can flow through the defect and any fluctuations in the leakage current are visible in the transistor output. Region Two is located near the surface in the base region of the device, but not in the emitter-base

space-charge region, when the surface potential is zero (the so-called flat-band condition). Current flow through defects in this region, which occurs under surface inversion, is a strong function of surface potential. Region Three is located at the surface in the emitter-base space-charge region, which means the leakage in this area can also be a strong function of surface potential.

Hsu, Whittier, and Mead only considered defects in Region One. The more important regions are Two and Three since the probability of crystallographic defects and flaws is much more likely. We extend the results to include the defects which appear at the surface.

Experimental results are given on five devices to substantiate the theory. These devices are specially fabricated transistors with gates over the emitter-base junction to control the surface potential. Four samples are gold-doped in order to introduce known impurities into the device. The experimental data is used to find values, which compare well with published values, for the capture cross-section and energy levels of the impurity atoms.

A Theoretical Noise Model

A noise circuit model of the transistor such as the one shown in Fig. 2(a) is adequate to describe the terminal behavior of burst noise[2,8]. The resistor r_s is the series resistance of the defect and the resistor r_d is the incremental resistance of the barrier associated with the defect. The current generator I_{BN} represents the burst noise generator. For circuit measurements, the model of Fig. 2(b) is more convenient. The mean squared value of noise generator i_{BB} is found to be

$$\overline{i_{BB}^2} = \left(\frac{r_d}{r_s + r_d}\right)^2 \left[\frac{\Delta\alpha}{\sigma\nu}\right] \left[I_{BN}^2/(1 + \omega^2\alpha^2)\right] \qquad (1)$$

Where σ is the average time the burst noise waveform is in the more positive state and ν is the average time the burst noise waveform is in the more negative state and

$$1/\alpha = 1/\sigma + 1/\nu \qquad (2)$$

The frequency dependence of the spectrum is determined by the statistics of the simple two-time constant process[12,13]. The burst noise generator sees only a portion r_c of the total base resistance $r_b + r_c$ depending upon the physical location of the crystallographic defect. To characterize the burst noise, the quantities I_{BN}, r_s, r_d, r_c, r_b, σ and ν must be related to the physical parameters of the defect.

The energy band diagram shown in Fig. 3 shows a typical cross-section of a defect responsible for burst noise. The n-type silicon represents the emitter of the transistor and the p-type silicon represents the base of the transistor. Between these two regions lies the defect. We assume the defect acts as an effective high recombination region and is essentially metallic in nature. The flaw is assumed to be located in the barrier formed between the defect and n-type material. Since barriers on n-type material are typically larger than barriers on p-type material, current flow through the defect will be more sensitive to fluctuations in the barrier on n-type material. It is the random occupancy of the flaw or trap which is responsible for the current fluctuations. All of the applied voltage V_{BE} does not appear across the emitter barrier. Voltage is also dropped across the second barrier and, for larger currents through the defect, voltage is dropped across the series spreading resistance of the defect. For Schottky barriers typically found in semiconductors, the current is described by[13]

$$J = J_o \exp\left[\frac{qV_A - \phi_b}{mkT} + \frac{m-1}{mkT}\left(E_{Fn} - E_c(\infty)\right)\right] \tag{3}$$

The barrier height $_b$ includes any effects due to surface potential if the defect is in Regions Two or Three. The values of the voltages across each barrier are simply related since the current across the two barriers is equal. The resulting expression is

$$\frac{qV_{An} - \phi_{bn}}{m_1 kT} + \frac{m_1 - 1}{m_1 kT}\left[E_{Fn} - E_c(\infty)\right] = \frac{qV_{Ap} - \phi_{bp}}{m_2 kT} + \frac{m_2 - 1}{m_2 kT}\left[E_V(\infty) - E_{Fp}\right] \tag{4}$$

where V_{An} and V_{Ap} represent the values of applied voltage dropped across each barrier and m_1 and m_2 are constants associated with barriers on n-type and p-type material, respectively.

The effect of the random occupancy of the flaw in Fig. 3 is to alter the barrier height from its steady state value ϕ_{bn}. From Eq. (3), the fluctuation ΔJ in total current is

$$\Delta J = -\left(\frac{qJ}{m_1 kT}\right) \Delta\phi_{bn} \tag{5}$$

where $_{bn}$ is the fluctuation in barrier height. The potential fluctuation $_{bn}$ is

$$\Delta\phi_{bn} = \frac{q}{4\pi\varepsilon r} e^{-r/\lambda} \tag{6}$$

if the point charge associated with the g-r center is within a Debye shielding length. This expression represents the potential of a single electron modified by the shielding of other local charges. The magnitude of the burst noise current I_{BN} is found by combining Eqns. (5) and (6) and integrating over the surface area of the defect. Then,

$$I_{BN} \equiv \Delta I = \frac{q^2 J \lambda}{2\varepsilon m_1 kT} \left[e^{-x_o/\lambda} - e^{(-x_o^2 + A/\pi)^{1/2}/\lambda} \right] \tag{7}$$

where x_o is the approximate distance from the flaw to the potential barrier metallurgical junction.

The incremental resistance r_d of the barrier is also found from Eqn. (3) as

$$r_d = (\partial I/\partial V_{An})^{-1} = m_1 KT/qI \tag{8}$$

The total series resistance r_s includes the effect of the second potential barrier (the one without a flaw as shown in Fig. 3) and the series spreading resistance of the defect. For a hemispherical geometry the spreading resistance is

$$r_{sp} = \rho/\sqrt{2\pi A} \tag{9}$$

where ρ is the average bulk resistivity of the semiconductor in the vicinity of the defect. The small signal resistance associated with the second barrier is similar to that of the first except for the different value of the parameter m. This resistance is given by

$$\left(\frac{\partial I}{\partial V_{Ap}}\right)^{-1} = m_2 kT/qI \tag{10}$$

Tht total series resistance r_s of the burst noise model is then

$$r_s = m_2 kT/qI + \rho\sqrt{2\pi A} \tag{11}$$

These equations are useful in interpreting the terminal behavior of burst noise. Eqn. (7) indicates that I_{BN} is proportional to the defect current density; hence, a plot of $\ln(I_{BN})$ versus the applied voltage is directly proportional to the applied voltage and will give the sum of the parameters m_1 and m_2. Eqn. (4) and Kirchoff's voltage law are used to express V_{An} in terms of V_{BE}. The strong temperature dependence of the parameter m precludes obtaining the barrier height ϕ_{bn} from a plot of $\ln(I_{BN})$ versus 1/T. Eqn. (1)

indicates that the magnitude of the equivalent input burst noise current pulses will be equal to I_{BN} when $r_s << r_d$. This condition is true for small forward-bias. For large forward-bias, the resistance r_s can be much greater than r_d. For this case, the magnitude of the equivalent input burst noise pulses becomes independent of the applied voltage. Thus an experimental plot of $\ln(I_{BN})$ versus the applied voltage is a straight line for small values of V_{BE} and possibly saturates for larger values of the applied voltage.

The other parameters of Eqn. (1) for the spectrum of the burst current generator depend upon the frequency behavior of the spectrum. The frequency behavior is determined by the random occupancy of the flaw or trap and Shockley-Read-Hall (SRH) statistics give[13]

$$\sigma = \frac{1}{e_p + c_n n} \tag{12}$$

and

$$\nu = \frac{1}{e_n + c_p p} \tag{13}$$

The capture process for holes and electrons is the same as in bulk silicon except that the capture cross sections are now field dependent because of the large electric fields which exist in most space charge regions. The rate at which electrons (or holes) can leave a trap or be emitted from a potential well is proportional to the depth of the potential well. For the case of an electron trap this depth is $E = E_c - E_t$, neglecting the Poole-Frenkel effect and the emission process can be described by the SRH parameter e_n. For a quasi-equilibrium approximation the emission rates are $e_n = e_n n_1$ and $e_p = c_p p_1$ where p_1 and n_1 are the carrier concentrations that would exist if the Fermi level were located at the energy level associated with the flaw E_T. Emission from the potential well is enhanced by the Poole-Frenkel effect by effectively decreasing the depth of the potential well because of the presence of a large macroscopic field in the barrier. Thus $E = E_c - E_T - \Delta U$ where

$$\Delta U = 2q/\sqrt{qF/4\pi\varepsilon} \tag{14}$$

Typical values of ΔU are .1 eV or more.

The complete expressions for σ and ν are[3]:

$$\sigma = \frac{1}{c_p N_V \exp[(E_V - \Delta U - E_T)/kT] + c_n N_C \exp[(E_{Fn} - E_C/kT]} \tag{15}$$

and

$$\nu = \frac{1}{c_n N_C \exp[(E_T + \Delta U - E_C)/kT] + c_p N_V \exp[(E_V - E_{Fp})/kT]} \quad (16)$$

The first order properties of σ and ν are determined by the properties of the flaw. The second order properties of σ and ν occur through the field dependence of T_n, T_p, c_n, c_p, and ΔU.

Usually, the flaw acts as either an electron trap, a hole trap, or a generation-recombination center and one of the two terms in the denominator of Eqns. (15) and (16) is negligible. To a first order approximation, only one of the two terms in the denominators depends on applied bias for a forward-biased barrier. If the flaw is an electron trap then σ is a strong function of applied voltage through the term $E_{Fn} - E_c$ which is proportional to V_{BE} and ν is independent of applied voltage. If the flaw is a hole trap, then σ is independent of applied voltage and ν is a strong function of applied voltage through the term $E_V - E_{Fp}$ which is proportional to V_{BE}. If the flow is a recombination center, both σ and ν are strong functions of applied bias. If the flaw is located in a reverse-biased barrier, hole and electron emission is the dominant process. These emission probabilities are field dependent and show a strong dependence on the applied voltage. Plots of $\ln(\sigma)$ and $\ln(\nu)$ versus $1/T$ will give either the energy level of the trap or information about the hole and electron quasi-Fermi levels at the flaw.

Eqns. (15) and (16) neglect tunneling from the flaw into the metal. The probability of tunneling from the flaw into the metal or surface states at the metal-semiconductor interface depends strongly on the potential barrier seen by the trapped charge. Tunneling from the trap back onto the conduction band can also occur. For a forward-biased barrier, the electric field is decreased at the flaw and tunneling is likely to be a negligible effect.

The physical location of the burst noise source can be determined by two techniques. Jaeger and Brodersen have shown that the value of the resistance r_c locates the burst noise source in many cases[2]. This method is good for locating sources in the active base region. The method is not accurate if the resistance r_c is a small part of the total base resistance. Here, the dependence of the burst noise on gate voltage or surface potential is useful for locating the noise source.

Experimental Results

Experiments were performed on many transistors with burst noise. The

results compare favorably with the theoretical calculations. The devices were made by a planar diffusion process into silicon using boron and phosphorous as impurity dopants. Five transistors with burst noise are discussed here: two burst noise sources were located in the emitter-base space charge region away from the surface; two were located near the surface in Region Two; and one in Region Three. The transistor structure is shown in Fig. 4. The large circular device was designed so that surface properties, geometrical factors, and bulk properties could be studied. A device of a more typical size is also present, located very near the larger device. The smaller device was used to study geometry effects and to compare with results from the larger device.

The first device, No G-1-53, had a burst level in Region Two. This device was typical of many of the transistors processed which had burst noise. The device was fabricated using planar processing and careful procedures for cleanliness. Any mobile ions present in the passivating oxides were first drifted to the surface by a heat treatment with a large negative gate voltage for several minutes to minimize their effect on the base surface potential. This procedure allowed accurate correlation between changes in gate bias and surface potential. Sixty devices were tested on the wafer from which device G-1-53 was taken with the following results:

Number of units showing burst noise with zero gate voltage	2
Number of units showing burst noise with gate voltage greater than zero	8
Number of relatively low noise units	47
Number of units showing large 1/f noise	3
TOTAL	60

Figure 5 shows properties which are a strong function of gate voltage indicating that the burst noise source of device G-1-53 is located near the base surface. Sources located in Regions Two and Three are generally strong functions of gate voltage. The value of the resistance r_c, the amount of base resistance seen by the noise source, was small, substantiating the fact that the noise source was close to the base surface.

To interpret the experimental data the defect potential barrier must reflect its dependence on the gate voltage or surface potential. The band diagram of Fig. 6 shows a defect located near the base surface. We assume that defects near the base surface extend to the surface. Under base surface in-

version the defect will come into contact with the "inversion channel" and is consequently a source of burst noise. Because the potential into the device (away from the surface) is a function of position, the equation describing the potential becomes two dimensional. With the depletion approximation an approximate expression for the potential variation near the base surface is

$$\phi(x,y) = [\phi_s(x,y) - V_D]\ [1 - x/x_D]^2 + V_D \tag{17}$$

where

$$x_D = [2\varepsilon(\phi_s - V_D)/q\ n_{ch}]^{\frac{1}{2}} \tag{18}$$

and n_{ch} is the electron concentration in the inverted channel. The barrier height ϕ_{bn} is expressed as[13]

$$\phi_{bn} = [E_C(x_T,y_T) - E_F(x_T,y_T)] + [\phi_{so} - V_{An} - V_D \quad 1 - x_T/x_D]^2. \tag{19}$$

where (x_T,y_T) is the location of the flaw. The flaw is acting as an electron trap since, as shown in Fig. 7, ν is independent of the voltage V_{BE} and $\ln(\sigma)$ is linearly proportional to the voltage V_{BE}. Fig. 7 is then used to determine that[13]

$$E_C^o(x_T,y_T) - E_F^o(x_T,y_T) = .448 \text{ eV} \tag{20}$$

The superscript is used to denote thermal equilibrium. If $m_2 = 1$, Eqn. (19) becomes

$$\phi_{bn} = \frac{.448 + [\phi_{so} - E_G + kT/q\ \ln(N_V/N_{AS})]\ [1 - x_T/x_D]^2}{1 - [1 - x_T/x_D]^2} \tag{21}$$

The surface acceptor concentration N_{AS} is determined from the base surface inversion characteristic and MOS capacitance measurements. For device G-1-53, N_{AS} was determined to be $1.5 \times 10^{18} cm^3$. If the relative location of the trap x_T/x_D were known, one could determine ϕ_{bn}. A range of possible values can be computed from Eqn. (21) for device G-1-53, giving

$$.448 \text{ eV} \leq \phi_{bn} \leq 1.05 \text{ eV} \tag{22}$$

This range of values is typical of the metallic impurities commonly found in silicon[14].

The plots of Figs. 8 and 9 are used to estimate the electron capture

cross-section. Fig. 8 shows the temperature dependence of σ and ν. The slope of these curves gives values of $(E_C - E_T - \Delta U) = .33$ eV and $(E_{Fn} - E_C) = .348$ eV. Fig. 9, showing the magnitude of the bursts versus V_{BE}, can be used in conjunction with Eqns. (1), (3), (4) and (7), to find the sum of the parameters m_1 and m_2. The electron capture cross section is now found using Eqns. (14) and (15) as $8.6 \times 10^{-18} cm^2$.

The remaining unknowns to completely characterize the barrier are the bulk resistivity of the series spreading resistance ρ, the cross section area of the defect A, and the constant J_o. For cases in which ρ is large enough to cause I_{BB} to saturate for large values of V_{BE}, ρ can be estimated. Eqns. (1) and (7) are used to show that, if $\lambda >> x_o$ and $x_o^2 << A/\pi$, the magnitude of the input noise current pulse saturates to a value of

$$I_{BB} = q/[(2)^{\frac{1}{2}} \varepsilon\rho]. \tag{24}$$

In general, the resistivity may not be large enough to cause the saturation of the current pulses. Fig. 9, for instance, shows no saturation. For this case, the series spreading resistance is not important. A curve of I_{BB} versus the leakage current is useful in determining the defect cross section area A. If the defect leakage current is $I = JA$, the magnitude of the burst current pulse is

$$I_{BB} \equiv - I[q^2/2(m_1 + m_2)\ \varepsilon kTA]\ [x_o - (x_o^2 + A/\pi)^{\frac{1}{2}}]. \tag{25}$$

The slope of a plot of I_{BB} versus I would give A if a reasonable approximation for x_o is made. The value of $m_1 + m_2$ is known from experiment. If x_o^2 is small compared with A/π, then $A = 2.7 \times 10^{-8} cm^2$ for device G-1-53. This cross-sectional area agrees with previously reported values[10].

The experimental data obtained from device G-1-53 correlated in all respects with the proposed model. Data obtained from the model allowed the calculation of a possible range of values for various parameters describing the potential barrier at the defect even though exact values could not be obtained. The estimated electron capture cross section was of the same magnitude as several impurity elements, but was closest to reported values for tungsten[15]. Also, the energy level of the flaw in the energy gap was very close to reported values for tungsten[15]. The important parameters are summarized in Table 1.

The four burst noise sources to be discussed in the remainder of the

paper occurred in devices which were gold doped. The deliberate introduction of gold significantly increased the percentage of devices which had burst noise. A sample of fifty devices from a gold-doped silicon wafer, showed more than 40 percent with large amounts of burst noise compared with 16 percent from the previous example. The gold apparently increases the incidence of metallic precipitates. The behavior of these burst noise sources can be used to compute the properties of the gold defects and the values compared with the well-known properties of gold.

The first gold-doped device, device G-5-66, has a burst noise source located in Region Two. The plot of Fig. 10 shows that both σ and ν are exponential functions of the base-emitter voltage for V_{BE} greater than 0.5 V while only σ is an exponential function of V_{BE} for V_{BE} less than 0.5 V. Eqns. (15) and (16) would indicate that the flaw is acting as a generation-recombination center for V_{BE} greater than 0.5 V and as an electron trap for V_{BE} less than 0.5 V. The possibility of such an occurrence has been discussed by Hsu, Whittier, and Mead[10]. Temperature data similar to the plot of Fig. 9 was obtained. The important parameters obtained from the experimental data are given in Table 1.

Figure 11, a plot of the burst magnitude versus V_{BE}, shows a saturation of the burst magnitude which indicates that the series spreading resistance is limiting the magnitude of the bursts. In this case Eqn. (24) is used to estimate the resistivity of the spreading resistance as 12 Ωcm.

The value of the capture-cross sections is approximately an order of magnitude smaller than the reported values of gold. The capture cross-sections are field dependent and the location of the flaw is near the depletion region of the Schottky barrier. The magnitude of the field can be estimated since the impurity is likely gold. The reported values of the value of the gold-silicon barrier can be used, in conjunction with Eqns. (4), (17), (18) and (21), to locate the position of the flaw and the value of the electric field at that point. The distance from the flaw to the Schottky barrier is 28 Å and the field at this point is 2.5×10^5 V/cm. This high value of electric field easily accounts for the reduced values of the capture cross-sections.

Two devices showed burst noise which was independent of gate voltage. The results of the measurements are given in Table 1. Such sources are located in Region One of the device. Low frequency noise measurements offer a convenient method to locate physically the noise sources[2]. Figs. 12 and 13 show the two measurements which were used to determine the values of the resistances r_c and

r_b shown in Fig. 2(b). Fig 12 is a plot of equivalent input voltage bursts versus source resistance and the intercept for $E_{BB} = 0$ is simply the value of the resistance r_c. Fig. 13 is a plot of noise figure versus source resistance for a frequency at which 1/f noise dominated the noise performance. The 1/f noise current generator appears directly across r_π and the minimum noise figure occurs when $R_S = r_c + r_b$. For device G-7(4)-8, the measurements gave values of $r_c = 470\ \Omega$ and $r_c + r_b = 2\ k\Omega$. Recall, for sources located near the surface, the value of r_c was approximately zero. Therefore, since the burst source of device G-7(4)-8 sees spproximately 25 percent of the total base resistance, these measurements confirm the location of the burst noise source.

The plot of Fig. 14 shows that ν varies linearly with V_{BE} while σ is essentially constant. Thus, the flaw would appear to be acting as a hole trap. The temperature dependence of σ, in conjunction with Eqn. (15), is interpreted to give a value of the hole capture cross section very near reported low-field values for the gold acceptor level. In fact, the temperature dependence of ν gives a clue that the flaw is located in the low-field region away from the Schottky barrier. The temperature data shows that the quantity $E_{Fp} - E_V$ is 0.62 eV. Hence, the Fermi level for holes E_{Fp} lies well above mid band. Fig. 3 would indicate that the flaw must lie some distance from the metal-silicon junction for this to occur, since in the vicinity of the junction, the quasi-Fermi level for holes lies well below mid band. The value of the field becomes smaller as the distance from the junction decreases.

Device G-7(2)-23 gives further validity to this argument. This device is also a burst source located in Region One. Noise measurements gave equal values of 150 Ω for r_c and $r_c + r_b$ indicating the defect was located well beneath the emitter. In this case the plots of σ and ν indicate that the flaw is acting as a generation-recombination center. The values of the capture cross-section in Table 1 are an order of magnitude smaller than the reported value. The measurement of temperature dependence of σ and ν gave values of $E_{Fp} - E_V = 0.44$ eV and $E_C - E_{Fn} = 0.52$ eV. Both the quasi-Fermi levels for holes and electrons lie below the middle of the band gap indicating a p-type region. Fig. 3 now indicates that the flaw must be close to the barrier in a high field region for this to occur. Hence, the low values of the capture cross-section are attributed to the high electric field in the area of the Schottky barrier.

The final device exhibited a burst noise which was dependent on the gate

voltage. However, for large gate voltages which caused the base surface to invert, the burst noise disappeared. This type of characteristic can be attributed to defects in Region Three since the conducting path across the base-emitter region disappears for strong base inversion.

Fig. 16 shows the dependence of σ and ν on the applied bias. The plot indicates that the flaw is acting as a generation-recombination center. Compared with the previous devices, σ and ν vary more rapidly with V_{BE}. Other discrepancies with previous data also were found: the parameter m had a negative temperature coefficient, and the values of and were very strong functions of temperature. The theory developed for the forward-biased Schottky barrier did not fit the observed data.

The data did, however, match with reported values for reversed-biased Schottky barriers. How this barrier could occur in the device is not at all clear. The possibility exists for defects both across the base-emitter junction and in the surface inversion layer, and the metallurgical situation is complex in this region. The dependence of σ and ν on V_{BE} is due to the strong field dependence of the thermal emission probabilities for holes and electrons and the possibility of a large amount of tunneling from the trap. Sah has recently published data on the field dependence of the thermal emission rates of carriers at gold centers in reverse-biased silicon pn junctions, and has indicated that for such cases the detailed field dependence of the emission probability does not follow the conventional Poole-Frenkel formula[16]. The emission probabilities computed from the temperature dependence of σ and ν in conjunction with Eqns. (15) and (16) ($e_p = 1/\sigma$, $e_n = 1/\nu$) are the same magnitude of those reported by Sah for the gold acceptor level. At 30°C, $e_n = 140$ sec^{-1} and $e_p = 178$ sec^{-1}.

Conclusions

This work has developed and substantiated a burst noise theory for the bipolar junction transistor. From previous experimental results and models reported in the literature, the mechanism for burst noise was postulated to be the modulation of leakage current, flowing through a defect across the emitter-base junction, by the random occupancy of a single flaw located near the defect. From the proposed mechanism, a burst noise model was developed. The model was presented in terms of the physical processes occurring at the defect and flaw.

Experiments were performed on five typical test transistors with large amounts of burst noise from a single noise source. The resulting experimental data were analyzed based on the proposed burst noise model. Some of the important data on the flaw and defect, obtained from an analysis of the burst noise source, are summarized in Table 1. Agreement with reported values was good.

The addition of gold into the silicon increased the percentage of transistors showing a large amount of burst noise from 3 to 15 percent to better than 40 percent. Similarly, the percent of units with a large amount of burst noise increased when the base surface was inverted because typically many more defects are located near the surface of a device than in the bulk. The devices which showed burst noise under base surface inversion always showed large excess currents under inversion. Although cases were found in which leakage occurred with the base surface inverted and the devices showed no burst noise, no cases were found in which a device showed burst noise under base surface inversion and no excess leakage current. Based on the proposed model for burst noise, the above effects were expected.

References

1. C. A. Bittman, G. H. Wilson, R. J. Whittier, R. K. Waits, "Technology for the Design of Low Power Circuits," IEEE Journal of Solid State Circuits, SC-5, No. 1, 29-37, February 1970.

2. R. C. Jaeger, A. J. Brodersen, "Low Frequency Noise Sources in Bipolar Junction Transistors," IEEE Transactions on Electron Devices, ED-17, No. 2, 128, February 1970.

3. G. Giralt, J. C. Martin, F. X. Mateu-Perey, "Sur un Phénomène de Bruit dans les Transistors, Caracterisé par des Créneaux de Coupant D'Amplitude Constante," C. R. Acad. Sc. Paris, 261, 5350-5353, 1965.

4. J. C. Martin, D. Esteve, G. Blasquey, "Burst Noise in Silicon Planar Transistors," Conference on Physical Aspects of Noise in Electronic Devices, University of Nottingham, England, September 11-13, 1968.

5. G. M. Glasford, "Study of Low Frequency Noise and Deterioration in Semiconductor Devices," Technical Report No. RADC-TR-65-140, June 1965.

6. W. H. Card, A. Mauretic, "Burst Noise in Semiconductor Devices," 2nd Symposium on the Physics of Failure, Chicago, 1963.

7. W. H. Card and P. K. Chaudhari, "Characteristics of Burst Noise," Proceedings of the IEEE, 53, 652, 1965.

8. S. T. Hsu, R. J. Whittier, "Characterization of Burst Noise in Silicon Devices," Solid State Electronics, 12, 867-878, November 1969.

9. M. Conti, "Surface and Bulk Effects in Low Frequency Noise in NPN Planar Transistors," Societa Generale Semiconduttori, Spa. S.G.S., 1969.

10. S. T. Hsu, R. J. Whittier, and C. A. Mead, "Physical Model for Burst Noise in Semiconductor Devices," private communication.

11. D. Wolf, E. Holler, "Bistable Current Fluctuations in Reverse-Biased p-n Junctions," Journal of Applied Physics, 38, 189-192, 1967.

12. S. Machlup, "Noise in Semiconductors: Spectrum of a Two-Parameter Random Signal," Journal of Applied Physics, 25, No. 3, March 1954.

13. K. B. Cook, Jr., "Burst Noise in Bipolar Junction Transistors," Ph.D. Dissertation, University of Florida, 1970.

14. A. Y. C. Yu, "The Metal-Semiconductor Contact: An Old Device with a New Future," IEEE Spectrum, 7, No. 3, 83, March 1970.

15. E. Schibli, A. G. Milnes, "Lifetime and Capture Cross-Section Studies of Deep Impurities in Silicon," Materials Science and Engineering, 2, 229-241, 68, 1967.

16. C. T. Sah, "Thermal Emission Rates of Carriers at Gold Centers in Silicon," Applied Physics Letters, 15, No. 5, September 1969.

Notation

A	defect cross-sectional area
$E_c(\infty)$	conduction band energy far from the potential barrier of the defect
E_{co}	conduction band energy at surface when $s = 0$
$E_V(\infty)$	valence band energy far from the potential barrier of the defect
E_{Fn}	quasi-fermi energy for electrons
E_{Fp}	quasi-fermi energy for holes
E	electric field
i_{BN}	burst noise current across defect barrier
I_{BN}	magnitude of the burst noise pulse
i_{BB}	equivalent input burst noise current
m	a parameter used to describe current flow across the Schottky barrier
r_c	that part of the total base spreading resistance seen by the burst noise source
r_d	differential resistance of the defect potential barrier
r_s	series resistance in burst noise model
T_n	tunneling probability for electrons
T_p	tunneling probability for holes
ΔU	energy due to the Poole-Frenkel effect
V_{An}	amount of V_{BE} dropped across the barrier on n-type material
V_{Ap}	amount of V_{BE} dropped across the barrier on p-type material
V_D	diffusion potential (contact potential)
x_T	location of flaw
σ	width of the more positive burst noise pulse
ν	width of the more negative burst noise pulse
ϕ_{bn}	barrier height on n-type material
ϕ_{bp}	barrier height on p-type material
λ	Debye shielding length
ρ	resistivity in vicinity of defect

Table 1 A Summary of the Experimental Results

Device	Location of Burst Noise Source	Flaw	Experimentally Determined Emission Rates and Capture Cross Sections	Exp. Determined Flaw Energy Level	Exp. Determined Defect Area, A	Possible Material Forming Defect
G-1-53	Region Two, near the base surface $r_c = 0$	electron trap	$\sigma_{cn} = 8.6 \times 10^{-18} cm^2$	$E_C - E_T - \Delta U = .33$ eV	$2.7 \times 10^{-8} cm^2$	tungsten
G-5-66 (gold doped)	Region Two, near the base surface $r_c = 0$	recombination center	$\sigma_{cn} = .72 \times 10^{-16} cm^2$ $\sigma_{cp} = .15 \times 10^{-16} cm^2$		$10^{-9} cm^2$	gold
G-7(4)-8 (gold doped)	Region One, $r_c = 470\ \Omega$ $r_x = 2k\ \Omega$	hole trap	$\sigma_{cp} = 3 \times 10^{-14} cm^2$	$E_C - E_T - \Delta U = .535$ eV		gold
G-7(2)-23 (gold doped)	Region One, $r_c = r_x = 150\ \Omega$	recombination center	$\sigma_{cn} = 6.2 \times 10^{-16} cm^2$ $\sigma_{cp} = 10^{-17} cm^2$			gold
G-5-48 (gold doped)	Region Three, in the e-b scr near the base surface	generation center	at 30°C $e_n = 140\ sec^{-1}$ $e_p = 178\ sec^{-1}$			gold

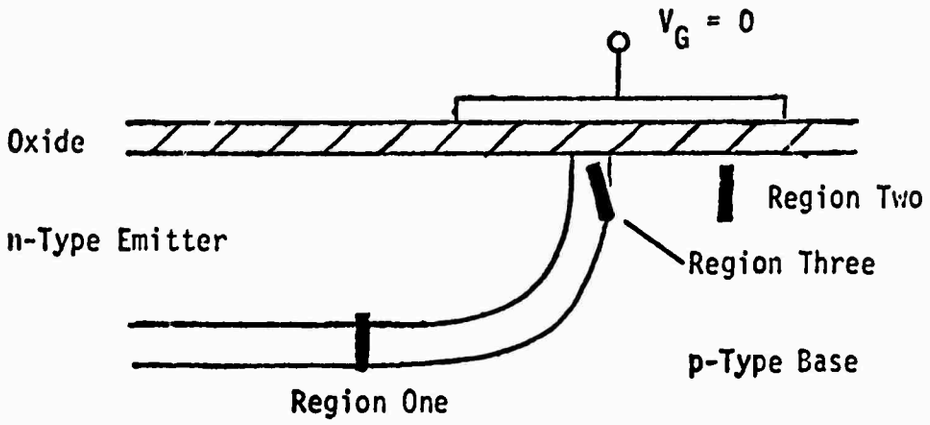


Figure 1 - Possible locations of defects which could cause **burst** noise.

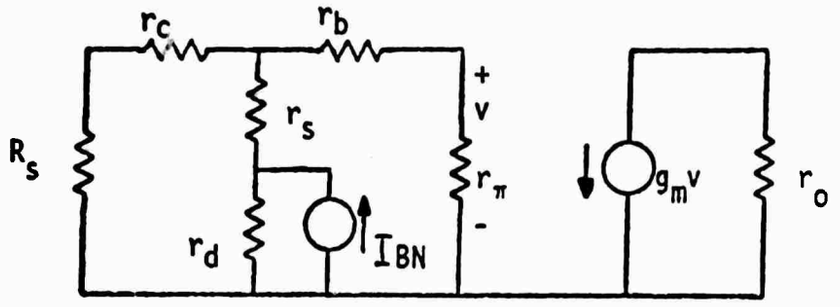


Figure 2a - An equivalent burst noise model for the bipolar junction transistor.

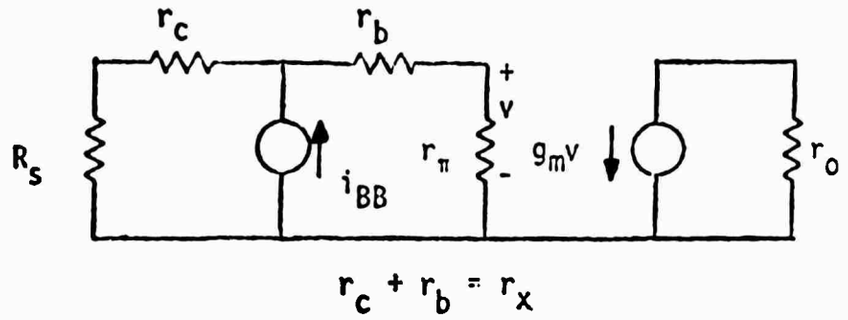


Figure 2b - Burst noise model for the bipolar junction transistor.

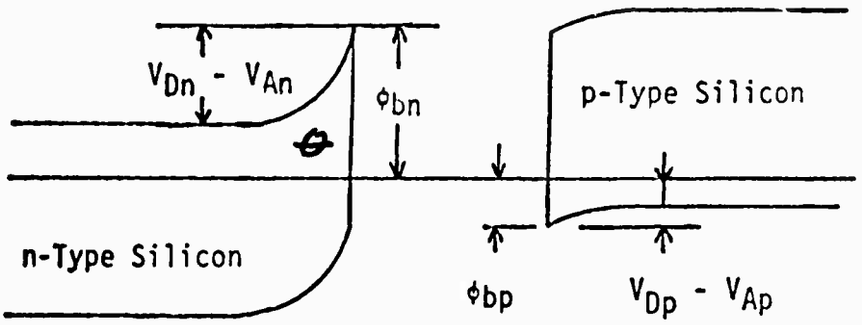


Figure 3 - Potential barriers on n-and p-type material with an applied voltage.

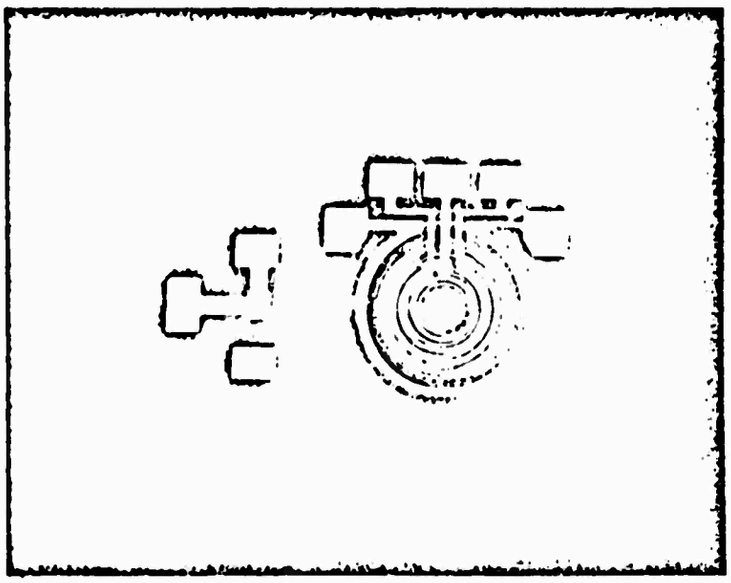


Figure 4a - Bipolar junction transistor test structure.

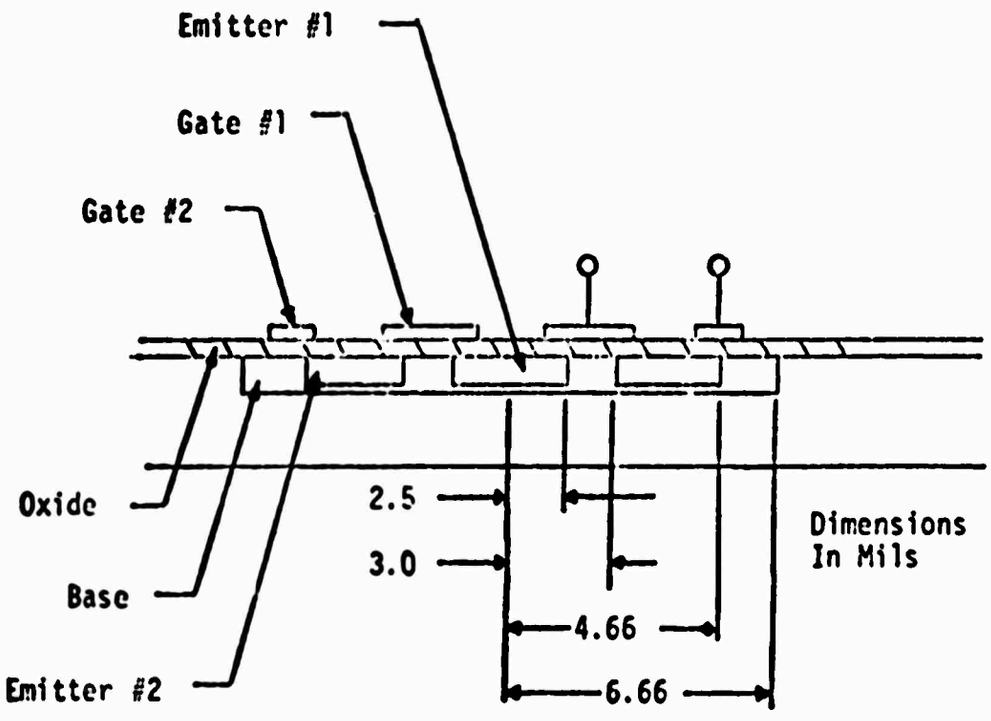


Figure 4b - Cross section of test transistor indicating important dimensions.

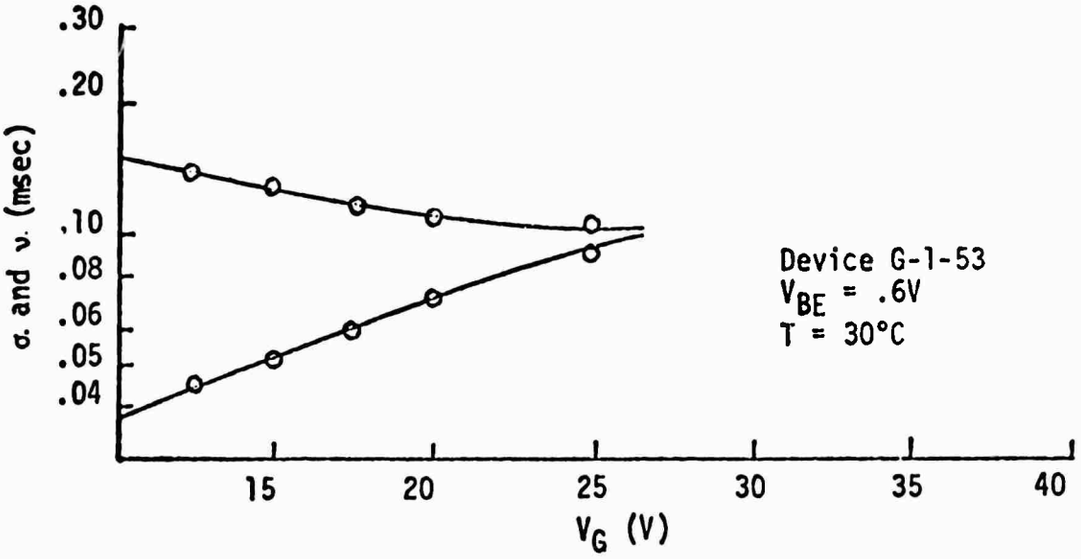


Figure 5a - Plot of σ and ν versus V_G for device G-1-53.

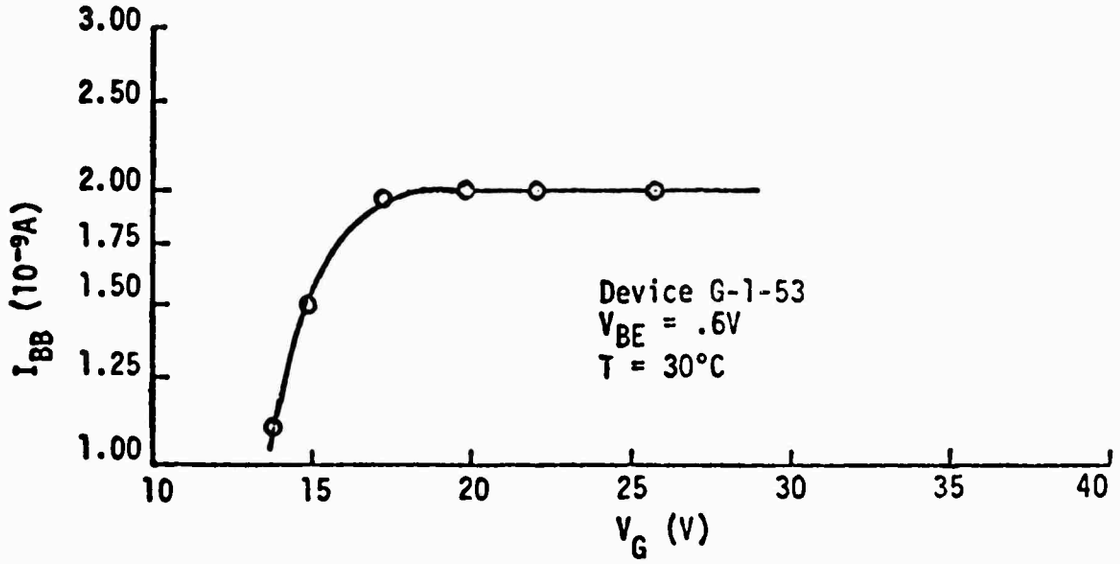


Figure 5(b) - Plot of I_{BB} versus V_G for device G-1-53.

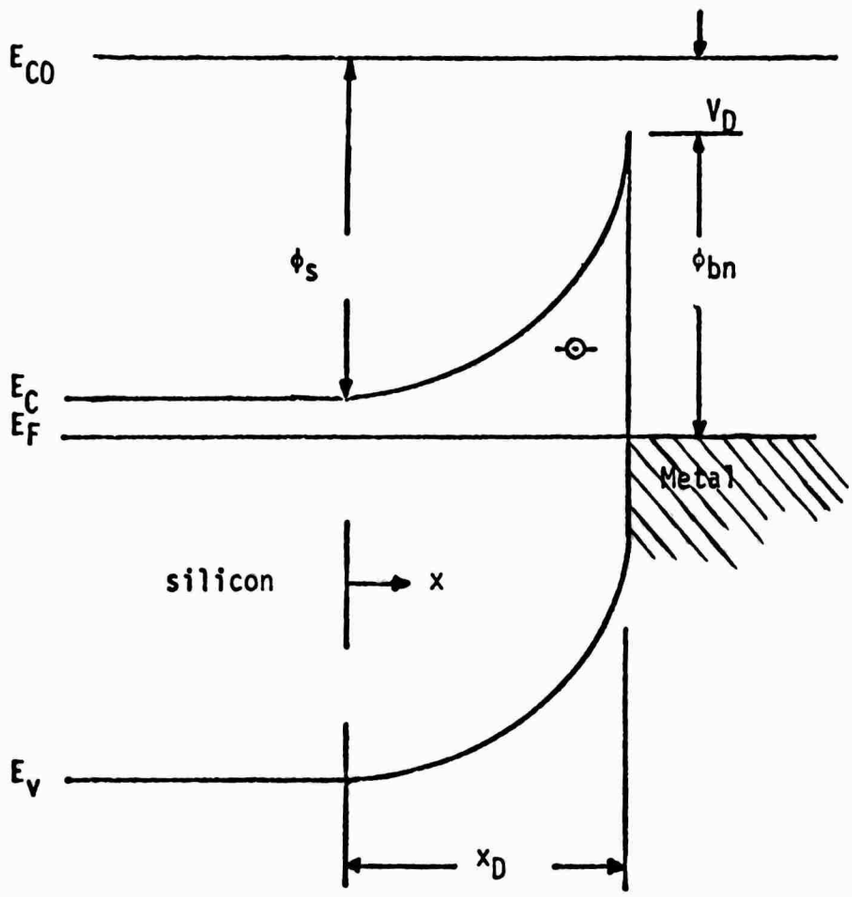


Figure 6 - Potential barrier associated with a surface defect, looking from the surface.

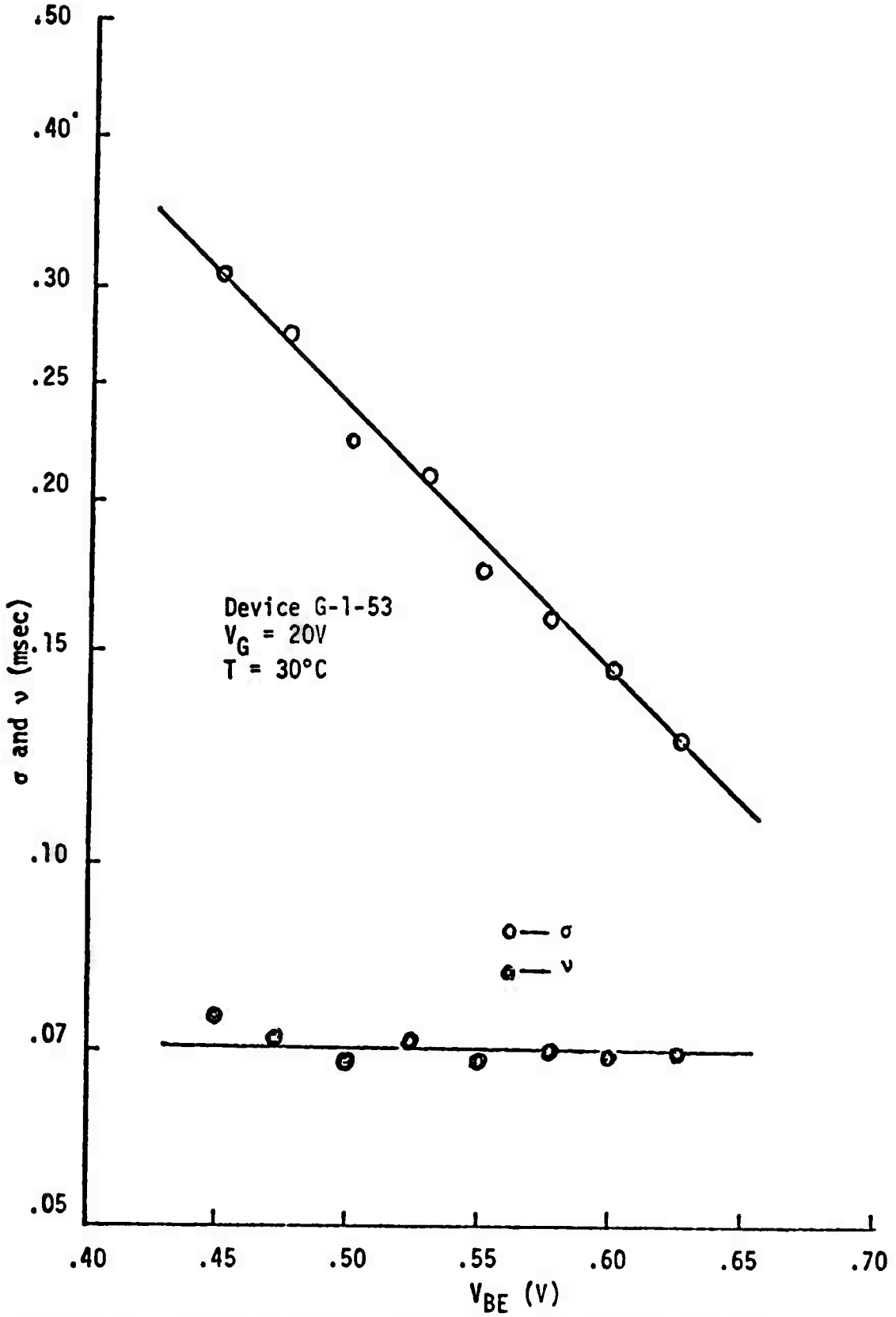


Figure 7 - Plot of σ and ν versus V_{BE} for device G-1-53.

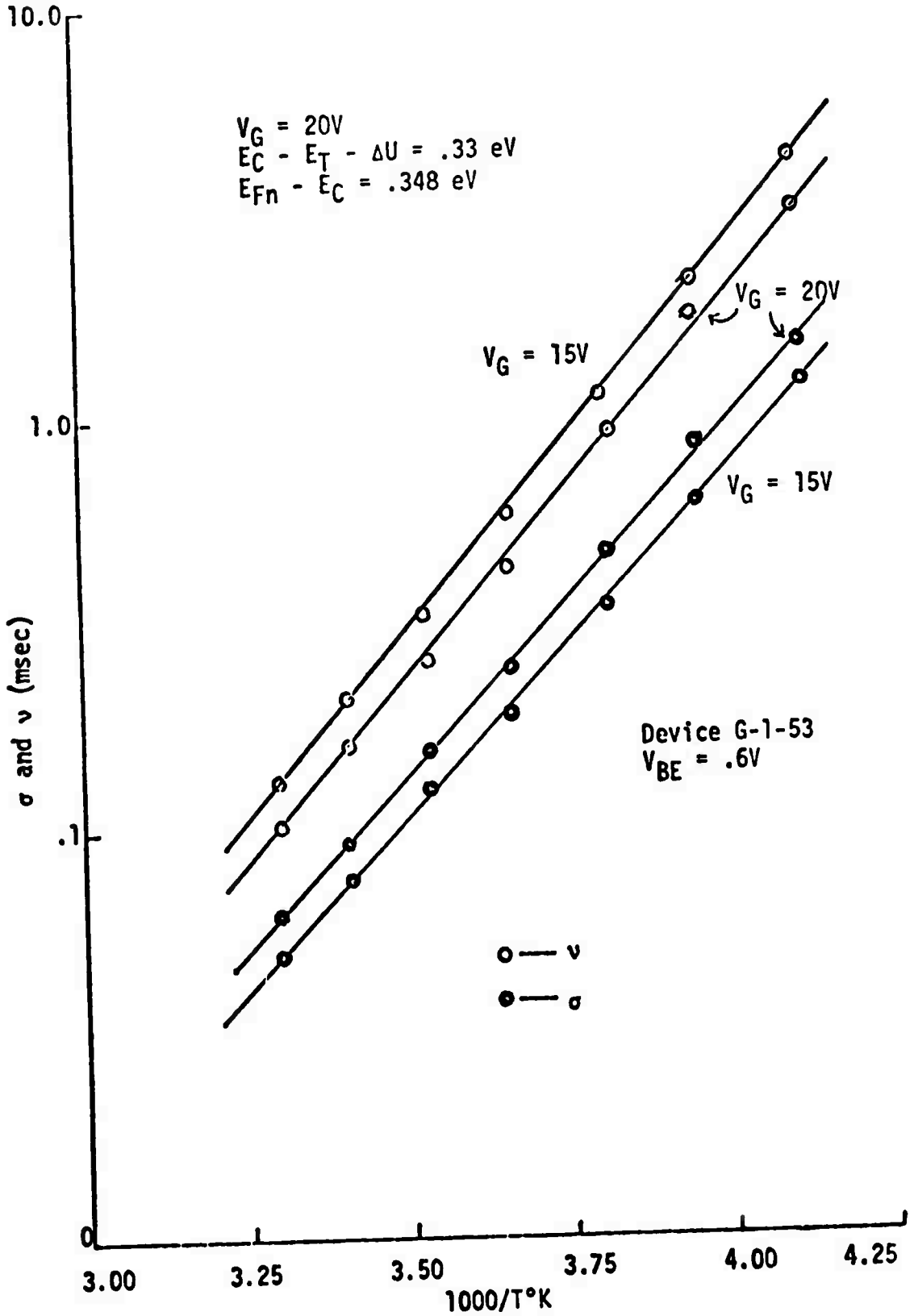


Figure 8 - Plot of σ and ν versus 1000/T°K for device G-1-53.

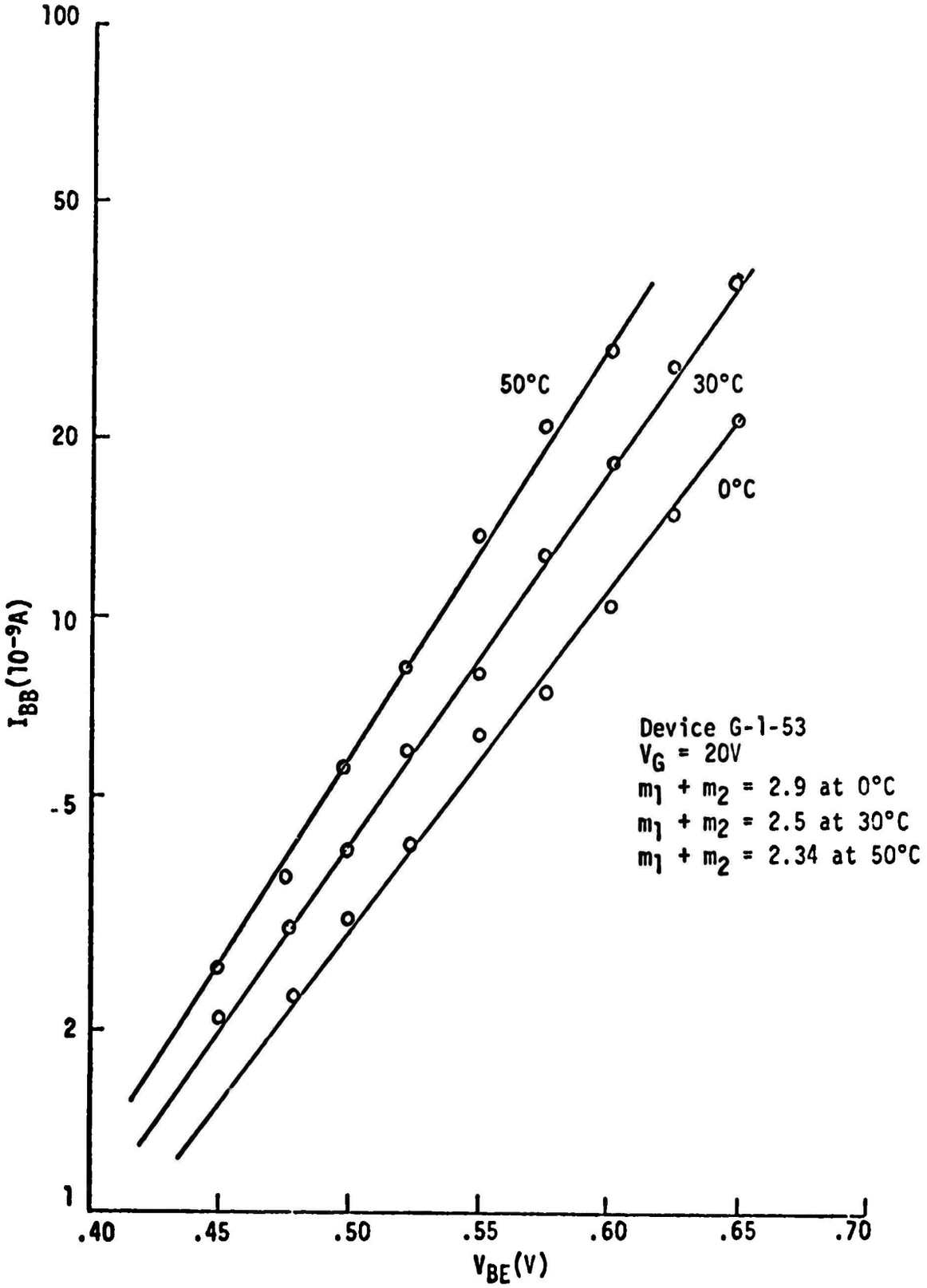


Figure 9 - Plot of I_{BB} versus V_{BE} for device G-1-53.

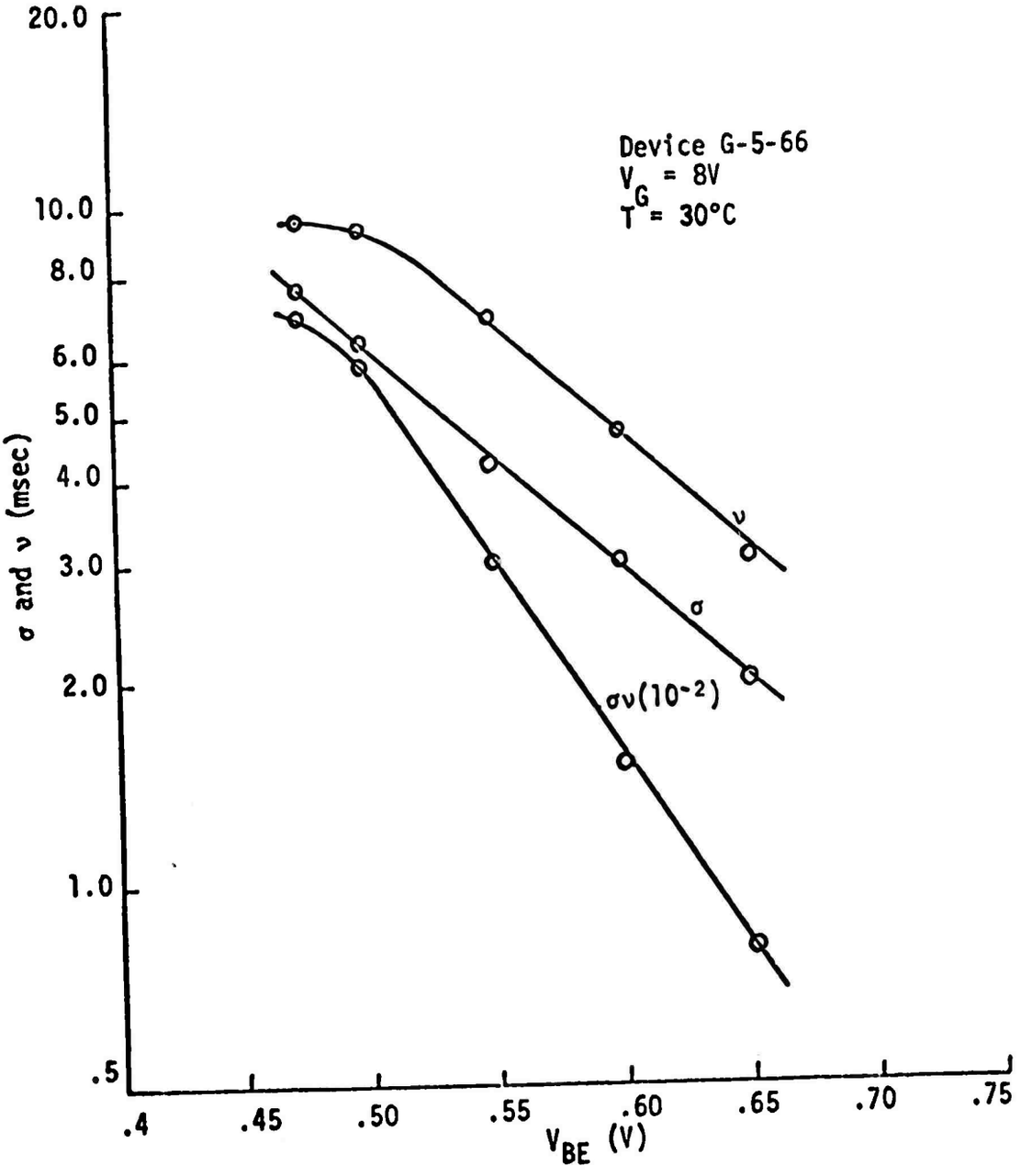


Figure 10 - Plot of σ and ν versus V_{BE} for device G-5-66.

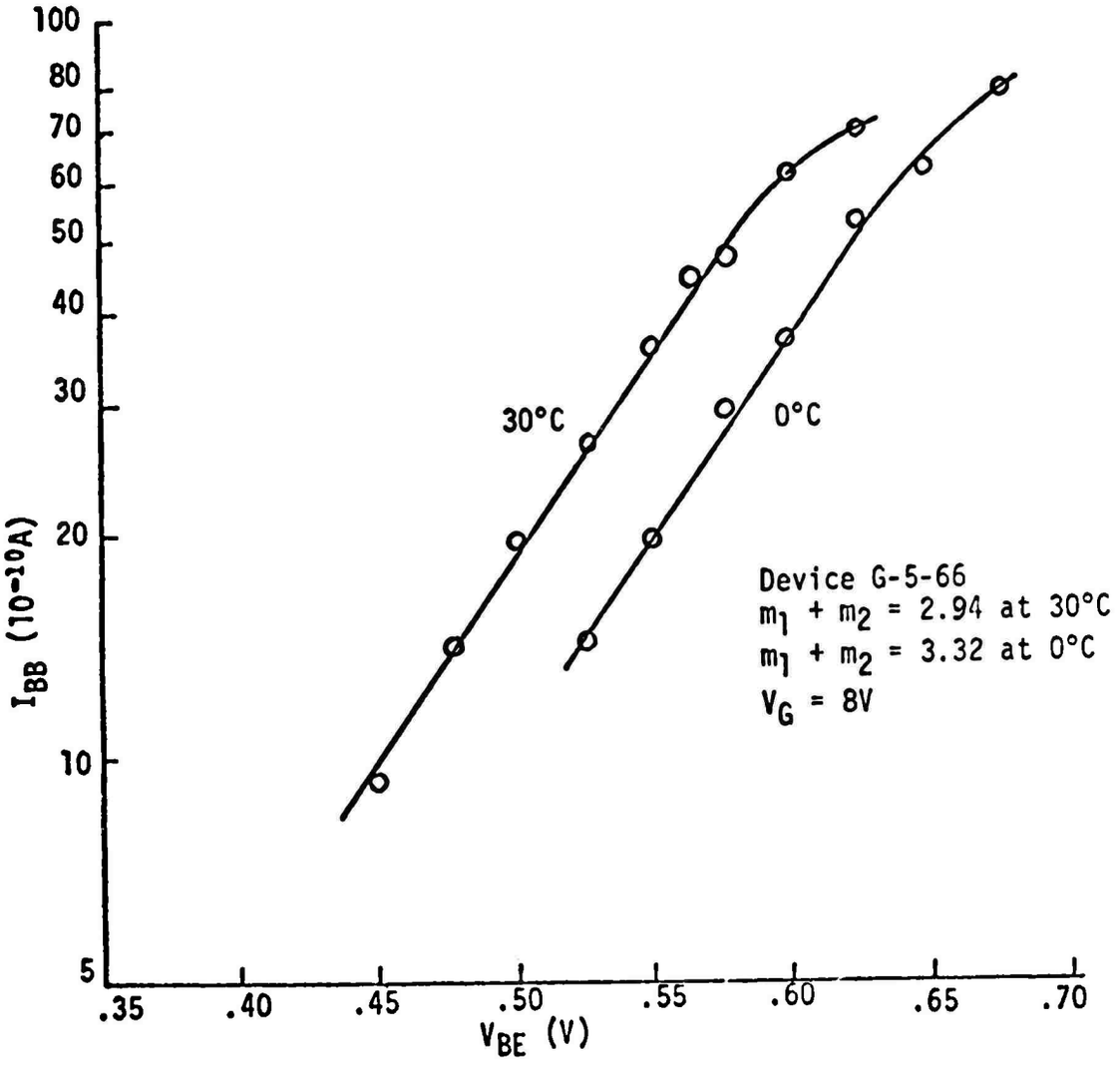


Figure 11 - Plot of I_{BB} versus V_{BE} for device G-5-66.

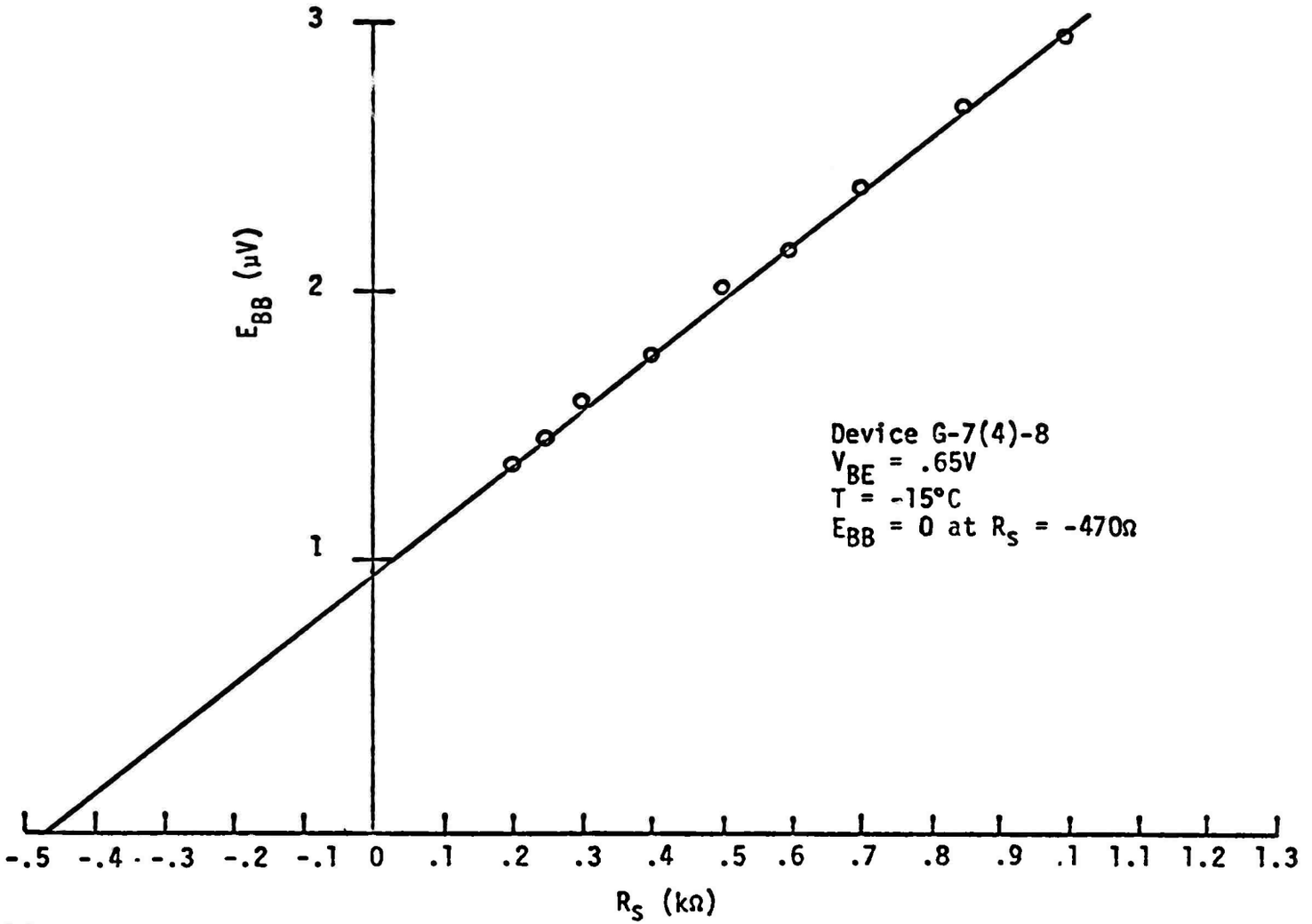


Figure 12 - Equivalent burst voltage versus source resistance.

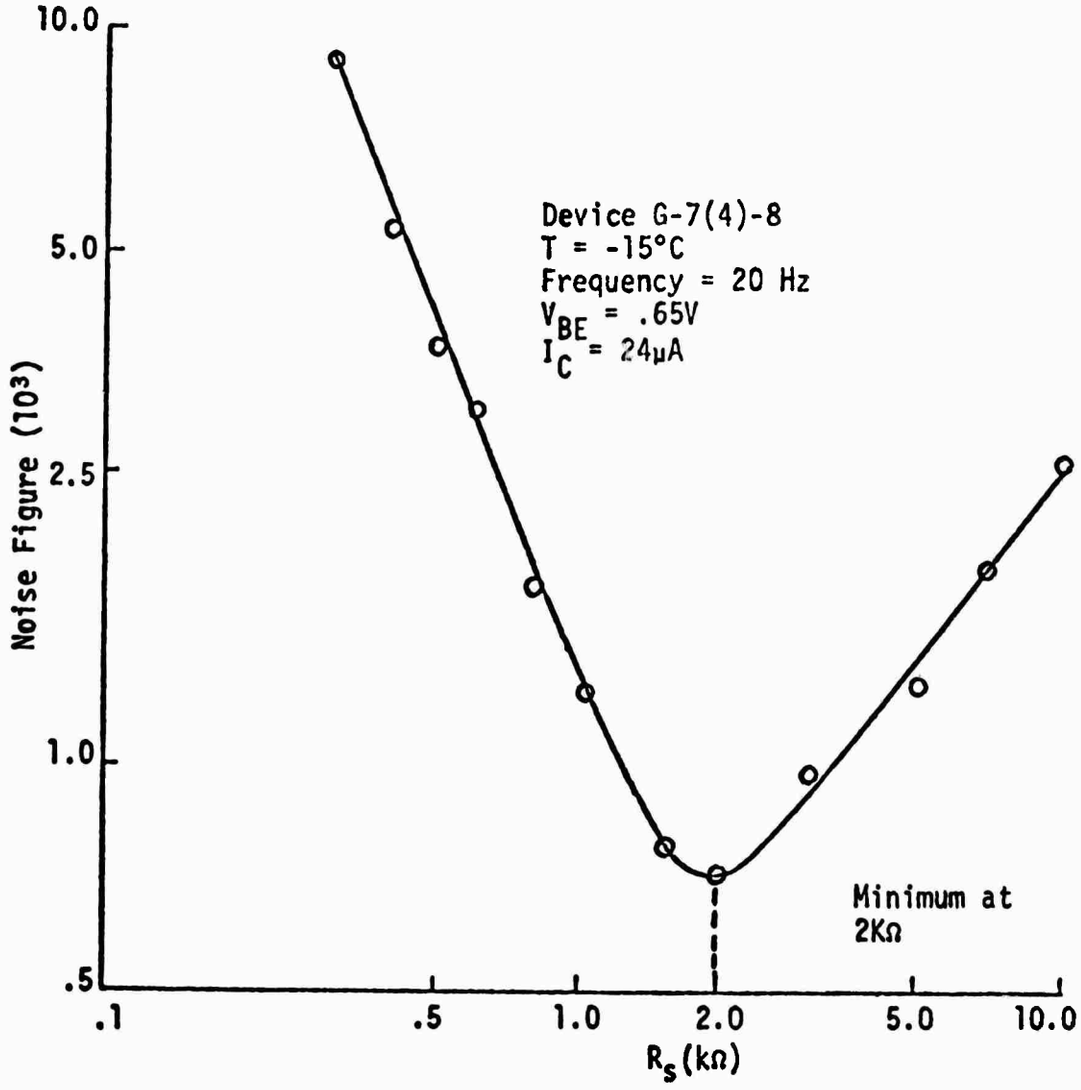


Figure 13 - Plot of noise figure versus source resistance for device G-7(4)-8.

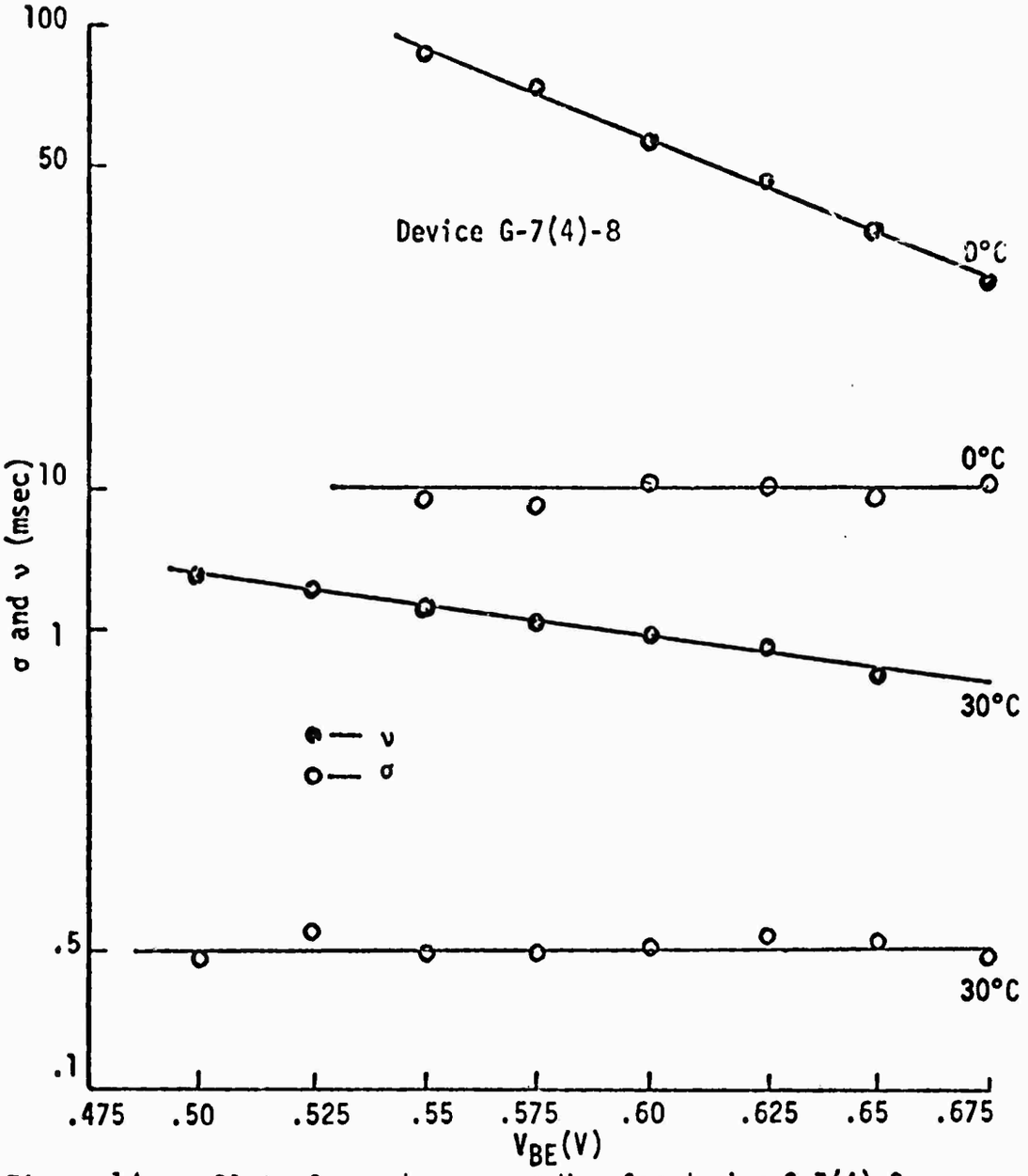


Figure 14 - Plot of σ and ν versus V_{BE} for device G-7(4)-8.

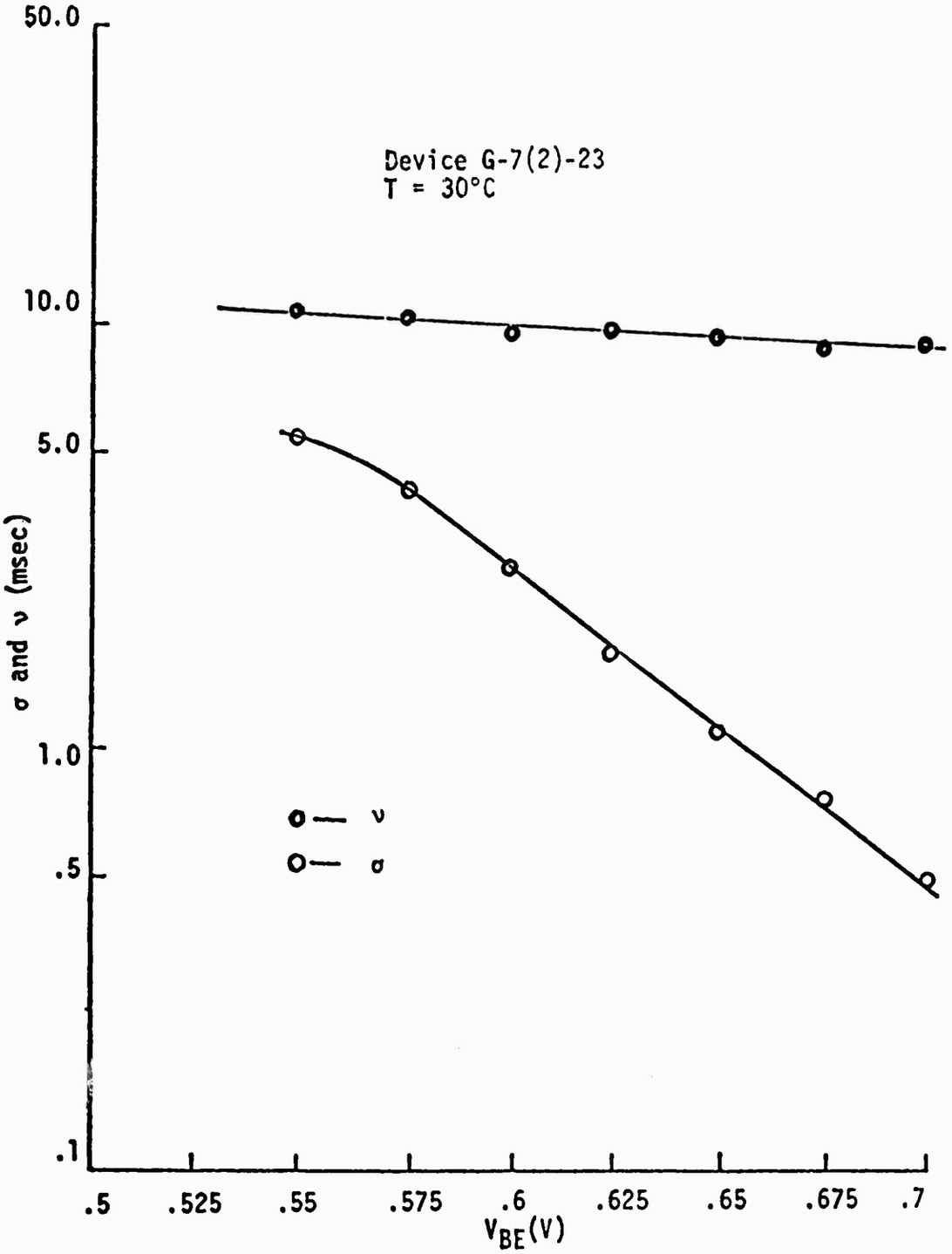


Figure 15 - Plot of σ and ν versus V_{BE} for device G-7(2)-23 at 30°C.

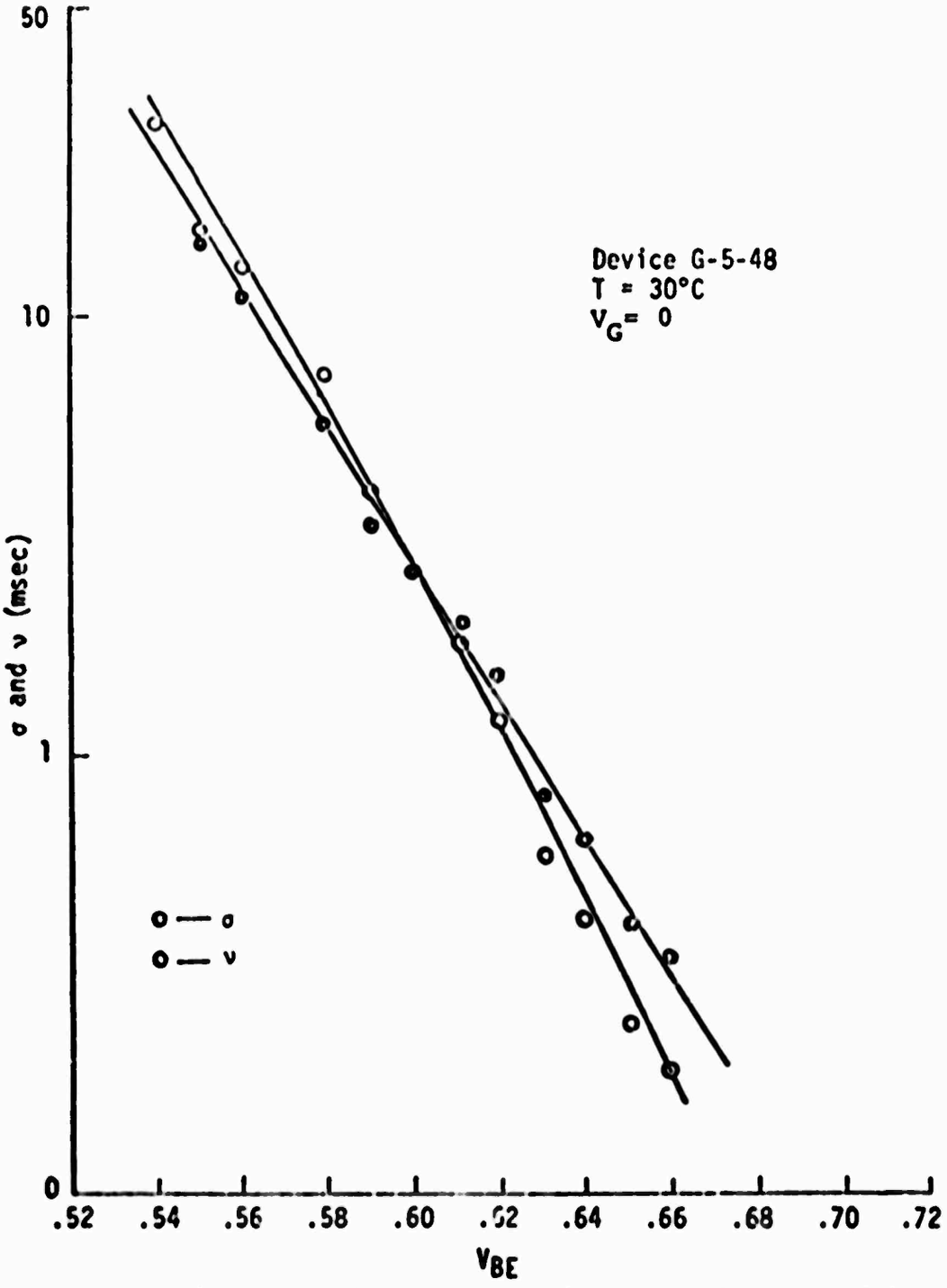


Figure 16. - Plot of σ and ν versus V_{BE} for device G-5-48 at 30°C.

B. NOISE IN PHOTOTRANSISTORS AND PHOTODIODES (F. de la Moneda, E. R. Chenette, and A. van der Ziel)

This paper presents the results of some recent measurements of the spectral density of the noise current at the output of planar silicon phototransistors and a phototransistor noise model which agrees well with experimental results.

Figure 1 shows a schematic diagram of the basic system used to determine the spectral density of the noise.

Representative spectra of commercial and University of Florida devices are shown on Figures 2a and 2b, respectively. The values of I_{eq} of these curves are in good agreement with those obtained from the analytical expression of the spectral density of the output noise.

$$S_o(f) = 2qI_{CEO}\left[1 + \frac{2h_{fe}^2/h_{FE}}{1 + (f/f_\beta)^2}\right] \qquad (1)$$

Here I_{CEO} is the d.c. collector current caused by photoexcitation with the base floating, h_{FE} is the d.c. current gain, h_{fe} is the incremental current gain, and f_β is the beta cut-off frequency.

The low frequency plateau of Equation (1) is clearly visible although it is masked by 1/f noise at the lowest frequencies. The high frequency plateau corresponding to full shot noise of I_{CEO} is also observed for the lower bias currents in the frequency range shown. It is thus possible to determine h_{fe}^2/h_{FE} and f_β from noise data.

Since the devices had an external base lead small signal measurements[1] of h_{fe} and f_β could be carried out. Another determination of f_β using an amplitude modulated GaAs light emitting diode was also made. Finally, the d.c. parameter h_{FE} was also measured. The ratio[2] $m(I_E) = h_{fe}/h_{FE}$ which is a slow function of I_E, is useful in the discussion of these data.

Figure 3 shows h_{FE}, obtained from the d.c. characteristics, $h_{fe} = mh_{FE}$, obtained from the small signal measurements, and $h_{fe}^2/h_{FE} = m^2h_{FE}$ obtained from the low frequency noise data. This gives three methods for determining m. The values of m thus obtained agreed within the limit of experimental error. Figure 4 shows f_β, obtained from the noise data, from the small signal data, and from the modulated GaAs diode; these data also agreed within the limit of experimental error.

The derivation of Equation (1) involves setting up a model which represents the different sources of noise in a phototransistor. These sources can be traced back to:

(a) The incoming photons which induce generation fluctuations.

(b) Mechanisms inherent in the device such as fluctuations in the generation of free carriers by phonons, diffusion fluctuations, and recombination fluctuations in the bulk and on the surface.

Van Vliet[3] has shown that the fluctuations in (a) have full shot noise. Most of the fluctuations in (b) can be represented by random shot noise generators. One can then study the total output noise without paying any regard to the specific origin of its components. This is the underlying assumption made in the following analysis.

Figure 5 shows the familiar "Hybrid-pi" equivalent circuit with the shot noise generators $\overline{i_b^2}$ and $\overline{i_c^2}$. As pointed out by van der Ziel[4], this is a good and yet simple noise representation for the common emitter bipolar transistor. The spectrum given by Equation (1) is derived from this model in conjunction with the expressions for the d.c. collector and emitter currents, and the constraint that the net base current is zero. This constraint gives h_{FE} as the relationship between I_{CEO} and $(I_{PH} + I_{CB})$. I_{PH} is the photo-generated current, I_{CB} is the collector-base leakage current.

$$(h_{FE})^{-1} = (1 - \gamma_N \alpha_F) + \frac{\gamma_N \alpha_F I_{ET} \exp(q V_{BE}/m'kT}{I_{CEO}} \qquad (2)$$

Here, $I_{ET} \exp(qV_{BE}/m'kT)$ is the emitter transition region leakage current, and γ_N the emitter injection efficiency factor.

Equation 1 can be used to compute the noise equivalent power, NEP, the figure of merit conventionally used to characterize the noise performance of radiation detectors. It is given by[3]

$$NEP = h\nu_s j_s A_b \left(\sqrt{i_o^2} / \sqrt{i_s^2} \right) \qquad (3)$$

where h is Planck's constant, ν_s is the optical frequency of the incoming radiation signal, j_s is the a.c. photon flux per unit area, A_b is the phototransistor base area, and i_s the a.c. output current response due to j_s; i.e., $i_s = A_b (\beta+1)\eta q j_s$. With the aid of this relationship and Equation (1), the

NEP can be written as

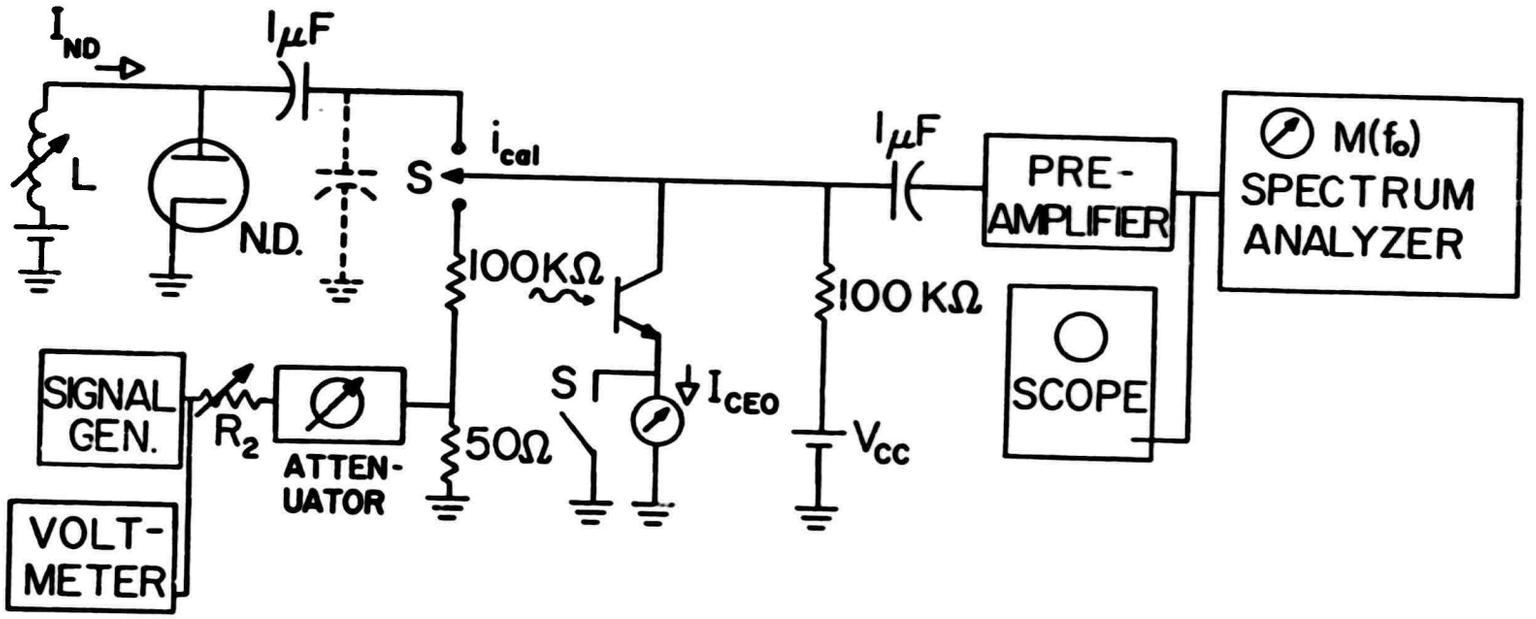
$$\text{NEP} = h\nu_s \left[\frac{(2qI_{CEO})^{\frac{1}{2}} \left[1 + \frac{2h_{fe}^2/h_{FE}}{1 + (f/f_\beta)^2} \right]^{\frac{1}{2}} (\Delta f)^{\frac{1}{2}}}{|\beta + 1|\, \eta q} \right]$$

In summary, we have characterized the phototransistor noise performance both experimentally and analytically. The validity of our analytical results has been established by a set of three independent measurements. The agreements between the results of these measurements is excellent. As a byproduct, we have also that noise measurements on a phototransistor can be used for its small signal characterization in terms of $m^2 h_{FE}$ and f_β. This result may be of great usefulness, particularly when the base lead of the device is not available.

A more detailed treatment of this research will appear in the next scientific report.

REFERENCES

1. M. B. Das and A. R. Boothroyd, IEEE Transactions Electron Devices, ED-8, 15 (1961).
2. B. Schneider and M.J.O. Strutt, Proc. IRE, 47, 546 (1959).
3. K. M. van Vliet, Journal of Applied Optics, 6, 1145 (1967).
4. A. van der Ziel, Proc. IEEE, 57, 1211 (1969).



SPECTRUM ANALYZER

FREQUENCY	INSTRUMENT	Δf
3.5 Hz-80 KHz	G-R 1921 ANALYZER	.255 f_o
20 Hz-50 KHz	H-P 302A ANALYZER	5.23 Hz
1KHz-1.5 MHz	H P 310A ANALYZER	2.54 KHz
1MHz-30 MHz	COLLINS 51S-1 RECEIVER	

Figure 1 Schematic diagram of system for measurement of noise in phototransistors.

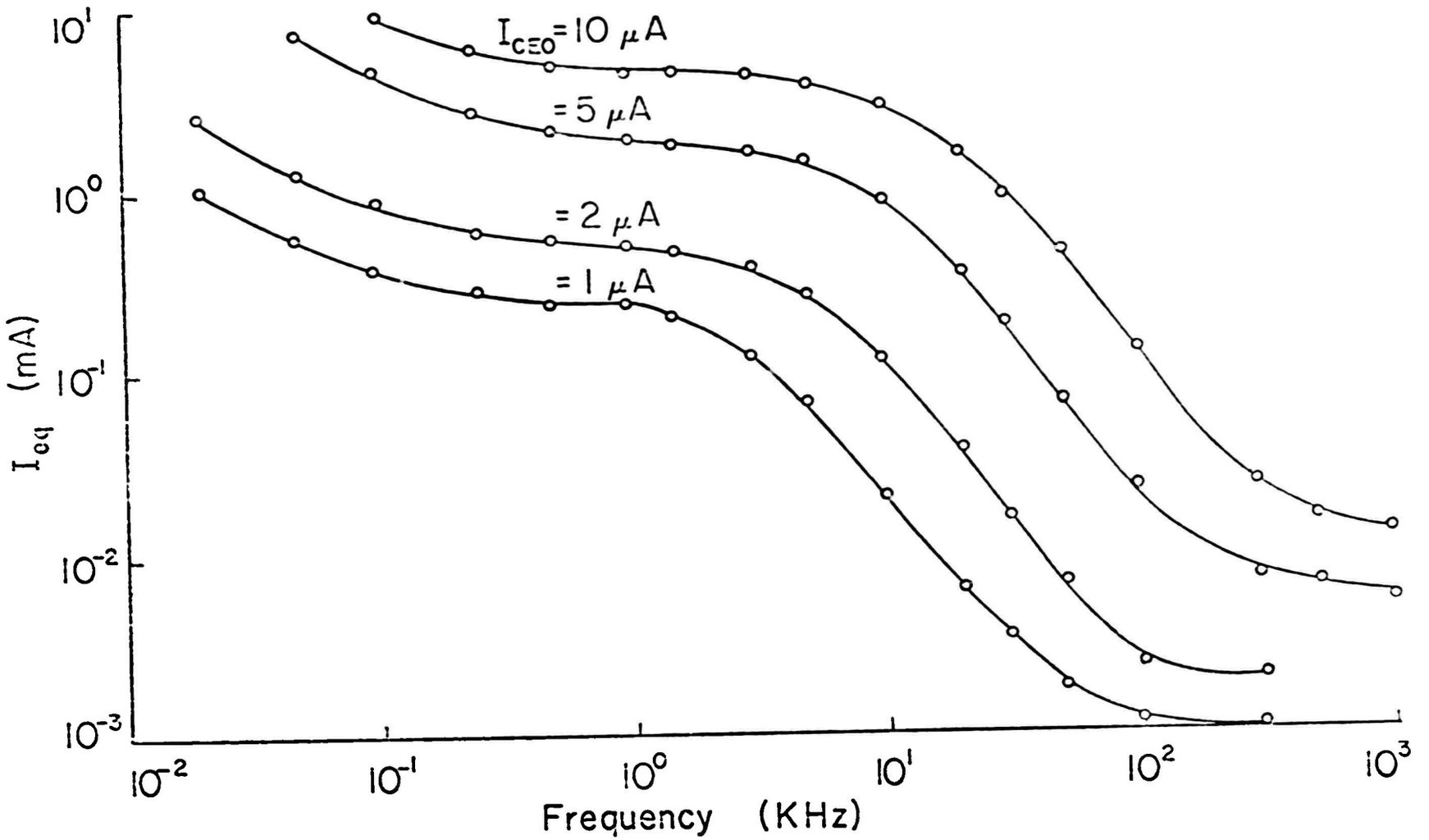


Figure 2a Noise spectra with bias current as a parameter, MOTOROLA MRD-5.

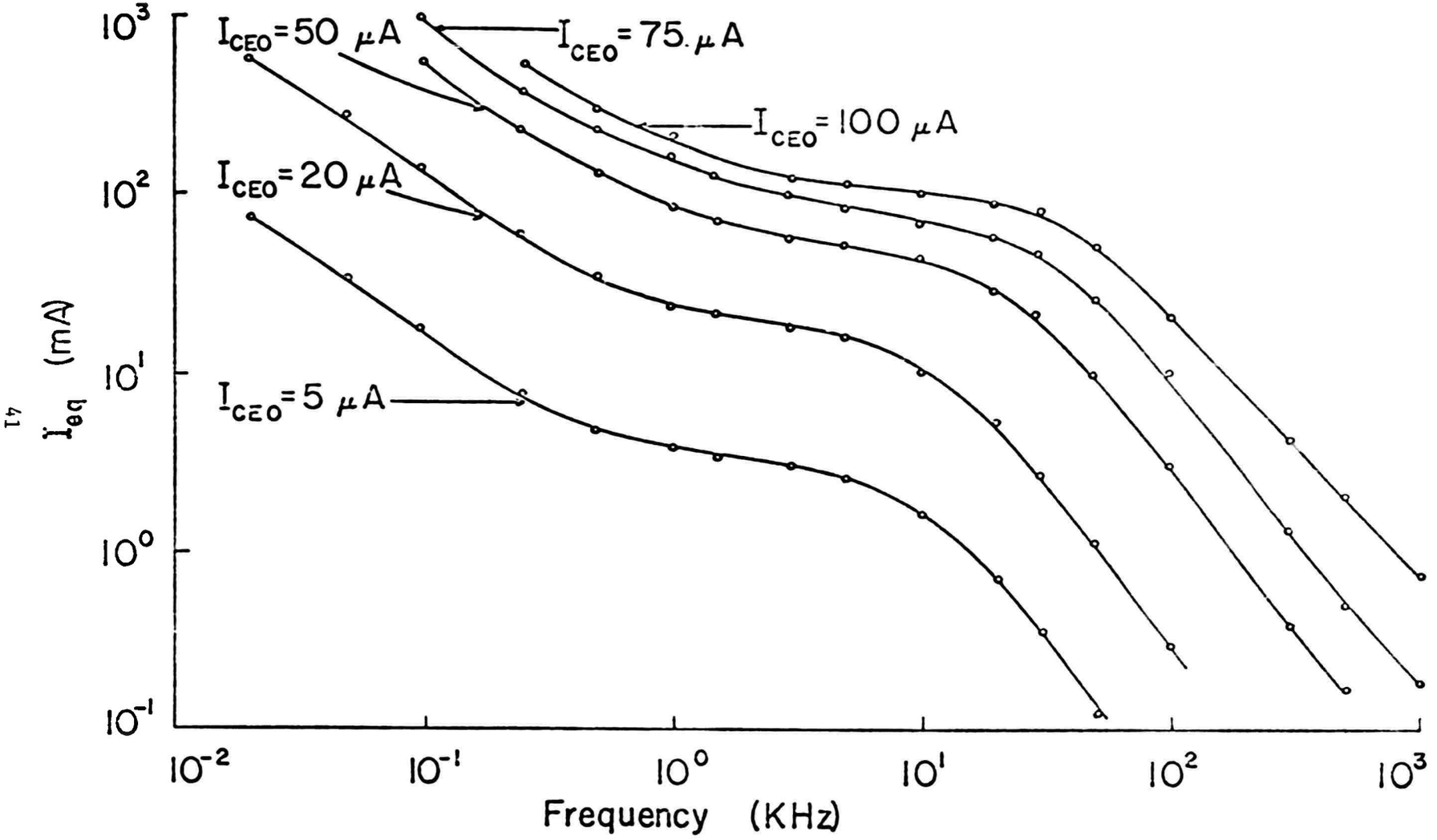


Figure 2b Noise spectra with d.c. bias current as a parameter, University of Florida phototransistor.

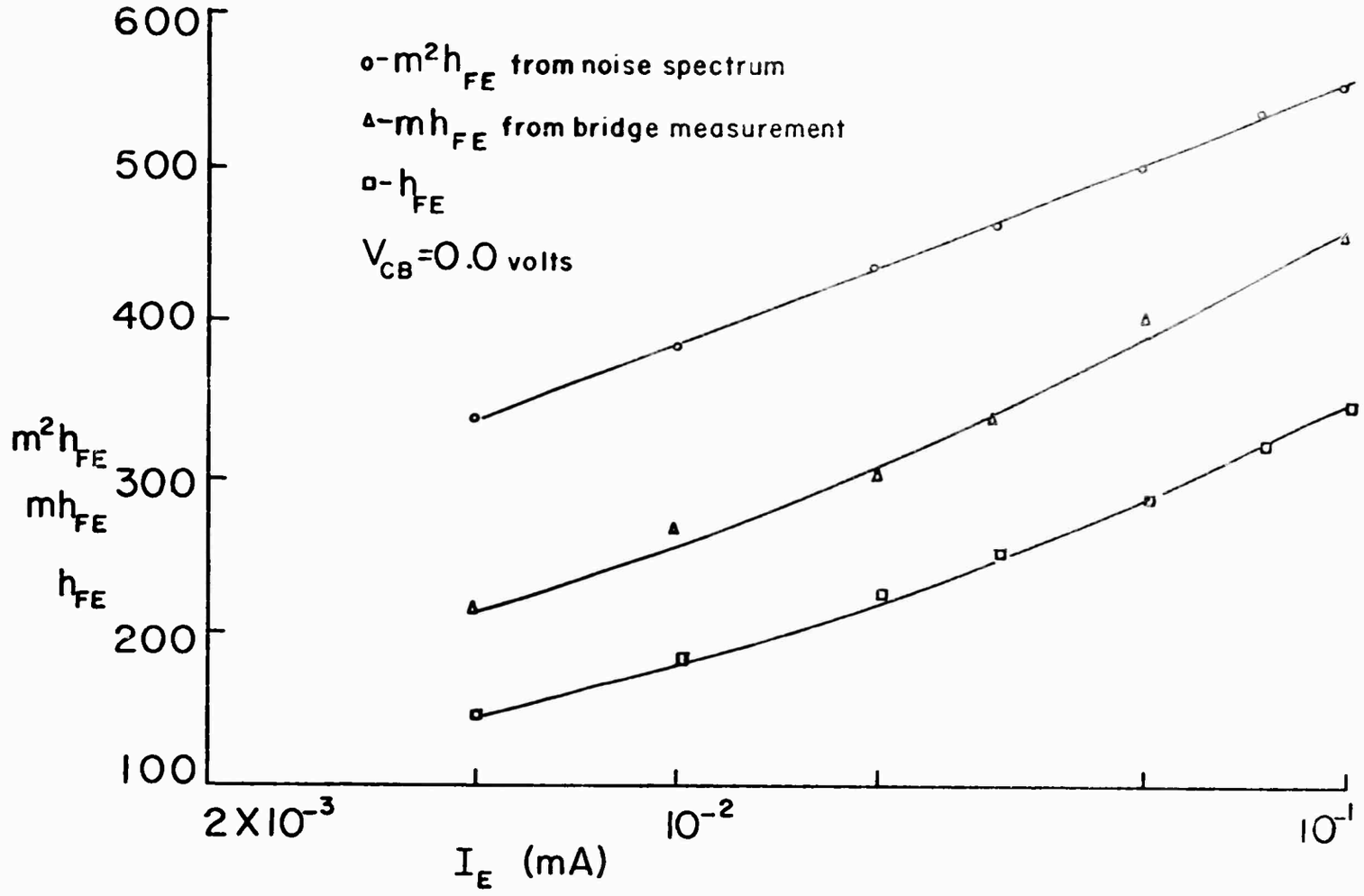


Figure 3 Comparison of h_{FE} as a function of bias current for device of Fig. 2a. m^2h_{FE} from noise measurements is compared with mh_{FE} from bridge measurements, and with h_{FE} from d.c. measurements.

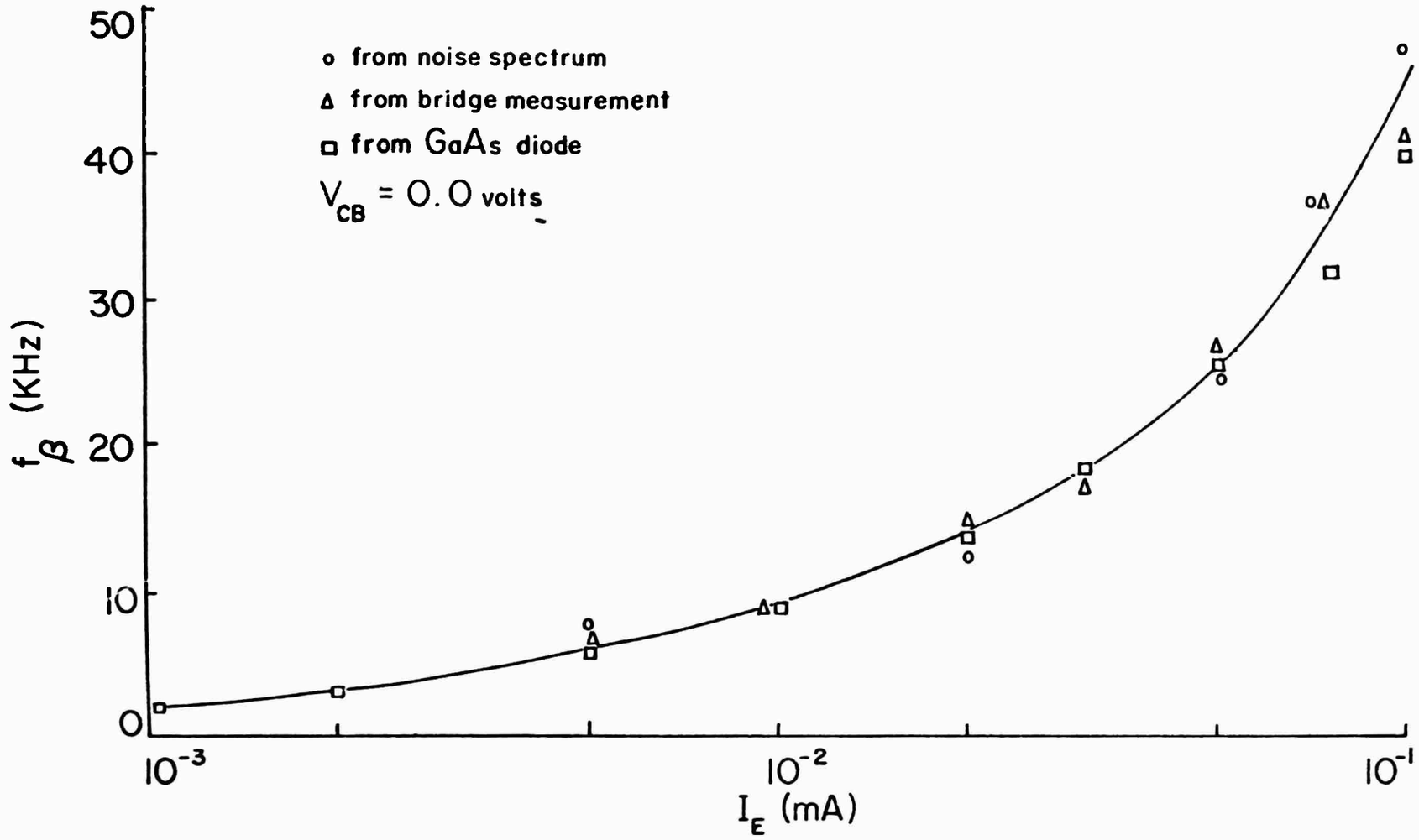
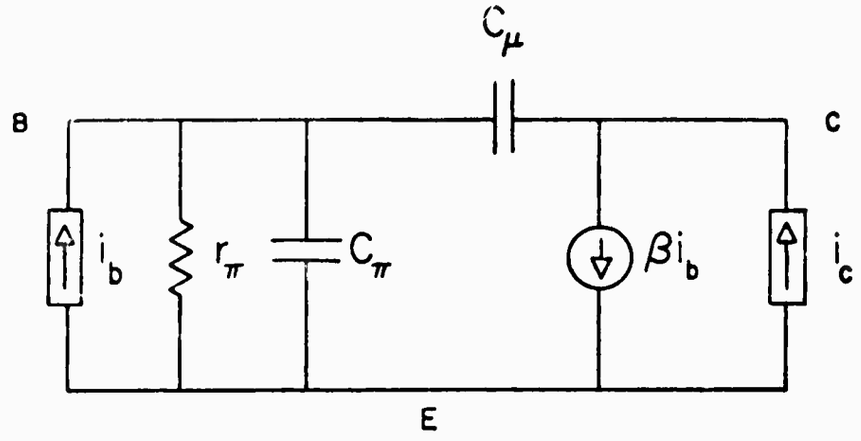


Figure 4 Comparison of f_β as a function of bias current for device of Fig. 2a. Values of f_β obtained from noise, and small signal measurements are compared with those obtained from the frequency response of phototransistor to an amplitude modulated GaAs light diode.



$$\overline{i_c^2} = 2q\left[\alpha_F I_D \exp(qV_{BE}/kT) + (I_{BC} + I_{PH})\right]\Delta f$$

$$\overline{i_b^2} = 2q\left[2(I_{BC} + I_{PH})\right]\Delta f$$

$$\overline{i_c i_b^*} = 2q(I_{BC} + I_{PH})\Delta f$$

Figure 5 Phototransistor noise model. Shot noise current generators i_b and i_c are imposed on a "Hybrid-pi" small signal equivalent circuit.

C. IMPURITY CONCENTRATION DEPENDENT DENSITY OF STATES AT THE FERMI LEVEL FOR SILICON (D. D. Kleppinger and F. A. Lindholm)

I. INTRODUCTION

Experimental results, ranging from a variable impurity ionization energy in bulk semiconductors[1] to an anomalous temperature sensitivity of transistor gain[2,3], indicate that the density of quantum states in a semiconductor depends on the impurity concentration. The conventional engineering theory of semiconductors[4,5,6], does not include this dependence. In a recent paper[7], however, we proposed an approach that enabled its inclusion over the range of impurity concentration that is of technological interest. Application of the approach to zinc doped (p-type) gallium arsenide, as an example, yielded theoretical predictions that compared satisfactorily to an experimentally determined density of states. The current paper continues to explore the applications of the approach by studying its predictions for silicon.

The discussion proceeds as follows. Section II summarizes the approach. Section III deals with the calculated density of states in phosphorus doped silicon and discusses the observed variation of the impurity ionization energy with impurity concentration. Furthermore, this section discusses the consistency between the predicted density of states in arsenic doped silicon and the existence of three empirically observed modes of conduction. The dependence of the Fermi level on temperature and impurity concentration for phosphorus doped silicon are discussed in Section IV. Section V indicates applications of the results reported here.

II. A SUMMARY OF THE APPROACH

Two kinds of evidence exist to indicate the impurity concentration dependence of the density of states: that relating to its effects in bulk semiconductor material[1,8-10], and that illustrating its manifestations in the properties of semiconductor junction devices[2,3,11]. To describe this dependence in silicon, we shall use an approach described in detail in a recent paper by the authors[7]. Here we provide a brief summary of the approach.

Qualitatively, one can explain as follows the origin of the concentration dependence of the density of states. The impurity concentration affects both the density of states associated with the host lattice and the density of states associated with the impurity atoms. As the impurity concentration increases,

the electrons* donated to the crystal become so numerous that their presence essentially alters the periodic potential in the crystal. This causes the density of states in the host lattice bands to deviate from the conventional dependence on the square root of energy. Both the valence and conduction bands are deformed by this mechanism. Additionally, the states associated with the impurity atoms split and form a quasi-continuous band of energies as the impurity concentration increases. This splitting of states occurs because of the increasing interaction between impurity atoms as the impurity concentration increases (or the inter-impurity distance decreases).

The equations used in the approach to describe the above mechanisms come from the works of Bonch-Bruyevich[12] and Morgan[13]. Two equations due to Bonch-Bruyevich are used to describe the density of states in the host lattice bands. For energies near the intrinsic band edge, E = 0 ev,

$$\rho(E) = 1.309 \times 10^{17}(m^*/m)^{11/8}\ \varepsilon^{3/8}\ N^{5/24} + 6.267 \times 10^{25}\ \varepsilon^{3/8}(m^*/m)^{13/8}\ N^{-5/24}\ E, \quad (1)$$

and

$$\rho(E) = 6.8125 \times 10^{21}(m^*/m)^{3/2}\ E^{1/2} - 4.065 \times 10^{3}(m^*/m)\ \varepsilon^{-3/2}\ N^{5/6}\ E^{-3/2}, \quad (2)$$

at energies farther into the allowed band. The density of states associated with the impurities is described by Morgan:

$$\rho(E) = 3.651 \times 10^{8}\varepsilon^{3/4}\left(\frac{m^*}{m}\right)^{1/4} N^{.5833}\ \exp\left[\frac{-(E-E_D)^2\varepsilon^{3/2}(m^*/m)^{1/2}}{9.551 \times 10^{-18}\ N^{.8334}}\right] . \quad (3)$$

Although the separate theories of Bonch-Bruyevich and Morgan were derived for restricted ranges of impurity concentration, we made and justified[7] the approximation that the two theories could be superposed and applied over the range of impurity concentration of technological interest. The next section presents the density of states obtained for n-type silicon by applying this approach.

III. THE DENSITY OF QUANTUM STATES IN SILICON

A. Phosphorus doped silicon

Equations (1) through (3) will give the density of states for silicon if the appropriate material parameters are used: $m^* = 1.08\ m$[14] and $\varepsilon = 11.7$[15]. Figures 1 through 8 display the predicted density of quantum states as a func-

* For concreteness, we will consider n-type material throughout the paper.

tion of energy in n-type (phosphorus doped) silicon at various impurity concentrations in the range $10^{16}cm^{-3}$ through $3 \times 10^{20}cm^{-3}$. The conventional, parabolic curve is plotted on each figure as a reference.

The significant features of the sequence of graphs are the growth of the impurity band with its eventual overlapping of the conduction band and the alteration of the conduction band itself.

At an impurity concentration of $10^{16}cm^{-3}$(see Fig. 1), the impurity states are described by a Gaussian whose standard deviation is 0.001575 ev. This spread in energy is sufficiently small that the impurity states may be considered well localized in energy. The description of the conduction band states, afforded by Bonch-Bruyevich's equations, is essentially coincident with the parabolic reference curve. Thus the usual conventional assumptions of a parabolic conduction band and a delta function distribution of impurity states are valid for an impurity concentration of $10^{16}cm^{-3}$.

An examination of Figs. 2 and 3, for concentrations of $10^{17}cm^{-3}$ and $10^{18}cm^{-3}$, reveals the growth in height and width of the Gaussian curve describing the impurity states. The density of states in the conduction band is beginning to deviate from the simple, parabolic form.

At an impurity concentration of $3 \times 10^{18}cm^{-3}$ (see Fig. 4), the impurity band and the conduction band have overlapped to a significant extent. The effects of this overlapping will be discussed later in this section.

As the impurity concentration increases from $10^{19}cm^{-3}$ through $3 \times 10^{20}cm^{-3}$, the two bands merge even more and tend to become indistinguishable. Only very small structural features occur to give evidence of the origin of the states as they exist at these doping levels. This explains, perhaps, why experiments often fail to detect structure in the density of states[16].

Figures 1 through 8, make evident that the conventional assumptions of a parabolic conduction band and a delta function of impurity states, both independent of impurity concentration, fail at the higher concentrations. We now compare some of the experimental evidence for this failure to the predictions of the approach for the density of states in silicon.

As was noted above, an impurity concentration of approximately $3 \times 10^{18}cm^{-3}$ phosphorus atoms is just large enough to cause the impurity and conduction bands to overlap. There is experimental evidence to support this prediction.

Pearson and Bardeen[1] found that the ionization energy of phosphorus doped silicon monotonically decreases toward zero with increasing impurity concen-

tration and becomes zero at a concentration of approximately $1.25 \times 10^{18} cm^{-3}$. This dependence on the impurity concentration is consistent with the growth of the impurity band and the distortion of the conduction band as shown in Figs. 1 through 8. The predicted overlapping at a concentration of approximately $3 \times 10^{18} cm^{-3}$ is consistent with Pearson and Bardeen's value of $1.25 \times 10^{18} cm^{-3}$ for the vanishing of the ionization energy because, when the density of states functions overlap, the energy gap designated as the ionization energy disappears.

B. Conduction Modes in Arsenic Doped Silicon

Swartz[10] has studied the resistivity of n- and p-type silicon samples containing impurities in the concentration range of approximately $10^{17} cm^{-3}$ to $3 \times 10^{18} cm^{-3}$. One subset of his data is particularly useful in lending validity to our theoretical predictions. This subset consists of resistivity measurements on four arsenic doped samples whose doping levels were $5.75 \times 10^{17} cm^{-3}$, $9.13 \times 10^{17} cm^{-3}$, $1.56 \times 10^{18} cm^{-3}$, and $2.96 \times 10^{18} cm^{-3}$. These doping levels correspond to the moderate concentration range for which the impurity atoms interact sufficiently that there is a small but finite probability of transferring electrons between impurity states. At temperatures near 0°K, electrical conduction occurs by "hopping" between impurity states. However, at the upper end of this moderate concentration range a further possibility for conduction arises. Because the impurities assume random, substitutional positions in the host lattice, spatial paths of relatively large impurity concentrations may exist. Along these paths, threadlike energy bands develop and enable conduction in a partially filled band.

Let us now discuss Swartz's resistivity data in some detail and show the consistency between the data, the calculated density of states, and our qualitative expectations for the conduction modes.

Swartz's study of the dependence of the resistivity on reciprocal temperature shows three distinct types of behavior. Consider first the sample doped with $1.56 \times 10^{18} cm^{-3}$ atoms. This sample displays all three types of behavior. Above 20°K, the resistivity varies exponentially with 1/T at a rate expected for normal conduction. Between 10°K and 20°K, it varies exponentially but at a slower rate than that found for temperatures above 20°K. The third type of behavior occurs below 10°K and is characterized by a power law relationshop between resistivity and 1/T. Swartz identifies the conduction mechanism occurring between 10°K and 20°K as the hopping of carriers from one impurity

atom to another. Such a hopping process is characterized by an activation energy. This is consistent with the exponential relationship between the resistivity and 1/T in this temperature range. He describes the hopping mechanism as a non-band conduction mode to distinguish it from the mode found below 10°K. This latter mode he calls impurity band conduction. In this lowest temperature range, the temperature dependence of the resistivity is moderate compared to the exponential dependence in the other ranges. This suggests that conduction is occurring in a partially filled band of states containing a number of electrons which is constant and independent of temperature in this range.

The sample containing $9.13 \times 10^{17} cm^{-3}$ of arsenic atoms definitely displays the normal and hopping mechanisms and appears to be on the verge of entering the impurity band mode below approximately 10°K. The least heavily doped sample ($5.75 \times 10^{17} cm^{-3}$) does not enter the impurity band mode, although it does display hopping conduction. Finally, the most heavily doped sample ($2.96 \times 10^{18} cm^{-3}$) does not display the non-band mechansim but rather passes directly from the normal mechanism to the impurity band mechansim.

The final observation needed from Swartz's data is that the resistivity at any temperature decreases with increasing impurity content. Now we are able to analyze his data in terms of the density of states we have calculated.

By plotting the $\pm 2\sigma$ limits from Morgan's theory, we show, in Fig. 9, the development of the impurity band with increasing impurity concentration in arsenic doped silicon. The energy, E_0, at which the density of conduction band states predicted by Bonch-Bruyevich's equation, Eq. 1, vanishes is also plotted. Again, the merger of the conduction and impurity bands occurs at a concentration of approximately $3 \times 10^{18} cm^{-3}$. The vertical arrows on the figure indicate the doping levels of Swartz's samples.

Examine first the sample containing an impurity concentration of $1.56 \times 10^{18} cm^{-3}$. Figure 9 shows that the impurity and conduction bands have not merged significantly at this concentration. The existence of the normal conduction mode requires no explanation. The existence of the other two conduction modes in this sample can be explained as follows. Between 10°K and 20°K, the thermal environment is not able to excite many electrons from the impurity states to the conduction band; but it can provide sufficient energy to some electrons to allow them to participate in phonon activated hopping

between impurity atoms. Below 10°K, however, there is an energetically more favorable and therefore more probable conduction mechanism available in this sample. Conduction below 10°K occurs along the spatially localized paths of relatively higher impurity concentration where thread-like impurity bands have developed.

For Swartz's most heavily doped sample ($2.96 \times 10^{18} cm^{-3}$), the resistivity data indicate that the impurity atoms apparently interact to such an extent that impurity hopping never dominates and conduction occurs, at low temperatures, in a partially filled band. Figure 9 shows that the impurity band is wider at $2.96 \times 10^{18} cm^{-3}$ than at $1.56 \times 10^{18} cm^{-3}$. It is consistent, therefore, to expect that the impurity band has developed more generally throughout the material in the more heavily doped sample. With such a band available, we would expect it to determine the low temperature conduction mechanism.

Similarly, the development of a small indication of impurity band conduction is consistent with the spreading of the impurity band as the impurity concentration increases from $5.75 \times 10^{17} cm^{-3}$ to $9.12 \times 10^{17} cm^{-3}$. This spreading is illustrated in Fig. 9.

The resistivity of the least heavily doped sample exceeds that of the most heavily doped sample because its carriers are never able to enter the energetically favorable impurity band mode of conduction. Fewer and fewer carriers are able to participate in hopping conduction as a decrease in thermal energy makes this process less probable. Thus, the resistivity increases faster with decreasing temperature than for the impurity band mode which requires no activation energy.

In addition to Swartz, other workers have considered the transitions between the three modes of conduction mentioned. Miller and Abrahams[17] state that hopping will occur in n-type silicon doped at a level below approximately $2 \times 10^{17} cm^{-3}$. Note that the concentration limit for hopping will depend on the particular impurity atom used. An impurity with a higher ionization energy (at small concentrations) will require a larger concentration to leave the hopping mode than will an impurity with a lower ionization energy. This occurs because the wave functions of the former will be more localized and will require a smaller separation between atoms to interact and form bands. Thus, phosphorus doped silicon would be expected to leave the hopping mode and enter the impurity band mode of conduction at a lower concentration than arsenic doped silicon. The statement by Miller and Abrahams is thus qualitatively

consistent with Swartz's data. Longo et al.[18] state that the transition between the hopping and impurity band modes of conduction occurs between $10^{18}cm^{-3}$ and $5 \times 10^{18}cm^{-3}$ boron atoms in silicon. This is also consistent.

This completes the interpretation of the empirically observed ionization energies and conduction modes in terms of the predictions of our calculated density of quantum states in silicon. In the next section, we examine the results of applying the approach to finding the Fermi level as a function of impurity concentration and temperature. The material studied is phosphorus doped silicon.

IV. THE FERMI LEVEL IN SILICON

The Fermi level of a system of electrons can be obtained by integrating the product of the density of states and Fermi-Dirac distribution functions over all energies and then requiring that this integral equal the total density of electrons actually present in the system. We carried out this procedure for phosphorus doped silicon using the density of states functions given by Eqs. (1) through (3). The calculation did not consider compensating impurities and assumed that only the electrons introduced by the donor impurity atoms (in excess of their core electrons and the number required to satisfy lattice bonds) would be considered as the system. In other words, only the "fifth" or excess electrons of the impurities were considered. Intrinsic carriers were also neglected. The intrinsic carrier pairs may be neglected in the ranges of impurity concentration ($10^{15}cm^{-3}$ to $3 \times 10^{20}cm^{-3}$) and temperature (0°K to 300°K) of interest because their density is always at least a factor of 10^{-5} less than the density of excess electrons introduced by the impurities.

Note that by solving for the Fermi level in this way we do not need to specify whether the excess electrons are bound to the impurity atoms or are free. This is especially helpful at the higher concentrations where the distinction between the impurity and host lattice states is unclear.

The results of the procedure are presented in Fig. 10 which shows the Fermi level as a function of temperature at several impurity concentrations.

The Fermi level at 0°K occurs at the conventional ionization energy for the plotted impurity concentrations of $10^{15}cm^{-3}$ through $10^{19}cm^{-3}$. However, because of the significant deformation of the conduction band at a concentration of $10^{20}cm^{-3}$, a large density of states exists below the conventional ionization energy. This causes the Fermi level at 0°K to occur below the

conventional ionization energy for this doping.

The temperature dependence of the Fermi level is greatest for the material with the lowest impurity concentration. At the highest concentration, $10^{20}cm^{-3}$, the Fermi level is nearly temperature insensitive, changing by only 0.008 ev from 0°K to 300°K.

For uncompensated semiconductors with light doping, conventional theory[14] predicts that the Fermi level tends toward the energy $-E_D/2$ as the temperature decreases toward 0°K. This prediction appears to conflict with our calculation for an impurity concentration of $10^{15}cm^{-3}$, which as shown in Fig. 10 yields that the Fermi level at 0°K tends toward $-E_D$. The conflict arises because in the conventional calculation one describes the impurity states as a delta function in energy of strength N. For truly localized impurity states, this description is valid; the effective density of the impurity states indeed equals the density of impurity atoms, despite the effects of spin degeneracy, because once an impurity atom captures an electron in one of its two states in the forbidden band the other state becomes unavailable. Morgan's result, given in (3), however, describes a Gaussian-shaped impurity band that contains 2N available states[7]. When an impurity band exists, the factor of 2 must be included to account properly for spin degeneracy. Thus, in the limit of dilute impurity concentration, the number of available states contained in Morgan's result fails to merge with the number appropriate for truly localized impurity states. This failure causes the disagreement cited above, for low impurity concentrations, between the position of the Fermi level at 0°K given by conventional calculations and that shown in Fig. 10. In an actual crystal with dilute doping, the location of the Fermi level relative to the edge of the host lattice band is a mixture of these two results. Randomness in the location of the impurity atoms produces regions of relatively high doping, within which the Fermi level resides approximately at energy $-E_D$, and regions of relatively low doping, within which the Fermi level is located at approximately energy $-E_D/2$.

Figure 11 shows the Fermi level as a function of impurity concentration at 300°K. The additional curves plotted on Fig. 11 depict the growth of the impurity band and the deformation of the conduction band, each of which occurs with increasing impurity concentration. The former is represented by plotting the $\pm 2\sigma$ limits of the describing Gaussian function, and the latter is shown by the energy, E_0, at which Bonch-Bruyevich's density of states vanishes.

The combination of curves shows that the Fermi level, at 300°K, enters a significant density of allowed states at a concentration of approximately

$3 \times 10^{18} cm^{-3}$. Note however, that, although the Fermi level enters a region of allowed states, it does not rise above the conduction band edge in the conventional sense. In the conventional theory, the Fermi level enters the conduction band at approximately $4 \times 10^{20} cm^{-3}$ when the parabolic density of states is assumed for the conduction band and Fermi-Dirac statistics are used[20].

It is known that the Fermi level lies within a region of allowed states when tunneling is possible. Since the Fermi level shown in Fig. 11 does fall within a region of allowed states at high concentrations, it is consistent with the observed existence of tunneling, as in tunnel diodes.

V. DISCUSSION

By applying an approach that accounts for the impurity concentration dependence of the density of states, we have calculated the density of states and the Fermi level for the technologically important example of silicon. The results are in agreement with experimental observations. The approach used is by no means restricted to silicon, but it can be applied to other semiconductors.

The concentration dependent density of states and Fermi level calculated in this paper, while extending the basic description of semiconductor material, furnish also the means for calculating various aspects of semiconductor device behavior. For example, it relates directly to device behavior associated with interband transitions and thus to the temperature dependence of device characteristics, to the behavior of semiconductor detectors, etc. As a tool for the design and analysis of pn junction devices, however, the theory presented here is overly complex. The complexity arises from the need to simultaneously include Fermi-Dirac statistics and the complicated relationship between the density of states and the impurity concentration. A tractable theory can be obtained by considering separately the statistics and the density of states and then combining the results. The conventional approach considers Maxwell-Boltzmann statistics and concentration independent density of states functions in order to develop a tractable theory. An extension of this theory can be made that includes Fermi-Dirac statistics while retaining the conventional density of states functions. This extended theory retains tractability by using a mathematical approximation (as opposed to a physical approximation like the Maxwell-Boltzmann statistics) to the Fermi-Dirac statistics or, more precisely, to the Fermi-Dirac integral. The final step then is to include, by approximation, the essence of the complicated impurity concentration dependence of the density of states. This yields a tractable theory suitable for the study of pn junction devices. We shall describe this theory in a future paper.

SYMBOL LIST

The number in parentheses following the definition indicates the equation in which the symbol is first found.

E	a general energy (ev) (1)
E_0	the energy at which $\rho(E=E_0) = 0$ from Eq. (1) (ev) (used in discussing Fig. 9)
E_D	the energy of the donor level in the forbidden gap (ev) (3)
E_F	the Fermi energy (ev)
m	free electron rest mass (g) (1)
m^*	the density of states effective mass (g) (1)
N	impurity atom concentration (cm^{-3}) (1)
T	the temperature (°K)
ε	the relative dielectric constant (1)
ρ	the density of quantum states ($ev^{-1}cm^{-3}$) (1)
σ	the standard deviation of the Gaussian function describing the impurity states (ev) (used in discussing Fig. 9).

REFERENCES

1. G. L. Pearson and J. Bardeen, Phys. Rev. 75, 865 (1949).
2. W. L. Kauffman and A. A. Bergh, IEEE Trans. on Elec. Devices ED-15, 732 (1968).
3. D. Buhanan, IEEE Trans. on Elec. Devices ED-16, 117 (1969).
4. R. D. Middlebrook, An Introduction to Junction Transistor Theory. Wiley, New York (1957).
5. A. K. Jonscher, Principles of Semiconductor Device Operation. Wiley, New York (1960).
6. A. B. Phillips, Transistor Engineering and Introduction to Integrated Semiconductor Circuits. McGraw-Hill, New York (1962).
7. D. D. Kleppinger and F. A. Lindholm, Solid State Elec., in press.
8. P. P. Debye and E. M. Conwell, Phys. Rev. 93, 693 (1954).
9. E. M. Conwell, Phys. Rev. 103, 51 (1956).
10. G. A. Swartz, J. Phys. Chem. Solids 12, 245 (1960).
11. E. O. Kane, Phys. Rev. 131, 79 (1963).
12. V. L. Bonch-Bruyevich, The Electronic Theory of Heavily Doped Semiconductors. American Elsevier, New York (1966).
13. T. N. Morgan, Phys. Rev. 139, A343 (1965).
14. J. S. Blakemore, Semiconductor Statistics. Macmillan, New York (1962).
15. A. S. Grove, Physics and Technology of Semiconductor Devices. Wiley, New York (1967).
16. See for example: F. D. Shepherd, Jr., "A Study of Band Edge Distortion in Heavily Doped Germanium," Air Force Cambridge Research Laboratory Physical Sciences Research Papers, No. 150.
17. A. Miller and E. Abrahams, Phys. Rev. 120, 745 (1960).
18. T. A. Longo, R. K. Ray, and K. Lark-Horovitz, J. Phys. Chem. Solids 8, 259 (1959).
19. F. Reif, Fundamentals of Statistical and Thermal Physics. McGraw-Hill, New York (1965).
20. D. D. Kleppinger and F. A. Lindholm, to be published.

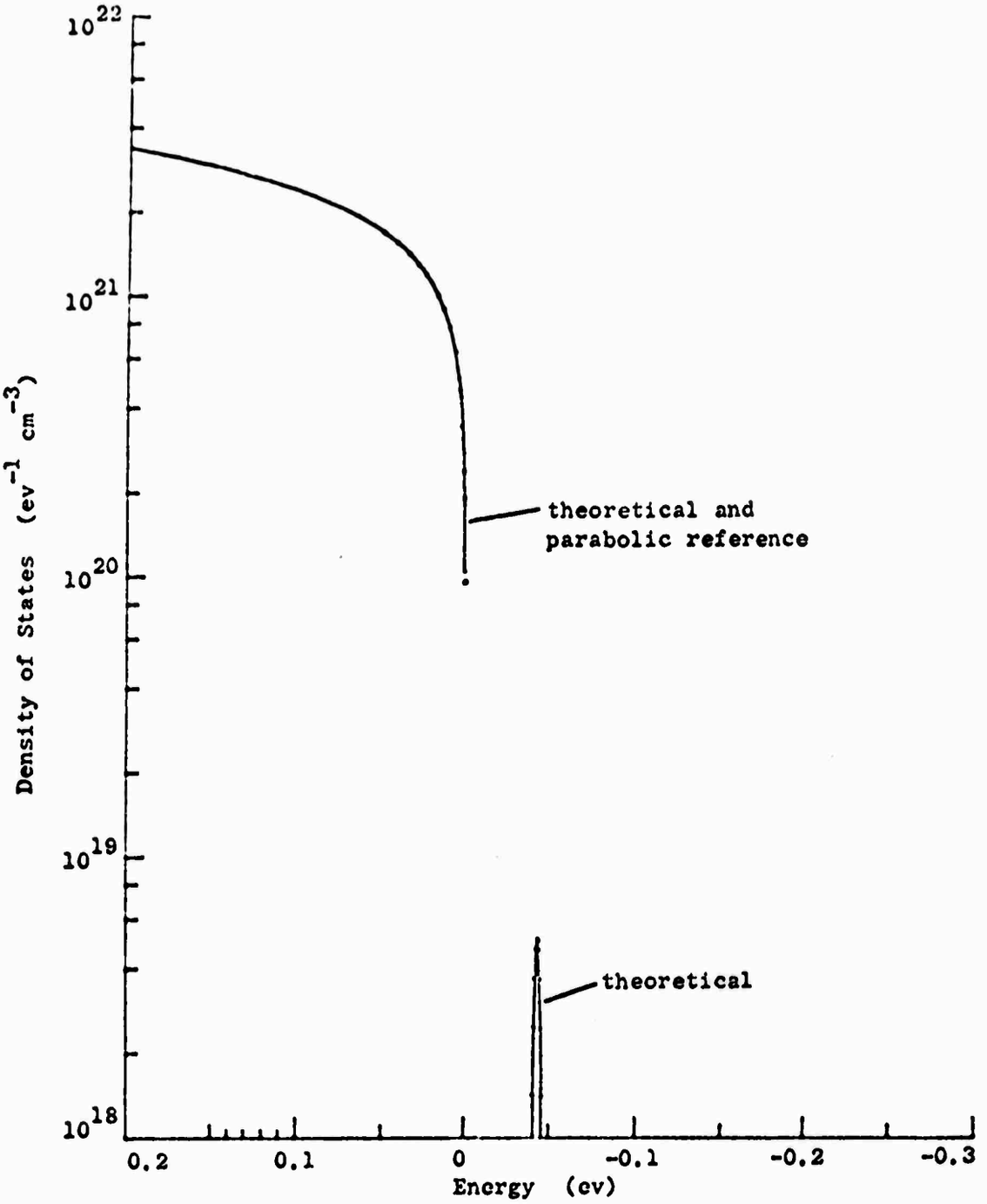


Figure 1. The density of quantum states in silicon doped with $1x10^{16}$ cm^{-3} of phosphorus atoms.

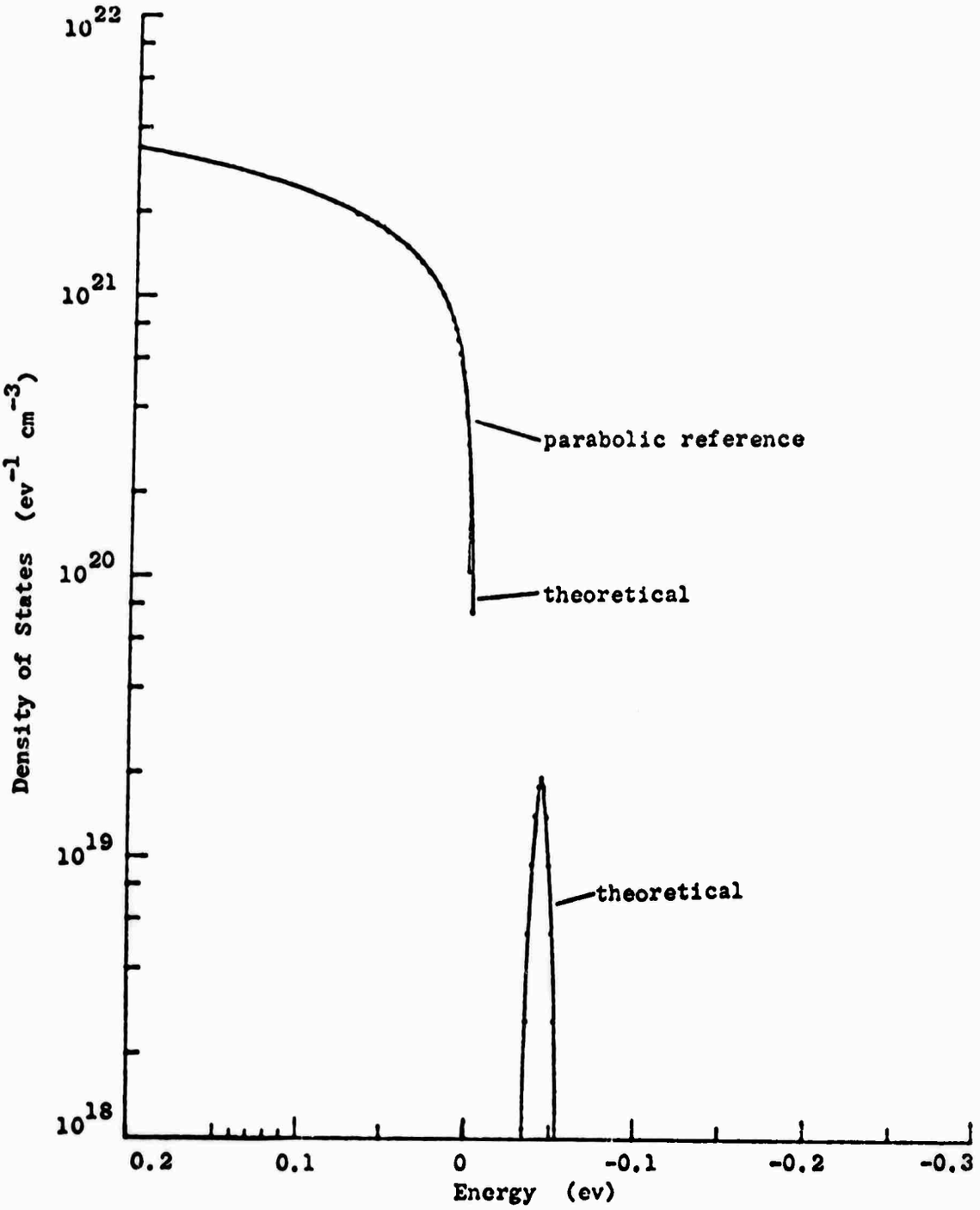


Figure 2. The density of quantum states in silicon doped with $1x10^{17}$ cm^{-3} of phosphorus atoms.

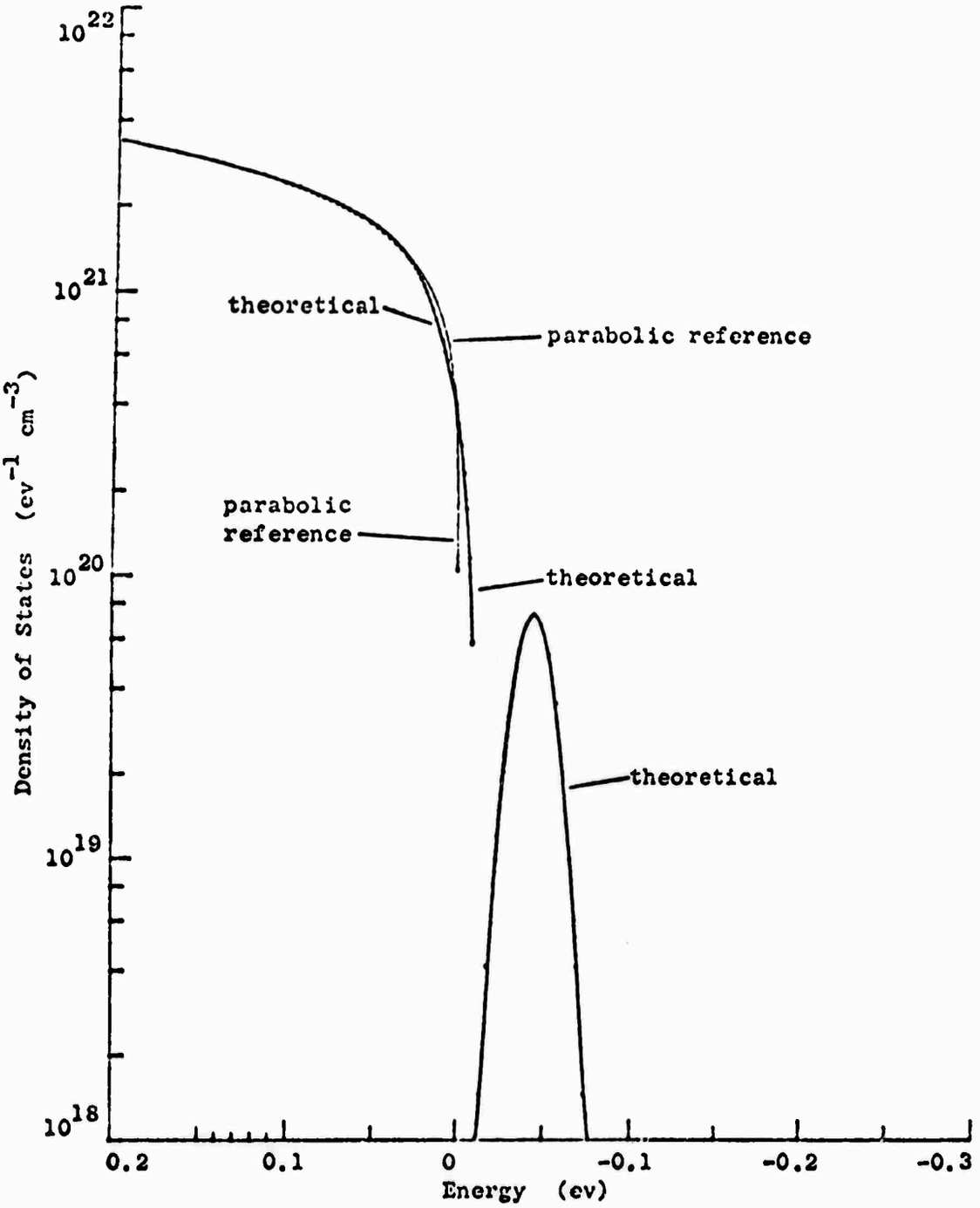


Figure 3. The density of quantum states in silicon doped with $1x10^{18}$ cm^{-3} of phosphorus atoms.

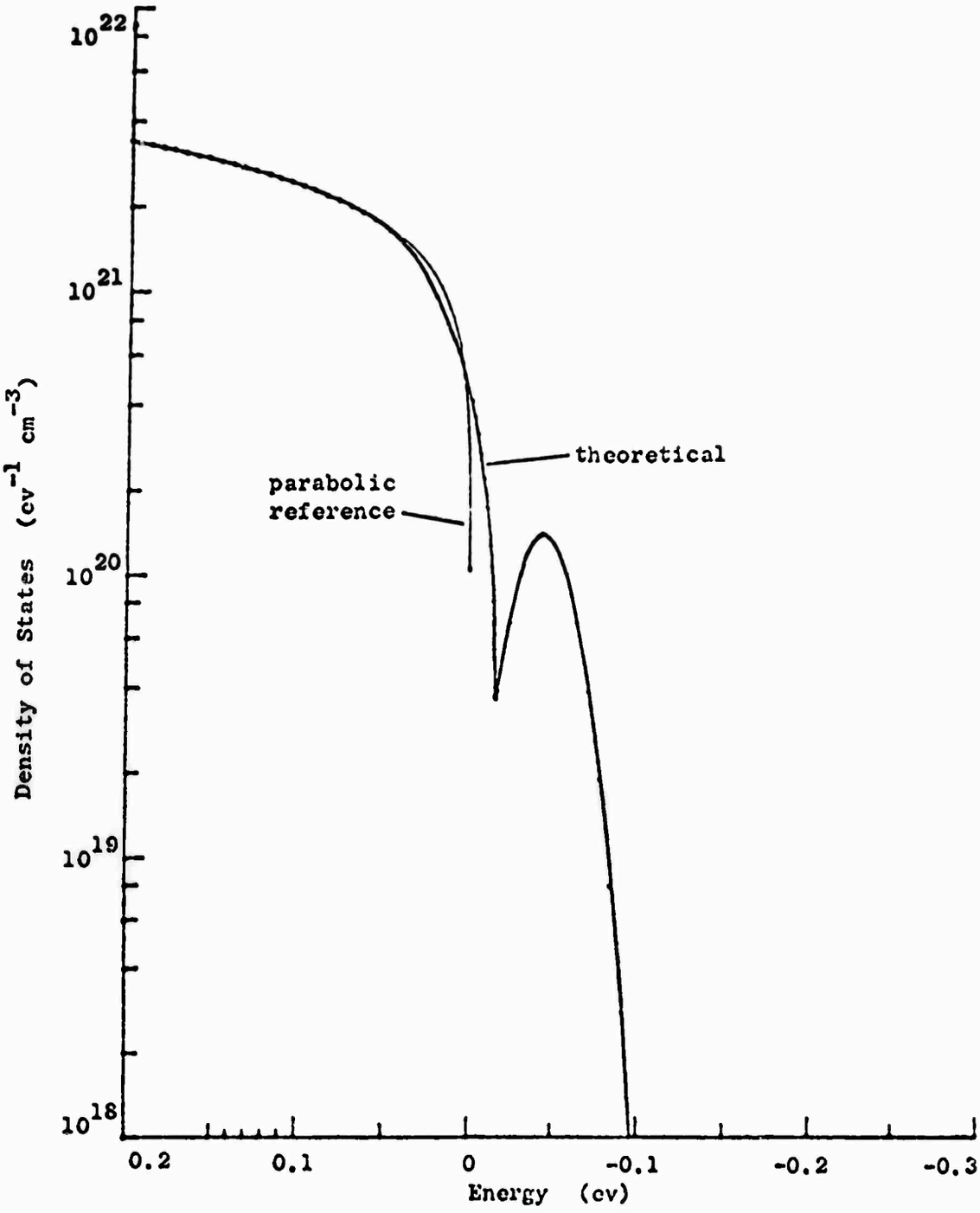


Figure 4. The density of quantum states in silicon doped with 3×10^{18} cm^{-3} of phosphorus atoms.

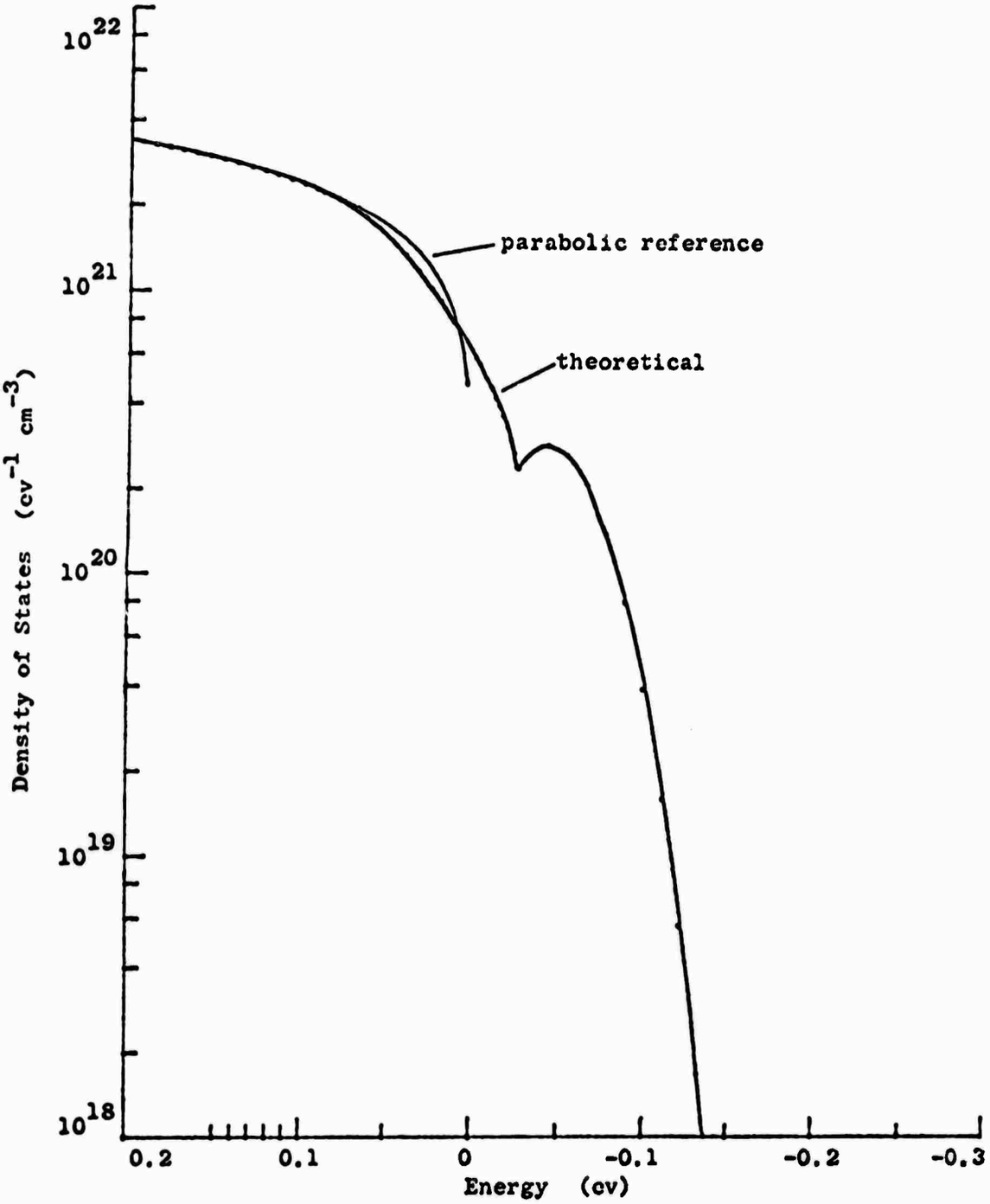


Figure 5. The density of quantum states in silicon doped with $1x10^{19}$ cm^{-3} of phosphorus atoms.

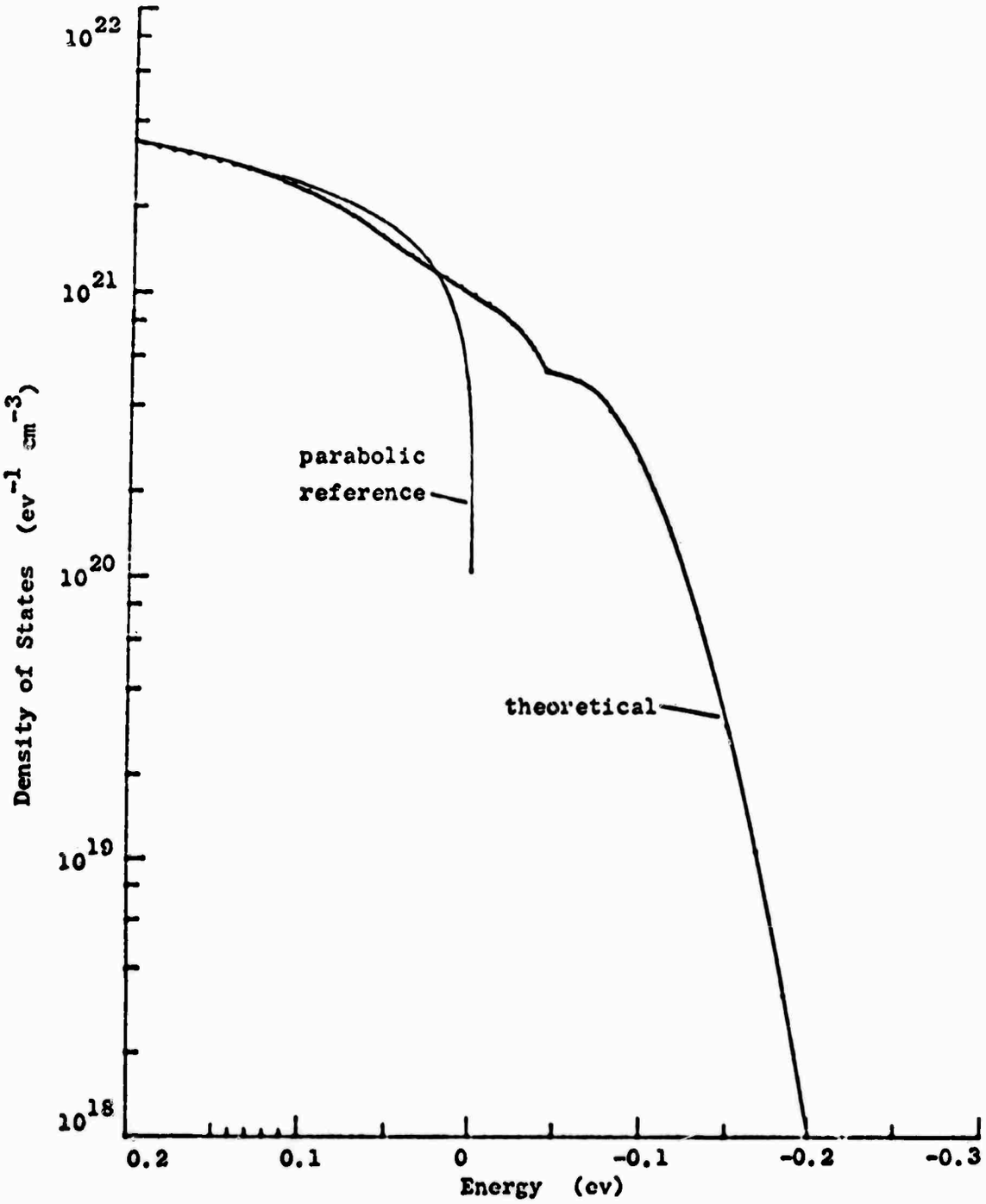


Figure 6. - The density of quantum states in silicon doped with 3×10^{19} cm^{-3} of phosphorus atoms.

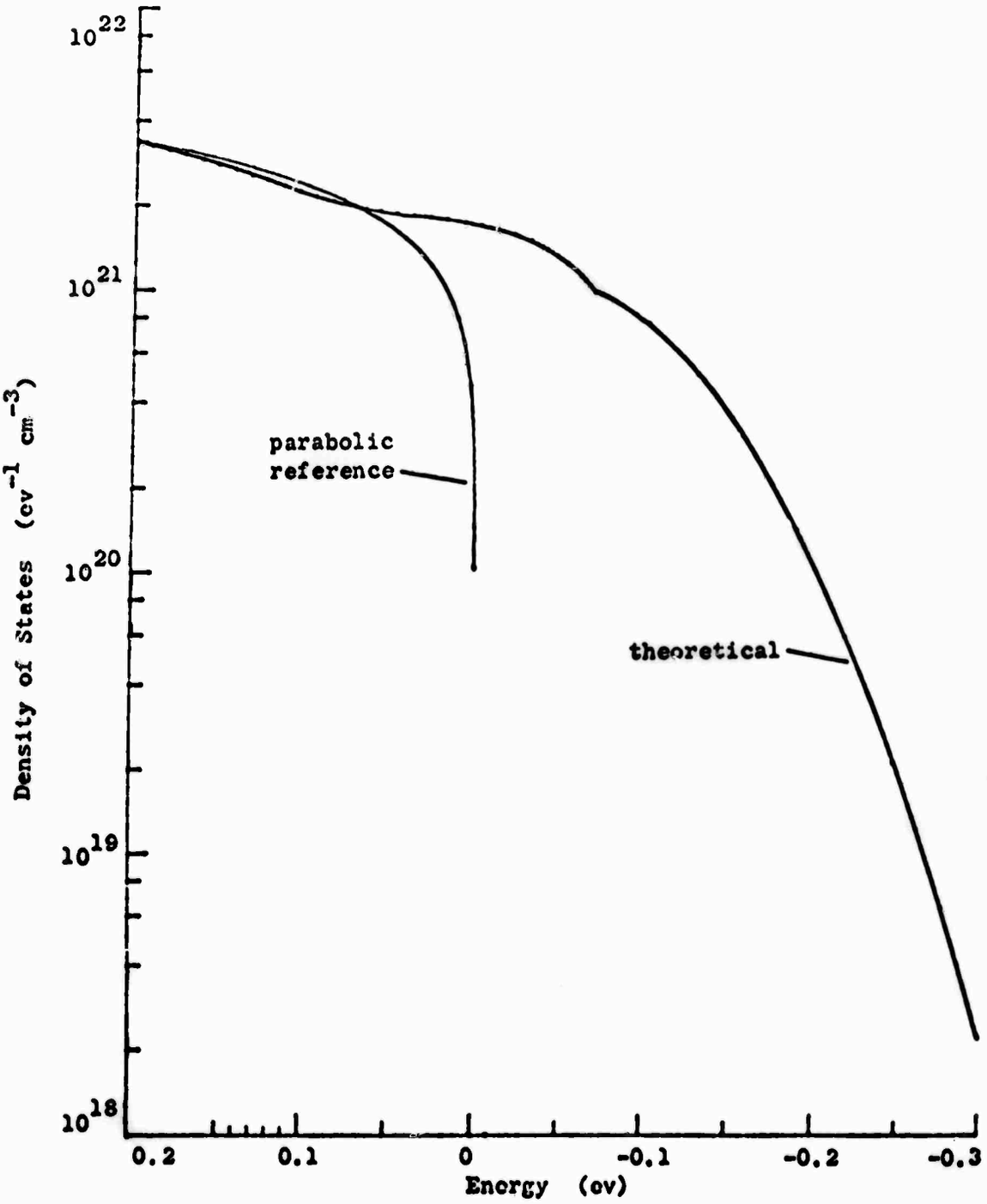


Figure 7. The density of quantum states in silicon doped with $1x10^{20}$ cm^{-3} of phosphorus atoms.

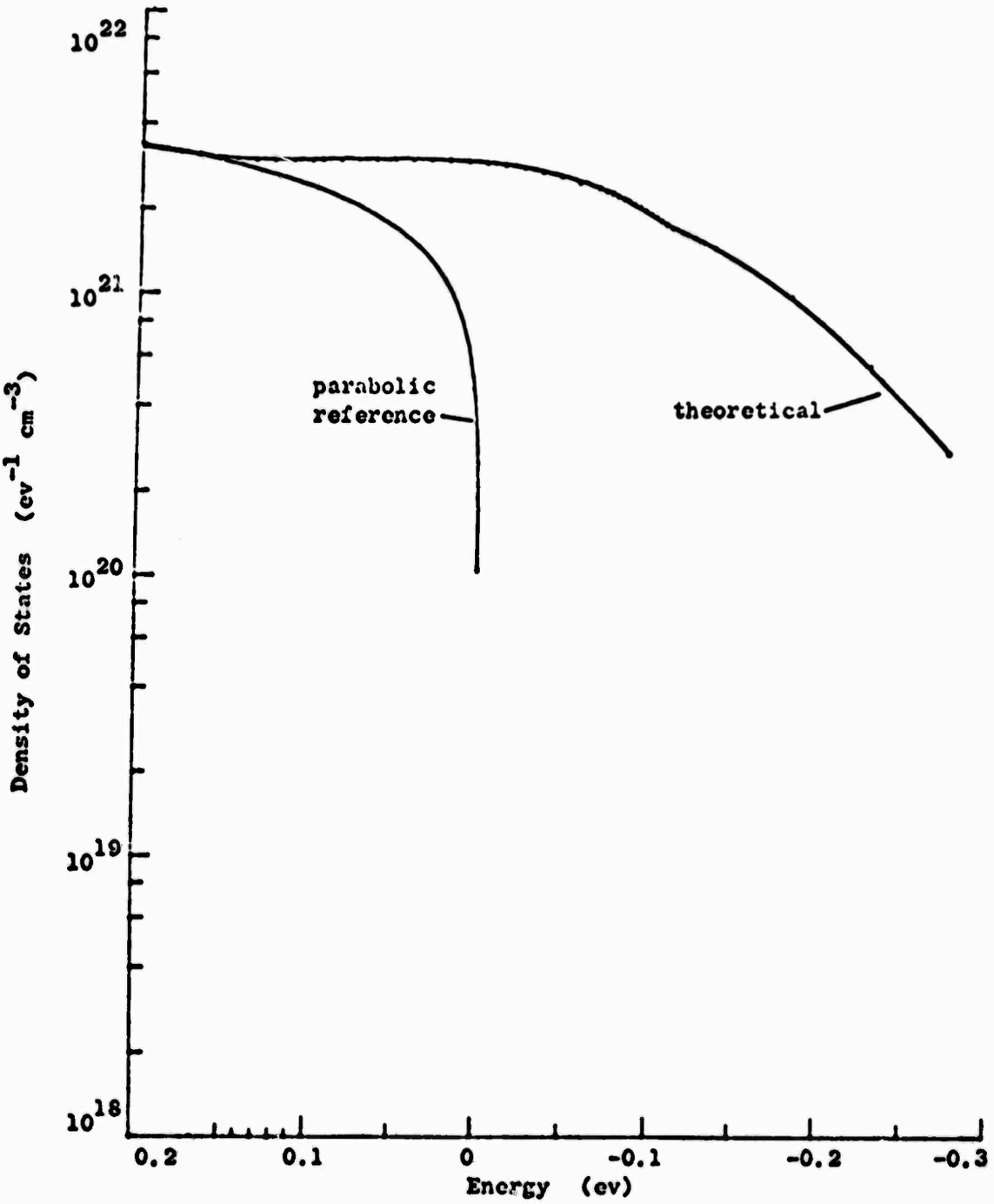


Figure 8. The density of quantum states in silicon doped with 3×10^{20} cm^{-3} of phosphorus atoms.

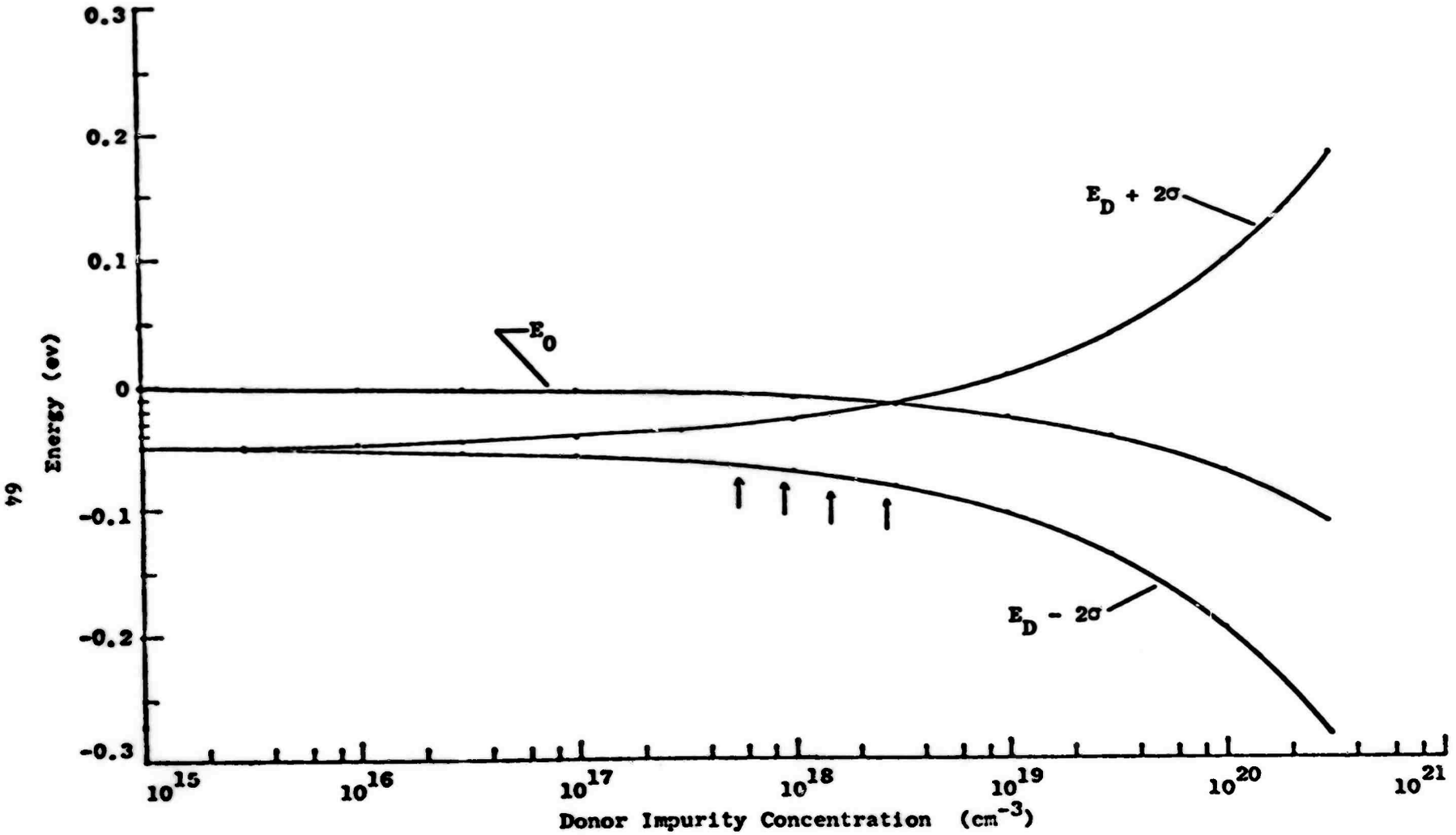


Figure 9. The development of the impurity band and the alteration of the conduction band edge in arsenic doped silicon.

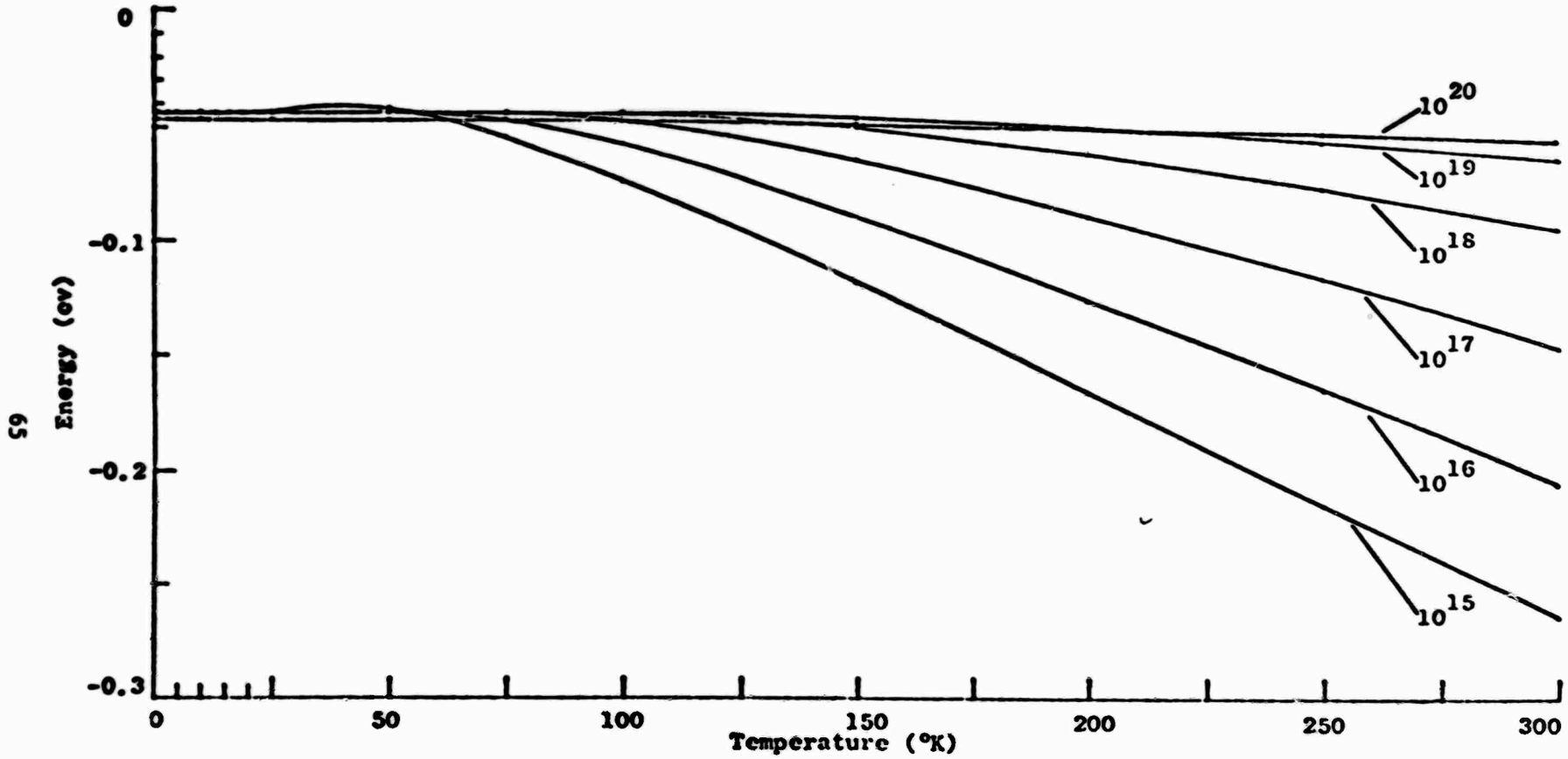


Figure 10. The variation of the Fermi level as a function of temperature for phosphorus doped silicon. The parameter labelling the curves is the phosphorus concentration per cubic centimeter.

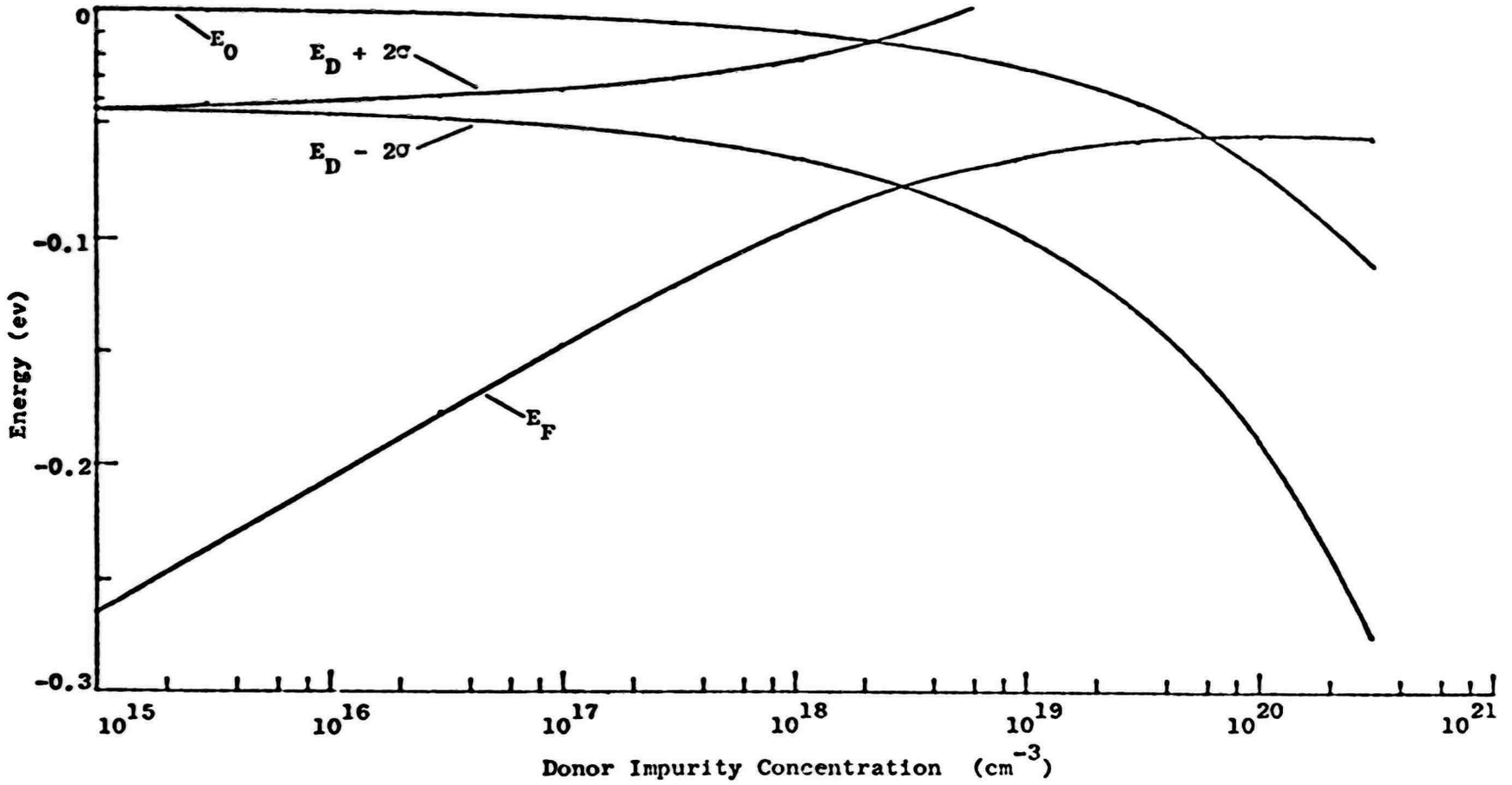


Figure 11. The variation of the Fermi level as a function of phosphorus concentration in silicon at 300°K. The curves for $E_D \pm 2\sigma$ display the development of the impurity band and the curve labeled E_0 shows the alteration of the conduction band edge.

D. RECOMBINATION AND TRAPPING OF THE PHOTO-EXCITED CARRIERS IN GOLD-DOPED SILICON (J. Agraz and S. S. Li)

I. INTRODUCTION

In steady state injection, the emisssion and capture of carriers (electrons and holes) by impurity centers cause the charge states in these centers to change from its thermal equilibrium value. Therefore, the density of carriers trapped in the flaws is in general a function of the injection. If the density of impurity centers is larger than the injected carrier density, then the charge in them will play an important role in preserving charge balance under steady state injection. In the present paper, we shall show that the charged impurity centers are dominant in controlling charge-neutrality except at very high injection.

The Shockley-Read[1] (S-R) and the Sah-Shockley[2] (S-S) statistics are used to interpret the interaction of the phosphorus and gold impurity centers with the injected carriers in silicon at low temperatures (below 100°K). The trapping of the injected carriers in these centers is included in a generalized charge neutrality equation for arbitrary injection levels and flaw densities. The formulation yields the relationship between injected hole and electron densities that preserve charge neutrality.

Specific results of the trapping and recombination processes as a function of injection are obtained for the case of n-type silicon doped with gold impurity at low temperatures. It is shown that a simple power law relationship between injected electrons and holes can be defined over specific injection range.

II. GENERALIZED CHARGE BALANCE

The assumption of charge neutrality is valid when the actual net charge density necessary to satisfy the requirements of transport is very small compared with the total opposing variable charge densities. In order to describe these, we need to consider first the details of the models for the impurities.

In Fig. 1 we show the energy band model for silicon. The impurity centers for gold and phosphorus are shown in their possible charge states.

The emission and capture statistics of the phosphorus level can be characterized by the Shockley-Read (SR) model[1]. In their development, Shockley and Read found the probability of occupancy of a single level impurity under steady state injection. This gives the density of electrons

in the phosphorus level as[1]

$$n_D = N_D \frac{C_{nD} n + C_{pD} p_{1D}}{C_{nD}(n + n_{1D}) + C_{pD}(p + p_{1D})} \tag{1}$$

There is experiment evidence[3] suggesting that the cross-section of electron capture by phosphorus is about 1000 times larger than that of hole capture. This is reasonable since electron capture by phosphorus is by a Coulomb attractive center requiring small energy transfer, whereas hole capture is by a neutral center requiring a large energy transfer. These considerations allow us to neglect the interaction of the phosphorus level with the valence band by setting C_{pD} to zero. The resulting density of electrons in that level is then reduced to

$$n_D = N_D \frac{n}{n + n_{1D}} \tag{2}$$

Sah and Shockley's statistics[2] give the ratios of the densities of a multi-level impurity in consecutive charged states. Applied to gold in silicon, this gives

$$\frac{N_{Au}^-}{N_{Au}^x} = \frac{nC_n^x + p_{1/2}C_p^-}{n_{1/2}C_n^x + pC_p^-} \tag{3}$$

$$\frac{N_{Au}^+}{N_{Au}^x} = \frac{n_{-1/2}C_n^+ + pC_p^x}{nC_n^+ + p_{-1/2}C_p^x} \tag{4}$$

These two ratios, along with the requirement that

$$N_{Au}^+ + N_{Au}^- + N_{Au}^x = N_{Au} \tag{5}$$

can be solved for the densities of the individual charged states of the gold centers. These densities are functions of the injected carrier densities.

We can now write the generalized charge balance equation for gold- and phosphorus-doped silicon:

$$p + (N_D - n_D) + N_{Au}^+ - N_{Au}^- - n = 0 \tag{6}$$

or, subtracting the thermal equilibrium contributions, we have excess density

$$\Delta p + \Delta N_{Au}^{+} = \Delta n + \Delta n_D + \Delta N_{Au}^{-} \tag{7}$$

Each term in Eq. (7) is a function of Δp and Δn only, therefore it gives an implicit relationship that the injected carrier densities have to satisfy for charge balance.

The procedure outlined above applies to n-type silicon either over- or undercompensated by gold. In fact, this approach could be applied in any case in semiconductors where a shallow impurity is compensated by a deep single or multilevel impurity. The details of the results are very sensitive to the level of compensation and the capture parameters. There is a general tendency which can be very useful. Under certain conditions, one can define ranges of injection where the excess carrier densities satisfy the charge balance equation by obeying a power law. In other words, one can define the parameters Γ and α such that

$$\Delta p = \Gamma \Delta n^{\alpha} \tag{8}$$

holds with Γ and α approximately constant within a specified range of Δn and Δp, and α is either one or two.

III. RESULTS AND DISCUSSION

A. Undercompensated Case

Consider the undercompensated case at temperatures low enough that the material is extrinsic n-type and the equilibrium electron density is small compared to the gold concentration. This can be achieved with $N_{Au} > 10^{14} cm^{-3}$, and $N_D > N_{Au}$, at temperatures below 30°K. At these temperatures, the equilibrium carrier density is given by[4]:

$$n_o \simeq N_C \frac{N_D - N_{Au}}{N_{Au}} \exp\left(\frac{E_D - E_C}{kT}\right) \tag{9}$$

Here we let E_D represent the degeneracy of the donor level in the manner described by Shockley and Read[1]. Also, at these low temperatures $n_{1/2}$, $p_{1/2}$, $n_{-1/2}$, $p_{-1/2}$ and p_o are negligibly small and n_{1D} can be expressed in terms of n_o. Using the definition of n_{1D} and Eq. (9) we find that

$$n_{1D} = \frac{N_{Au}}{N_D - N_{Au}} n_o \tag{10}$$

From Eqs. (3), (4) and (5) we find that

$$N_{Au}^{x} = N_{Au} \frac{n\Delta p}{\Delta p^2 \gamma_{-\frac{1}{2}} + \Delta pn + n^2/\gamma_{\frac{1}{2}}} \tag{11}$$

Finally, realizing that at thermal equilibrium

$$(N_D - n_D)_{th.eq.} = N_{Au} \tag{12a}$$

and

$$(N_{Au}^{-})_{th.eq.} = N_{Au} \tag{12b}$$

the charge balance equation can be rewritten as

$$\Delta p + N_{Au} \frac{2\Delta p^2 \gamma_{-1/2} + n\Delta p}{\Delta p^2 \gamma_{-1/2} + n\Delta p + n^2/\gamma_{1/2}} = \Delta n + N_{Au} \Delta n \left(\frac{N_D}{N_D - N_{Au}} n_o + \Delta n\right)^{-1} \tag{13}$$

where $\gamma_{1/2} = \frac{C_p^-}{C_n^x}$ is the ratio of the hole and electron capture rate at negative and neutral gold acceptor levels, respectively.

and $\gamma_{-1/2} = \frac{C_p^x}{C_n^+}$ is the ratio of the hole and electron capture rate at the neutral and positive gold centers.

Equation (13) can be solved for Δp as a function of Δn, with the impurity concentrations and the capture rate ratios as the parameters. This solution has been computed for assumed values of the γ's and the results are plotted in Fig. 2. In this graph, we can identify five ranges of injection as follows:

1. At low injection, the equilibrium distribution is perturbed only slightly, and we obtain a linear range.
2. The low injection nonlinear range occurs when the injected carrier densities are comparable or greater than the equilibrium density. In this range the density of electrons trapped in the flaws changes strongly with injection
3. The intermediate linear range results when the donors become saturated with electrons. In this range the change in the gold centers remains invariant with respect to injection, maintaining a constant ratio of free carriers.

4. The high injection nonlinear range results as the injected carriers themselves begin to contribute directly to charge balance.

5. In the high injection linear range, the carrier densities are simply so large that they dominate and the charge balance condition is $\Delta p = \Delta n$.

The case of compensation by a single deep level impurity can be obtained by setting $\gamma_{-1/2}$ to zero. Curve 4 is plotted in Fig. 2 with this assumption. This case differs from the two level impurity in that the intermediate linear range is diminished. The high injection linear range remains unchanged showing that the gold donor level has negligible effects in this range of injection.

The assumed values of the γ's were chosen according to the following criterion. We assume that $\gamma_{1/2} >> 1$ because it is the ratio of an attractive to a neutral capture rate. Similarly, we assume that $\gamma_{-1/2} \leq 1$.[6] However, the experimental data available do not establish these ratios accurately. Therefore, we arbitrarily chose a low and a high value for the γ's ($\gamma_{1/2}$= 500, 100; $\gamma_{-1/2}$ = 1, 0.1) in our solutions. The γ's determine the boundaries and the factors of proportionality for each range of injection. This can be observed in Fig. 2 and will be substantiated in the power laws obtained later in this paper. The exponents of the power laws are not affected except for the extreme case where the range is too restricted to make a power law valid. This situation can be observed in Fig. 2 for $\gamma_{1/2}$ = 100 and $\gamma_{-1/2}$ = 1; the nonlinear range following the low injection linear range is not a power law of 2 as in the other cases.

B. The Overcompensated Case

Let us consider the overcompensated case ($N_{Au} > N_D$) at low temperatures. The thermal equilibrium densities are:

$$n_D|_{th.eq.} = 0 \tag{14a}$$

$$N_{Au}^{-}|_{th.eq.} = N_D \tag{14b}$$

$$N_{Au}^{x}|_{th.eq.} = N_{Au} - N_D \tag{14c}$$

and

$$N_{Au}^{+}|_{th.eq.} = 0 \tag{14d}$$

With this we write the charge balance equation for this case as follows:

$$\Delta p + \frac{(N_{Au} + N_D)\Delta p^2/\gamma_{-1/2} - (N_{Au} - N_D)\Delta n^2/\gamma_{1/2} + N_D\Delta n\Delta p}{\Delta p^2\gamma_{-1/2} + \Delta n\Delta p + \Delta n^2/\gamma_{1/2}} \tag{15}$$

$$= \Delta n + N_D\Delta n(n_{1D} + \Delta n)^{-1}$$

The equilibrium carrier densities are essentially zero because the Fermi level is locked at the gold acceptor level and it is not practical to represent n_{1D} in terms of n_o in this case.

We have obtained numerical solutions of Eq. (15) for Δp vs. Δn for assumed values of the parameters. The results are plotted in Fig. 3. Again five ranges of injection can be identified and the γ's have the same effect as discussed for the undercompensated case.

The low injection range in this case is limited by n_{1D}. For Δn small compared to n_{1D}, the equilibrium distribution is only slightly disturbed, resulting in the low injection linear range. When Δn becomes comparable to and larger than n_{1D} the trapped carrier densities change strongly with injection and the low injection nonlinear range is obtained. For higher injection the donors saturate. The explanation for the intermediate-linear and the high injection ranges follows the same reasoning as for the undercompensated case.

We can obtain approximate solutions of the charge balance equations, giving Δp in terms of Δn in some of these ranges. These approximations take the form of power laws, i.e.,

$$\Delta p = \Gamma\Delta n^{\alpha} \tag{16}$$

The values of the constants Γ and α are listed in Table 1, along with the boundaries of the corresponding range injection.

The small signal quadratic ranges have a single representation valid from zero injection to the upper bound of the quadratic range. These are:

$$\Delta p = \frac{N_D - N_{Au}}{N_D\gamma_{1/2}}\left(\Delta n + \frac{\Delta n^2}{n_o}\right) \quad \text{for } N_D > N_{Au} \tag{17}$$

and

$$\Delta p = \frac{N_{Au} - N_D}{N_D \gamma_{1/2}} \Delta n + \frac{N_{Au}}{N_D n_{1D} \gamma_{1/2}} \Delta n^2 \quad \text{for } N_D < N_{Au} \tag{18}$$

We have obtained the relationship between the injected electron and hole densities for charge balance and have shown that there exist different ranges of injection where this relationship reduces to a simple power law (i.e., $\Delta p = \Gamma \Delta n^{\alpha}$). The exponent in this law varies between about 1/2 to 2 and under certain conditions a power law can extend over more than one order of magnitude of injection.

Recombination

The recombination rate through the gold centers is given by[2]:

$$R = (np - n_i^2) \left[\frac{(N_{Au}^- + N_{Au}^x) C_n^x C_p^-}{C_n^x (n + n_{1/2}) + C_p^- (p + p_{1/2})} + \frac{(N_{Au}^x + N_{Au}^+) C_n^+ C_p^x}{C_n^+ (n + n_{-1/2}) + C_p^x (p + p_{-1/2})} \right] \tag{19}$$

In the previous charge balance analysis, we found the density of impurities in each charged state. Therefore, charge balance gives all the information we need to evaluate the recombination rate. At low temperatures, Eq. (19) reduces to:

$$R = \frac{\Delta n}{\tau_n} = n \Delta p \left[\frac{(N_{Au}^x + N_{Au}^-)/N_{Au}}{n/\gamma_{1/2} + \Delta p} \frac{1}{\tau_n^x} + \frac{(N_{Au}^x + N_{Au}^+)/N_{Au}}{n + \Delta p \gamma_{-1/2}} \frac{1}{\tau_p^x} \right] \tag{20}$$

From the charge neutrality equations (13) and (15), the density of the charged and neutral gold centers can be computed as a function of excess electron density. This is demonstrated in Figs. 4 and 5, assuming that the hole and electron capture rate ratio at the gold acceptor center, $\gamma_{1_2} = 100$ and the hole and electron capture rate ratio, $\gamma_{-1_2} = 0.2$ for both under- and over-compensated cases. The curves in these figures were obtained numerically during the computations leading to Figs. 2 and 3. Note that the charged and neutral gold density versus injection curves (Figs. 4 and 5) are insensitive to the change in the values of γ_{1_2} and γ_{-1_2}.

The electron lifetime, τ_n as a function of excess electron density can now be computed by using Eq. (20). The result is shown in Figs. 6 and 7, assuming that $\gamma_{1_2} = 100$ and $\gamma_{-1_2} = 0.1$ and for different τ_n^x and τ_p^x ratio. It is interesting to note that the five injection ranges that appeared in Figs. 2 and 3 are also apparent in Figs. 4, 5, 6 and 7. In the low injection linear range most of the gold centers are in the thermal equilibrium configuration and the lifetime is constant. In the quadratic range, the charge in the gold centers changes and most of these become neutral in the upper end of the range. The decrease of the lifetime is mostly due to the increase in hole density. In the intermediate linear range the distribution of charge in the gold centers and the lifetime are essentially constant. In the high injection region the charge distribution changes to the high injection linear range values and the electron lifetime again reaches a constant value. It is also noted from Figs. 6 and 7 that the normalized electron lifetime τ_n/τ_n^x versus injection (Δn) depends strongly on the ratio of τ_n^x and τ_p^x.

IV. CONCLUSIONS

The Shockley-Read and the Sah-Shockley statistics were used to describe the interaction of gold and phosphorus centers with injected carriers in silicon. The most significant effect of these centers on the injected electron and hole densities is that their equality is destroyed by the charge balance requirement. The relationship between the injected carriers for charge neutrality is established in the form of power laws (i.e., $\Delta p = \Gamma \Delta n^{\alpha}$) that hold in different ranges of injection.

The power law relationship between Δp and Δn is linear (i.e., $\alpha = 1$) in three ranges: at very low or very high injection and in an intermediate injection range. The low and intermediate ranges are joined by a well-defined quadratic range (i.e., $\alpha = 2$). We have found the constants Γ and α that

define these power laws in four defined ranges of injection. Also, numerical analysis was used to plot Δp vs Δn for all injection satisfying the neutrality condition (Figs. 2 and 3). Application of the present theory to the photomagnetoelectric effect in gold-doped silicon at low temperatures has shown good agreement with experimental observation[5]. Another application of the present theory to the diffusion of photoexcited carriers in semiconductors will be reported in a later publication.

Table 1. Approximate solutions of the charge balance equation in n-type silicon compensated by gold. The constants below define the relationship $\Delta p = \Gamma \Delta n^{\alpha}$ which the injected carrier must satisfy for charge balance

Doping and Equilibrium Carrier Density	Γ	α	$\Delta n >>$	$\Delta n <<$
Undercompensated $N_D > N_{Au}$ $n_o << N_{Au}$ $n_i = 0$	$\frac{N_D - N_{Au}}{N_D \gamma_{1/2}}$	1	0	n_o
	$\frac{N_D - N_{Au}}{N_D \gamma_{1/2} n_o}$	2	n_o	$\frac{N_D n_o}{N_D - N_{Au}} \sqrt{\gamma_{1/2}/\gamma_{-1/2}}$
	$(\gamma_{1/2}\gamma_{-1/2})^{-1/2}$	1	$\frac{N_D n_o}{N_D - N_{Au}} \sqrt{\gamma_{1/2}/\gamma_{-1/2}}$	N_{Au}
	1	1	N_{Au}	∞
Overcompensated $N_D < N_{Au}$ $n_o = 0$ $n_i = 0$	$\frac{N_{Au} - N_D}{N_D \gamma_{1/2}}$	1	0	$\frac{N_{Au} - N_D}{N_{Au}} n_{1D}$
	$\frac{N_{Au}}{N_D \gamma_{1/2} n_{1D}}$	2	$\frac{N_{Au} - N_D}{N_{Au}} n_{1D}$	$\frac{N_D n_{1D}}{N_{Au}} \sqrt{\gamma_{1/2}/\gamma_{-1/2}}$
	$(\gamma_{1/2}\gamma_{-1/2})^{-1/2}$	1	$\frac{N_D n_{1D}}{N_{Au}} \sqrt{\gamma_{1/2}/\gamma_{-1/2}}$	N_{Au}
	1	1	N_{Au}	∞

REFERENCES

1. W. Shockley and W. T. Read, Jr., Phys. Rev. 87, 835 (1952).
2. C. T. Sah and W. Shockley, Phys. Rev. 109, 1103 (1958).
3. J. M. Brown and A. G. Jordan, J. of Appl. Phys. 37, 337 (1966).
4. J. Balkemore, Semiconductor Statistics (pergamon Press, New York, 1962).
5. J. Agraz-G and S. S. Li, Phys. Rev. B-15, September (1970).
6. J. M. Fairfield and B. V. Gokhale, J. Solid State Elec., 8, 685 (1965).

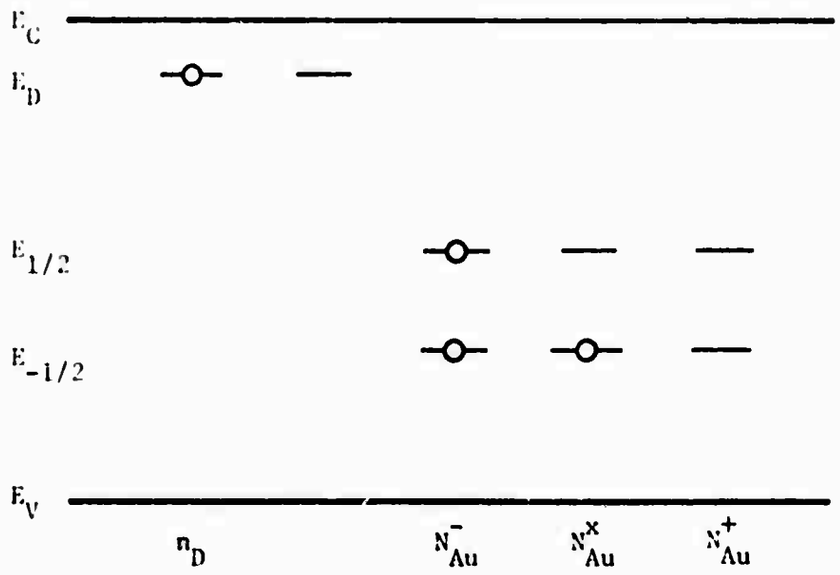


Fig. 1 Localized states introduced by gold and phosphorus in the energy band gap for silicon. N_{Au}^- denotes density of negative gold acceptor centers, N_{Au}^x for neutral gold centers and N_{Au}^+ for positive gold donor levels.

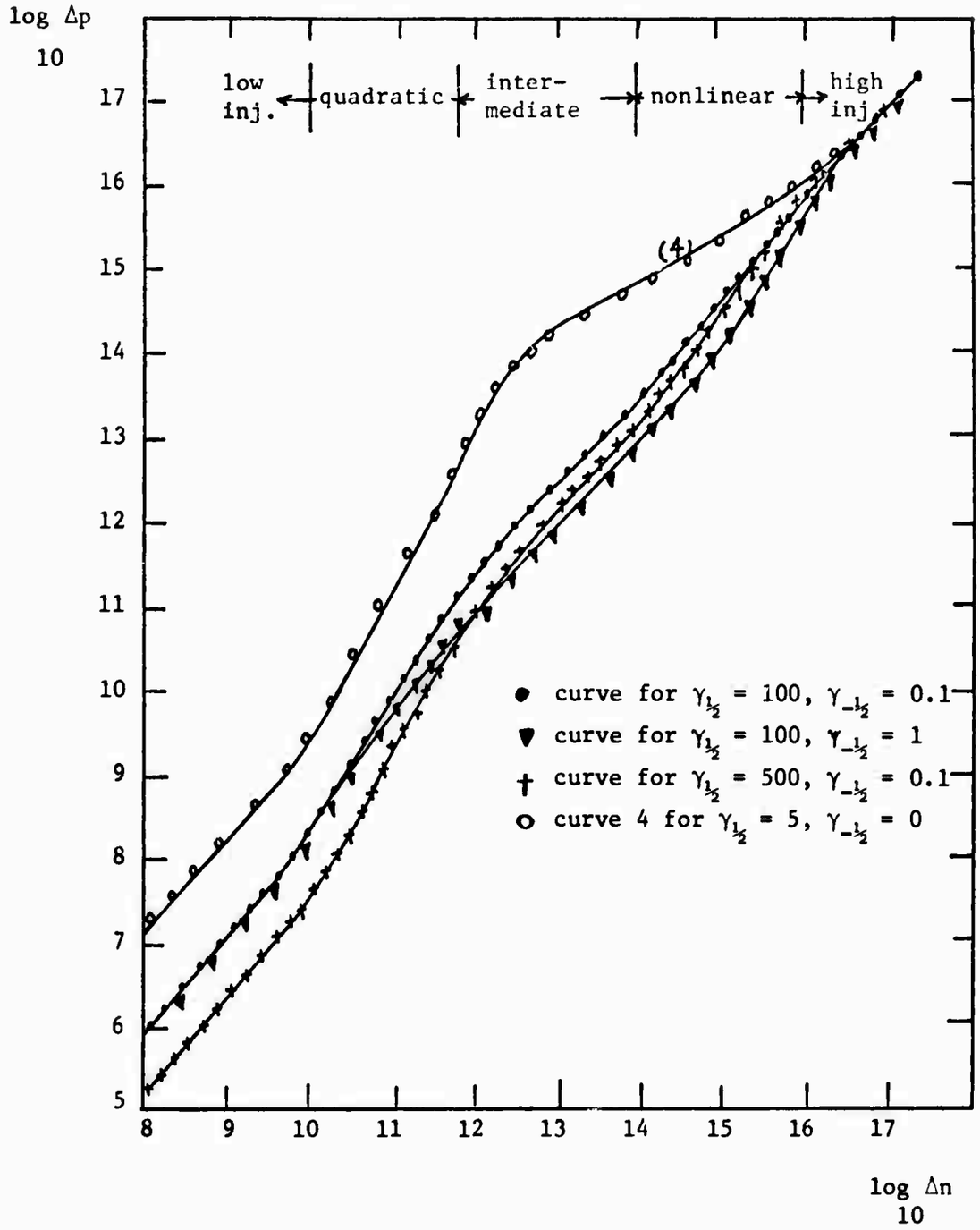


Fig. 2 Δp versus Δn obtained from the charge balance condition in n-type silicon undercompensated with gold at low temperatures (i.e., $n_o << N_{Au}$). The assumed values are: $N_D = 2 \times 10^{17} cm^{-3}$, $N_{Au} = 10^{16}$ and $n_o = 10^{10} cm^{-3}$.

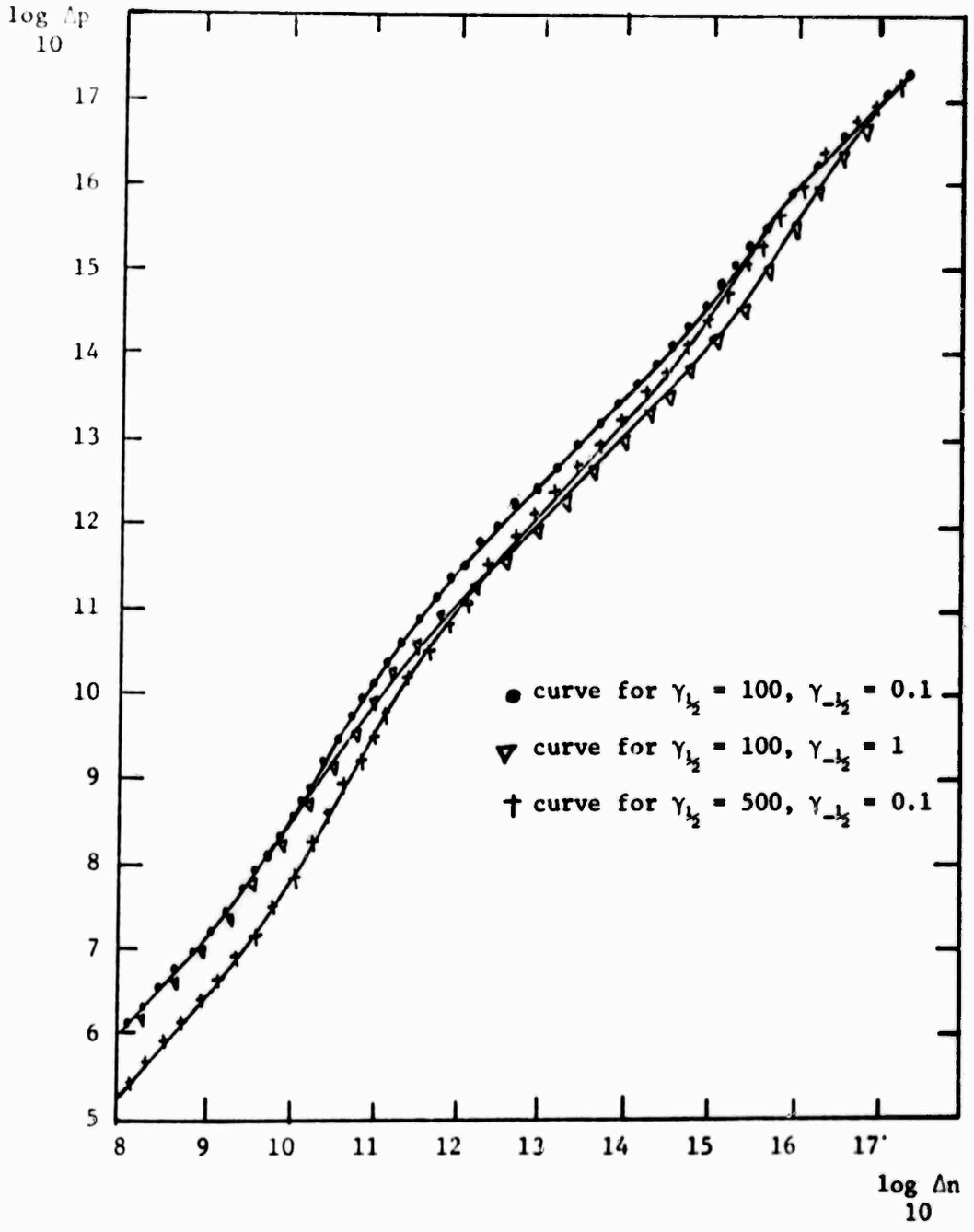


Fig. 3 Δp versus Δn in n-type silicon overcompensated with gold at low temperatures. The assumed values are: $N_D = 10^{16}cm^{-3}$, $N_{Au} = 2 \times 10^{16}cm^{-3}$ and $n_{1D} = 10^{10}cm^{-3}$.

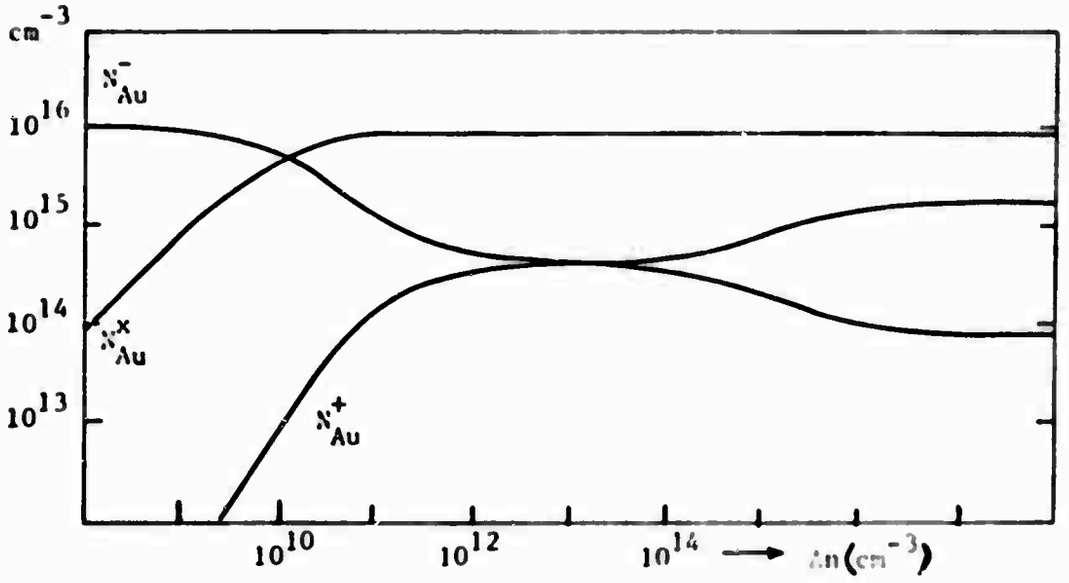


Fig. 4 Density of the charged and neutral gold centers as a function of excess electron concentration for the undercompensated case. The assumed values are: $N_D = 2 \times 10^{17} cm^{-3}$, $N_{Au} = 10^{16} cm^{-3}$, $n_o = 10^{10} cm^{-3}$, $\gamma_{\frac{1}{2}} = 100$ and $\gamma_{-\frac{1}{2}} = 0.2$.

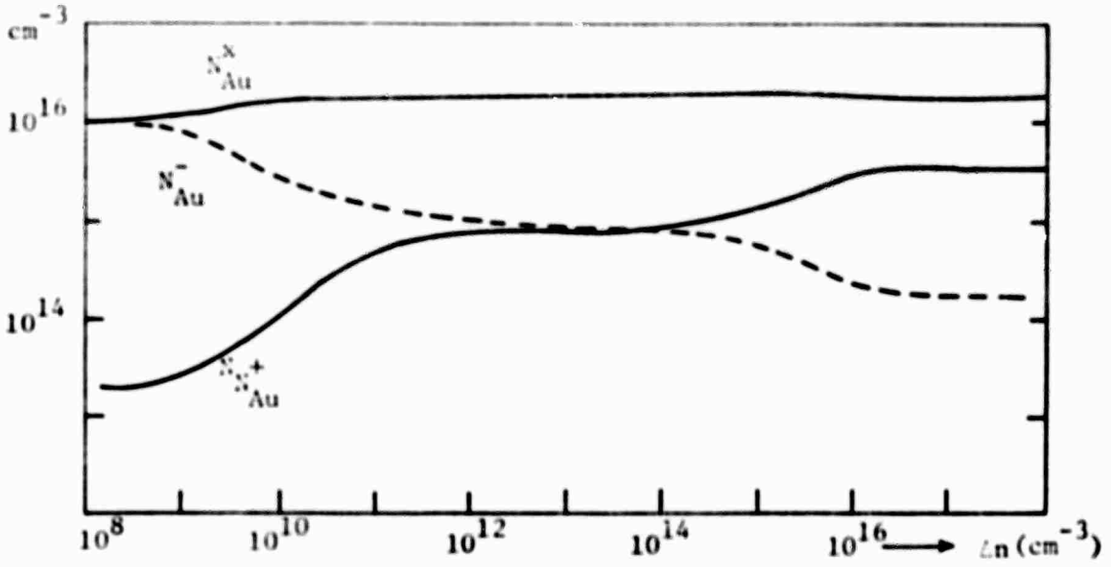


Fig. 5 Density of the charged and neutral gold centers as a function of excess electron concentration for the overcompensated case. The assumed values are: $2N_D = N_{Au} = 2 \times 10^{16} cm^{-3}$, $n_{1D} = 10^{10} cm^{-3}$, $\gamma_{1/2} = 100$ and $\gamma_{-1/2} = 0.2$.

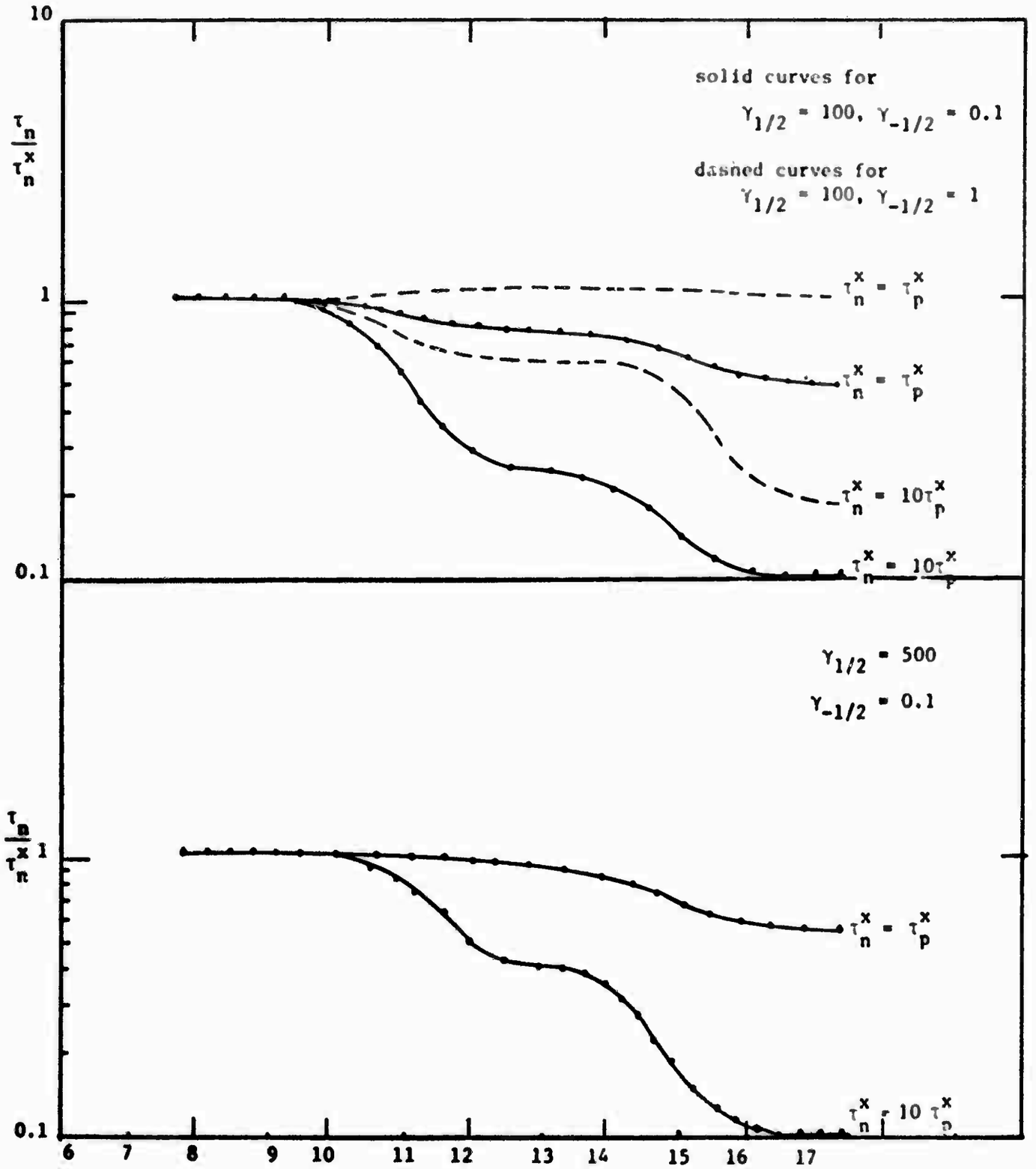


Fig. 6 Normalized electron lifetime, τ_n/τ_n^x as a function of excess electron density for undercompensated case and for different values of τ_n^x/τ_p^x, $\gamma_{\frac{1}{2}}$ and $\gamma_{-\frac{1}{2}}$.

The assumed values are: $N_D = 2 \times 10^{17}$, $N_{Au} = 10^{16}$ and $n_o = 10^{10} cm^{-3}$.

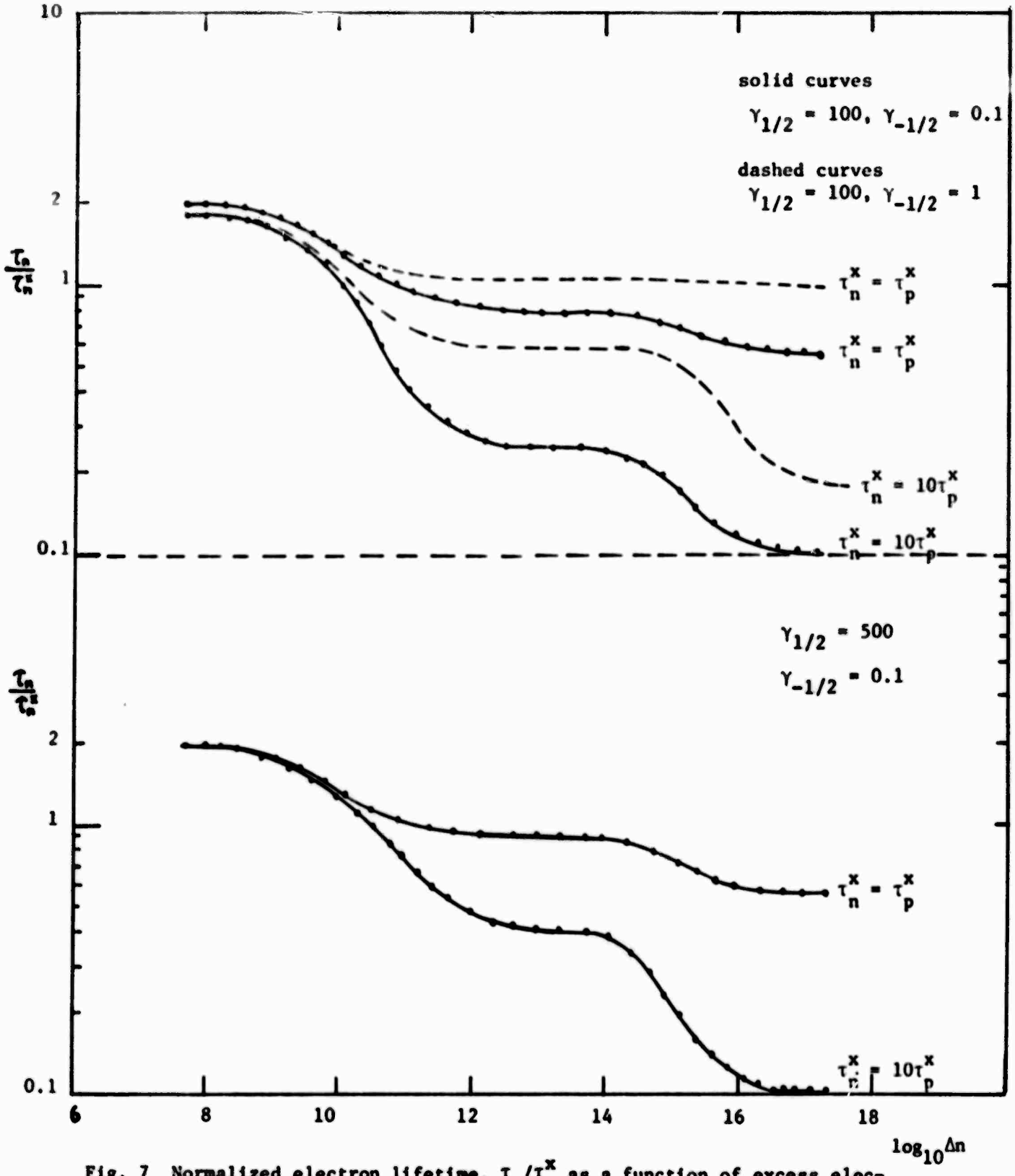


Fig. 7 Normalized electron lifetime, τ_n/τ_n^x as a function of excess electron density for overcompensated case and for different values of τ_n^x/τ_p^x, $\gamma_{\frac{1}{2}}$ and $\gamma_{-\frac{1}{2}}$.

The assumed values are: $N_{Au} = 2$, $N_D = 2 \times 10^{16}\text{cm}^{-3}$, $n_{D1} = 10^{10}\text{cm}^{-3}$.

E. UNIFIED MODELING OF FIELD-EFFECT DEVICES (F. A. Lindholm)

I. INTRODUCTION

The emergence in recent years of the computer as a tool for the analysis and design of semiconductor devices and circuits has focused renewed attention on the models that characterize the terminal behavior of devices. In particular, because the computer has afforded for the first time a general means for accomplishing the detailed analysis of large-signal transient behavior, it has awakened special interest in large-signal (nonlinear) dynamic models. The bipolar junction transistor has received most attention, it being the active device most widely used in the electronic technology; but attention also has centered on the field-effect transistor, which ranks second in usage. This paper will deal with models for the two major types of field-effect transistor and for closely related devices, with emphasis placed on models that describe the dynamic response to large-signal excitation.

In 1968 Shichman and Hodges[1] proposed a large-signal dynamic model for the metal-oxide-semiconductor transistor (MOST), also called the insulated-gate field-effect transistor (IGFET), formulated for use in the computer analysis of MOS digital integrated circuits. Shortly afterwards, Frohman-Bentchkowsky and Vadasz[2] proposed another model for the same purpose. For the JFET the main contribution purposefully related to large-signal computer analysis is due to Roberts and Harbourt[3].

The present paper describes a unifying approach to the derivation of models for field-effect devices. The most general model yielded by this approach contains, as special cases, models for both the small- and large-signal behavior of the two major types of field-effect transistor and of other related devices, such as the semiconductor current limiter[4] and the pinch resistor[5]. From the most general model emerge, by application of specific constraints and approximations, several models proposed earlier; among these are the models mentioned above for the MOST, due to Shichman and Hodges and to Frohman-Bentchkowsky and Vadasz, and for the JFET, due to Roberts and Harbourt. The unifying approach thus demonstrates the relationships borne among various of the existing models for field-effect devices.

Additionally, it provides a wide diversity of new models of different complexities and degrees of accuracy.

The central idea of the modeling approach is to resolve the device under study into two parts: an intrinsic part which is basic, being the same whether the device is an MOST, a JFET, or any of the related devices; and an extrinsic part which depends on the details of the device structure. The main approximation of the approach is the extrapolation of the static characteristics of a device to find its response to time-varying excitation. For any given device, the approach yields many different models; the differences owe their existence to the complexity chosen in the static characterization and to the choice of elements added to represent the extrinsic part of the device structure.

For the most part, consideration will center on devices having the type of structure employed in integrated circuits. The approximations undergirding the modeling approach receive attention as does the consequent accuracy of the resulting models.

II. BASIC INTRINSIC STRUCTURE

Fig. 1 shows in cross-section the essential features of the type of MOST and JFET structures employed in integrated circuits. As the figure indicates, either device consists basically of two regions: a gate and a semiconductor substrate. The gate, being non-conductive, provides a means for controlling, by applied voltage, the number of mobile carriers in the semiconductor substrate; it hence provides a means for controlling, by voltage applied at the gate, the current that flows at the drain.

For both the MOST and the JFET, the dashed lines in Fig. 1 define the boundaries of the basic intrinsic part, which is shown separately in Fig. 1(c) to emphasize that it is the same for either device. As was noted earlier, we shall model first the basic intrinsic part and later add elements to represent the extrinsic part of the device.

The structure of the basic intrinsic part comprises a gate and a semiconductor substrate. For the MOST the gate is an insulator, often silicon dioxide; for the JFET it is a semiconductor, of impurity type opposite to that of the semiconductor substrate. To some extent, the boundary defining the underside of the basic intrinsic structure is arbitrary and is chosen for convenience; as will become evident in the development below, the main requirement is that the boundaries include essentially all of the charge present in the substrate. The voltages at the terminals of the basic intrinsic structure

relative to its source are shown primed; they differ from those at the actual device terminals because of potential differences occurring in the extrinsic part of the device.

The actual structure of the JFET contains two gates, as opposed to the single gate contained in the basic intrinsic structure. Negligible loss in generality results, however, from basing treatment, as is done in this paper, on a single-gate structure. This one can show by using the following symmetry argument. In many JFET's the two gates are shorted, giving electrical symmetry. In most JFET's, however, the doping of the gates is asymmetrical, the upper gate being more heavily doped than the lower. But, because the current-voltage characteristics depend insensitively on the doping profile[6], this material asymmetry occasions current-voltage characteristics that differ only moderately from those of a symmetrically-doped device or hence from those of a device having a single gate. (If desired, the detailed effects of any doping profile existing in a device can be included[7], though at the cost of increased complexity of expression.)

If the gates are not shorted, the JFET operates as a four-terminal field-effect transistor (FTFET), the characteristics of which have been fully described[8,9]. As is discussed in Section IV, treatment of a single-gate structure provides an adequate description of FTFET operation in the modes of most interest in the integrated-circuit technology.

III. CHARGE CONTROL MODELING OF THE BASIC INTRINSIC STRUCTURE

To begin the modeling, we note that a net flow of current into the substrate of the basic intrinsic structure shown in Fig. 1(c) contributes a time-rate change of the charge Q present within the substrate:

$$i'_s + i'_d = \frac{dQ}{dt} \tag{1}$$

Charge neutrality of the overall intrinsic structure imposes the constraint

$$i'_s + i'_d = -i'_G \tag{2}$$

which demonstrates the equality between the rate of change of charge in the substrate and the displacement current that determines the current at the gate terminal.

Being analogous to an expression of Kirchhoff's current law, Eqn. (1) would describe a lumped electric network if Q were to be expressed as a function of the voltages at the terminals of the intrinsic structure. Strictly

speaking, however, Q at instant t depends not only on the values of v'_G and v'_D at that instant but also on all the values occurring at earlier times; thus it depends not only on $v'_G(t)$ and $v'_D(t)$, but also on $\frac{dv'_G(t)}{dt}$, $\frac{dv'_D(t)}{dt}$, $\frac{d^2v'_G(t)}{dt^2}$, etc. Determination of this dependence requires solution of Maxwell's equations and the pertinent (largely nonlinear) equations of semiconductor physics under auxiliary conditions imposed by the boundaries of the intrinsic structure and by the excitation and the external circuit. The determination thus poses an intractable problem.

To evade this difficulty, we solve a simpler related problem -- that of determining the dependence of Q on the terminal voltages for the dc steady state -- and make the approximation

$$Q \simeq Q_{SS}(v'_G,\ v'_G-v'_D) \triangleq \hat{Q} \qquad (3)$$

Here Q_{SS} denotes the functional dependence in the dc steady state and the symbol $\hat{Q}$ denotes what we term, for obvious reasons, the quasi-static dependence of the charge.

Combining (3) with (1) and (2) and using the chain rule produces

$$-i'_G = i'_S + i'_D = \frac{\partial \hat{Q}}{\partial v_G}\frac{dv'_G}{dt} + \frac{\partial \hat{Q}}{\partial (v'_G-v'_D)}\frac{d(v'_G-v'_D)}{dt}$$

$$= C^i_{GS}\frac{dv'_G}{dt} + C^i_{GD}\frac{d(v'_G-v'_D)}{dt}$$

which corresponds to the symbolic model of Fig. 2. To specify the functional dependence of the current source appearing in this figure, we use an approximation similar to that employed above: we assume that the x-directed current, averaged over the length of the substrate to remove dependence on the spatial variables, depends on the terminal voltages under general dynamic conditions according to the same functional relationship as that describing this current in the dc steady state. Thus

$$i \simeq (1/L)\int_0^L I_X(v'_G,v'_D)\,dx = I_D(v'_G,v'_D) \triangleq \hat{i} \qquad (5)$$

Here I_X describes the functional dependence of the x-directed substrate cur-

rent in the dc steady state; and $\hat{i}$, in analogy with the terminology used above for the charge, denotes the quasi-static dependence of the current.

Note that a symbolic model different from that of Fig. 2 would result if one chose to specify that the quasi-static charge, for example, depends on v'_G and v'_D rather than on v'_G and $(v'_G - v'_D)$. Examination of the dominant physical processes occurring in the various parts of the intrinsic structure, however, shows that the choices of independent variables made in eqns. (3) and (5) are the appropriate ones; for they lead to a symbolic model that is consistent with the conditions that no conduction current flows across the gate and that no displacement current flows in the intrinsic structure between drain and source.

The quasi-static approximations expressed in eqns. (3) and (5) are similar to those used in modeling the small-signal behavior of the bipolar transistor by the hybrid-pi model and those used in modeling the large-signal behavior of the bipolar transistor by the charge control method and its equivalents. In a later section, we discuss the error introduced by the employment of these approximations. For most field-effect transistors in typical applications, the error is small.

IV. STATIC CHARACTERIZATION OF THE INTRINSIC PART

The model of Fig. 2 describes the behavior of the intrinsic part of a field-effect transistor. It is a general model, requiring no restriction on the applied signal, neither on its size (provided breakdown does not occur) nor on its variation with time. The element values of this model depend on the quasi-static functions, $\hat{Q}$ and $\hat{i}$; determination of the element values thus requires specifying the dependence, for the dc steady state, of the substrate charge and the drain current on the terminal voltages and the device make-up.

This section treats these static dependences, for which past research has yielded various characterizations of differing complexity and accuracy. To each characterization there corresponds a different model for the intrinsic part of the field-effect transistor, each a special case of the model of Fig. 2. Combined with the many different representations that one may use to portray the behavior of the extrinsic part, this variety available in static characterization provides for a given device a wide choice in model accuracy and complexity.

A. Square-Law Characterization

The square-law characterization[6,10] applies to the general intrinsic structure itself and thus at once to both JFET's and MOST's:

$$\hat{i} \triangleq I_D(v'_G, v'_D) \simeq \frac{\mu C_G}{L} v_D Z[(v'_G - V_P) - \frac{1}{2} v'_D] \qquad (6)$$

for operation in the non-saturated mode, and

$$i = I_D(v_G', v_D') \simeq \frac{\mu C_G}{2L} Z(v_G' - V_P)^2 \tag{7}$$

for saturated operation (see List of Symbols). These equations hold for either a JFET or an MOST, the distinction between the two devices being given by

$$|v_G'/V_P| \geq 1 \tag{8}$$

for an (enhancement mode) MOST, and by

$$0 < |v_G'/V_P| \leq 1 \tag{9}$$

for a JFET. If a bias voltage is applied to the bottom gate of a JFET or to the deep portion of an MOST substrate, the effect can to a good approximation be included by specifying a functional dependence between the bias voltage and V_p, the pinch-off voltage. This is discussed further in a later section.

To this static dependence of the drain current corresponds a static dependence of the charge Q in the channel. It takes less space, however, to show the corresponding expression[11] for the derivative $\frac{dQ}{dt}$ of the quasi-static charge, which is the expression actually relevant in the modeling:

$$\frac{d\hat{Q}}{dt} = C_{GD}^i \frac{d(v_G' - v_D')}{dt} + C_{GS}^i \frac{dv_G'}{dt} \tag{10}$$

where

$$C_{GD}^i = \frac{2ZLC_G}{3} \frac{(3v_{GT} - v_D')(v_G' - v_D')}{(2v_G' - v_D)^2} \tag{11}$$

$$C_{GS}^i = \frac{2ZLC_G}{3} \frac{(3v_{GT} - 2v_D')}{(2v_{GT} - v_D')^2} \tag{12}$$

$$v_{GT} = v_G - V_P \tag{13}$$

for non-saturation; and

$$C_{GD}^i = 0; \; C_{GS}^i = \frac{2ZC_GL}{3} \tag{14}$$

for saturation.

The square-law characterization is much the simplest of those to be discussed. Moreover, the other characterizations described below are less general; each applies either to the MOST or the JFET alone, not to both devices at once. These more complicated characterizations do provide, however, a more accurate description of device behavior, though at the cost of greatly increased complexity of expression.

B. More Detailed Characterizations of the MOST

In contrast to the square-law characterization, the basic detailed characterization for the MOST (due to Sah and Pao[12], and extended by Cobbold[13]) takes into account the dependence on terminal voltages of the charge of the uncovered impurity ions in the substrate. For an n-channel enhancement MOST, for example, the characterization for non-saturation is:

$$\hat{i} \triangleq I'_D(v_G, v'_D) = \frac{C_o Z\mu}{L} \{ (v''_G - V_{FB} - 2\phi_F - v''_D/2)\, v''_D + (v'_G - V_{FB} - 2\phi_F + V_{SS}/2)\, V_{SS} - \frac{2}{3C_o} (2K_S\varepsilon_o q\, N_A)^{\frac{1}{2}} [(v''_D + 2\phi_F)^{3/2} - (2\phi_F - V_{SS})^{3/2}] \} \qquad (15)$$

where

$$v''_G = v'_G - V_{SS}$$

$$v''_D = v'_D - V_{SS}$$

and V_{SS} denotes the back bias applied to the substrate.

Many alternative expressions have been proposed to describe the current in saturation. The simplest of these is the value $i'_{D_{sat}}$ resulting from setting v'_D in Eqn. (15) equal to

$$v'_{D_{sat}} = v''_G - V_{FB} - 2\phi_F + \frac{K_s\varepsilon_o q\, N_A}{C_o^2} \left[1 - \sqrt{1 + \frac{2C_o^2(v'_G - V_{FB})}{K_s\varepsilon_o q\, N_A}} \right] \qquad (16)$$

Notice that $i'_{D_{sat}}$ is independent of v'_D and that the corresponding current-

voltage characteristic exhibits perfect saturation and zero incremental drain conductance. Actual devices show imperfect saturation, however, with the drain current increasing with increasing drain voltage. Reddi and Sah[14] give a simple description of this phenomenon:

$$\hat{i} \triangleq i'_{D_{sat}} \left[1 + (2K_s\varepsilon_o/qN_AL^2)^{\frac{1}{2}}(v'_D - v'_{D_{sat}} + V_{CD})^{\frac{1}{2}}\right], \quad v'_D > v'_{D_{sat}} . \tag{17}$$

Using a different scheme of approximation, Hofstein[15] combines his result with square-law characterization to find that

$$\hat{i} \triangleq \frac{\mu C_o}{2L} Z(v'_G - V_p)^2 (1 + \ell_C v'_D) \tag{18}$$

Here ℓ_C denotes the channel-length modulation constant, which is determined empirically to have a value for many devices of approximately 0.1 volt^{-1}. Frohman-Bentchkowsky and Grove[16] describe the effect in terms of the output conductance, which can be integrated to give

$$\hat{i} \triangleq \int g_{d_{sat}} \, dv'_D \, , \quad v'_D \geq v'_{D_{sat}} \tag{19}$$

where

$$g_{d_{sat}} \simeq \frac{i'_{D_{sat}}}{L(1 - \Delta L/L)^2} \frac{d\Delta L}{dv_D} \tag{20}$$

and the reduction in channel length ΔL is given by

$$\frac{d\Delta L}{dv_D} \simeq \frac{K_s x_o}{K_o} \frac{0.4(v''_G + Q_{ss}/C_o)}{[0.2(v'_D - v''_G + Q_{ss}/C_o) + 0.6(v''_G - v'_{D_{sat}})]^2} \tag{21}$$

The constants appearing in Eqn. (21) are empirically derived. Among the many other characterizations proposed[17], Chiu and Sah's detailed two-dimensional analysis yields apparently the most accurate description to date[18]. As one might expect, however, this accuracy is gained at the cost of greatly increased complexity of expression.

C. More Detailed Characterizations of the JFET

The simplest of the detailed characterizations for non-saturation is due to Shockley[19]:

$$i \triangleq [2\sigma aZV_p/3L]\left\{\frac{3(v'_G - v'_D + V_C)}{V_p} - 3\left[\frac{v'_G + V_C}{V_p} - 2\left(\frac{v'_G - v'_D + V_C}{V_p}\right)\right]^{3/2} + 2\left[\frac{v'_G + V_C}{V_p}\right]^{3/2}\right\} \quad (22)$$

for operation in non-saturation. The pinch-off voltage V_p is given by

$$V_p = \frac{qN}{2K_s\varepsilon_o} \quad (23)$$

Strictly, this description holds only for JFET's with abrupt junctions and homogeneous doping in channel and gate. It holds adequately for most doping profiles occurring in actual devices, however, because the current-voltage behavior depends insensitively on the profile[6]. If desired, one may employ expressions that explicitly describe a double-diffused JFET[7] or those that describe a device with an arbitrary distribution of impurities[7,20]. Though derived for a symmetrical device with an electrical short connecting the two gates, Eqn. (23) further gives with good accuracy the current-voltage dependence of a device whose two gates are biased independently, provided the impurity concentration in one gate is much less than that in the channel or in the other gate. Modern fabrication techniques often yield a device having this material asymmetry. Full descriptions are available, however, of the effect of independently biased gates, and, indeed, of devices having distinct signals applied to the two gates[8,9].

Eqn. (21), for non-saturation, holds provided

$$v'_D \leq v'_G + V_C - V_P \triangleq v'_{D_{sat}} \quad (24)$$

For larger drain voltages, saturation occurs. As for the MOST, the simplest description of the current in saturation derives from setting $v'_D = v'_{D_{sat}}$ in Eqn. (22). This yields a current that is independent of the drain voltage and hence predicts zero incremental drain conductance in saturation. To describe the increase of drain current that in actual JFET's accompanies an increase in the drain voltage, again many expressions have been proposed, each based on different approximations[21]. The simplest is essentially that given in Eqn. (17), which applies also to the MOST. For short-channel devices, failure of the gradual approximation apparently precludes analytical treatment and necessitates numerical study[21,22].

D. Expressions for the Quasi-Static Charge

For the square-law characterization of the current, Eqns. (11)-(14) show

the functional dependence of the time rate of change $\frac{d\hat{Q}}{dt}$ of the quasi-static charge. For the more detailed characterizations of current in the MOST and the JFET, however, we gave no corresponding expressions for $\frac{d\hat{Q}}{dt}$. These expressions are complex. But, as is demonstrated in the Appendix, they can be derived in a straightforward manner from the assumed quasi-static dependence of the current.

V. MODELS FOR THE EXTRINSIC PART

The extrinsic part of a field-effect transistor is that part of the device lying outside of the dashed lines shown in Fig. 1. To model the extrinsic part, we must focus on the MOST and the JFET separately.

A. The MOST

Fig. 3(a) shows the model for the extrinsic part, obtained by inspection of the dominant processes occurring within the device of Fig. 2(a). Notice that the top capacitance of each RC transmission line derives from the overlapping of the gate metal above the source or drain islands. Being a metal-oxide-semiconductor capacitance involving a semiconductor of high conductivity, the top capacitance has practically a constant value, independent of the terminal and internal voltages. On the other hand, the bottom capacitance is a junction capacitance, which exhibits well-known voltage dependence[23]. The other elements of the model can be treated as constants; their values can be determined either from terminal measurements or from calculation based on the device make-up or from a combination of the two. To obtain a description entirely in terms of ordinary differential equations, thereby gaining computational advantage, one can approximate the RC transmission lines in various ways, among which are the simple approximations shown in Fig. 3(c)-(e). Because of the low values of distributed resistance and capacitance involved, the dominant natural frequency is sufficiently high that these approximations will often suffice[24].

B. The JFET

Fig. 4(a) shows the model for the extrinsic part of the JFET, obtained by inspection of the dominant processes occurring within the device of Fig. 1(b). The capacitances associated with the RC transmission lines and the substrate junction are all junction capacitances. The two other capacitances represent the coupling of flux lines between the metal associated with the drain, source and gate terminals; they therefore have constant values, which one can determine by measurements at the device terminals. To determine the values of the other elements one may employ terminal measurements or calculations based on the device structure or a combination.

VI. RELATION TO EXISTING MODELS

Combining the intrinsic models developed in Sections III and IV with the extrinsic models of Section V furnishes a wide variety of models of different complexity and accuracy to represent the MOST and the JFET. Among these, some have appeared previously in the literature. This section treats these special cases.

A. Models for the MOST

THE SHICHMAN-HODGES MODEL[1]: This model combines the square-law characterization embodied in Eqns. (6), (7) and (17) with an approximate description of the effect of substrate bias V_{SS} on the threshold voltage V_p. In describing the quasi-static dependence of the current, the model neglects voltage drops occurring in the extrinsic part of the device; thus in the above equations v_D', the voltage between intripsic drain and intrinsic source, becomes $v_D - v_S$, the voltage between the drain and source terminals, etc. The extrinsic part of the device is represented as shown in Fig. 3. The approximation of Fig. 2(c) is used to represent that nearest the drain. The model treats capacitances that connect to the gate as constants, ignoring the voltage dependence given by Eqns. (10)-(14); this is a valid approximation if the gate metal overlaps the source and drain islands to a sufficient extent. The symbolic Shichman-Hodges model appears in Fig. 5.

THE FROHMAN-BENTCHKOWSKY & VADASZ MODEL[2]: This model employs the quasi-static dependence of current given in Eqns. (15) and (19). To achieve greater accuracy, Frohman-Bentchkowsky and Vadasz include the voltage dependence of the mobility by using an empirical approximation (their Eqn. 3). The displacement currents to the gate flow through capacitances that are nonlinear, in contrast to the Shichman-Hodges model; the nonlinear dependence of the capacitances is not stated. For this voltage dependence, there are several alternatives including that given in Eqns. (10)-(14), which is associated with the square-law characterization, and the more complex dependence directly associated with Eqn. (15). Of the extrinsic elements, source and drain resistances are included, but the island substrate capacitances and the substrate resistances are apparently neglected.

B. Models for the JFET

THE ROBERTS-HARBOURT MODEL[3]: This model describes the quasi-static dependence of the current by the square-law characterization of Eqns. (6) and (7). Displacement currents to the gate flow through nonlinear capacitances whose

dependence on voltage is accounted for by empirical approximation rather than by the dependence contained, for example, in Eqns. (10)-(14). The model includes extrinsic source and drain resistances but neglects the other extrinsic elements shown in Fig. 4, which is justifiable because Roberts and Harbourt consider only the discrete JFET. In symbolic form, the model comprises three piecewise linear parts to facilitate its use in ECAP.

C. Diversity of Models Available

The variety available in quasi-static characterization for the intrinsic part combined with the numerous alternative complexities available to represent the extrinsic part provides a wide diversity of large-signal models for the MOST and the JFET. As special cases emerge the two models described above for the MOST, but the general model of Figs. 2 and 3 contains many others. For the JFET, the numerous models implied in Figs. 2 and 4 are almost wholly new, the Roberts-Harbourt model being a primitive, though useful, special case for a discrete device.

VI. SMALL SIGNAL MODELS

To each large-signal model for the JFET or the MOST corresponds a model that describes the response to small-signal variations about bias values, obtained by expanding the equations describing the large-signal model in a Taylor series about the bias values and retaining only terms of the first-order in the variation. Figure 6 shows an example of the resulting small-signal models. Notice that the nonlinear dependence of the quasi-static current $\hat{i}(v_G', v_D')$ becomes in the small-signal model the sum of two terms

$$\frac{\partial \hat{i}}{\partial v_G'} v_g' + \frac{\partial \hat{i}}{\partial v_D'} v_d' = g_m v_g' + g_d v_d' ,$$

which are linear in the increments v_g' and v_d'. Notice further that the nonlinear capacitances lose their dependence on signal voltage, their values being determined by the bias voltages.

VII. SEMICONDUCTOR CURRENT LIMITER AND PINCH RESISTOR

The semiconductor current limiter and the pinch resistor are devices closely related to the JFET. Large- or small-signal models for them hence derive in a straightforward manner from the JFET models developed above.

The semiconductor current limiter[4] is basically a JFET with the gate electrically shorted to the source. Application of the constraint $v_{GS} = 0$ (see Fig. 1) to the JFET models therefore yields models for the current limiter.

The pinch resistor[5] is basically a JFET used as an integrated-circuit resistor. Its channel is usually fabricated during the base diffusion of bipolar transistors, and the emitter diffusion forms the top gate. In contrast with conventional diffused resistors, the pinch resistor, by the presence of its top gate, provides greatly increased sheet resistance: 10,000 to 50,000 ohms/square in comparison with 100 to 250 ohms/square typically available for conventional diffused resistors. A pinch resistor therefore requires much less area to achieve a given value of resistance; this reduces leakage current, junction capacitance, and, in many cases, processing cost. To decrease further the area required, the gates are sometimes left floating. In this floating-gate mode, the gate junctions near the source assume in effect a very small forward bias[9]. Therefore, to a good approximation, application of the constraint $v'_G = 0$ in the JFET models will yield appropriate models for floating-gate pinch resistors. For gates biased independently, the pertinent constraints to be applied exist in the literature concerning the FTFET[8,9].

VIII. APPROXIMATIONS AND MODEL VALIDITY

For a given device application, the validity of any of the models derived in this paper depends on the approximations upon which the model is based. This section will review the approximations underlying the modeling and briefly discuss their limits of validity.

In the modeling of the extrinsic part of field-effect devices, the main approximation consists in replacing the RC transmission lines associated with the drain and the source by simple lumped models, as indicated in Figs. 3 and 4. For any such replacement, one can easily compute the natural frequencies of the lumped approximation and compare these to the dominant natural frequencies of the RC line, a device that has been extensively studied[25]. Experiment suggests the validity of this approximation for representative circuit applications[1-3], and the simple comparison referred to above provides an analytical means for confirming this suggestion.

In the modeling of the basic intrinsic part, the validity of the approach depends on the validity of the quasi-static approximations used to specify the functional dependence of the charge and the current in the substrate. This in turn depends on two related conditions: that the time required for the charge density in the substrate to redistribute to the form obtaining in the steady-state, which is of the order of the carrier transit time, be small compared with the reciprocal of the maximum significant frequency contained in the signal;

and that the displacement current carried by the intrinsic capacitances be small compared with the current flowing from drain to source. The validity of the approximation is therefore clearly most threatened by the combination of small v_D and large $\frac{dv_G}{dt}$. For modern MOST's and JFET's in representative circuit applications, however, calculations based on self-consistency reasoning will indicate the validity of the quasi-static approximation. Because the transit time increases approximately in proportion to the channel length, the validity is more in question, in general, for long-channel devices. Long-channel devices include high-value pinch resistors, certain semiconductor current limiters, and the LOMOST, a long-channel MOS transistor used as a current stabilizer[26]. For devices and applications requiring increased accuracy, it can be shown that the modeling approach described in this paper can be used to generate more accurate, though more complex, models; basically these are derived by sectioning the device into subregions of sufficiently small length, modeling each subregion, and connecting the subregion models in tandem[11].

The final approximations underlying the validity of the models are those underlying the static characterizations. A full discussion concerning the validity of these approximations has appeared in the literature[27].

IX. DISCUSSION

One of the critical problems in the computer-aided design of integrated circuits, or of any circuit containing many active devices, is the choice of a device model that is sufficiently accurate for the application under study yet not so overly complex as to threaten computer capability in memory and numerical accuracy. Design by optimization theory compounds this problem because a single design may require thousands of circuit analyses. To enable the choice of a model that gives an effective compromise between accuracy and simplicity, the designer or analyst must have available a wide diversity of models of varying complexities and degrees of accuracy.

As its major contribution, the present paper has provided, for field-effect devices, this wide diversity of models. Figure 2, together with Fig. 3 for the MOST, and together with Fig. 4 for the JFET, contain many potentially useful models, each of different complexity and accuracy. By application of appropriate constrainst, these figures contain models for other related devices: the semiconductor current limiter and the pinch resistor. Moreover, to each of the large-signal models there corresponds a model describing the device response to small-signal excitation; a representative example appears in Fig. 6.

The complexity and accuracy of each model derives from the choice of elements used to represent the extrinsic part of the device and from the choice of static characterization employed to generate the model for the intrinsic part. Many combinations are possible. If new, more compact static characterizations become available in the future, as apparently is now happening in bipolar transistor theory[28], the modeling approach described in this paper can make immediate use of them to generate new, more compact, and superior models for field-effect devices.

APPENDIX

This appendix describes a general method for finding the functional dependence of $\hat{Q}$, the quasi-static charge in the substrate of the basic intrinsic part of a JFET or a MOST.

Referring to Fig. 1(c) and assuming an n-type channel, we see that $Q = Z \int\int_0 dydx = \int Z[\int q(p-N_S)dy]\, dx - \int Z[\int qndy]\, dx = \int Q_B\, dx - \int Q_n\, dx$. For operation in the non-saturated mode, the integration covers the entire substrate between the intrinsic source and drain. For operation in the saturated mode, it covers that part of the substrate between the intrinsic source and the pinch-off point. The symbol N_S relates to the impurity concentration of the substrate: $N_S = N_A$ for an n-channel MOST and $N_S = -N_D$ for an n-channel JFET.

For either an MOST or a JFET, the functional dependences of Q_B and Q_n are known for the dc steady state in terms of the material and structural parameters, the potentials at the intrinsic terminals, and the channel potential V(x) [27]. To enable integration, therefore, we change variables by employing the quasi-static expression

$$\hat{i} = \mu Z \frac{dv}{dx} \hat{Q}_n$$

which derives from invoking the gradual approximation[19,27] and assuming constant effective mobility in the channel. This leads to

$$\hat{Q} = \mu Z\{\int (1/\hat{i})\ \hat{Q}_B\hat{Q}_n\, dv - \int (1/\hat{i})(\hat{Q}_n)^2\, dv\}$$

Because the integrand is a known function of v, the integration can be done and $\hat{Q}$ determined. The result depends on the particular approximations chosen to express the functional dependences of $\hat{i}$, $\hat{Q}_B$, and $\hat{Q}_n$.

LIST OF MAIN SYMBOLS

a	half-width of JFET channel
C_G	effective gate capacitance for JFET or MOST, used in square-law characterization
C_o	parallel-plate capacitance across oxide of MOST
C_{GD}^i	gate-to-drain capacitance of basic intrinsic part of device
C_{GS}^i	gate-to-source capacitance of basic intrinsic part of device
g_d	incremental drain conductance
g_m	transconductance
i_S, i_S', i_D', i_D, i_G', i_G	total instantaneous currents defined in Fig. 1
k	Boltzmann's constant
K_s, K_o	relative dielectric constant of substrate, of oxide
L	length between source and drain of intrinsic part of device
N	impurity concentration in channel of JFET
N_A	impurity concentration in substrate of MOST
q	magnitude of electron charge
Q	charge in substrate of basic intrinsic part of device
$\hat{Q}$	quasi-static charge in substrate of basic intrinsic part of device
Q_I	fixed positive charge at the Si-SiO_2 interface of MOST
T	absolute temperature
v_G, v_G', v_D, v_D', etc.	total instantaneous voltages defined in Fig. 1
V_{FB}	flat-band voltage of MOST = $\phi_{MS} - Q_I/C_o$
V_P	pinch off voltage of JFET; threshold voltage of MOST
Z	width of basic intrinsic part of device
ε_o	dielectric permittivity of free space
ϕ_F	Fermi potential of MOST substrate = $(kT/q)\, n(N_A/n_i)$
ϕ_{MS}	Metal-semiconductor work function difference for MOST
μ	effective inversion layer mobility for MOST; effective mobility of JFET channel
σ	conductivity of JFET channel

REFERENCES

1. H. Shichman and D. A. Hodges, "Modeling and simulation of insulated-gate field-effect transistor switching circuits," IEEE J. Solid State Circuits, vol. SC-3, pp. 285-289, September 1968.

2. D. Frohman-Bentchkowsky and L. Vadasz, "Computer-aided design and characterization of digital MOS integrated circuits, IEEE J. Solid State Circuits, vol. SC-4, pp. 57-64, April 1969.

3. B. D. Roberts and C. O. Harbourt, "Computer models of the field-effect transistor," Proc. IEEE, vol. 55, pp. 1921-1929, November 1967.

4. R. M. Warner, Jr., "A new passive semiconductor component," 1958 IRE National Convention Record, vol. 6, pp. 43-48, 1958.

5. R. P. O'Grady, "The pinch resistor in integrated circuits," Microelectronics and Reliability, vol. 7, pp. 233-236, 1968.

6. I. Richer, "Basic limits on the properties of field-effect transistors," Solid State Electronics, vol. 6, pp. 539-542, September-October, 1963.

   R. D. Middlebrook and I. Richer, "Limits on the power-law exponent for field-effect transistor transfer characteristics," Solid State Electronics, vol. 6, p. 542, October 1963.

7. R. R. Bockemuehl, "Analysis of field-effect transistors with arbitrary charge distribution," IEEE Trans. on Eelctron Devices, vol. ED-10, pp. 31-34, January 1963.

   I. Richer, "Properties of an arbitrarily doped field-effect transistor," California Institute of Technology, Pasadena, Solid-State Electronics Lab Report, May 1964.

   R. S. C. Cobbold, "Theory of four-terminal double-diffused field-effect transistors," IEEE Trans. on Electron Devices, vol. ED-12, pp. 302-307, June 1965.

   F. A. Lindholm and P. R. Gray, "Large-signal and small-signal models for arbitrarily-doped four-terminal field-effect transistors," IEEE Trans. on Electron Devices, vol. ED-13, pp. 819-829, December 1966.

8. D. C. Latham, D. J. Hamilton and F. A. Lindholm, "New modes of operation for field-effect devices," Proc. IEEE (Correspondence), vol. 51, p. 226, January 1963.

   P. H. Hudson, F. A. Lindholm and D. J. Hamilton, "Transient performance of four-terminal field-effect transistors," IEEE Trans. on Electron Devices, vol. ED-12, pp. 399-407, July 1965.

   R. S. C. Cobbold and F. N. Trofimenkoff, "Four-terminal field-effect transistors," IEEE Trans. on Electron Devices, Vol. ED-11, pp. 246-247, May 1965.

   R. S. C. Cobbold, "High-frequency properties of four-terminal field-effect transistors, Proc. IEE, vol. 113, pp. 73-82, January 1966.

9. D. C. Latham, F. A. Lindholm, and D. J. Hamilton, "Low-frequency operation of four-terminal field-effect transistors," IEEE Trans. on Electron Devices, vol. ED-11, pp. 300-305, June 1964.

10. R.S.C. Cobbold, and F.N. Trofimenkoff, "Theory and application of the field-effect transistor: Part 1," Proc. IEE (London), vol. 111, pp. 1981-1992, December 1964.

I. Richer and R. D. Middlebrook, "Power-law-nature of field-effect transistor experimental characteristics," Proc. IEEE (Correspondence), vol. 51, Pp. 1145-1146, September 1963.

11. F. A. Lindholm, to be published.

12. C. T. Sah, "Characteristics of the metal-oxide-semiconductor transistors," IEEE Trans. on Electron Devices, vol. ED-11, pp. 324-345, July 1964.

C. T. Sah and H. C. Pao, "The effects of fixed bulk charge on the characteristics of MOS transistors," IEEE Trans. Electron Devices, vol. ED-13, pp. 393-409, April 1966.

13. R.S.C. Cobbold, "The MOS transistor as a four-terminal device," Electronic Letters, vol. 2, pp. 189-190, June 1966.

14. V.K.G. Reddi and C. T. Sah, "Source to drain resistance beyond pinch-off in MOS transistors," IEEE Trans. on Electron Devices, vol. ED-12, pp. 139-141, 1965.

15. J. T. Wallmark and H. Johnson, Eds., FIELD EFFECT TRANSISTORS, Edgewood Cliffs, N. J., Prentice-Hall, 1966, pp. 113-155.

16. D. Frohman-Bentchkowsky and A. S. Grove, "Conductance of MOS transistors in saturation," IEEE Trans. on Electron Devices, vol. ED-16, pp. 108-113, January 1969.

17. see references cited in 14-16 above and 18 below.

18. T. L. Chiu and C. T. Sah, "Correlation of experiments with a two-section model theory of the saturation drain conductance of MOS transistors," Solid State Electronics, vol. 11, pp. 1149-1163, 1968.

19. W. Shockley, "A unipolar 'field-effect' transistor," Proc. IRE, vol. 40, pp. 1365-1376, November 1952.

20. I. Richer, "The equivalent circuit of arbitrarily doped field-effect transistors," Solid State Electronics, vol. 8, pp. 381-393, April 1965.

21. J. Grosavelt, C. Motsch, and R. Tribes, "Physical phenomenon responsible for saturation current in field-effect devices," Solid State Electronics, vol. 6, pp. 65-67, January 1963.

J. R. Hauser, "Characteristic of junction field-effect devices with small channel length-to-width ratios," Solid State Electronics, vol. 10, pp. 577-587, June 1967.

S. Y. Wu and C. T. Sah, "Current saturation and drain conductance of junction-gate field-effect transistors," Solid State Electronics, vol. 10, pp. 593-609, June 1967.

F. N. Trofimenkoff and A. Nordquist, "FET operation in the pinch-off mode," Proc. IEE, April 1967.

C. Kim and E. S. Yang, "An analysis of current saturation mechanism of junction-gate field-effect transistors," IEEE Trans. on Electron Devices, vol. ED-17, pp. 120-127, February 1970.

22. D. P. Kennedy, "Two-dimensional computer solution of JFET characteristics," Symposium on Models for Solid State Devices, Madison, Wisc., November 1969.

23. F. A. Lindholm, "Integrated-circuit transistor and diode models for network analysis programs," IEEE Trans. on Circuit Theory, January 1971.

24. see experimental confirmation cited in reference 1.

25. D. J. Hamilton, F. A. Lindholm and A. H. Marshak, PRINCIPLES AND APPLICATIONS OF SEMICONDUCTOR DEVICE MODELING, Holt, Rinehart and Winston, New York, 1971.

M. S. Ghausi and J. J. Kelly, "INTRODUCTION TO DISTRIBUTED-PARAMETER NETWORKS, Holt, Rinehart and Winston, New York, 1968.

26. J. Kassabov, "Long channel MOS transistor (LOMOST) as a current stabilizer," Solid State Electronics, vol. 13, pp. 161-164, February 1970.

27. see the references cited above and such reference books as:
S. M. Sze, PHYSICS OF SEMICONDUCTOR DEVICES, Wiley, New York, 1969.

28. H. F. Gummel, "A charge control relation for bipolar transistors,"BSTJ, vol. 49, pp. 115-120, January 1970.

H. K. Gummel and H. C. Poon, "A compact bipolar transistor model," 1970 International Solid State Circuits Conference, Philadelphia, Penna., February 1970.

H. K. Gummel and H. C. Poon, "An integral charge control model of bipolar transistors," BSTJ, vol. 49, May 1970.

LIST OF MAIN SYMBOLS

a	half-width of JFET channel
C_G	effective gate capacitance for JFET or MOST, used in square-law characterization
C_o	parallel-plate capacitance across oxide of MOST
C_{GD}^i	gate-to-drain capacitance of basic intrinsic part of device
C_{GS}^i	gate-to-source capacitance of basic intrinsic part of device
g_d	incremental drain conductance
g_m	transconductance
i_S, i_S', i_D', i_D, i_G', i_G	total instantaneous currents defined in Fig. 1
k	Boltzmann's constant
K_s,K_o	relative dielectric constant of substrate, of oxide
L	length between source and drain of intrinsic part of device
N	impurity concentration in channel of JFET
N_A	impurity concentration in substrate of MOST
q	magnitude of electron charge
Q	charge in substrate of basic intrinsic part of device
$\hat{Q}$	quasi-static charge in substrate of basic intrinsic part of device
Q_I	fixed positive charge at the $Si-SiO_2$ interface of MOST
T	absolute temperature
v_G,v_G',v_D,v_D', etc.	total instantaneous voltages defined in Fig. 1
V_{FB}	flat-band voltage of MOST = $\Phi_{MS} - Q_I/C_o$
V_P	pinch off voltage of JFET; threshold voltage of MOST
Z	width of basic intrinsic part of device
ϵ_o	dielectric permittivity of free space
Φ_F	Fermi potential of MOST substrate = $(kT/q)\, n(N_A/n_i)$
Φ_{MS}	Metal-semiconductor work function difference for MOST
μ	effective inversion layer mobility for MOST; effective mobility of JFET channel
	conductivity of JFET channel

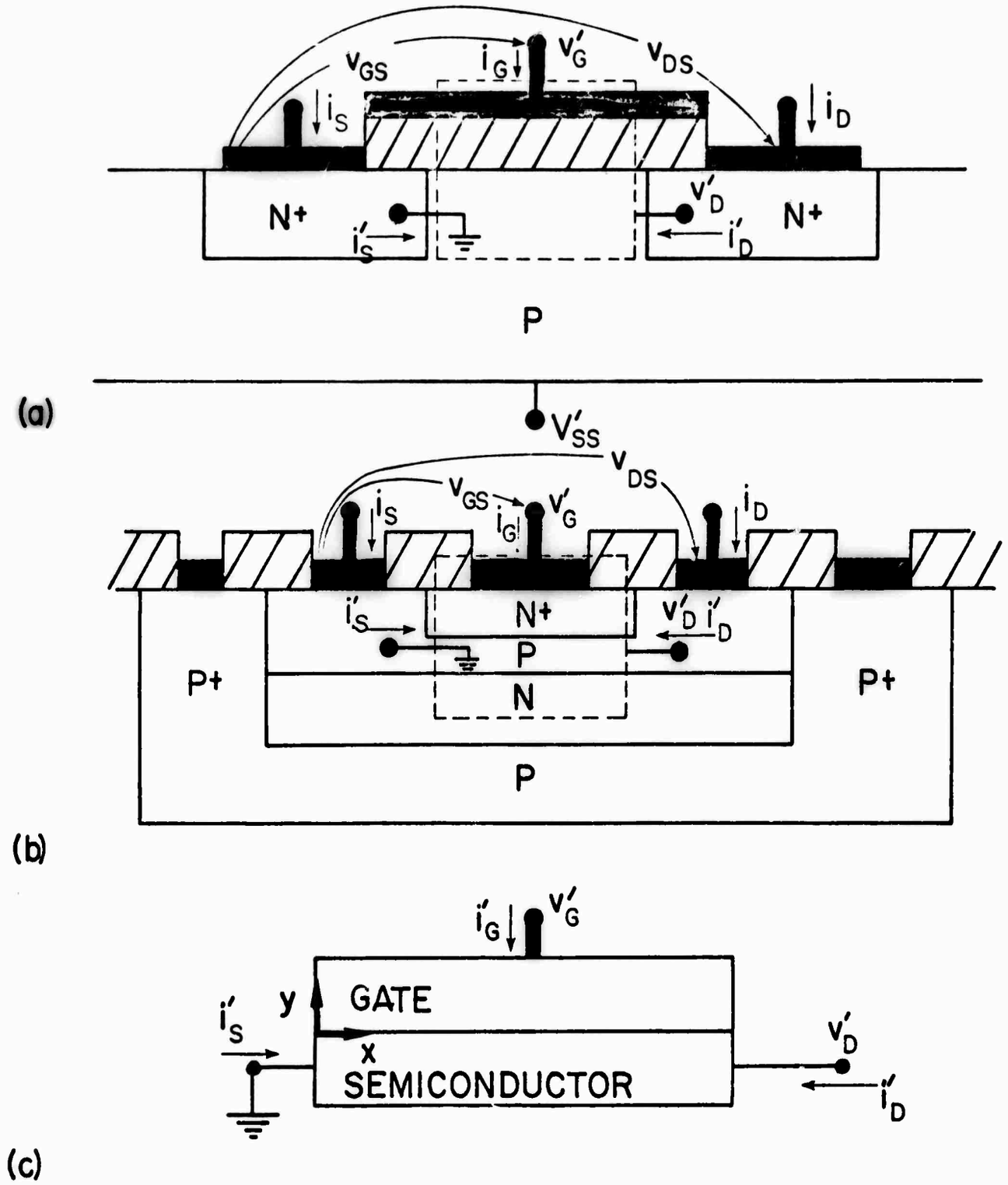


Fig. 1 Illustrating the resolution of the MOST and the JFET into a basic intrinsic part and an extrinsic part. (a) the MOST; (b) The JFET; (c) The basic intrinsic part, which is the same for either the MOST or the JFET (or for the semiconductor current limiter or the pinch resistor.

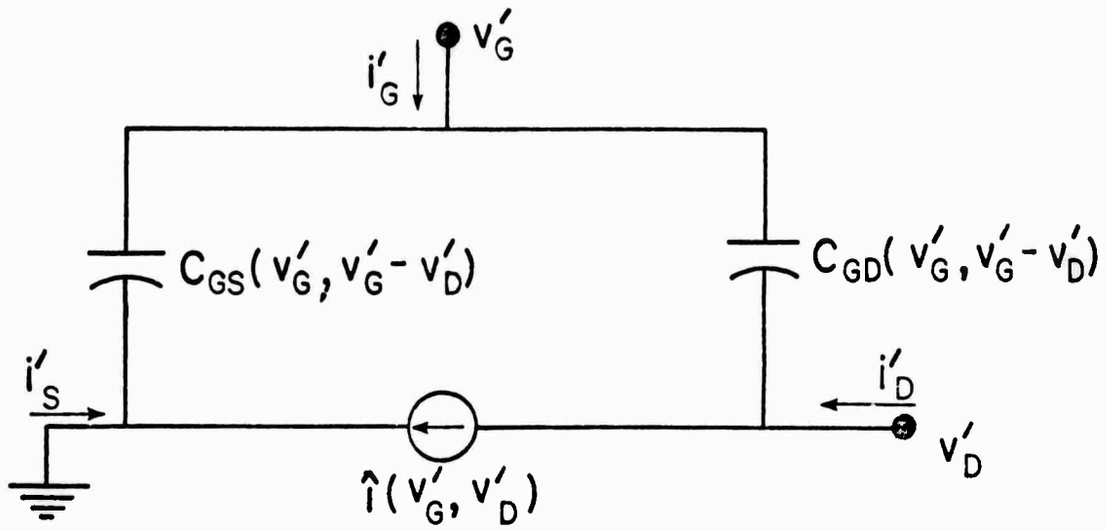


Fig. 2 Large-signal model for the basic intrinsic part. The material and structural parameters of the device, together with the voltages v'_G and v'_D, determine the functional dependence of each element in the model. As discussed in Sec. IV, one has a wide variety of possible dependences from which to choose.

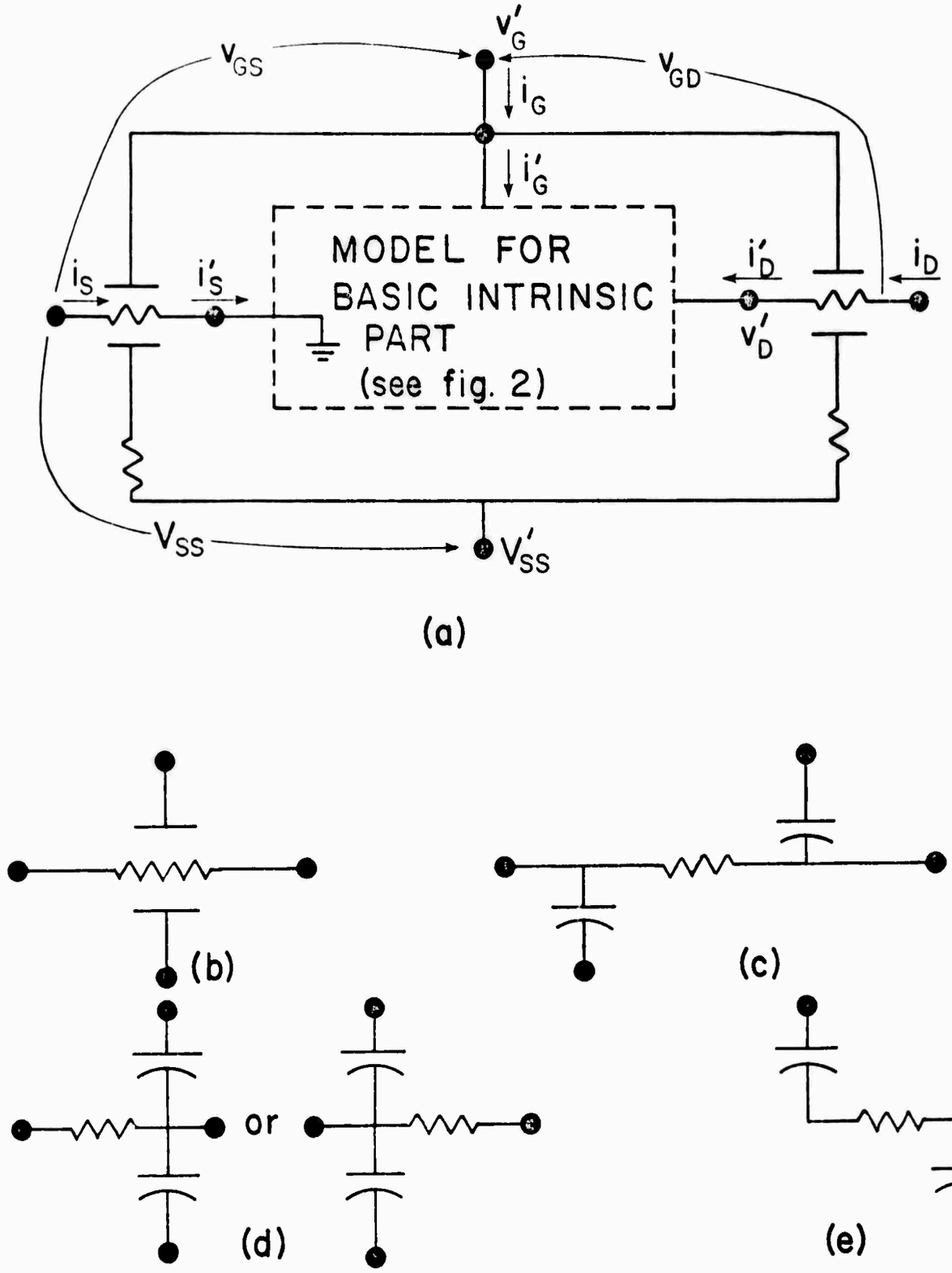


Fig. 3 (a) Illustrating the large-signal model for the extrinsic part of the MOST or IGFET; (b) Symbol for a two-capacitor RC transmission line; (c)-(e) Various lumped approximations of the RC transmission line.

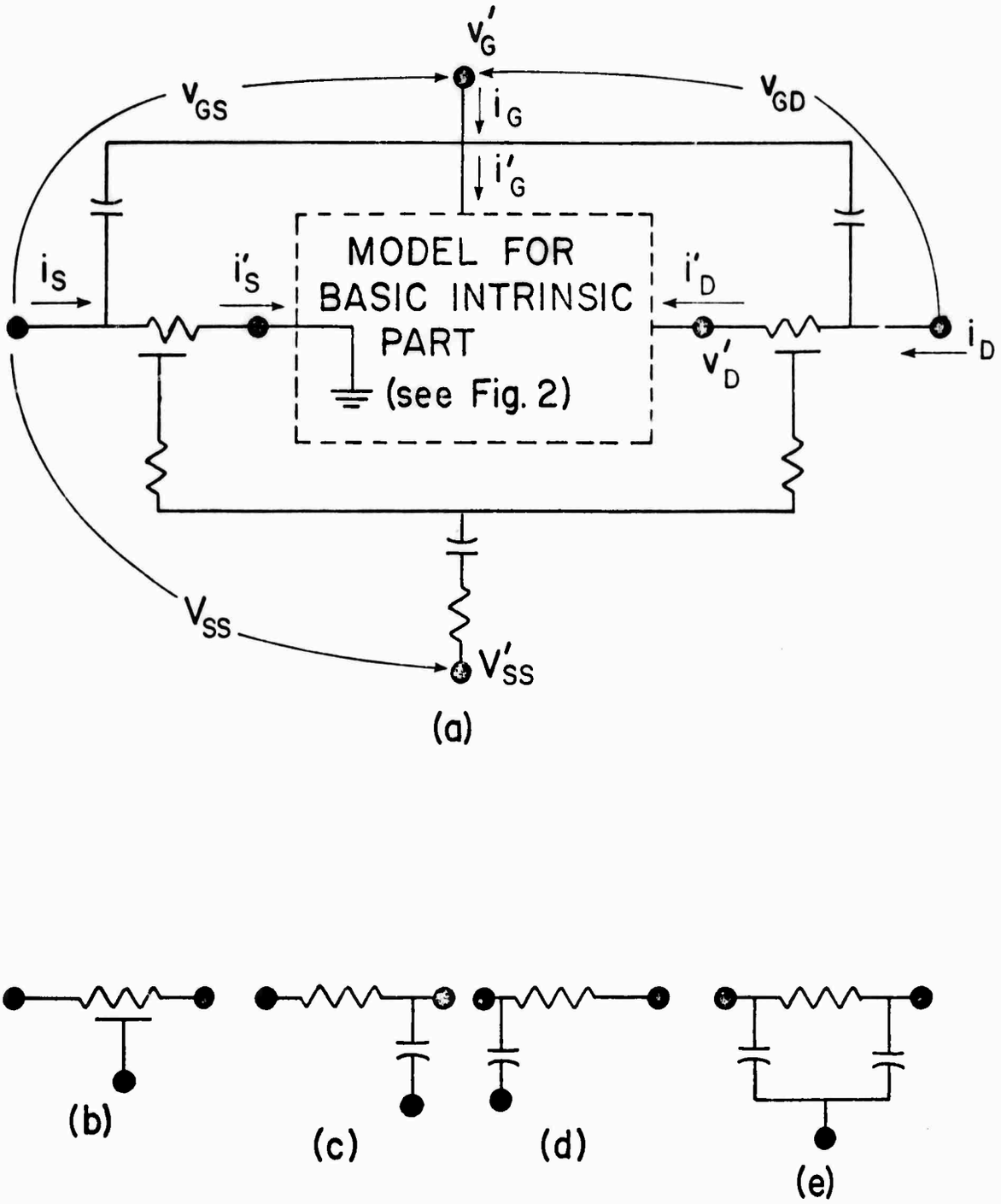


Fig. 4 (a) Illustrating the large-signal model for the extrinsic part of the JFET; (b) Symbol for a one-capacitor RC transmission line; (c)-(e) Various approximations of the RC transmission line.

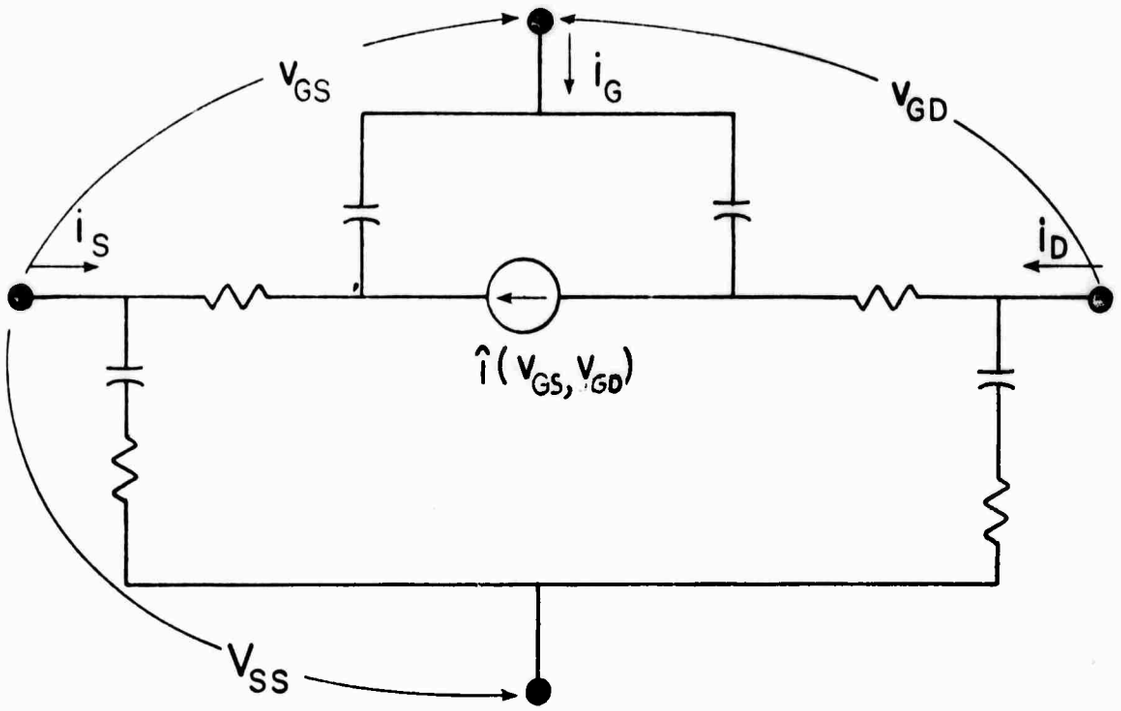


Fig. 5 The symbolic form of the Shichman-Hodges model. As discussed by Shichman and Hodges, the element values of this model can be determined by terminal measurements or from knowledge of the material and structural parameters of the device.

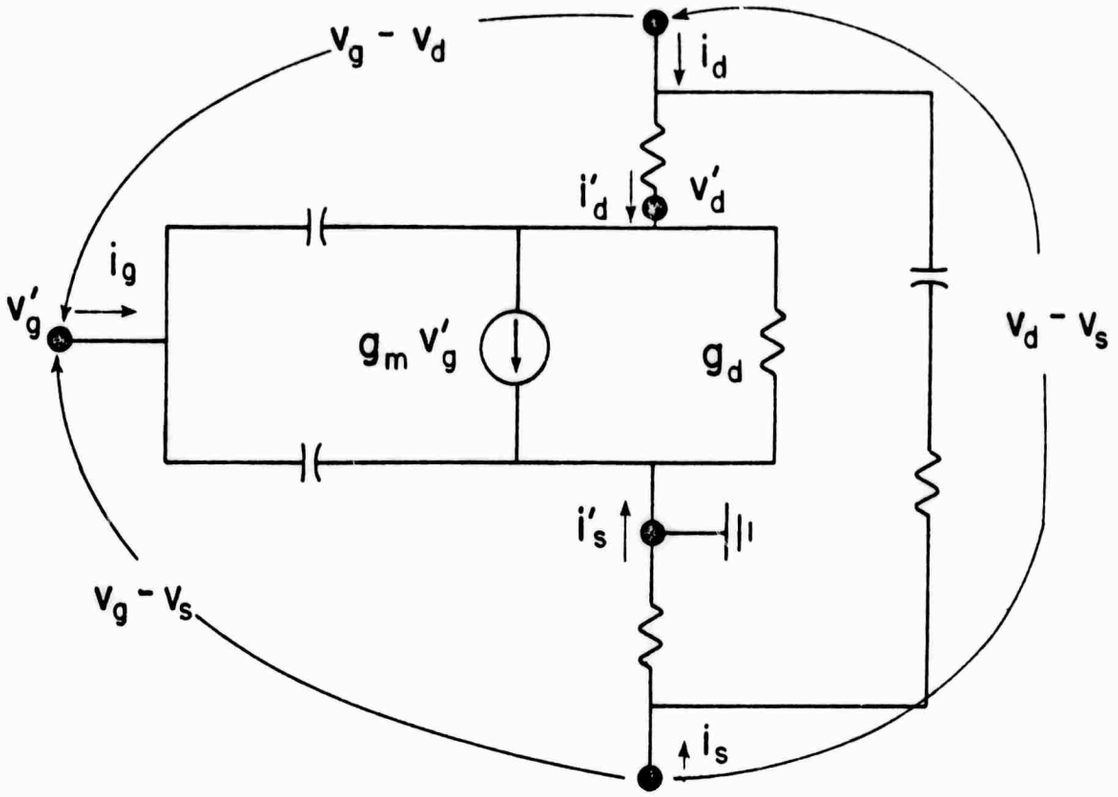


Fig. 6 An example of a small-signal model. The model shown represents an MOS transistor, and its extrinsic elements are the small-signal counterparts of those of the Shichman-Hodges model. The parameter values are known functions of the bias voltages.

F. SIMULTANEOUS AUTOMATED AC AND DC DESIGN OF LINEAR SEMICONDUCTOR INTEGRATED CIRCUIT AMPLIFIERS (A. J. Brodersen, S. W. Director and W. A. Bristol)

I. INTRODUCTION

An automated frequency domain design algorithm[1] has been developed for networks whose designable elements include linear, time-invariant resistors, capacitors and dependent sources. (Other elements such as inductors and transformers are not directly realizable as integrated circuit elements and therefore will be excluded from consideration.) This algorithm is based on the minimization of a performance function which suitably reflects the design objectives. The algorithm has been employed for the ac design of linear integrated circuit amplifiers in which transistor parameters and transistor connections are predetermined and resistors and capacitors are designed to achieve the desired response. Moreover, the algorithm is capable of altering the network structure by introducing or "growing" resistors and capacitors. To avoid excessive core storage requirements and execution time, a set of possible growth locations is usually specified a priori. Designs resulting from this procedure compare favorably with more conventional design techniques.

Subsequently this automated design algorithm was extended to the dc design of nonlinear networks which contain diodes and transistors as non-designable elements[2]. By separate use of both the frequency domain and dc design algorithms the realistic design of symmetric or differential amplifiers is possible[3]. For this class of amplifiers, the ac or frequency domain design can be separated from the dc or bias design. Reference (3) describes an iterative design procedure for symmetric amplifiers in which the transistor connections and parameters are held fixed. This design approach is unnecessarily restrictive. For example, the bias points of the transistors must be specified a priori and cannot be used to realize the ac specifications.

The present paper presents a method whereby the ac and dc design requirements can be realized by a single design step. In addition, the transistor is considered a designable element and can be "grown" in arbitrary locations to meet the ac specifications. Two design examples are presented to show the feasibility of the approach: one illustrates the simultaneous ac and dc design capability; the other illustrates the ability to grow transistors to meet the ac design specifications.

II INTEGRATED DEVICE DESIGN PARAMETERS

All the elements (resistors, capacitors, diodes, and transistors) of a monolithic integrated circuit are ultimately specified in terms of planar dimensions (the masks) and a series of diffusions, oxidations, etc. (the processing steps). Typically for economic reasons the processing, or impurity profiles, are fixed and the mask geometries varied to realize the designable elements. The only constraint on the geometries is dictated by the resolution capability of the processing. In particular, a resistor with resistance R is realized as a base diffusion whose length to width ratio, ℓ/w, is such that

$$R = \frac{\ell}{w} \rho_s$$

where ρ_s, the sheet resistivity, is determined by the processing. This equation neglects any effects due to non-uniform current distribution. If such effects are important, well known techniques[7] can be used to correct the value of the resistance. Thus we can consider R to be the designable electrical parameter of a resistor and ℓ/w the corresponding designable mask parameter. A capacitor of capacitance C is realized as a reverse biased base-collector or emitter-base junction or a metal-oxide-silicon structure whose area A is such that

$$C = \alpha A$$

where α is a processing dependent constant. The designable electrical parameter of a capacitor is C while the corresponding designable mask parameter is A.

It is not apparent what the designable electrical parameters for a transistor are, but a suitable choice for designable mask parameters is the emitter area A_E and perimeter P_E. Actually any two independent specifications of length and width of the emitter suffice; these two prove most convenient for what follows. Once the emitter geometry is given, the base and collector geometries are determined by the masking tolerances. The corresponding designable electrical parameters are found from consideration of the transistor model. The transistor model consists of two parts: the intrinsic transistor model and the extrinsic transistor model. The parameters of the extrinsic model, such as base resistance and base-collector capacitance, are present because of the physical structure and, in general, degrade the signal gain-bandwidth of the transistor. For determining the desirability of growing a transistor, our attention must necessarily focus on the intrinsic model. Once the structure is determined, the ultimate circuit performance depends more on the extrinsic elements of base resistance and collector-base capacitance, are present because of the physical

structure and, in general, degrade the signal gain-bandwidth of the transistor. For determining the desirability of growing a transistor, our attention must necessarily focus on the intrinsic model. Once the structure is determined, the ultimate circuit performance depends more on the extrinsic elements of base resistance and collector-base capacitance, and the designable parameters should necessarily reflect this dependence[4].

A suitable characterization of the small signal behavior of the intrinsic transistor in the forward active region is given by the two-port equations

$$I_B(\omega) = Y_{11}\, V_{BE}(\omega) + Y_{12}\, V_{CE}(\omega) \tag{1a}$$

and

$$I_C(\omega) = Y_{21}\, V_{BE}(\omega) + Y_{22}\, V_{CE}(\omega) \tag{1b}$$

where I_B is the base current phasor, I_C the collector current phasor, V_{BE} the base-emitter voltage phasor and V_{CE} is the collector-emitter voltage phasor. (The two-port representation is chosen for the convenience of the nodal analysis subroutine.) The hybrid-pi circuit model[5] of Fig. 1 can be used to show the dependence of the two-port y-parameters on the dc collector current I_C, the current gain β_o, and gain-bandwidth product, ω_t:

$$Y_{11} = \frac{q}{kT}\,|I_C|\left(\frac{1}{\beta_o} + j\,\frac{\omega}{\omega_t}\right) \tag{2a}$$

$$Y_{12} = -\eta\,\frac{q\,|I_C|}{kT}\left(\frac{1}{\beta_o} + j\,\frac{\omega}{\omega_t}\right) \tag{2b}$$

$$Y_{21} = \frac{q\,|I_C|}{kT} \tag{2c}$$

$$Y_{22} = \eta\,\frac{q}{kT}\,|I_C|\left(1 + \frac{1}{\beta_o} + j\,\frac{\omega}{\omega_t}\right). \tag{2d}$$

The quantity η is the base-width modulation factor and is dependent on the processing. Both β_o and ω_t are completely specified for a given A_E, P_E, I_C, and the processing parameters a_i and b_j[6]:

$$\frac{1}{\beta_o} = a_1 + a_2 \left(\frac{P_E}{A_E}\right) + a_3 \frac{I_C}{A_E} + a_4\left(\frac{1}{I_C}\right) \tag{3}$$

$$\frac{1}{\omega_t} = b_1 + b_2\left(\frac{A_E}{I_C}\right) + b_3 A_E. \tag{4}$$

It is apparent from expressions (2), (3) and (4) that the designable electrical parameters for a transistor are $\frac{1}{\beta_o}$, $\frac{1}{\omega_t}$, and I_C. For many design situations such as wideband amplifiers, the parameters β_o and ω_t may seek their maximum allowable values. The choice of β_o and ω_t as designable parameters allows us to determine the desirable locations for "growing" transistors.

III INTEGRATED AMPLIFIER DESIGN SPECIFICATIONS

The design algorithm is based on the minimization of a performance function which adequately reflects the desired network response. Usually a given gain-frequency response is the desired frequency domain specification for an amplifier operating at room temperature. An ac performance function ε_{ac}, which describes this behavior is the integral of the squared error between the actual and desired response over the frequency range of interest:

$$\varepsilon_{ac} = \int_{\Omega} 1/2[W(\omega)\,|\,V(\omega) - \hat{V}(\omega)\,|]^2\, d\omega \tag{5}$$

where $W(\omega)$ is a real nonnegative weighting function, $\hat{V}(\omega)$ is the desired voltage gain phasor, $V(\omega)$ is the actual voltage gain phasor under unit excitation and Ω is the frequency range of interest. Additional ac specifications, such as gain sensitivity, could also be included if desired.

The dc specifications must be similarly reflected. In order to guarantee that transistors used for amplification operate in the forward active region, the dc collector-emitter voltage of each transistor used for amplification can be required to be some positive dc voltage, the exact value depending upon the required dc voltage swings necessary for a given application. In addition, the dc collector current I_C of each transistor is required to remain constant over some temperature range (T_o, T_f). The following dc performance measure results:

$$\varepsilon_{dc} = \sum_{\substack{\text{all amplifying}\\ \text{transistors}}} 1/2[W_1(V_{CE} - \alpha\, V_{CC})]^2$$

$$+ \sum_{\text{all transistors}} \int_{T_o}^{T_f} 1/2[W_2\,(I_C(T) - I_C(298°))]^2 dT,$$

where $\alpha > 0$. The reference collector current $\hat{I}_C(298°)$ for each transistor is a designable parameter. (More precisely, $\hat{I}_C(298°)$ is a designable specification.) W_1 and W_2 are nonnegative weighting functions which are adjusted to insure that a 1 V error in V_{CE} and a 1 mA error in I_C have the same relative effect in the error ε_{dc}, i.e.,

$$W_2 \approx 1000\, W_1.$$

It is convenient to associate all desired responses with independent sources, i.e., voltages across current sources and currents through voltage sources. Therefore, each transistor that has a desired V_{CE} associated with it is assumed to have a zero-valued current source between its collector and emitter; V_{CE} is the voltage across this source. Similarly, each transistor that has a desired collector current associated with it is assumed to have a zero-valued voltage source in its collector lead; I_C is the current through this source. The introduction of these zero-valued sources does not alter the network performance.

The ac and dc performance functions are combined to form an overall performance measure:

$$\varepsilon = \varepsilon_{ac} + \varepsilon_{dc}$$

The design that results from the minimization of this performance function will represent a compromise between all the specifications involved. (Such a compromise is illustrated by the first example.)

IV DESIGN ALGORITHM

Many iterative techniques exist for the minimization of a function of several variables. Either the conjugate gradient method or the method of Fletcher and Powell prove to be useful. These methods adjust the designable parameters along a direction which is dependent on the gradient of the performance function. Therefore the gradient of the performance function with respect to all the designable parameters of the network must be computed.

An efficient method for the calculation of the gradient of the performance function with respect to all designable elements results from the application of Tellegen's theorem[9] to the original network and the topologically equivalent adjoint network[10]. Derivation of the gradient components is shown in the Appendix. the results are

$$\frac{\partial \varepsilon_{ac}}{\partial R} = \int_{\Omega} \mathrm{Re}\{-I_R\, \phi_R\}\, d\omega, \quad \text{for resistances;}$$

$$\frac{\partial \varepsilon_{ac}}{\partial G} = \int_{\Omega} \mathrm{Re}\{V_G \Psi_G\} \, d\omega, \text{ for conductances;}$$

$$\frac{\partial \varepsilon}{\partial C} = \int_{\Omega} \mathrm{Re}\{j\omega V_C \Psi_C\} d\omega, \text{ for capacitances;}$$

$$\frac{\partial \varepsilon}{\partial S} = \int_{\Omega} \mathrm{Re}\left\{-\frac{1}{j\omega} I_S \Phi_S\right\} d\omega, \text{ for elastances}$$

and

$$\frac{\partial \varepsilon_{ac}}{\partial(\frac{1}{\beta_o})} = \int_{\Omega} \mathrm{Re}\left\{\frac{q|I_C|}{kT} [(\eta+1) V_{BE} - \eta V_{CE}]\Psi_{BE} + \eta[V_{CE} - V_{BE}]\Psi_{CE}\right\} d\omega,$$

$$\frac{\partial \varepsilon_{ac}}{\partial(\frac{1}{\omega_t})} = \int_{\Omega} \mathrm{Re}\left\{j\omega \frac{q\, I_C}{kT} [(\eta+1)V_{BE} - \eta V_{CE}]\Psi_{BE} + \eta[V_{CE} - V_{BE}]\Psi_{CE}\right\} d\omega,$$

and

$$\frac{\partial \varepsilon_{ac}}{\partial |I_C|} = \frac{q}{kT} \int_{\Omega} \mathrm{Re}\left\{\left(\frac{1}{\beta_o} + \frac{j\omega}{\omega_t}\right)[(\eta+1)V_{BE} - \eta V_{CE}]\Psi_{BE} + \left[1 - \eta\left(\frac{1}{\beta_o} + \frac{j\omega}{\omega_t}\right)\right)V_{BE} + \eta\left(1 + \frac{1}{\beta_o} + \frac{j\omega}{\omega_t}\right)V_{CE}\right] \Psi_{CE}\right\} d\omega$$

for transistors. Ψ and Φ are used to denote the branch voltages and currents, respectively, of the adjoint network.

Observe that all these gradient components are obtained after two analyses -- one on the original network and one on the topologically equivalent adjoint network. Moreover these gradient components exist for the zero-valued, or non-existent passive elements. Thus gradient components can be computed for zero-valued resistances and elastances, "in" short circuits and zero-valued conductances and capacitances "in" open circuits. We may envision a zero-valued transistor as one with $(\frac{1}{\beta_o}) = (\frac{1}{\omega_t}) = |I_C| = 0$ (see Fig. 2). In this case the only gradient component not identically equal to zero is

$$\frac{\partial \varepsilon_{ac}}{\partial |I_C|} = \frac{q}{kT} \int_{\Omega} \mathrm{Re}\{V_{BE} \Psi_{CE}\} d\omega$$

(η = 0 for a nonexistent transistor). It has been shown[1] that the gradient of a zero-valued voltage-controlled current source between nodes E, B, and C is

$$\frac{\partial \varepsilon_{ac}}{\partial g_m} = \int_{\Omega} R_e \{V_{BE} \Psi_{CE}\} \, d\omega. \tag{6}$$

But $g_m = \frac{q}{kT} |I_C|$ so that the gradient computation associated with a zero-valued transistor is the same as that for a voltage-controlled current source.

At this juncture it is important to assess the practicality of what has been indicated above. The desirability of growing a voltage-controlled current source between three distinct nodes is easily determined. Since the voltages V_{BE} and Ψ_{CE} can always be calculated, even for the case of $g_m = 0$, the gradient expression (6) may be evaluated. But for many situations it proves more convenient to consider the possibility of growing transistors simultaneously in a short circuit and across an open circuit. For instance, consider the single common emitter stage of Fig. 3a. An additional common emitter stage might be required if more gain is needed. If so, another common emitter stage should then be grown as indicated in Fig. 3b.

In general, there are three possibilities for transistor growth as illustrated in Fig. 4. Observe that for the common-emitter configuration, the gradient component (6) reduces to that for a conductance since $V_{CE} = V_{BE}$ and $\Psi_{CE} = \Psi_{BE}$. Moreover the gradients associated with the common-base and common-collector configurations are identically zero (since $V_{CE} = V_{BE} = 0$ and $\Psi_{CE} = \Psi_{BE} = 0$, respectively). Thus the choice of a voltage-controlled current source to indicate transistor growth is a poor one. We are therefore forced to alter our approach. It is reasonable to ask what causes a need for transistors, the answer being the desirability for voltage gain and/or current gain. In the ideal situation, the common-base configuration yields only voltage gain, the common-collector configuration yields only current gain, and the common-emitter configuration yields both voltage and current gain. Therefore transistor growth should be dependent on the need for voltage and/or current gain.

A voltage-controlled voltage source connected as shown in Fig. 5a simulates the voltage gain of the common-base configuration while a current controlled current source connected as shown in Fig. 5b simulates the current gain of the common collector configuration. Both a voltage-controlled voltage source and a current-controlled current source connected as shown in Fig. 5c simulate the voltage and current gains of the common-emitter configuration. The gradient

components associated with each of these configurations are[1]

$$\frac{\partial \varepsilon_{ac}}{\partial \mu} = \int_{\Omega} \mathrm{Re}\{V_1 \phi_2\}\, d\omega$$

for the voltage gain,

$$\frac{\partial \varepsilon_{ac}}{\partial \beta} = \int_{\Omega} \mathrm{Re}\{I_1 \psi_2\}\, d\omega$$

for the current gain, and

$$\frac{\partial \varepsilon_{ac}}{\partial \mu \beta} \equiv \frac{\partial \varepsilon_{ac}}{\partial \mu} + \frac{\partial \varepsilon_{ac}}{\partial \beta} = \int_{\Omega} \mathrm{Re}\{V_1 \phi_2\} + \mathrm{Re}\{I_1 \psi_2\}\, d\omega$$

for both voltage and current gain. Observe that for the "zero-valued" case ($\mu = \beta = 0$) each of these gradients exist and, moreover, each of the models of Fig. 5 reduce to the appropriate initial model of Fig. 4a. Furthermore, after two circuit analyses, one on the original and the other on the adjoint, it is not only possible to determine if a transistor should be grown, but the appropriate configuration is also indicated as shown in Table 1 below.

Table 1

$\frac{\partial \varepsilon_{ac}}{\partial \mu}$	$\frac{\partial \varepsilon_{ac}}{\partial \beta}$	Transistor configuration
≥ 0	≥ 0	no transistor needed
< 0	≥ 0	grow common base stage
≥ 0	< 0	grow common collector stage
< 0	< 0	grow common emitter stage

If transistor growth is indicated, the hybrid-pi model of Fig. 1 should be inserted.

Resistors are the only designable elements available to meet dc specifications. The corresponding dc resistor gradient components are calculated in a manner similar to that outlined above for the ac case. In particular

$$\frac{\partial \varepsilon_{dc}}{\partial R} = -\int_{T_o}^{T_f} I_R(T)\, \phi_R(T)\, dT$$

for resistances and

$$\frac{\partial \varepsilon_{dc}}{\partial G} = \int_{T_o}^{T_f} V_G(T)\, \phi_G(T)\, dT$$

for conductances, where ϕ and ψ denote branch currents and voltages, respectively,

in the dc adjoint network. It can be shown that the adjoint network associated with a nonlinear network is a linear network[3]. The nodal admittance matrix of the adjoint network is the transpose of the nodal admittance matrix of the small-signal equivalent network obtained from the linearization of each nonlinear element in the original network about its operating point[10] The excitations associated with the adjoint network are

$$\phi_{CE} = W_1(V_{CE} - \alpha V_{CC})$$

for all current sources used for measurement of transistor collector to emitter voltage, and

$$\psi_C(t) = W_2(I_C(T) - I_C(298°))$$

for all voltage sources used for dc collector current measurements.

In summary, the gradient components of the overall performance function (6) are found as follows:

i) Perform a dc analysis on the original nonlinear network to determine the collector currents of each transistor. If this analysis is accomplished by repeated linearizations of the nonlinear elements, store the nodal admittance matrix, $\underset{\sim}{Y}$, of the final linearized network.

ii) Use the dc collector currents to establish the two-port parameters of the small-signal model of each transistor.

iii) Perform an ac analysis on the small-signal equivalent network under unit excitation and an ac analysis on the adjoint network under weighted response error excitation. The ac gradient components can now be computed.

iv) The desired 298° collector current for each transistor should now be adjusted according to

$$I_C \rightarrow I_C - \alpha \frac{\partial \varepsilon}{\partial I_C}$$

where α is a real nonnegative constant. Now the weighted dc error excitations can be calculated from the results of step (i) and used to excite the adjoint dc network. The remaining portion of the gradient component for resistors can now be computed and added to the partial resistor gradient calculated in step (iii).

V DESIGN EXAMPLES

The first example, the complete design of a simple differential amplifier, illustrates the use of a performance function which reflects both ac and dc design requirements. The initial network configuration is shown in Fig. 6a. Design specifications are placed on the magnitude of ac gain:

$$|A_v(j\omega)| = 40 \text{ dB for } |\omega| \leq 6.28 \text{ M rad/sec;}$$

and on the dc operating point:

$$I_1(t) = I_2(t) = I_3(T) = 1 \text{ mA}$$

and

$$V_1(t) = V_2(T) = 3V \qquad \text{for } 250°K \leq T \leq 350°K.$$

The designable parameters include the transistors Q_1 and Q_2 and the resistors R_1 thru R_5. It is reasonable to employ engineering judgment and limit the choice of possible "growth" locations to the above. The circuit which results from the application of the design algorithm is shown in Fig. 6b. The network responses are shown in Figs. 6c, 6d and 6e. These results were obtained after only 30 sec. of IBM 360/65 computer time.

Observe that the final network does not meet all the design specifications exactly. A compromise has been made because the requirements posed by the ac gain specification and the dc bias point specifications conflict. If the ac gain specification was to be met exactly the load resistor R_L would have to decrease; however this decrease would further increase the error in the collector-emitter voltages V_1 and V_2. In other words, the specifications of ac gain and V_1 and V_2 tend to pull the value of R_L in opposite directions and the design algorithm reaches a realistic engineering compromise. If it were necessary, in an actual design, to meet any single requirement exactly the weights associated with that component of the error function could be increased.

The second example is much more interesting from the standpoint of "growing" transistors. An initial amplifier configuration is chosen and specifications are placed on the input admittance, output impedance and voltage gain. It is convenient to express these specifications in terms of the hybrid g-parameters so that the port variables of the amplifier are given by

$$I_1(\omega) = g_{11}(\omega)\ V_1(\omega)$$

and

$$V_2(\omega) = g_{21}(\omega)\ V_1(\omega) + g_{22}(\omega)\ I_2(\omega)\ .$$

The simple common-emitter stage of Fig. 7a is a good starting network. The desired responses $\hat{g}_{11}(\omega)$, $\hat{g}_{21}(\omega)$, and $\hat{g}_{22}(\omega)$ along with the actual responses of the common-emitter stage are shown in Figs. 7b-7d. The letters A, B, and C in Fig. 7a indicate a set of possible transistor growth sites. Gradients for voltage gain, A_V, and current gain, A_I, are calculated at each of these locations. Resistance gradients are also computed. If a gradient component is negative, growth is indicated. So as to prevent unlimited growth, the computer program

allows only one transistor to grow at a time. After a local minimum has been reached, additional transistors are allowed to grow. The largest gradient component is associated with a current gain at location C. Therefore a common-collector stage is grown at C and the designable parameters adjusted until a (local) minimum is reached. The network structure and responses at this point are shown in Figs. 8a-d. Observe that the output impedance requirement, g_{22}, has been met. The gradients are then computed at the remaining possible growth sites. The desirability of a common base stage is indicated at location B. The network structure and responses which result after a (local) minimum has been reached are shown in Figs. 9a-d. Observe that both the voltage gain and output impedance specifications, g_{21} and g_{22}, have been closely met.

Finally, a common-collector stage is grown at location A. The final network structure and responses are shown in Figs. 10a-d. Observe that all specifications have been met satisfactorily. The total design required 55 seconds of IBM 360/65 computer time.

For a complete design we still need a suitable dc bias network to supply the necessary dc collector currents. This design typically requires the ability to grow additional transistors to meet the dc specifications. The present algorithm is capable only of adjusting resistance values to realize the dc specifications. Hence, the circuit designer must decide upon the location of any additional transistors. Further investigations will hopefully eliminate the necessity of this step. This cascade amplifier can be realized in a differential configuration in which the dc collector currents can be supplied by simple current sources.

APPENDIX

The derivation of the gradient components is outlined below. Consider frequency domain performance first. The weak form of Tellegen's theorem is

$$\sum_b [V_b \Phi_b - I_b \Psi_b] = 0 \tag{A1}$$

where the summation is taken over all network branches b, V_b and I_b are branch voltage and current phasors in the original network and Ψ_b and Φ_b are branch voltage and current phasors in the adjoint network. After parameter changes are made, (A1) becomes

$$\sum_b [(V_b + \Delta V_b)\Phi_b - (I_b + \Delta I_b)\Psi_b] = 0$$

so that

$$\sum_b (\Delta V_b \Phi_b - \Delta I_b \Psi_b) = 0. \tag{A2}$$

The element branch relations are now considered. For purposes that will become clear later, resistor elements are considered to be separated into resistance elements described by the branch relationship

$$V_R = RI_R,$$

and conductance elements described by the branch relationship

$$I_G = GV_G,$$

where the subscripts are used to denote element type. Similarly capacitor elements are considered to be separated into capacitance elements described by the branch relationship

$$I_C = j\omega C V_C,$$

and elastance elements described by the branch relation

$$V_S = \frac{1}{j} S I_S.$$

Then we have that

$$\Delta V_R \Phi_R - \Delta I_R \Psi_R = (R\Phi_R - \Psi_R)\Delta I_R + I_R \Phi_R \Delta R \tag{A3}$$

for resistance elements

$$\Delta V_G \Phi_G - \Delta I_G \Psi_G = (\Phi_G - G\Psi_G)\Delta V_G - V_G \Psi_G \Delta G \tag{A4}$$

for conductance elements;

$$\Delta V_C \Phi_C - \Delta I_C \Psi_C = (\Phi_C - j\omega C \Psi_C)\Delta V_C - j\omega V_C \Psi_C \Delta C \qquad (A5)$$

for capacitance elements;

$$\Delta V_S \Phi_S - \Delta I_S \Psi_S = (\frac{1}{j\omega S} \Phi_S - \Psi_S)\Delta I_S + \frac{1}{j\omega} I_S \Phi_S \Delta S \qquad (A6)$$

for elastance elements;

$$\begin{aligned}
&\Delta V_{BE}\Phi_B - \Delta I_B\Psi_{BE} + \Delta V_{CE} \Phi_C - \Delta I_C\Psi_{CE} \\
&= (\Phi_B - Y_{11}\Psi_{BE} - Y_{21} \Psi_{CE})\Delta V_{BE} + (\Phi_C - Y_{12} \Psi_{BE} - Y_{22} \Psi_{CE})\Delta V_{CE} \\
&- [(\frac{\partial Y_{11}}{\partial \gamma} V_{BE} + \frac{\partial Y_{12}}{\partial \gamma} V_{CE}) \Psi_{BE} + (\frac{\partial Y_{21}}{\partial \gamma} V_{BE} + \frac{\partial Y_{22}}{\partial \gamma} V_{CE})\Psi_{CE}]\Delta\gamma \\
&- [(\frac{\partial Y_{11}}{\partial \tau} V_{BE} + \frac{\partial Y_{12}}{\partial \tau} V_{CE}) \Psi_{BE} + (\frac{\partial Y_{21}}{\partial \tau} V_{BE} + \frac{\partial Y_{22}}{\partial \tau} V_{CE})\Psi_{CE}]\Delta\tau \\
&- [(\frac{\partial Y_{11}}{\partial |I_C|} V_{BE} + \frac{\partial Y_{12}}{\partial |I_C|} V_{CE})\Psi_{BE} + (\frac{\partial Y_{21}}{\partial |I_C|} V_{BE} + \frac{\partial Y_{22}}{\partial |I_C|} V_{CE})\Psi_{CE}]\Delta |I_C|
\end{aligned} \qquad (A7)$$

for transistors where $\gamma \equiv \frac{1}{\beta_o}$ and $\tau \equiv \frac{1}{\omega_t}$; if all second order terms are neglected.

Characterization of the adjoint elements is obtained by requiring that the coefficients of the ΔV_b and ΔI_b terms be zero. Hence the adjoint element which corresponds to a resistor is described by the branch relation

$$\Psi_R = R \Phi_R$$

and is itself a resistor. Similar results hold for conductances, capacitances, and elastances. The adjoint element which corresponds to the transistor is a linear two-port with admittance matrix

$$\begin{bmatrix} Y_{11} & Y_{21} \\ Y_{12} & Y_{22} \end{bmatrix}$$

The only other elements that are present in the original network are independent voltage and current sources, some of which are zero-valued and introduced solely for the prupose of measuring a response. We assume that the values of the independent sources remain constant so that

$$\Delta V_V = \Delta I_I = 0.$$

Expression (A2) reduces to

$$-\sum_{V} \Delta I_V \Psi_V + \sum_{I} \Delta V_I \Phi_I = \sum_{\substack{\text{designable} \\ \text{elements } \hat{b}}} G_{\hat{b}} \Delta P_{\hat{b}} \qquad (A8)$$

where $P_{\hat{b}}$ is the designable parameter of the $\hat{b}^{th}$ branch and $G_{\hat{b}}$ the corresponding coefficient which can be found in Expressions (A3) - (A7). We can associate independent sources in the original network with independent sources in the adjoint network of the same type. The values of these sources are still to be found.

Now consider the ac performance measure (5). To first order, the change in error due to a change in the designable elements is

$$\Delta\varepsilon_{ac} = \int_{\Omega} \mathrm{Re}\{W(\omega)(V^*(\omega) - \hat{V}^*(\omega)\} \Delta V(\omega)\, d\omega \qquad (A9)$$

Comparison of (A8) and (A9) reveals that if all sources in the adjoint network are set to zero except the current source which corresponds to the zero-valued current source in the original network across which the response is measured, and the value of this current source is set to

$$\Phi(\omega) = W(\omega)\ (V^*(\omega) - V^*(\omega)\ )$$

then

$$\Delta\varepsilon_{ac} = \sum_{\hat{b}} \int_{\Omega} \mathrm{Re}\{G_{\hat{b}}\}\, d\omega\ \Delta P_{\hat{b}}\ .$$

Since

$$\Delta\varepsilon_{ac} = \sum_{\hat{b}} \frac{\partial\varepsilon}{\partial P_{\hat{b}}}\ \Delta P_{\hat{b}}$$

the desired gradient components result.

REFERENCES

1. S. W. Director and R. A. Rohrer, "Automated network design -- the frequency domain case," IEEE Transactions on Circuit Theory, vol. CT-16, no. 3, pp. 330-337, Aug. 1969.

2. R. Dowell, Automated Network Biasing, M.S. Thesis, Univ. of Calif., Berkeley, 1968.

3. S. W. Director, Network Design by Mathematical Optimization, presented at the Western Electronics Show and Conference (WESCON), San Francisco, Calif., Aug. 1969.

4. B. A. Wooley, "The computer-aided optimization of integrated broadband amplifiers," 1970 ISSCC Digest of Technical Papers, Vol. XIII, pp. 84-85, Feb. 1970.

5. C. L. Searle et al, Elementary Circuit Properties of Transistors, SEEC Vol. 3, John Wiley & Sons, Inc., 1964.

6. Warner and Fordemwalt, editors, Integrated Circuits: Deisgn and Fabrication, McGraw-Hill Book Co., 1965.

7. H. R. Camenzind, Circuit Design for Integrated Electronics, Addison-Wesley, 1968, pp. 46-53.

8. D. J. Wilde and C. S. Beightler, Foundations of Optimization, Prentice-Hall, Inc., Englewood Cliffs, New Jersey: 1967.

9. B.D.H. Tellegen, "A general network theorem, with applications," Phillips Research Reports, vol. 7, no. 4, pp. 259-269; Aug. 1952.

10. S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," IEEE Transactions on Circuit Theory, vol. CT-16, no. 3, pp. 318-323, Aug. 1969.

FIGURE CAPTIONS

Fig. 1 Hybrid-pi model used to model the ac behavior of transistors.

Fig. 2 A "zero-valued" transistor.

Fig. 3 (a) A single stage transistor amplifier indicating a possible transistor growth site. (b) The growth of an additional common-emitter stage.

Fig. 4 (a) Transistors usually grow simultaneously in a short circuit (between terminals 1 and 2) and across an open circuit (between terminals 1 or 2 and 3). The three possible ways in which a transistor may be grown are as a (b) common-emitter stage, (c) common-base stage, (d) common-collector stage.

Fig. 5 Controlled source configurations which represent (a) voltage gain, (b) current gain, (c) both voltage and current gains.

Fig. 6 (a) The initial differential amplifier configuration with zero-valued transistors; (b) final differential amplifier network; (c) desired (dotted line) and actual (solid line) gain vs frequency response; (d) desired and actual collector-emitter voltages; (e) desired and actual collector currents.

Fig. 7 (a) Initial network configuration for second example; (b), (c), (d) actual and desired responses.

Fig. 8 (a) A common collector stage has grown in location C: (b), (c), (d) actual and desired responses.

Fig. 9 (a) A common base stage has grown in location B; (b), (c), (d) actual and desired responses.

Fig. 10 (a) The final amplifier configuration; (b), (c), (d) actual and desired responses.

Figure Captions

Fig. 1. Hybrid-pi model used to model the ac behavior of transistors.

Fig. 2. A "zero-valued" transistor.

Fig. 3. (a) A single stage transistor amplifier indicating a possible transistor growth site. (b) The growth of an additional common-emitter stage.

Fig. 4. (a) Transistors usually grow simultaneously in a short circuit (between terminals 1 and 2) and across an open circuit (between terminals 1 or 2 and 3). The three possible ways in which a transistor may be grown are as a (b) common-emitter stage, (c) common-base stage, (d) common-collector stage.

Fig. 5. Controlled source configurations which represent (a) voltage gain, (b) current gain, (c) both voltage and current gains.

Fig. 6. (a) The initial differential amplifier configuration with zero-valued transistors; (b) final differential amplifier network; (c) desired (dotted line) and actual (solid line) gain vs frequency response; (d) desired and actual collector-emitter voltages; (e) desired and actual collector currents.

Fig. 7. (a) Initial network configuration for second example; (b), (c), (d) actual and desired responses.

Fig. 8. (a) A common collector stage has grown in location C; (b), (c), (d) actual and desired responses.

Fig. 9. (a) A common base stage has grown in location B; (b), (c), (d) actual and desired responses.

Fig. 10. (a) The final amplifier configuration; (b), (c), (d) actual and desired responses.

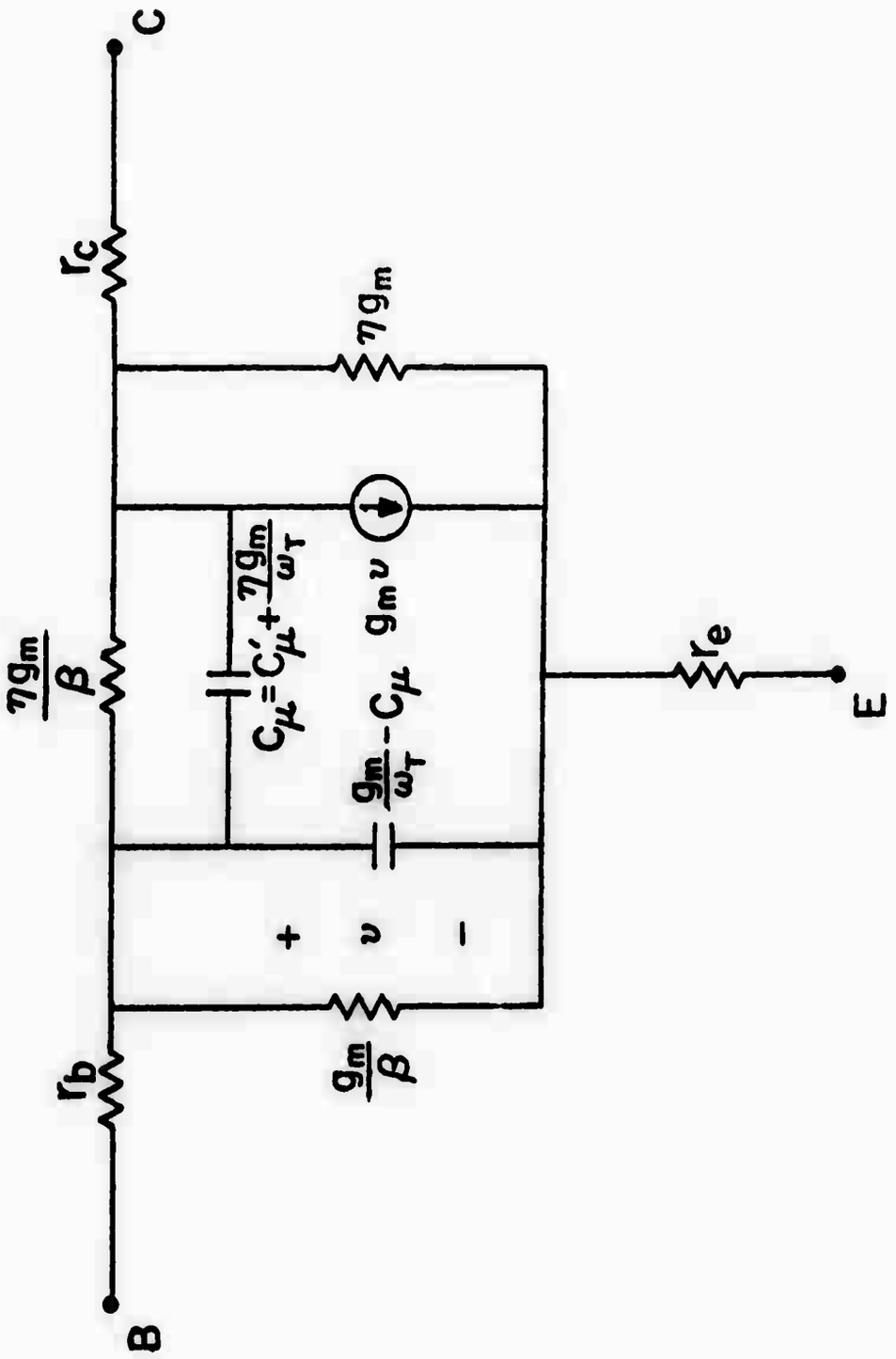
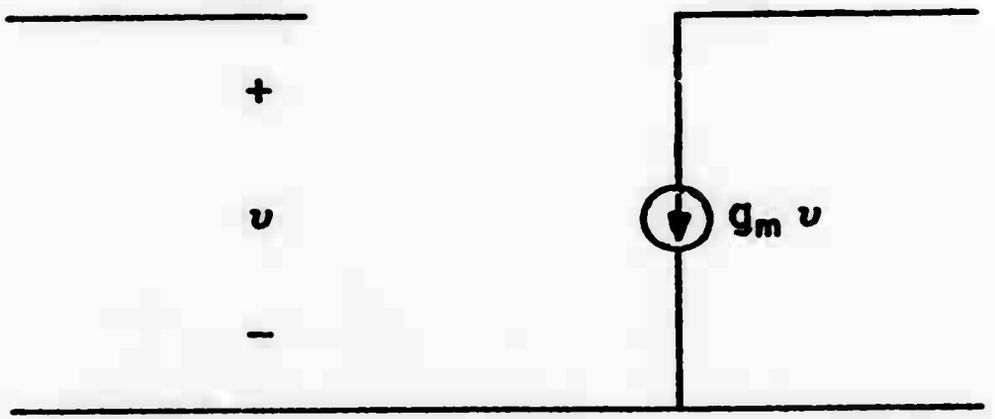

C
r_c
ηg_m
$\frac{\eta g_m}{\beta}$
$C_\mu = C'_\mu + \frac{\eta g_m}{\omega_T}$
$g_m v$
$\frac{g_m}{\omega_T} - C_\mu$
r_e
E
+
v
−
r_b
$\frac{g_m}{\beta}$
B

Figure 1



$g_m \approx 0$

Figure 2

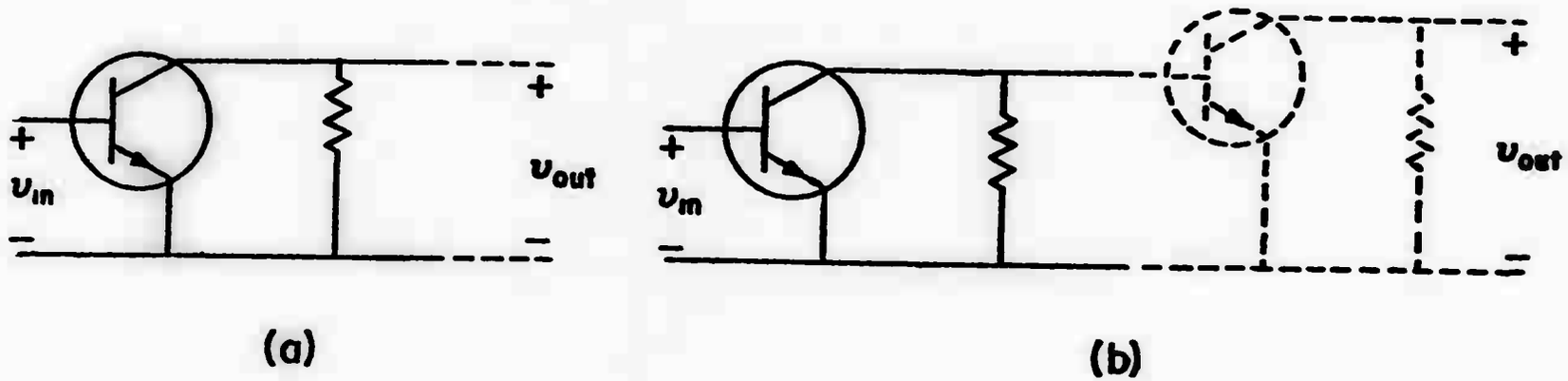
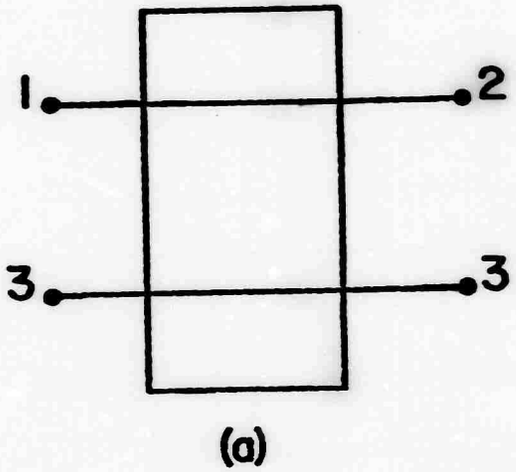
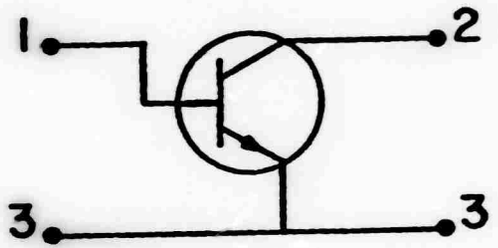


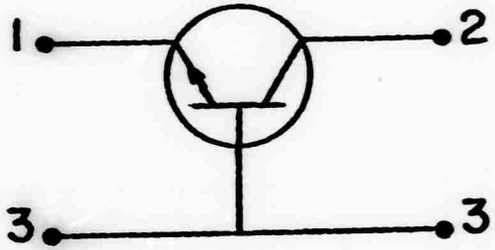
Figure 3



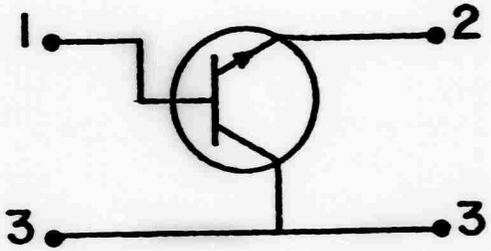
(a)



(b)



(c)



(d)

Figure 4

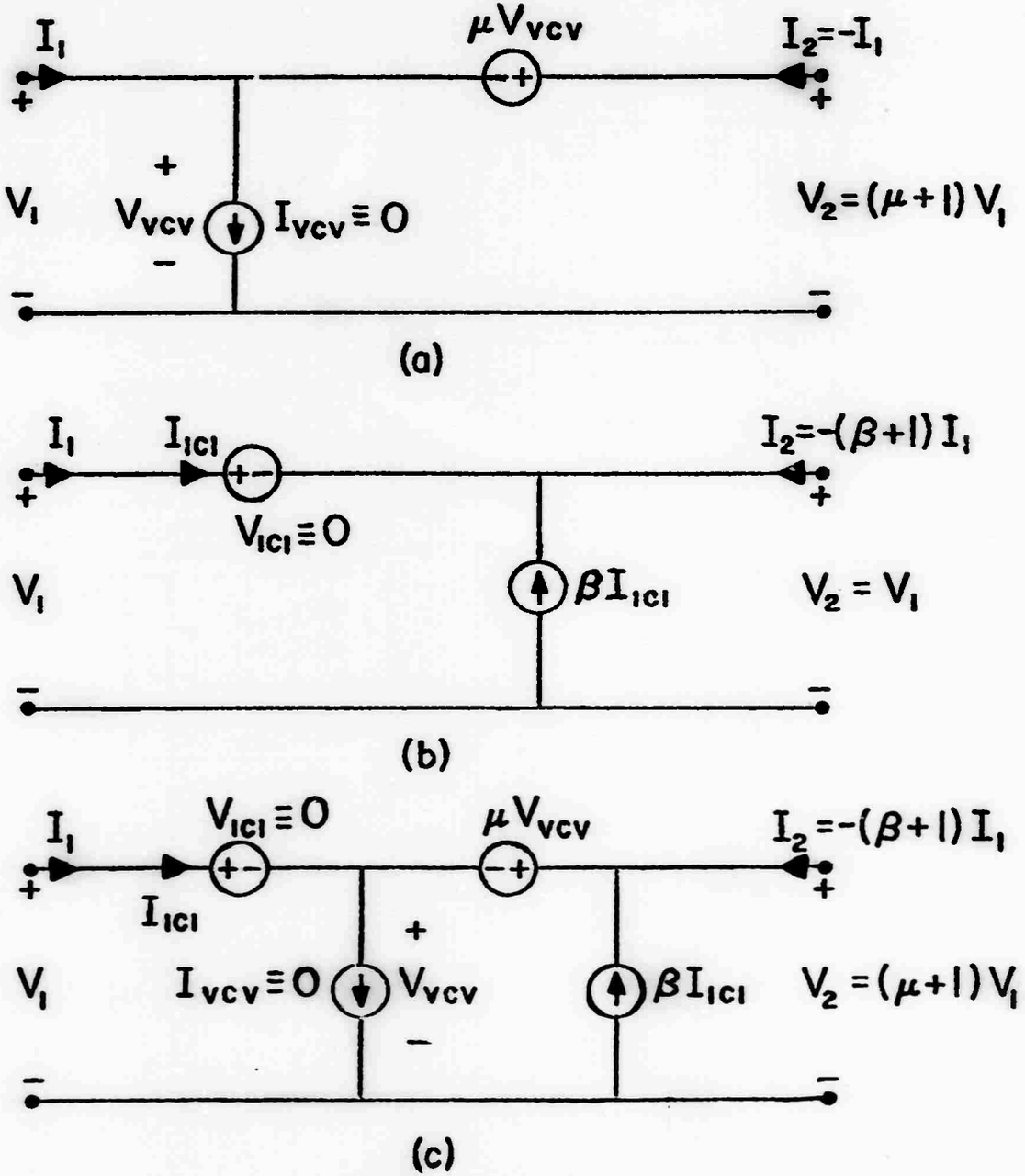

I1
μVvcv
I2=-I1
+
+
+
V1
Vvcv
Ivcv ≡ 0
V2=(μ+1) V1
−
−
−
(a)
I1
I1c1
I2=-(β+1) I1
+
+
+
Vic1 ≡ 0
V1
βIic1
V2 = V1
−
−
−
(b)
I1
Vic1 ≡ 0
μVvcv
I2 =-(β+1) I1
+
+
+
Iic1
+
V1
Ivcv ≡ 0
Vvcv
βIic1
V2 =(μ+1) V1
−
−
−
(c)

Figure 5

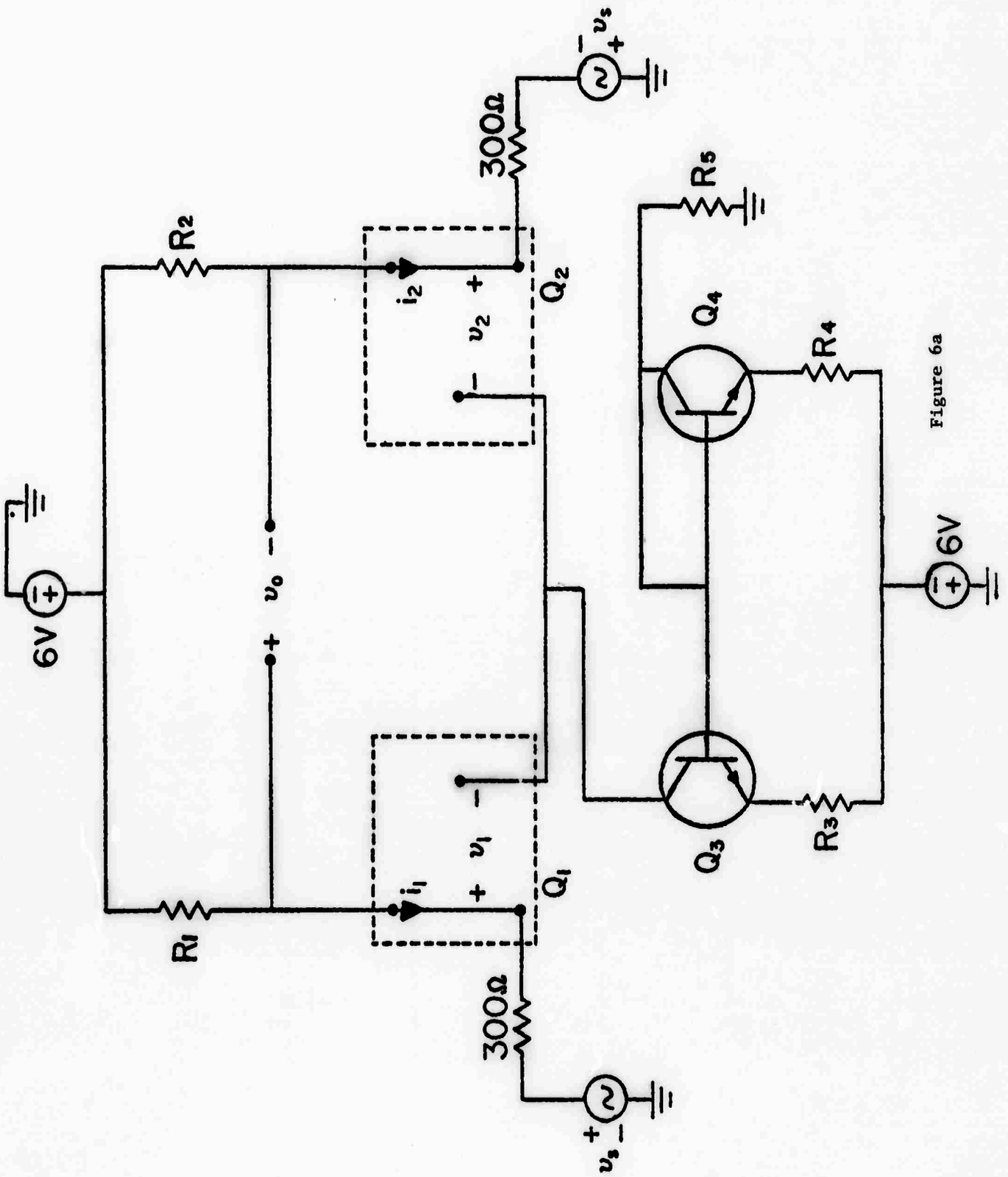


Figure 6a

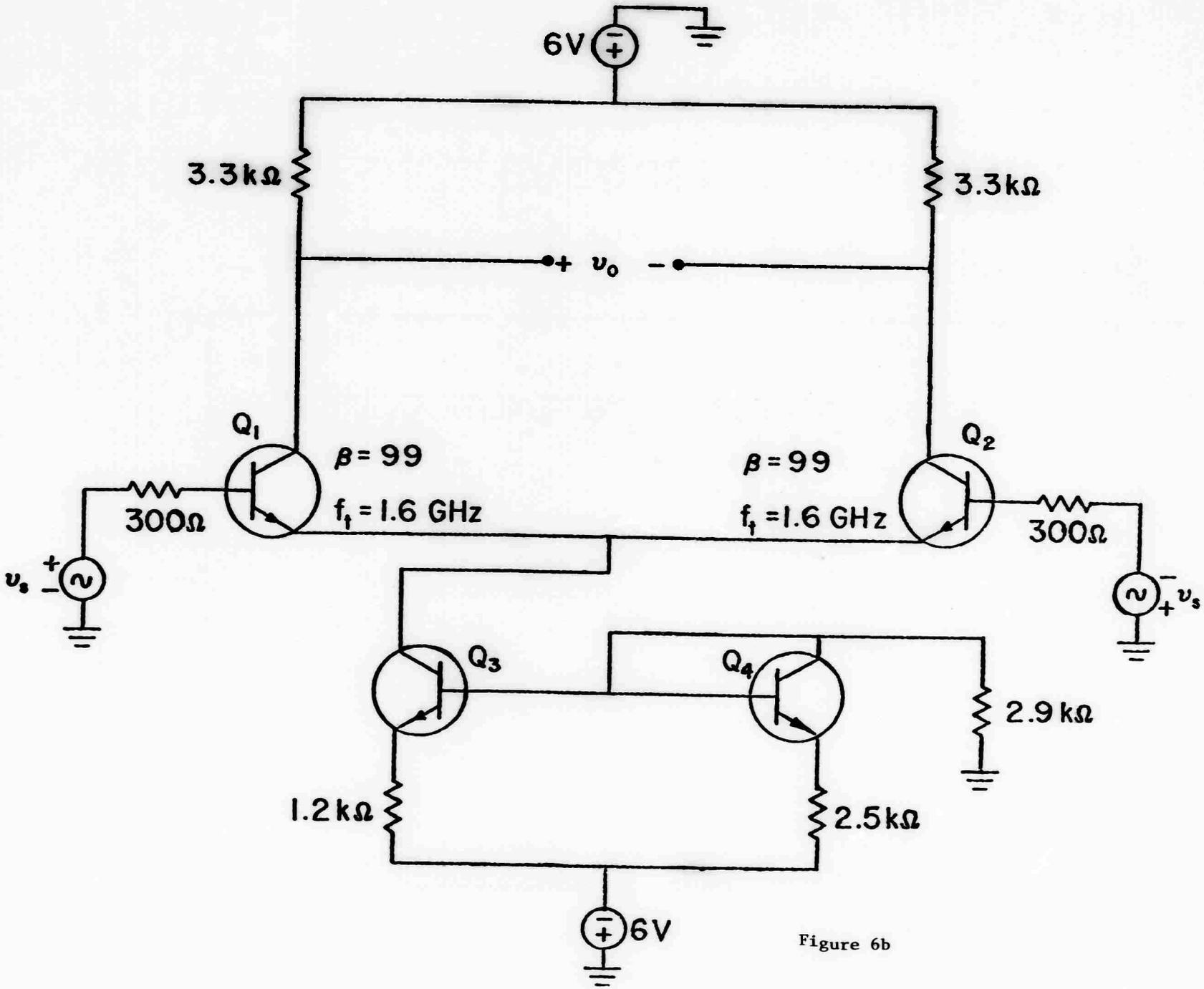

6V
3.3kΩ
3.3kΩ
+ v_o −
Q_1
$\beta = 99$
$f_t = 1.6$ GHz
300Ω
v_s
Q_2
$\beta = 99$
$f_t = 1.6$ GHz
300Ω
v_s
Q_3
Q_4
2.9kΩ
1.2kΩ
2.5kΩ
6V

Figure 6b

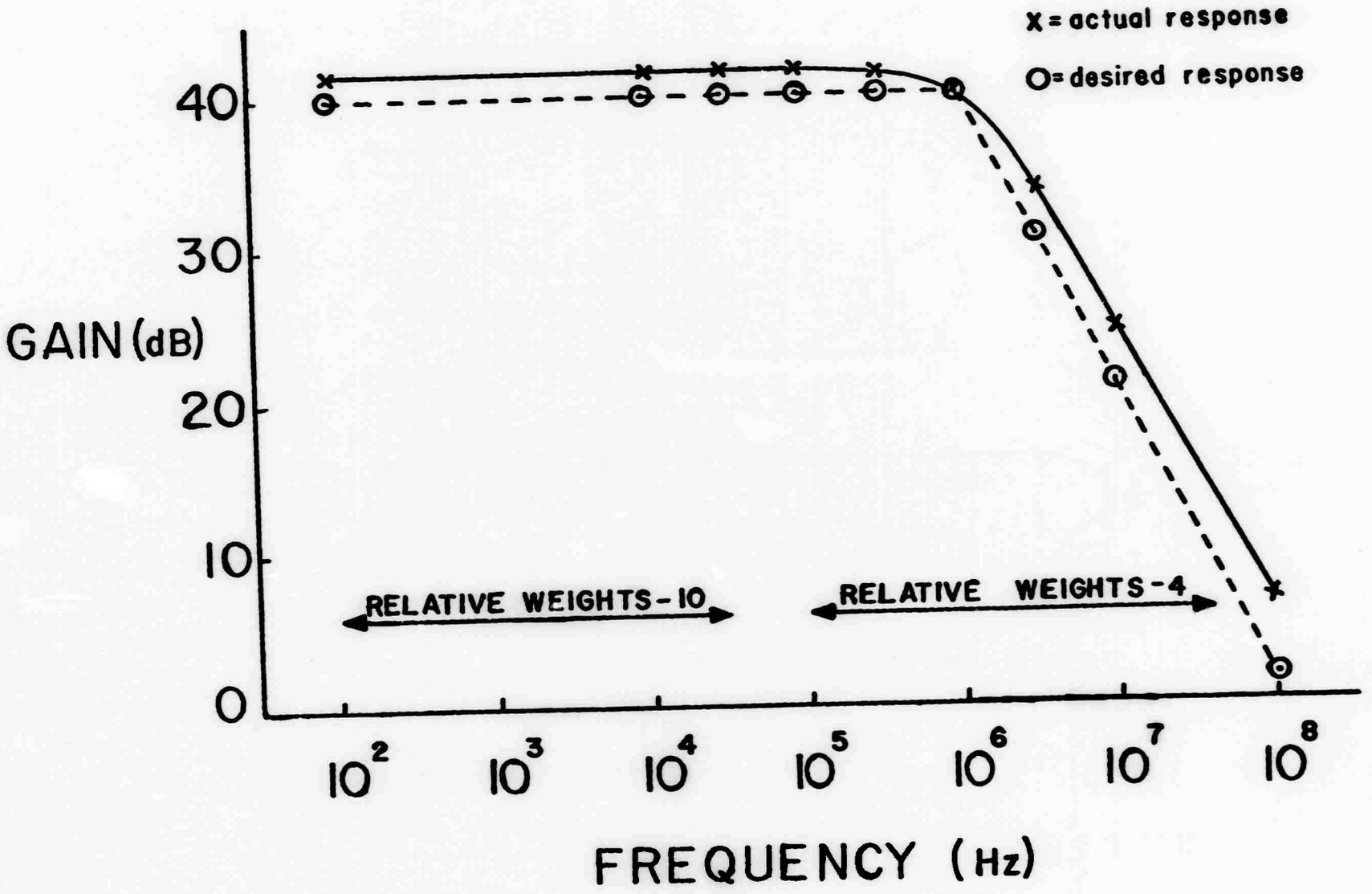

X = actual response
O = desired response
40
30
20
10
0
GAIN (dB)
RELATIVE WEIGHTS - 10
RELATIVE WEIGHTS - 4
10^2
10^3
10^4
10^5
10^6
10^7
10^8
FREQUENCY (Hz)

Figure 6c

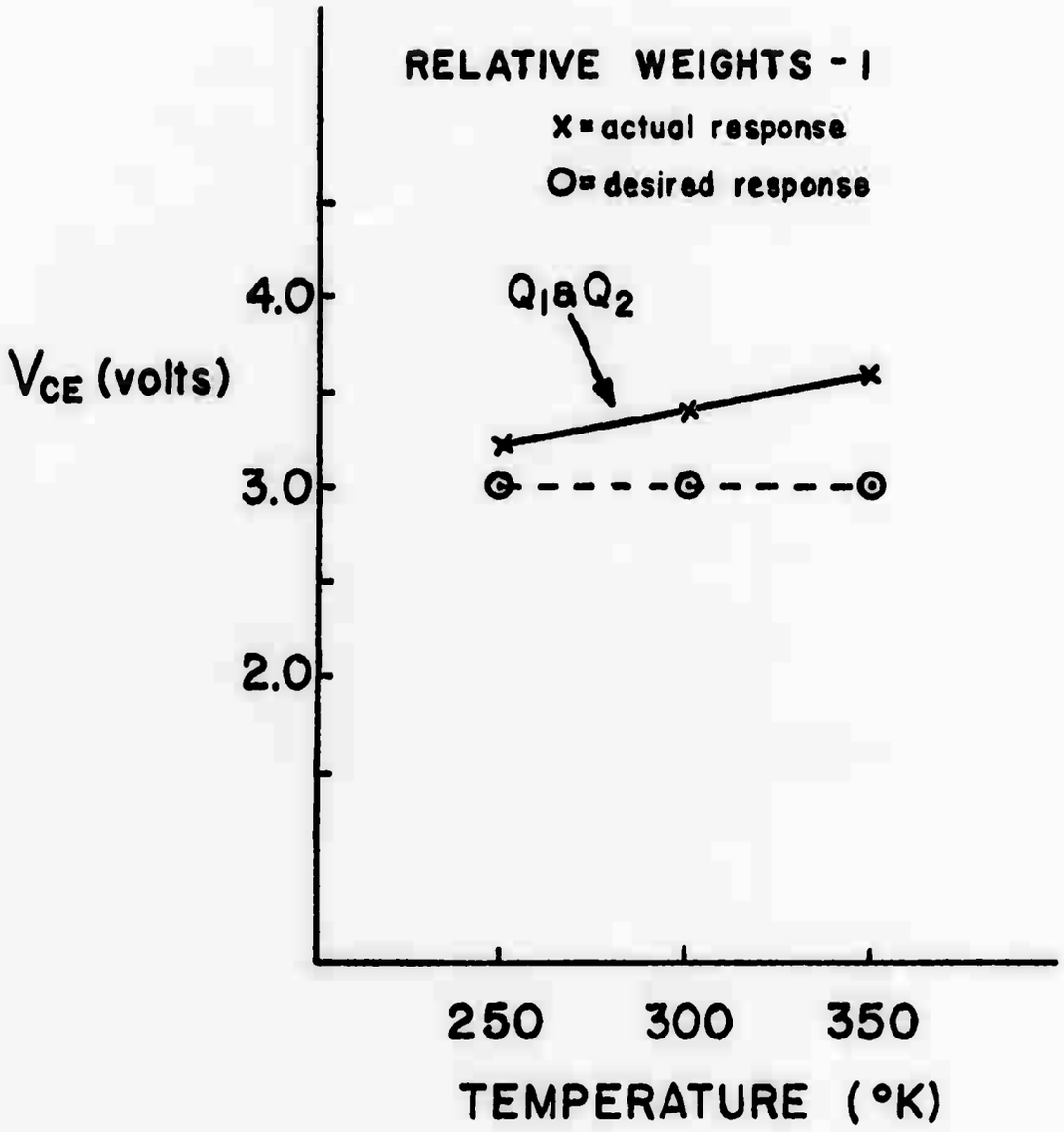

RELATIVE WEIGHTS - 1
X = actual response
O = desired response
Q_1 & Q_2
4.0
3.0
2.0
V_{CE} (volts)
250
300
350
TEMPERATURE (°K)

Figure 6d

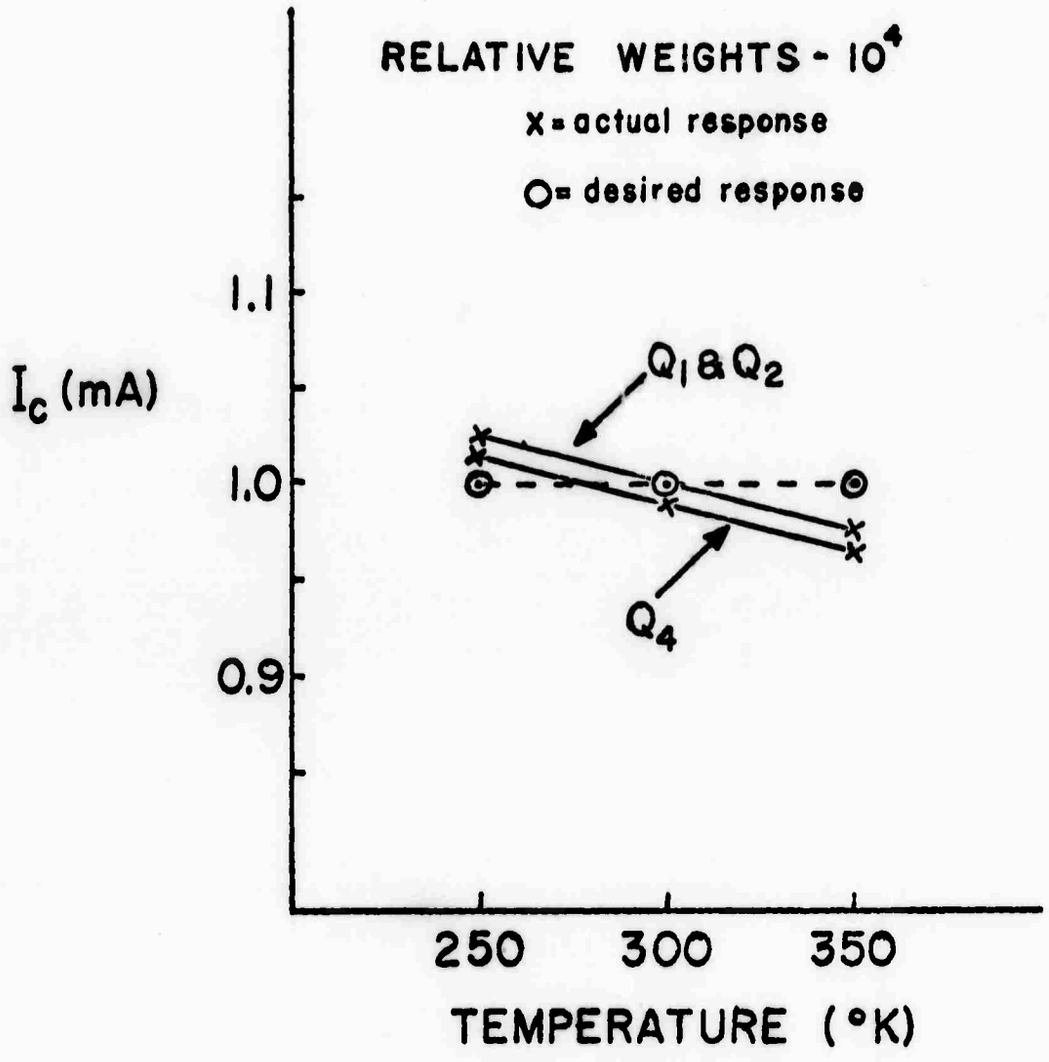
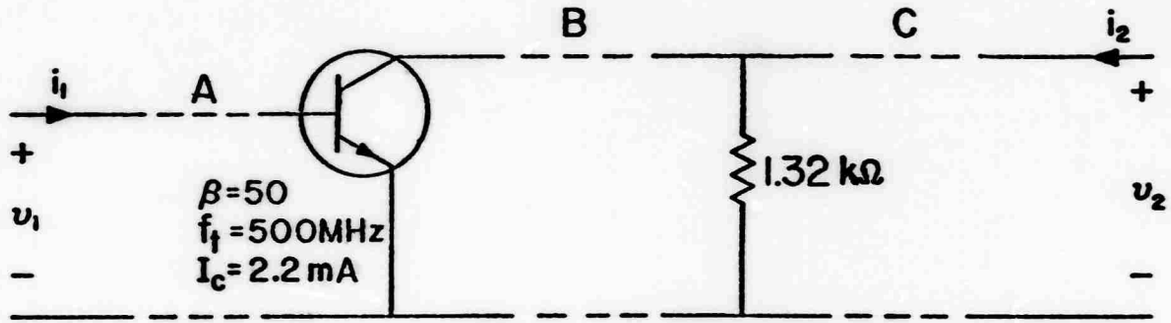

RELATIVE WEIGHTS - 10^4
X = actual response
O = desired response
I_c (mA)
1.1
1.0
0.9
Q_1 & Q_2
Q_4
250
300
350
TEMPERATURE (°K)

Figure 6e

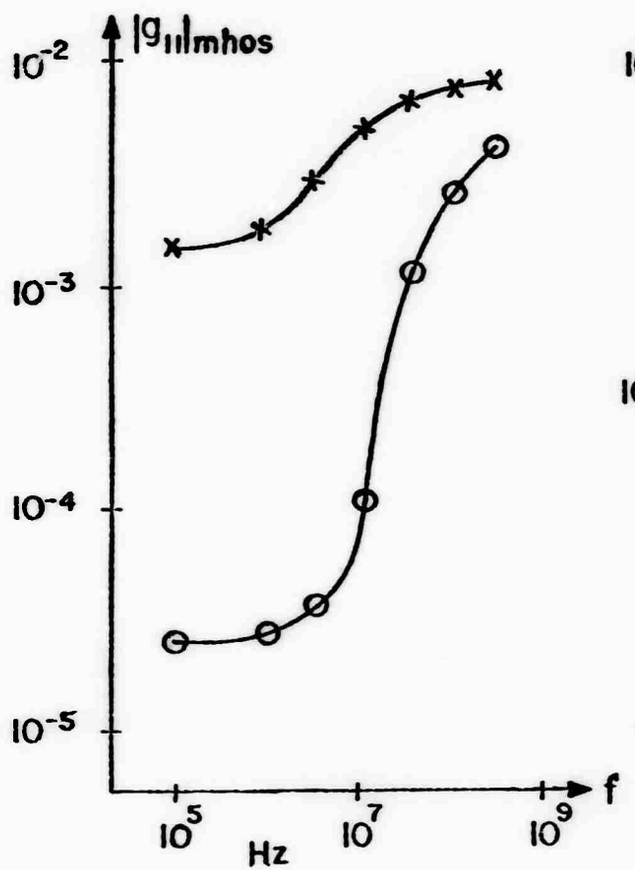


(a)

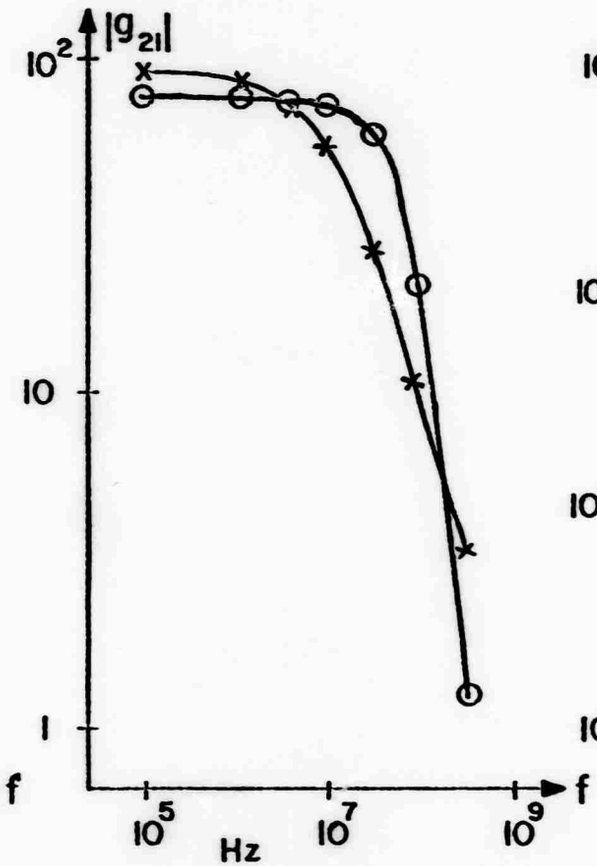
GRADIENTS

	A	B	C
A_v	1524	-597	2476
A_i	-110	2014	-1060

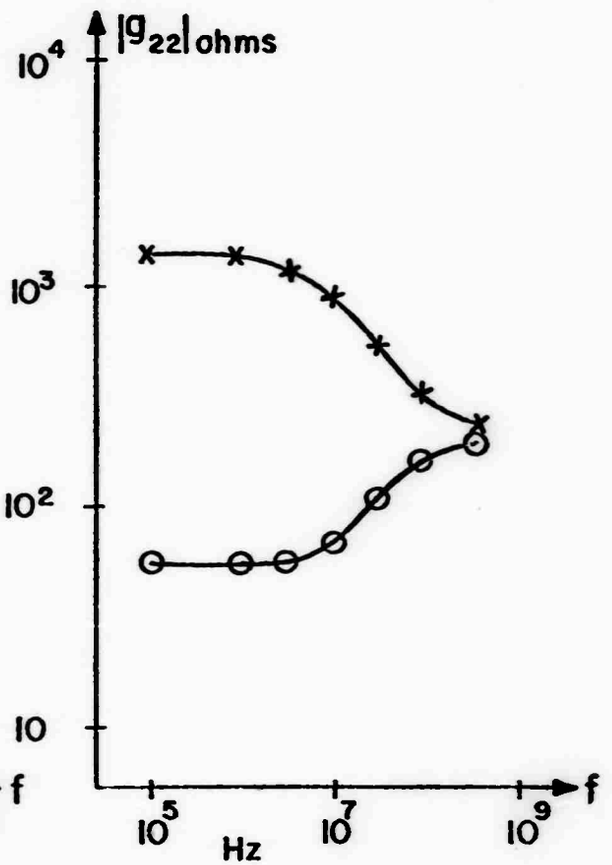
x - actual response
O - desired response



(b)

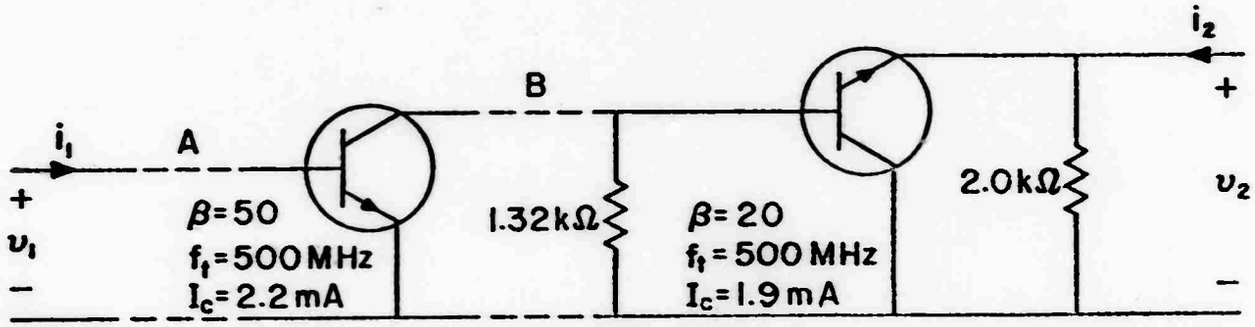


(c)



(d)

Figure 7

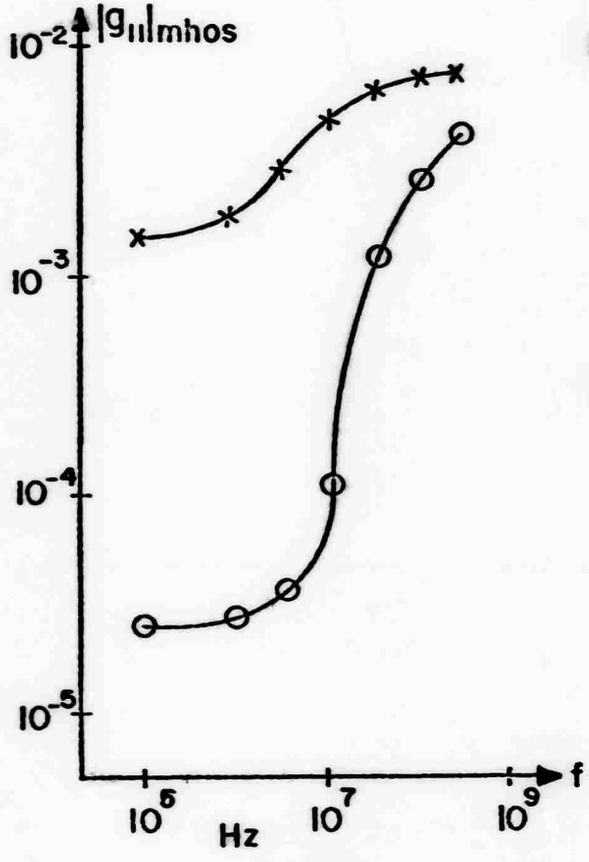


(a)

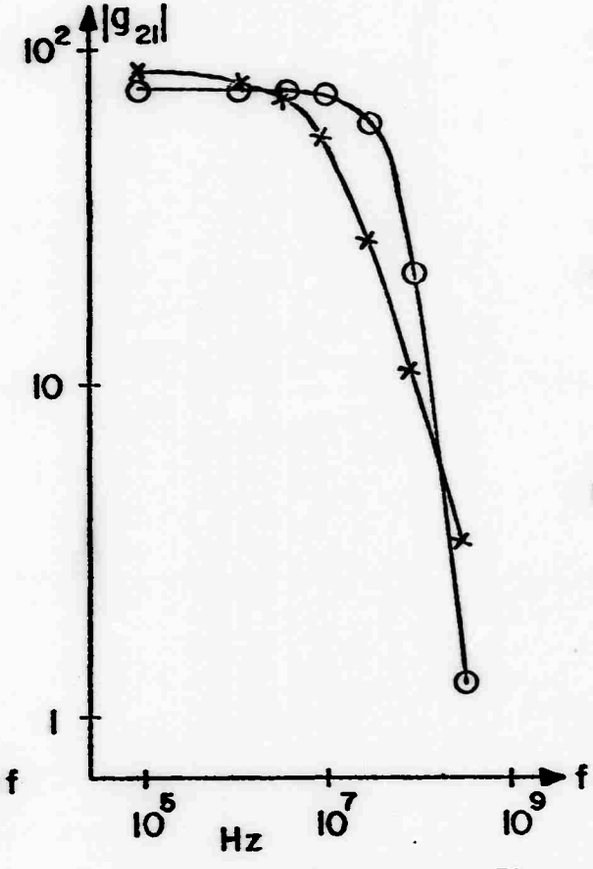
GRADIENTS

	A	B
A_v	-244	-681
A_i	-107	335

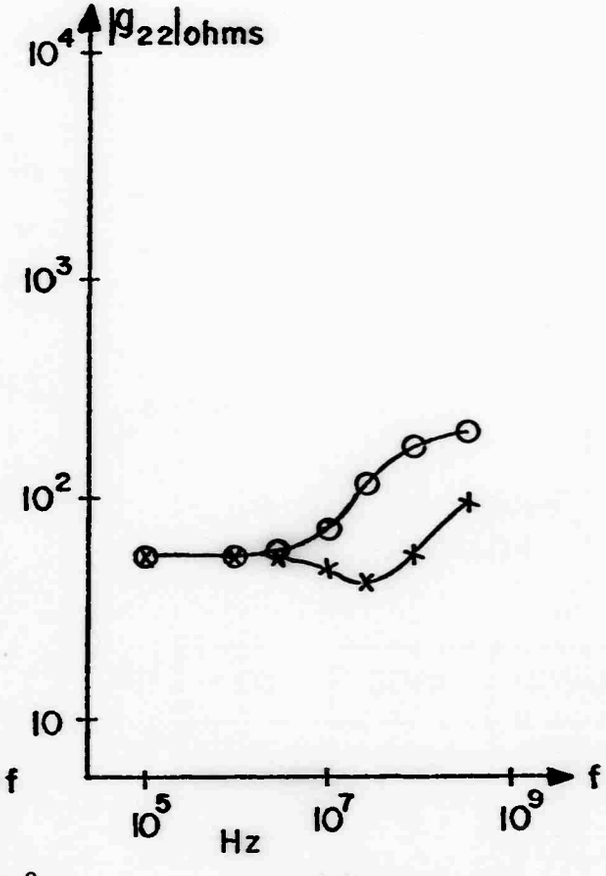
x - actual response
O - desired response



(b)



(c)



(d)

Figure 8

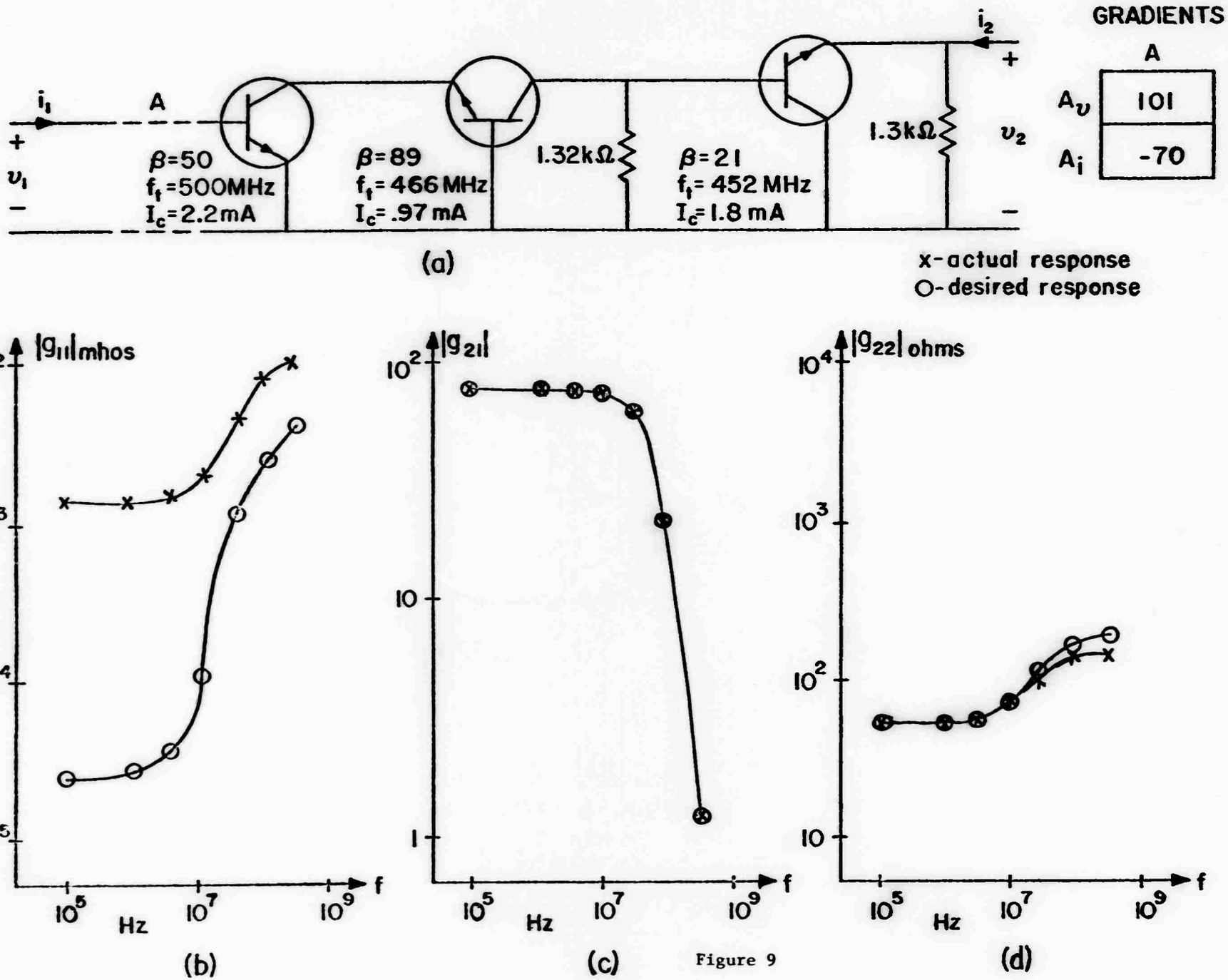

GRADIENTS
A
A_v 101
A_i -70
i_1
i_2
A
+
v_1
−
+
v_2
−
β=50
f_t=500MHz
I_c=2.2mA
β=89
f_t=466 MHz
I_c=.97 mA
1.32kΩ
β=21
f_t=452 MHz
I_c=1.8 mA
1.3kΩ
(a)
x-actual response
O-desired response
$|g_{11}|$ mhos
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^5
Hz
10^7
10^9
f
(b)
$|g_{21}|$
10^2
10
1
10^5
Hz
10^7
10^9
f
(c)
$|g_{22}|$ ohms
10^4
10^3
10^2
10
10^5
Hz
10^7
10^9
f
(d)

Figure 9

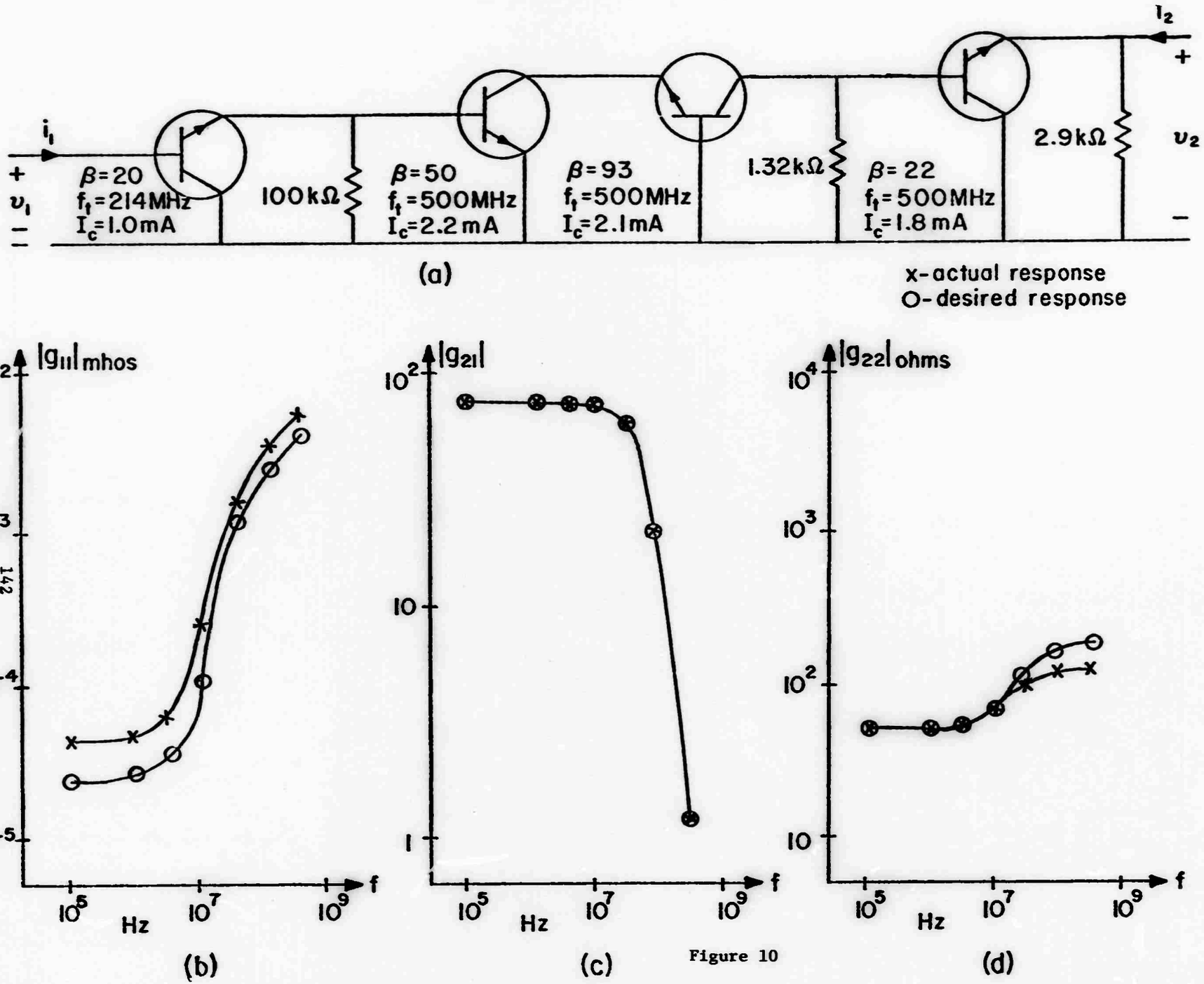


Figure 10

142

III. Insulating and Semiconducting Glasses (E. R. Chenette, R. W. Gould, L. L. Hench, J. J. Hren)

A. THE EFFECT OF CRYSTALLIZATION ON THE STRENGTH OF Li_2O-SiO_2 GLASS-CERAMICS (S. W. Freiman and L. L. Hench)

Introduction

One of the major reasons for the interest in glass-ceramics is the fact that these materials frequently possess mechanical properties superior to those of the glass from which they were formed. The mechanical property of primary interest is the fracture strength, and it is this property that will be discussed in this paper.

Several investigators have reported on the effect of crystallization on the mechanical properties of glass-ceramics.[1-6] Most of these studies, however, have been concerned only with the measurement of the fracture strength of the fully crystalline product. However, it is only through a study of the strength-microstructure relationship during the crystallization process, that one can fully characterize the effect of the second phase crystals on the strength of the material.

In order to measure independently the many variables which may affect the strength of polyphase materials, investigators have studied several glasses to which were added a given volume fraction of crystalline particles of a known size and shape. These composites were usually made by hot pressing the particles and glass.

Fulrath,[7] Grossman,[8] and Binns[9] showed that the presence of internal stresses may decrease the strength of two-phase compacts, although increases in strength should also be possible. It was demonstrated by Nason,[10] Studt and Fulrath,[11] and Jacobson[12] that chemical reactions at the glass-particle interface could influence the strength by changing the bond strength. Frey and Mackenzie[13] showed the influence of variations in elastic modulus on the strength of glass-particle composites. They also demonstrated that the sign of the thermal expansion coefficient difference between the crystals and the glass had no effect on the strength-volume fraction relationship although theoretically it should have. Hasselman and Fulrath[14] concluded that the strength of the glass-particle composites was dependent on the mean free path between particles, and also pointed out[15] that differences in the elastic properties of the glass and crystals would produce stress concentrations which would be deleterious to the strength.

The theoretical analyses obtained from these model glass-particle systems will be used to explain the experimentally measured strength dependencies of the partially crystalline glass-ceramics investigated in this program.

Experimental Procedure

The glass used in this investigation was melted for 24 hours in a platinum crucible. The melting temperatures, along with the subsequent nucleation, crystallization, and annealing conditions are given in Table 1. The kinetics of crystallization and representative micrographs of the non-Pt containing glasses are described in detail by Freiman and Hench.[16] Colloidal particles in the samples containing 0.2 wt.% Pt are evident in electron micrographs shown in Figure 1.

Specimens were cast as 1/2" diameter, 1" long cylinders. Sections, 1/2" long, were cut from the cylinders for use in the diametral compression test.[17] After crystallization, the sections were polished using a special polishing rig designed

Table 1

Sample Series	Composition (mole %) SiO_2	Li_2O	Pt	Melting Temp. (°C)	Annealing Treatment After Cast	After Cryst.	Nucleation Treatment at 475°C (hr)	Crystallization Temp. (°C)
A	66.7	33.3	0.0	1350	1 hr-300°C	20 min-430°C Furnace cool	3	575
B	66.4	33.1	~0.5	1350	"	"	0	"
C	66.7	33.3	0.0	1350	"	"	24	"
D	73.6	26.4	0.0	1415	"	"	3	"

to insure that the faces of the samples were parallel. Using this apparatus, it was possible to polish through 120, 180, 320, and 600 grit SiC papers while maintaining the parallelness of the sample ends within ±2 mils. The sections were then polished with 0.3 micron Al_2O_3 on Buehler microcloth.

The diametral compression test was developed by Caniero[17] as a test for the fracture strength of concrete. The test has been modified for use on porcelain whitewares,[18] and glass.[19] The test involves applying a compressive load along the diameter of a relatively short cylinder. Despite the surface compressive stress, samples always fail in tension along the highly stressed diameter. Loading pads of construction paper were placed between the specimens and the loading platens to prevent large stress concentrations from occurring. The expression relating applied load to fracture strength is given by equation (1).

$$\sigma = \frac{2P}{\pi dt} \tag{1}$$

where σ is the fracture strength, P is the applied load, d is the specimen diameter, and t is the specimen length. A preliminary study[20] showed that the testing method was nearly insensitive to loading rate and atmosphere. However, to eliminate any possible effects of static fatigue in the comparative analysis the loading rate was kept constant at 0.2 in/min, and all of the tests were conducted in air at room temperature under monitored temperature and humidity conditions.

The dynamic elastic modulus of the glass-ceramics were determined by a sonic resonance technique employing a Bruehl and Kjar apparatus. Sample densities were measured to an accuracy of 0.001 g/cc using water displacement techniques. Dilatometric softening points and thermal expansion coefficients were determined using an automatic recording Orton dilatometer with a heating rate of 10°C/min.

Volume fraction measurements were made on the partially crystallized specimens by means of quantitative x-ray diffraction as described previously by the authors.[16] Particle size

measurements were made from petrographic sections of the glass-ceramics. Pre-shadowed, direct carbon replicas were made from several of the etched fracture surfaces of the samples for examination in the electron microscope.

Fracture Strength

Three types of fracture were observed in these materials, when tested in diametral compression. At a low volume fraction of crystals, the samples exploded into many pieces at fracture following an elastic wave behavior discussed by Kenny.[19] The more highly crystalline samples either fractured along the diameter of the samples, or broke in the triple cleft pattern described by Rudnick, Hunter and Holden.[21] No evidence of shear or compressive failure at the sample surfaces was ever observed. Even though some unevenness of the cylindrical surfaces of the samples was present, observations of the loading pads after fracture had occurred showed that pressure had been applied uniformly over the length of the samples.

The fracture strengths of the various glass-ceramic compositions listed in Table 1 are plotted versus the volume fraction of $Li_2Si_2O_5$ crystals in Figure 2. Each point represents the arithmetic average of about twelve specimens. The error in the determination of the average strength was approximately ±10% within 95% confidence limits.

In order to establish the dependence of the strength of the glass-ceramics composite on the mean free path (λ) between second phase particles, fracture strength was plotted versus $\lambda^{-1/2}$ in Figure 3, where[22]

$$\lambda = \frac{4(k_1 \cdot b)(1-V_v)}{3V_v} \qquad (2)$$

V_v is the volume fraction of crystals, b is one-half the minor axis of the ellipsoidal crystal, and k_1 is a constant determined by the ratio of the major to the minor axis of the ellipse;[23] this ratio remained constant at 1.6 through all treatments.

Modulus of Elasticity

The measured values of the moduli of the partially crystalline Li_2O-$2SiO_2$ samples were compared with those predicted by theory.[24] The measured values of E and the values of the Poisson's ratio (ν) for the $Li_2Si_2O_5$ crystals determined by Bokin et al.,[25] were used in Hashin's equations to obtain the theoretical variation of modulus with crystallinity. This theoretical curve is shown as the dashed line in Figure 4. Considering the fact that the compositions of the initial glasses are different, the values of the moduli for the glass and the crystals agree fairly well with moduli determined by Bokin et al.[25]

Coefficient of Thermal Expansion

The coefficient of thermal expansion determined at several percentages of crystallization is plotted versus the volume fraction of crystals in Figure 5. The experimental curve is compared with that theoretically predicted by Kingery[26] for the coefficient of thermal expansion in multiphase bodies. The expression relating thermal expansion coefficient to the relative amount of the two phases is presented below.

$$\alpha = \frac{\alpha_1\omega_1K_1/\rho_1 + \alpha_2\omega_2K_2/\rho_2}{K_1\omega_1/\rho_1 + K_2\omega_2/\rho_2} \qquad (3)$$

where α, K, ω, and ρ are respectively, the thermal expansion coefficient, bulk modulus, weight fraction and density of a given phase.

If the Poisson's ratio of the two phases is nearly the same, as was true in this study, the Young's modulus can be used in place of bulk modulus. It can be seen that agreement between the theoretically predicted and the experimental values is reasonably good.

Density

Density measurements revealed that increasing the nucleation time increased the density of the glass from 2.30 gm/cc (3 hours) to 2.33 gm/cc (24 hours), to 2.34 gm/cc (48 hours), confirming that rearrangement processes occur during this treatment. It was found that the measured densities of partially crystalline samples were less than 100% theoretical density. The theoretical densities were determined on the basis of a glass density of 2.35 gm/cc, a crystal density of 2.45 gm/cc, and the assumption that density should vary directly as the volume fraction of crystals. The glass density value of 2.35 gm/cc was obtained by measuring a glass sample that had been heat treated at 575°C for a time just prior to the appearance of detectable crystallization.[16]

The decrease in sample density with decreasing nucleation treatment is also demonstrated by the micrographs given in Figures 6 and 7. The porosity takes the form of cracks which will later be shown to be a major factor in causing decreases in strength in certain specimens.

Discussion of Results

The strength of two phase composites is dependent on two classes of factors, those which will increase the strength, and chose which will decrease it. Each class will be considered separately.

Strength Increase

The ways in which the presence of second phase crystals can increase the strength of a composite are listed below.

1. Increase in effective Young's modulus with percent crystallinity; strength increases in proportion to $\sqrt{E}$.

$$\sigma \propto \sqrt{\frac{4E\gamma_o}{\pi c}} \qquad \begin{array}{l} \gamma_o = \text{fracture energy of glass} \\ c = \text{flaw size } (\mu) \end{array} \qquad (4)$$

2. Random reduction of flaw length.[14]

$$\sigma \propto \sigma_o (1-V_v)^{-1/2}$$

3. Decrease in particle size, p.[27]

$$\sigma \propto (p)^{-1/2} \qquad (5)$$

4. Reduction in flaw length due to reduction in mean free path between particles.[14]

$$\sigma \propto (\gamma)^{-1/2} \qquad (6)$$

5. Increase in energy needed to propagate a crack.[28]

$$\sigma \propto \sqrt{E(G_c)} \qquad (7)$$

G_c = fracture energy of the glass-ceramic

It would be desirable to consider each factor separately, but this is impossible since there is always some interaction between the variables. It will be shown later that this interaction can also work to cause a decrease in strength.

The first three factors will account for only part of the strengthening of the glass-ceramics shown in Figure 2. For example, at the point at which a strength of 11,300 psi at 10% crystallinity is measured, a combination of factors (1) and (2) predicts a strength of only 8,000 psi.

Considering factor (3), there is approximately a 60% decrease in particle size in comparing the 3-hour and 48-hour nucleation treatments which could produce an additional strengthening of 2,500 psi. However, there is no difference in average particle size between the specimens nucleated for 24 and 48 hours, and there is a marked difference in the strength obtained. Consequently, it can be concluded that strengthening from factor (3) must be relatively unimportant in these samples.

In order to compare the strengthening influence of factors (4) and (5) correctly, it is first necessary to separate factor

(1) from both Hasselman and Fulrath*[14] and Langes'[28] equations. When this is done equations (8) and (9) result:

$$\text{Hasselman and Fulrath:} \quad \sigma/\sigma_o = (c/\lambda)^{1/2} \tag{8}$$

$$\text{Lange:} \quad \sigma/\sigma_o = (1 + c/3\lambda)^{1/2} \tag{9}$$

σ and σ_o denote strength values for the crystallized and uncrystallized glass, respectively, at given mean free paths between particles, λ.

Both equations are plotted as functions of $\lambda^{-1/2}$ in Figure 3 using a constant glass flaw size of 85 microns. The glass flaw size was calculated from the Griffith relation, equation (4). In the calculation γ_o is assumed to be 3,000 ergs/cm^2,[29] E_o was measured to be 7.5×10^6 psi, and the strength of the uncrystallized glass was measured to be 7.2×10^3 psi.

Experimental strength values of the various Li_2O-SiO_2 glass-ceramics evaluated in the study were reduced by the quantity $(E_o/E)^{1/2}$, in order to compare them with equations (8) and (9) and the reduced values are plotted on Figure 3. Several observations are apparent. Neither theory, as presented in equations (8) and (9), fits the data well. Departure from the theories becomes more pronounced as the mean free path between the particles increases. Strength reduction eventually occurs for the glass-ceramics requiring a different viewpoint, discussed in the next section.

In general, the Lange theory appears to fit the results more closely. The fit of equation (9) would be especially good if the 3 in the denominator of equation (9) were 0.9 instead; i.e.,

$$\sigma/\sigma_o = (1 + \frac{c}{0.9\lambda})^{1/2} \tag{10}$$

---

*Elastic modulus changes associated with addition of Al_2O_3 particles to glass were not taken into account in the experimental evaluation of equation (8) in the previous work cited in reference 14.

Since equation (9) involves the assumption that the cracks between particles are spherical and since the micrographs in Figures 6 and 7 show that flaws in the glass-ceramics are elliptical in shape, it is not unreasonable that equation (10) would be appropriate.

The shape of the crack front as it passes between the inhomogeneities is most likely modified to an elliptical form by the residual stress field around the particles, as discussed in the next section. Consequently, less energy is required to form the increased crack front. The reduced fracture energy associated with the change in wave front shape is represented by the difference between equations (9) and (10).

If the flaw size of the glassy phase were underestimated, the slope of equation (9) would also be increased. However, Weiderhorn's work[29] and recent studies[30] suggest that the 3,000 ergs/cm^2 value of γ_o used in calculating the flaw size is on the high side if it is in error. Equation (8) does not appear to fit the data well under any reasonable modifications. A plot of equation (10) yields the dashed curve in Figure 3.

At a critical mean free path between the particles, the strength data for each of the series of glass-ceramics breaks from the modified Lange curve, as shown by the dotted curves. The critical mean free path becomes progressively smaller as the nucleation treatment increases. Phenomena which can account for the strength reduction and thermal history dependence of a critical mean free path is discussed in the next section.

Strength Decrease

As in the case of strengthening mechanisms, there are a number of factors which can contribute to the reduction in strength of the glass-ceramics at high volume fractions. These factors are listed below:

1. Stresses around particles due to thermal expansion coefficient differences between crystals and glass.

2. Stress concentrations around crystals due to a difference in elastic moduli between phases.
3. Stresses due to a difference in volume between crystals and glass.

The stresses due to thermal expansion coefficient differences were calculated from Selsing's[31] expression below:

$$\sigma_r = -2\sigma_t = -PR^3/r^3 \qquad (11)$$

$$= \frac{(\alpha_p - \alpha_g)\,\Delta T}{\left(\frac{1+\nu_g}{2E_g}\right) + \left(\frac{1-2\nu_p}{E_p}\right)}$$

where σ_r is the radial stress, σ_t the tangential stress, R the particle radius, and r the distance from the particle center. The subscript g refers to the glassy matrix, whereas subscript p refers to the particle. The measured values for E and α were employed along with values of ν_g and ν_p of 0.23 obtained by Bokin et al.[25] A maximum radial compressive stress of 8,000 psi and a tensile tangential stress of 4,000 psi was calculated. These stresses, however, are independent of particle size, and should also be independent of nucleation treatment, so that they cannot explain the differences in strengths observed.

The stress concentrations due to differences in elastic moduli were calculated from the expressions given by Goodier.[32] Because the elastic modulus changed by a factor of only 2 during crystallization, the stress concentration was calculated to be only 1.2. This stress is also independent of particle size.

It was concluded from the above relations that the only stresses which could produce effects which would be dependent on nucleation treatment were those stresses arising from the density difference between glass and crystals. Since a maximum density change of from 2.30 g/cc (3-hour nucleated glass) to 2.45 g/cc (crystals) can occur, a considerable radial

tensile stress can be produced during crystallization. Equation (11)[33] relates this stress to the $\Delta\rho$ between crystals and glass.

$$\sigma = \frac{E}{3(1-2\nu)} \left(1 - \frac{\rho_g}{\rho_p}\right) \qquad (12)$$

When the values for E and ν given previously and the maximum density difference given above are inserted into equation (12), a stress of about 3.3×10^5 psi is found, provided that no stress relaxation occurs. Theoretically, crystallization takes place at temperatures above the softening point of the glass, so stresses produced by crystallization should be instantaneously relieved. Experimental measurements of the dilatometric softening point, however, show that at crystal volume fractions as low as 0.3 the softening point of the specimen increased from 450°C to 600°C. This result implies that the rate of stress relief is reduced as the mean free path between crystals is decreased. Consequently, these volume differences, unlike the thermal expansion coefficient differences, can produce large radial tensile stresses in the glass which tend to cause the crystals to break away from the matrix. The cracks remain localized around individual crystals until larger volume fractions of crystals are attained. The overlap of the stress fields and crack interaction which occurs at large volume fraction (Figure 7) produces extensive cracking in the sample and gives rise to the observed decreases in strength.

As the nucleation time is increased, the degree of density mismatch between crystals and glass, and therefore the magnitude of the stress produced during the crystallization process, is reduced. Therefore, the localized cracks around the crystals are minimized as shown in Figure 6. A large strength reduction of the glass-ceramic consequently is not observed in Figure 2 at large volume fractions of crystals in well nucleated samples.

Conclusions

In summary, it has been shown that the elastic moduli of the $Li_2Si_2O_5$ glass-ceramics increased with progressive crystallization while the thermal expansion coefficient decreased. It was demonstrated that the crystals in the glass strengthened the composite by increasing the fracture energy of the material proportional to the reduction of the mean free path between particles. It was shown that the maximum degree of strengthening by the crystals is attained only in a 100% dense body, and that the lack of strengthening in the less nucleated samples can be attributed to cracking within the body. The cracking occurs because of stresses produced between the more dense crystals and the glass during crystallization.

BIBLIOGRAPHY

1. P. W. McMillan, Glass-Ceramics, Academic Press, London (1964), p. 132.

2. S. D. Stookey, "Catalyzed Crystallization of Glass in Theory and Practice," Glastech. Ber. 32K, V/1-8, Fifth International Congress on Glass (1959).

3. M. Watanabe, R. V. Caporali and R. E. Mould, "The Effect of Heat Treatment on the Strength and Abrasion Resistance of a Glass-Ceramic Material," Symposium on Nucleation and Crystallization in Glasses and Melts, M. K. Reser, G. Smith and H. Insley, eds., The American Ceramic Society, Inc., Columbus, Ohio (1962), p. 23.

4. M. Tashiro, "Chemical Compositions of Glass-Ceramics," Glass Ind., July 1966, p. 366.

5. P. W. Ownby, "A Preliminary Study of the Effect of Heat Treatment on the Strength and Microstructure of a Glass-Ceramic Material," M.S. Thesis, University of Missouri, Rolla, Missouri (1962).

6. R. F. Davis, "The Influence of Constitution and Microstructures on the Transverse Strength of a Crystallized Glass," Report AD 638-336 (1966).

7. R. M. Fulrath, "Internal Stresses in Model Ceramic Systems," J. Am. Ceram. Soc. 42, 423-29 (1959).

8. L. N. Grossman, "Fracture in Model Ceramic Systems," M.S. Thesis, University of California, Berkeley (1961).

9. D. B. Binns, "Some Physical Properties of Two Phase Crystal-Glass Composites," Science of Ceramics 1, 315-34 (1962), Academic Press.

10. D. O. Nason, "Effect of Interfacial Bonding on Strength of a Model Two-Phase System," M.S. Thesis, University of California, Berkeley (1963).

11. P. Studt and R. M. Fulrath, "Mechanical Properties and Chemical Reactivity in Mullite-Glass Systems," J. Am. Ceram. Soc. 45, 182-88 (1962).

12. L. A. Jacobson, "Strength of a Two-Phase Model System," M.S. Thesis, University of California, Berkeley (1962).

13. W. J. Frey and J. D. Mackenzie, "Mechanical Properties of Selected Glass-Crystal Composites," J. Materials Science 2, 124-30 (1967).

14. D. P. H. Hasselman and R. M. Fulrath, "Proposed Fracture Theory of a Dispersion Strengthened Glass Matrix," J. Am. Ceram. Soc. 49, 68-72 (1966).

15. D. P. H. Hasselman and R. M. Fulrath, "Micromechanical Stress Concentrations in Two-Phase Brittle-Matrix Ceramic Composites," J. Am. Ceram. Soc. 50, 399-404 (1967).

16. S. W. Freiman and L. L. Hench, "Kinetics of Crystallization in Li_2O-SiO_2 Glasses," ibid. 51, 382-87 (1968).

17. F. Caniero, Reunion des Laboratoires d'Essai des Materiaux, Paris (1947).

18. R. E. Moore, "Statistical Analysis of Fracture Stresses of Triaxial Porcelain Bodies," Ph.D. Thesis, University of Missouri, Rolla (1962).

19. W. J. Kenny, "Energy-New Surface Relationships in the Crushing of Solids," M.S. Thesis, University of Minnesota (1957).

20. S. W. Freiman, "The Relation of the Kinetics of Crystallization to the Mechanical Properties of Li_2O-SiO_2 Ceramics," Ph.D. Thesis, University of Florida (1968).

21. A. Rudnick, A. R. Hunter and F. C. Holden, "An Analysis of the Diametral-Compression Test," Materials Research and Standards 3, 283-89 (1963).

22. R. L. Fullman, "Measurement of Particle Sizes in Opaque Bodies," Trans. AIME 197 [3] 447-52 (1953).

23. R. T. DeHoff, "The Determination of the Geometric Properties of Aggregates of Constant-Size Particles X-ray Counting Measurements Made on Random Plane Sections," Trans. AIME 230, 764 (1964).

24. Z. Hashin, "Elasticity of Ceramic Systems," in Ceramic Microstructures, R. M. Fulrath and J. A. Pask, eds., J. Wiley and Sons, Inc., Newark (1966).

25. P. Ya. Bokin, A. I. Korelova, R. A. Govorova, O. S. Alekseeva and G. A. Nilsandrova, "Mechanical Properties and Microstructure of Lithium Silicate Glasses at Different Stages of Crystallization," in Structure of Glass 5, 126-33 (1965), Consultants Bureau Translation.

26. W. D. Kingery, Introduction to Ceramics, John Wiley and Sons (1960), p. 479.

27. F. P. Knudsen, "Dependence of Mechanical Strength of Brittle Polycrystalline Specimens on Porosity and Grain Size," J. Am. Ceram. Soc. 42, 376-87 (1959).

28. F. F. Lange, "Theory on Dispersion Toughening of Brittle Materials," Tech. Report No. 2, N00014-68-C-0323 (1969).

29. S. M. Weiderhorn, "Fracture Surface Energy of Soda-Lime Glass," Materials Science Research 3, 503-28 (1966), Plenum Press.

30. S. W. Freiman, G. Y. Onoda and A. G. Pincus, unpublished results.

31. J. Selsing, "Internal Stresses in Ceramics," J. Am. Ceram. Soc. 44, 419 (1961).

32. J. N. Goodier, "Concentration of Stress Around Spherical and Cylindrical Inclusions and Flaws," J. Appl. Mech. 1, 39-44 (1933).

33. S. Kumar and B. B. Nag, "Internal Stress in Crystallized Glasses in the System $MgO-Al_2O_3-SiO_2-TiO_2$," Trans. Indian Ceram. Soc. 22, 30-36 (1963).

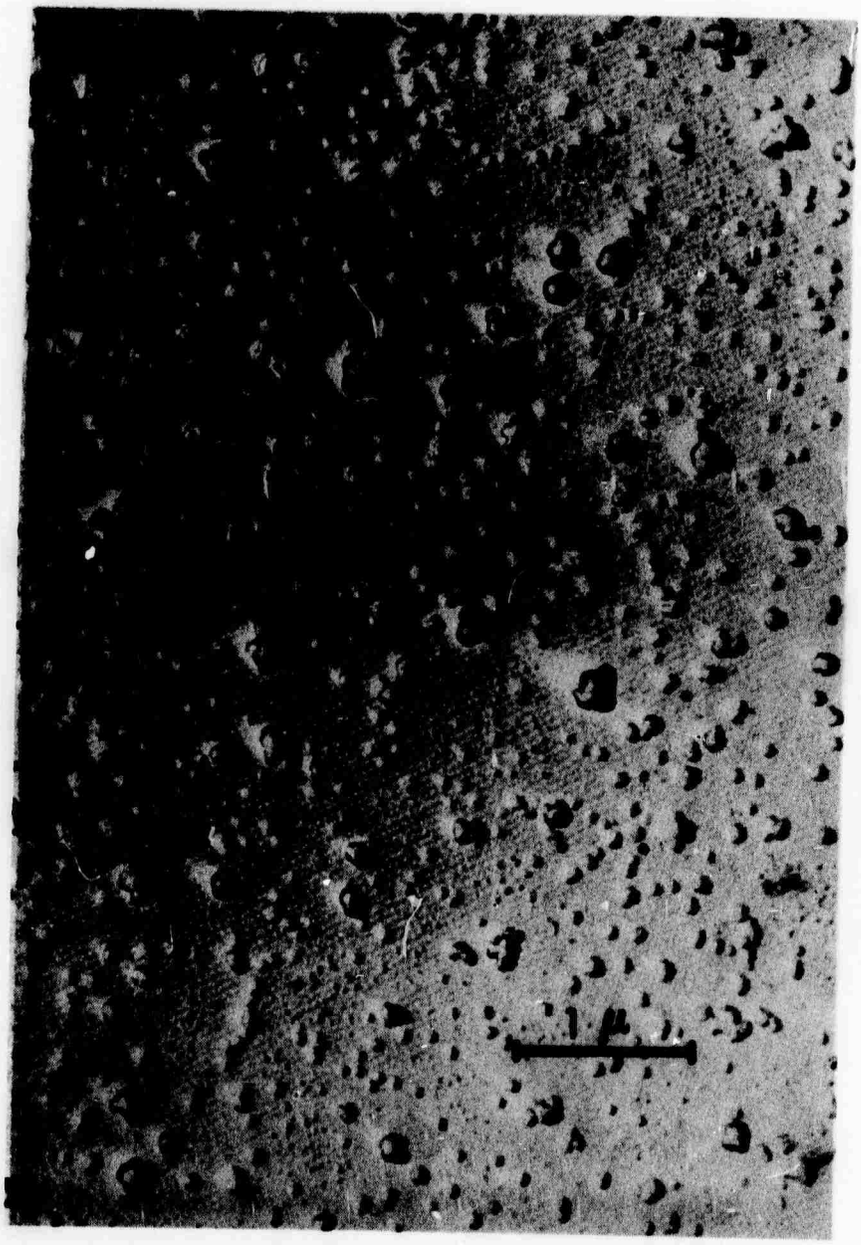


Figure 1. Fracture surface of an as cast Li_2O-$2SiO_2$ glass to which 0.25 % Pt has been added.

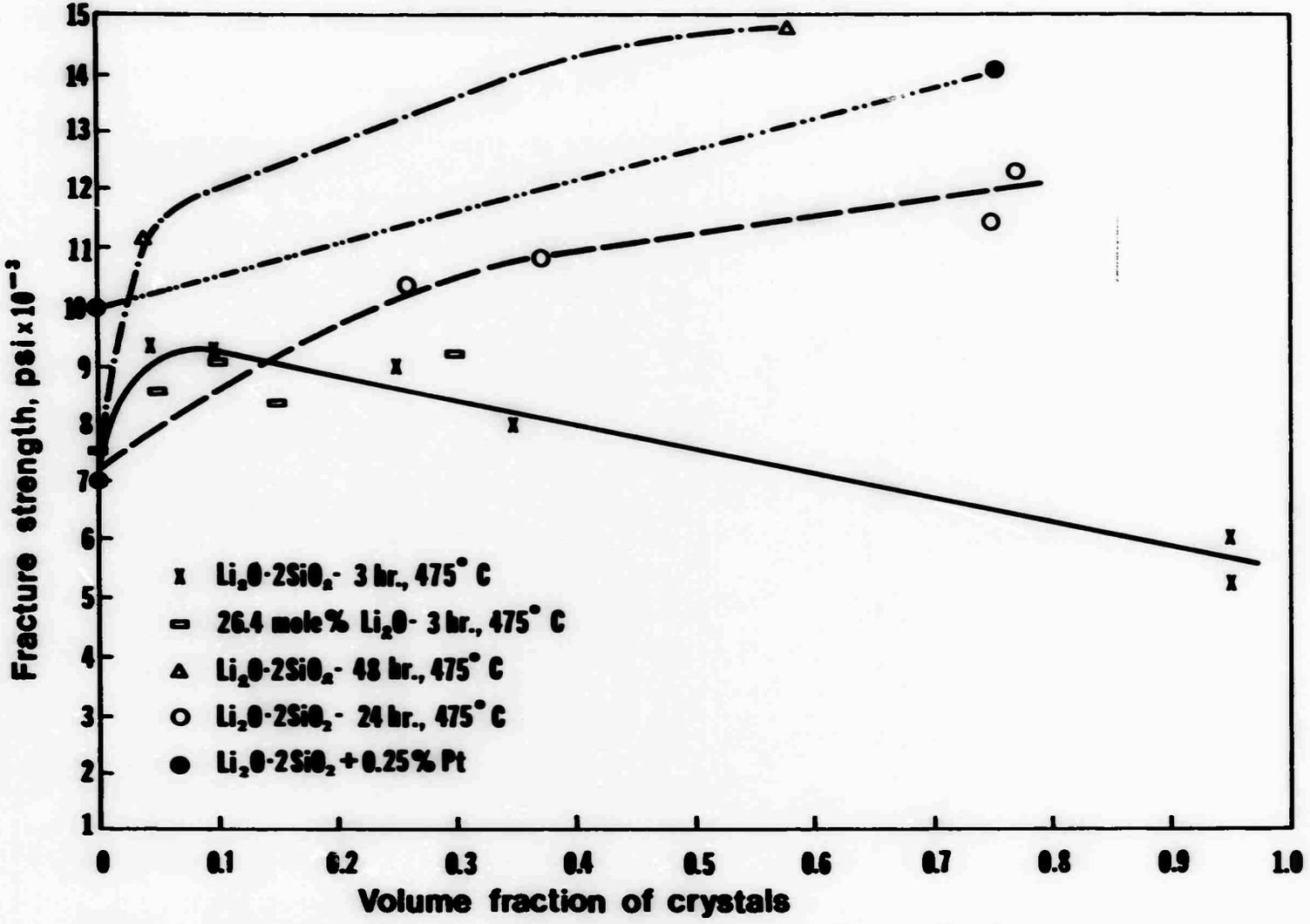


Figure 2. Variation in fracture strength of various Li_2O-SiO_2 glass-ceramics as a function of volume fraction of crystals present.

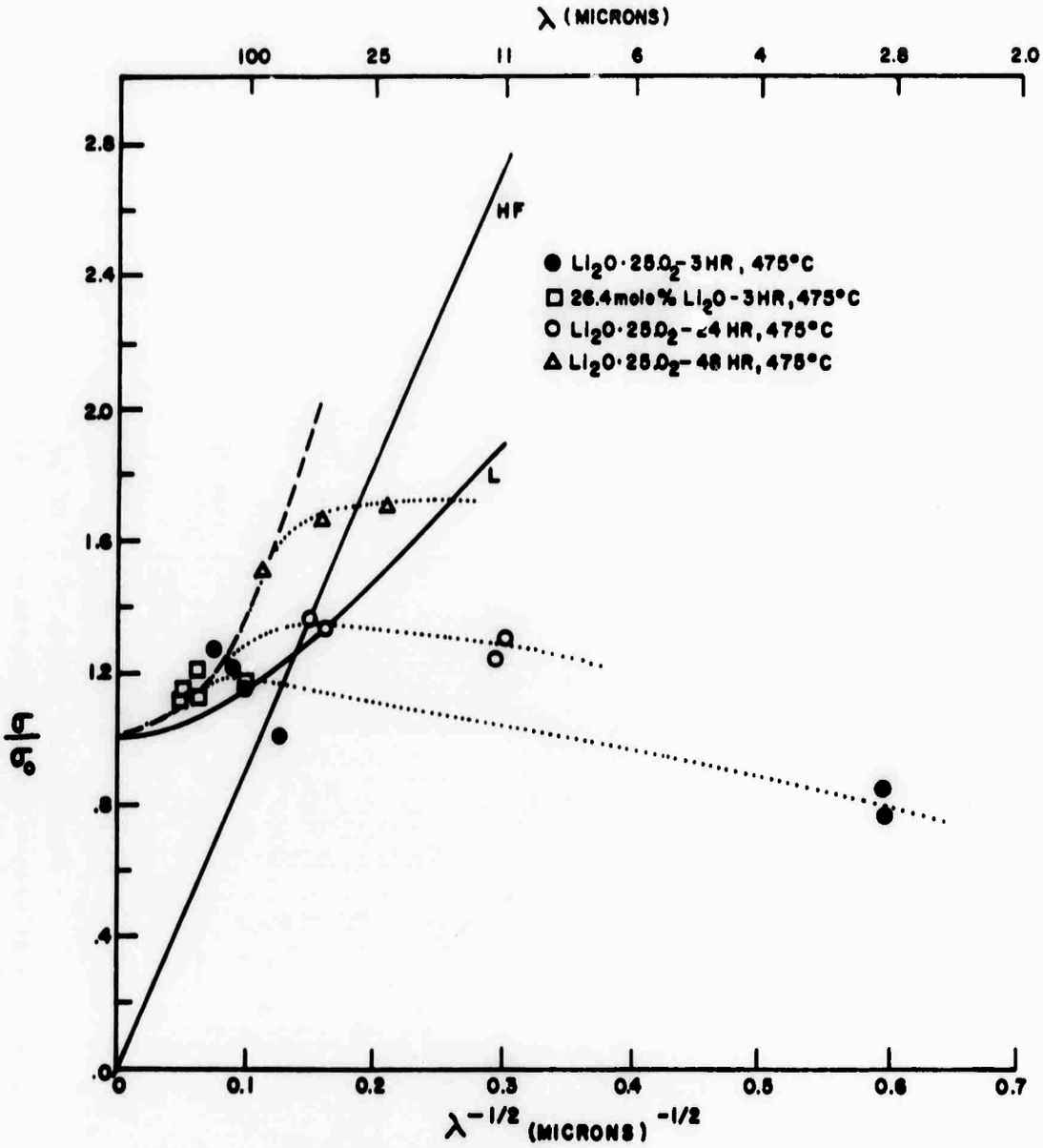


Figure 3. Comparison of the relative fracture strength of various glass-ceramics (σ glass-ceramic/σ_o glass) with the mean free path between the crystals (λ) and the Hasselman and Fulrath[14] and Lange theories.

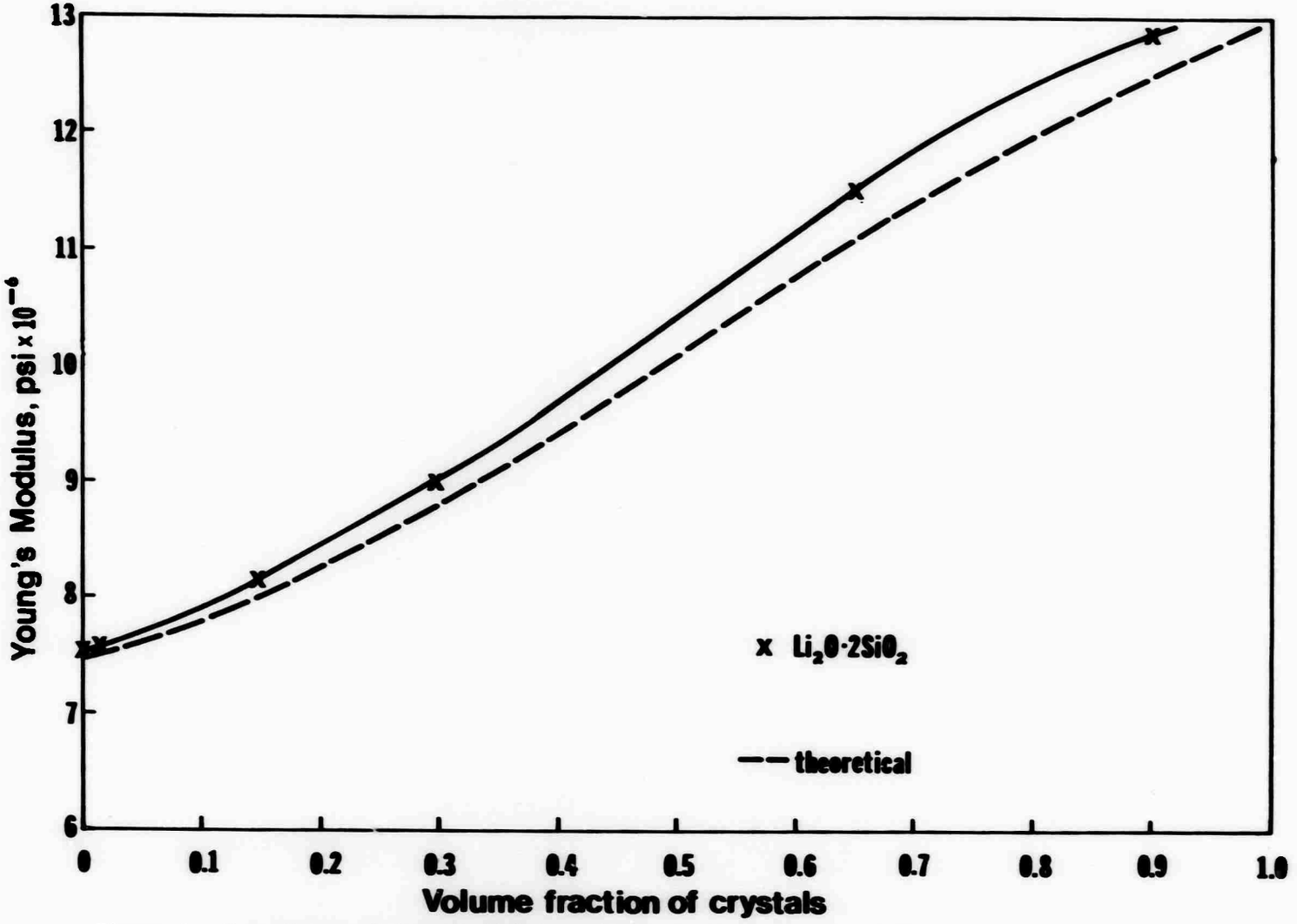


Figure 4. Dependence of Young's Modulus of a Li_2O-$2SiO_2$ glass-ceramic on volume fraction of crystals.

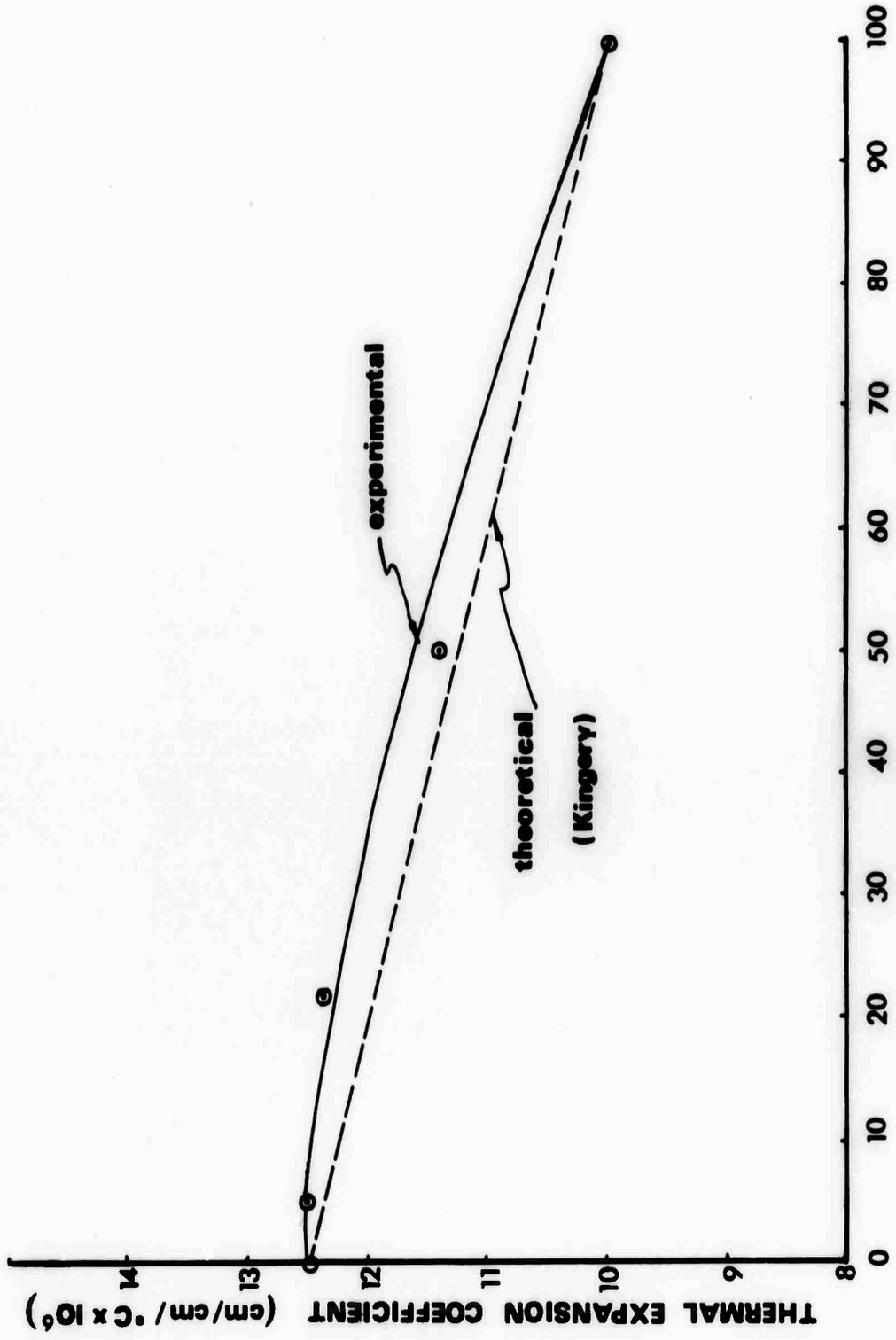


Figure 5. Change in thermal expansion coefficient of a glass-ceramic with percent crystallization.

Figure 6. Thin sections of partially crystalline glass which had been nucleated 48 hours at 475°C. (Uncrossed Nicols; 350X.)

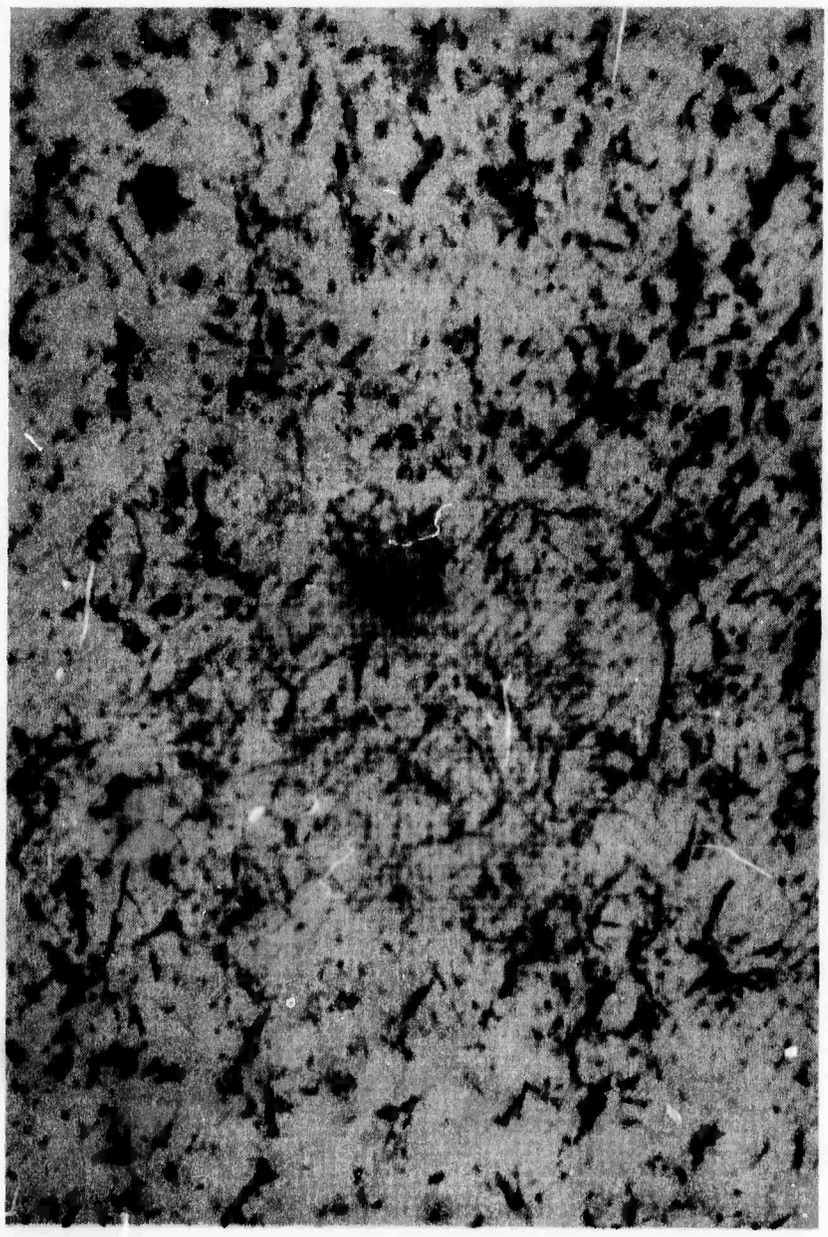


Figure 7. Thin section of partially crystalline glass which had been nucleated 3 hours at 475°C. (Uncrossed Nicols; 350X.)

B. THE EARLY STAGES OF CRYSTALLIZATION IN A Li_2O-$2SiO_2$ GLASS

(L. L. Hench, S. W. Freiman and D. L. Kinser)

Introduction

Recent investigations by the authors have established that a metastable Li_2SiO_3 crystalline phase is produced as a precursor to crystallization of the equilibrium $Li_2Si_2O_5$ phase when 30 mole % Li_2O-SiO_2 glasses are heat treated in the temperature range of 500°C.[1-3] The identification of the metastable phase was obtained using the techniques of dielectric relaxation, replica and transmission electron microscopy, D. C. polarization and x-ray diffraction employing a Guinier-DeWolff camera.

An objective of the present investigation is to obtain additional confirmation of the metastable reaction sequence and the crystallite sizes through the use of x-ray small angle scattering. A second objective is to determine the influence of the miscibility gap upon the metastable crystallization process by comparing the heat treatment sequence of a 33 mole % Li_2O-SiO_2 glass with the 30 mole % glass. Correlation of the metastable reaction sequence with the bulk crystallization kinetics in Li_2O-SiO_2 glasses previously reported by the authors[4] is an additional objective.

Experimental Procedure

Sample Preparation. The glass was prepared for these experiments by blending 99.9% silica (5μ Min-u-Sil) and 99.3% lithium carbonate* in a jar mill and melting for 18-70 hours at 1350°C for the 33 mole % and at 1400°C for the 30 mole % Li_2O-SiO_2 glasses in air in a covered Pt crucible. Samples 2.5 cm long and 1.3 cm in diameter or 1.8 cm in diameter and 0.5 cm thick were cast in a steel mold, placed in a furnace at a temperature range of 230°C - 300°C and heated for 1-4 hours to relieve stresses to a level sufficient to prevent sample breakage. The cast specimens were subsequently given nucleation and crystallization heat treatments in nichrome wound tube furnaces equipped with aluminum blocks to eliminate temperature gradients. Temperature control was maintained at ±1°C.

Dielectric Relaxation and D.C. Conductivity. Dielectric absorption spectra were measured on 0.1μ Al_2O_3 polished glass samples with vapor deposited gold electrodes over a frequency range of 20 Hz to 5 MHz using Wayne Kerr B221 and B601 bridges. The source and detector system, sample holder configuration, and environmental control system have been discussed in previous publications.[2,5] Dielectric loss peaks were measured as a function of temperature over a range of -80° to 200°C with a control of ±1°C. The oscilloscope, VTVM, and picoammeter system employed to make D.C. conductivity measurements within a range of times of 10^5 sec to 10^{-4} sec have also been described previously.[3]

X-ray Analysis. Efforts to obtain x-ray diffraction identification of the crystals produced in the early stages of the reaction sequence were made by using a vacuum Guinier-DeWolff camera using CuKα radiation.[1]

X-ray small angle scattering (XSAS) measurements were made on samples thinned to 0.1 to 0.2 mm thickness after thermal treatments in the form of 1.3 cm diameter by 2.5 cm long cylinders. The samples were polished sequentially to 1μ Al_2O_3.

---

*The major impurities consisted of 0.3% sulfate and 0.3% Na_2O.

Scattering data from the above surfaces were comparable with data obtained using Williams et al. polishing technique[6] establishing the independence of the glass surface on the results reported. A Siemans Kratky camera using CuKα radiation with balanced Ni and Co filters was used to measure the scattering intensity over the range of 0.01°-0.2° at a scanning speed of 125μ/min. The entrance and exit slits were 150μ and 50μ, respectively, and the detector slit width was 6 mm.

Experimental Results

The x-ray scattering curves obtained from the as cast and heat treated 30 mole % and 33 mole % Li_2O-SiO_2 glasses are given in Figures 1 and 2. The data is presented as log scattering intensity (log I) versus the square of the scattering angle (ε^2). The slope of the curves in the small angle region where the function is nearly a straight line gives a measurement of the Guinier radius (R_G) of the scattering centers.[7] Assuming spherical particles, the radii were calculated from the Guinier equation

$$I_{(s)} = (\rho-\rho_o)^2 V^2 \exp\left|-\frac{4}{5}\pi^2 R_G^2 \frac{\varepsilon^2}{\lambda^2}\right| \qquad (1)$$

where $(\rho-\rho_o)$ is the difference in electron density between the particle and the matrix, V is the irradiated volume, R_G is the Guinier radius, and λ is the x-ray wavelength. The values of R_G obtained are given in Table I. In cases where more than one straight line portion of the curve occurred, a range of radii are reported.

Porod[8] showed that the intensity in the tail of the scattering curve is related to the total surface area of the scattering particles. Assuming spherical particles, Gerold[9] derived an expression for the Porod radius

$$R_P = \frac{3}{8\pi(1-V_v)} \frac{M\lambda}{\varepsilon^2 I_{(\varepsilon)}} \qquad (2)$$

where V_v is the volume fraction of the scattering phase, and $M = Q_o/K'$ with Q_o = integrated scattering intensity and K' is a camera constant. Values obtained for the Porod radii of the samples are also included in Table I.

The parameter (M), which is directly proportional to the total integrated small angle x-ray scattering intensity (Q_o), is plotted as a function of thermal treatment of the lithia silicate glasses in Figure 3. Guinier showed that the magnitude Q_o is a function of the total volume of the scattering particles (V_p) and the electron density difference between the particles and the matrix;[7]

$$Q_o = V_p (\rho-\rho_o)^2 \qquad (3)$$

Figure 3 shows that the scattering intensity of both the 30 mole % and 33 mole % Li_2O-SiO_2 glasses increases to a maximum with thermal treatment at 5 hours at 500°C followed by a decline in scattering.

A comparison of the Guinier and Porod radii of the samples summarized in Table I reveals that the Porod radius of the particles is smaller than the Guinier radius. The difference between the radii can be largely attributed to differences in the definition of the two radii. The Porod radius is directly related to the surface area of the scattering particles[8] whereas the Guinier radius is defined in terms of the square root of the ratio of the 7th moment to the 5th moment of the size distribution of the particles.[10] If a widely dispersed system of scattering particles is present, the difference between the two radii will be very large and can be used to describe the size distribution of the system.[10] The differences reported in Table I indicate that the scattering particles are nearly monodisperse and remain so with thermal treatment.

Table I also shows that the size of the scattering particles does not change appreciably with the thermal treatment. The particle sizes are also approximately the same for the two compositions investigated.

X-ray diffraction results obtained from the Guinier-DeWolff camera for the 33 mole % lithia (Li_2O-$2SiO_2$) glass are given in Table II. The equilibrium lithium disilicate phase appears after a 50 hour heat treatment. No lithium metasilicate phase was observed within the limits of detection of the camera, 0.1 weight percent.[1]

Dielectric results for the 33 mole % lithia-silica glass are summarized in Figure 4. The dielectric loss angle (tan δ) measured at 79-84°C, is plotted as a function of log measuring frequency. Appreciable changes in the loss angle appear as a result of heat treating the glass at 500°C. A loss peak appears at approximately 4 hours ± 1 hour. (Data for other thermal treatment times are not shown in Figure 4 for the purpose of simplicity.)

The dielectric loss peak continues to increase in magnitude between 5 and 10 hours and then declines. After 50 hours of thermal treatment the loss peak disappears and further thermal treatment only serves to decrease the loss background.

D.C. conductivity data for the as cast and two of the heat treated 33 mole % Li_2O-SiO_2 glasses is presented in Figure 5 as a function of temperature. No detectable change in conductivity was noted until a 500°C heat treatment of 50 hours had been reached. The activation energy of the conduction process was measured to be 15 kilocalories per mole and it remained essentially constant with heat treatment.

Discussion of Results

Previous x-ray diffraction,[1] dielectric relaxation analysis,[1] and hot stage transmission electron microscopy[2] results from this laboratory have led to the conclusion that lithium metasilicate crystallites are a metastable precursor to the equilibrium crystallization of lithium disilicate phase in 30 mole % Li_2O-SiO_2 glasses. Independent investigations have also deduced the appearance of lithium metasilicate as a metastable phase in the Li_2O-SiO_2 based glass systems.[11,12] The

x-ray small angle scattering results of this study appear to confirm this conclusion. Consider the increase in scattering intensity with thermal treatment shown in Figure 3. In order for the scattering to increase, equation 3 indicates that either the total volume of the scattering particles must change with thermal history or the compositional difference between the particles and the matrix must change. The total volume of particles can increase by two alternative methods. The particles can grow during thermal treatment, causing the average radii to increase. Or, new particles may be grown during thermal treatment which would cause the size distribution of particles to change.

The stability of the measured particle size and the size distribution indicates that the change in scattering intensity is due to the compositional difference between the scattering particles and the matrix becoming greater during the first part of the thermal treatment, followed by a reduction in compositional differences in the later stages. Thus, the x-ray small angle scattering results are consistent with previous interpretations that metastable metasilicate particles precipitate in the 30 mole % Li_2O-SiO_2 glass and then resorb as the equilibrium disilicate crystal phase appears. It is the electron density difference between the metasilicate crystals and the silica rich glass which produces the small angle scattering.

Figure 3 reveals that the 33 mole % Li_2O-SiO_2 glass also undergoes an increase in x-ray scattering intensity with thermal treatment followed by a decrease at longer heat treatment times. The stability of particle size and size distribution indicated by the Guinier and Porod radii of Table I are similar to that reported for the 30 mole % glass. Consequently, it can be concluded that the metasilicate phase precipitates and later resorbs as equilibrium crystallization proceeds in the 33 mole % Li_2O-SiO_2 glass as well as the 30 mole % glass. This result indicates that the metastable precipitation and resorbtion process is independent of the presence of glass in glass phase separation, since no microscopic evidence of phase separation

was detected in the 33 mole % Li_2O-SiO_2 glass in this study as concluded by Charles.[13]

The dielectric results of Figure 4 reinforce the conclusion that a metastable phase precipitates during the heat treatment of the 33 mole % Li_2O-SiO_2 glass. The A.C. loss spectra is similar to that previously reported for a 30 mole % glass. Development of the loss peak with thermal treatment is interpreted to be due to the precipitation of second phase particles of higher conductivity which lead to a Maxwell-Wagner-Sillars type of interfacial polarization.[14] Longer heat treatments, >50 hours, apparently eliminate the compositional gradient required to maintain the interfacial potential gradient.

Changes in D.C. conductivity with heat treatment of the 33 mole % Li_2O-SiO_2, shown in Figure 5, are much slower than observed with the 30 mole % glass. The absence of a sharp decrease in D.C. conductivity during the 5-10 hour period of time suggests that a very low concentration of Li^+ ions is involved in the metastable reaction in the 33 mole % glass. A somewhat smaller volume fraction of metastable metasilicate phase can be expected from a lever-rule analysis of the Li_2O-SiO_2 phase diagram.[1] Apparently the decrease in volume fraction is sufficient that the metasilicate concentration is less than the 0.1% detection limit of the Guinier camera.

Relation of Results to Equilibrium Crystallization

An important question to consider is whether such a small volume fraction of a metastable will influence the equilibrium crystallization behavior of the glass.

A recent publication demonstrated that the crystallization kinetics of a 33 mole % Li_2O-SiO_2 glass which has had a three hour nucleation treatment at 475°C, insufficient to form significant concentration of the metasilicate phase, are appreciably slower than a glass of the same composition which has had a 24 hour nucleation treatment at 475°C.[4] X-ray diffraction employing the Guinier-DeWolff camera shows that the 24 hour treatment

produces a ~0.5% concentration of metastable metasilicate crystals. The accelerating effect of the metasilicate phase on the crystallization rate of the glass is shown in Figure 6. The rate of crystallization is increased by a factor of 3. However, it can also be noted that the incubation period is constant for the two samples at all temperatures. Therefore, the activation energy of the incubation process for both samples is independent of the metastable reaction process.

The activation energy for bulk crystallization is strongly dependent on prior nucleation treatment. The activation energy for samples with a 24 hour nucleation treatment corresponds to that of crystal growth in these systems (52 kcal/mole), while a sample with a three hour nucleation treatment exhibits a much higher activation energy (92 kcal/mole). Application of the Johnson-Mehl-Avrami (JMA)[15,16] and Hillert[17] analyses to the crystallization data reveals that the growth of the disilicate crystals is rod-like,[4,18] as is observed in the spherulites grown in the glass studied. The JMA analysis also showed that it is the activation energy of nucleation that is decreased when the sample has been given a heat treatment that produces metasilicate as a nucleation step.

Let us consider a model that can be used to describe the role of the metasilicate particles in controlling the kinetics of nucleation of the equilibrium lithium disilicate crystals.

Figure 7 depicts a two-dimensional structural projection, based on the work of Liebau,[19] of a $Li_2Si_2O_5$ crystal in the bottom half of the figure. The upper half of the figure is a similar projection of a Li_2SiO_3 crystal. In both crystals the Li^+ ions are tetrahedrally bonded with oxygen ions and are represented by triangles. If alternate Li_2O groups at the interface between the crystals, shaded in the figure, are removed from the metasilicate lattice a small displacive shift will be sufficient to produce the disilicate lattice. The interfacial region would be similar to the $Li_2Si_2O_5$ solid solutions discussed in detail by Glasser[20] except Li^+ ions would be

substituting for Si^{4+} in the present case instead of vice versa. The presence of the interfacial solid solution would perhaps account for the reported lattice parameter variations in x-ray diffraction results in the metasilicate reaction sequence[1,2] similar to those encountered by Glasser in the disilicate sequence.[20]

Thus, it is proposed that metasilicate crystallites form in the glass by an ordering of extremely small regions in the glass containing randomly distributed Li_2O-SiO_2 chains. The ordering probably also involves short range migration of some Li^+ ions to be incorporated into the metasilicate structure. With additional heating, Li_2O groups apparently diffuse out of the metasilicate crystallites, probably as ionic species, producing the complex layered disilicate structure from the SiO_2-rich matrix at the metasilicate interface. It is this diffusional step which is proposed to be responsible for the incubation period reported earlier.[4] When a critical size of disilicate region is reached, the growth proceeds rapidly towards completion. Because of the difficulty of Li_2O groups diffusing through the disilicate structure, it is energetically easier for the diffusion to occur at the ends of the dendrite arms, thus producing the spherulitic structures occurring in these glasses.

When a glass has been given a nucleation treatment which develops a large concentration of the highly ordered metasilicate regions, then the period of rapid crystal growth proceeds from many sites. The final spherulite size is small, the number per unit volume is large, and the size distribution is small, as reported.[21] As the spherulites envelop the ordered regions developed around each metasilicate site, the nucleation barrier in the growth of new dendrite arms is lowered,[4] and the overall bulk crystallization rate is accelerated (Figure 6). This process could perhaps be referred to as internally activated nucleation and is similar in effect to that suggested by Glasser.[20]

The basis for the efficiency of internally activated nucleation may be seen thermodynamically if one looks at the

equation given for heterogeneous nucleation rate.[22]

$$I = \sum_{j=0}^{j=max} \frac{N_j}{V} K_j \exp\left[-\frac{\Delta F_j^* + \Delta F_a}{kT}\right] \qquad (4)$$

I is the nucleation rate; N_j/V is the number of nucleation sites per unit volume; K_j is the frequency factor; ΔF_j^* is the energy necessary to form a stable nucleus at a particular site j; ΔF_a is the activation energy for a molecule to detach itself from the glass and attach itself to the growing nucleus. At large degrees of supercooling, such as occurs in the crystallization of glass, ΔF_a is the predominant term and, therefore, governs the rate of crystallization. It is essentially this energy that is measured as the activation energy for nucleation. Ordering at the metasilicate interfacial regions should reduce the energy barrier to transfer a molecule from the glass to the crystalline state, thereby reducing the activation energy for nucleation.

When the nucleation treatment has been insufficient to develop a large concentration of ordered metasilicate regions, the incubation time for crystal growth remains the same as the internally activated process because the diffusional step must still occur. The incubation activation energy remains the same, 100 kcal/mole, as observed.[4] Much fewer sites reach critical size, however, and the bulk crystallization rate is therefore slow. The spherulites do not envelop ordered regions during growth and the nucleation activation energy is large.

Conclusions

X-ray small angle scattering and dielectric relaxation results indicate that a metastable lithium metasilicate phase appears as a precursor to equilibrium crystallization in the 33 mole % Li_2O-SiO_2 glasses as well as 30 mole % glasses. Correlations with crystallization kinetics studies establish that the metastable phase provides an efficient aid to crystallization, by lowering the nucleation activation energy and

increasing the number of nuclei per unit volume without requiring a heterogeneous addition.

The results also indicate that glasses which have been given an extended exposure to high temperatures in annealing or during property measurements may yield data characteristic of a heterogeneous glass-crystal morphology and not that intrinsic to the glass itself.

BIBLIOGRAPHY

1. Kinser, D. L. and Hench, L. L., J. Am. Ceram. Soc. 51, 8 (1968).

2. Kinser, D. L. and Hench, L. L., "Hot Stage Transmission Electron Microscopy of a Lithia-Silica Glass," J. Materials Science, in press.

3. Kinser, D. L. and Hench, L. L., J. Am. Ceram. Soc. 52, 12 (1969) 638.

4. Freiman, S. W. and Hench, L. L., J. Am. Ceram. Soc. 51, 7 (1968) 382-387.

5. Hench, L. L., "Applications of Dielectric Relaxation in Materials Analysis," Society of Aerospace and Mechanical Process Engineers Fourteenth Symposium, Nov. 5, 1968, Cocoa Beach, Florida.

6. Williams, J. A., Rindone, G. E. and McKinstry, H. A., "Influence of Surface Preparation on the Small Angle X-ray Scattering by Glass," 1969 Am. Ceramic Society Meeting, May 4-6, Washington, D. C.

7. Guinier, A., X-Ray Diffraction, p. 338, W. H. Freeman and Company, San Francisco, 1963.

8. (a) Porod, G., Kolloid, Z. 124 (1951) 83.
   (b) Porod, G., Kolloid, Z. 125 (1952) 51, 109.

9. Gerold, V. K., Physica Stat. Solidi 1 (1961) 37.

10. (a) Harkness, S. D., Gould, R. W. and Hren, J. J., Phil. Mag. 19, 157 (1969) 115.
    (b) Bauer, R. and Gerold, V. K., Acta Met. 12 (1964) 1449.

11. McMillian, P. W., Phillips, S. V. and Partridge, G., J. of Materials Science 1 (1966) 269-279.

12. Kalinina, A. M. and Filipovich, V. N., "The Structure of Glass," 3 (1964) 53. (Consultants Bureau)

13. Charles, R. J., J. Am. Ceram. Soc. 50, 12 (1967) 631.

14. (a) Maxwell, J. C., Electricity and Magnetism, Vol. 1, Clarendon Press, London (1892), p. 328.
    (b) Wagner, K. W., Arch. Electrotech. 2, 9 (1914) 371.
    (c) Sillars, R. W., J. Inst. Elec. Engrs. (London) 80 (1937) 378.

15. Johnson, W. A. and Mehl, R. F., Trans AIME 135 (1939) 416.

16. Avrami, M., J. Chem. Phys. 1 (1939) 1103.

17. Hillert, M., Acta Met. 7 (1958) 653.

18. Freiman, S. W. and Hench, L. L., J. Am. Ceram. Soc. 52, 2 (1969) 111.

19. Liebau, F., Acta Crystallogr. 14 (1961) 389.

20. Glasser, F. P., Phys. Chem. Glasses 8 (1967) 224.

21. Gokularathnam, C. V., Freiman, S. W., DeHoff, R. T. and Hench, L. L., J. Am. Ceram. Soc. 52, 6 (1969) 327.

22. Fine, M. E., Phase Transformations in Condensed Systems, MacMillan Company, New York (1964) p. 43.

TABLE I

Particle Sizes Obtained from X-ray Small Angle Scattering

Guinier Radius R_G(Å)			Porod Radius R_P(Å)		
Heat Treatment Time-Hours	30 Mole % Li_2O	33.3 Mole % Li_2O	Heat Treatment Time-Hours	30 Mole % Li_2O	33.3 Mole % Li_2O
2	265-302	---	2	159	---
3	---	256	3	---	162
5	228-272	233	5	149	196
10	237-310	242	10	143	194

TABLE II

Summary of X-ray Lattice Spacings and Relative Intensities for 33 Mole % Lithia-Silica Glasses

Standard Crystal $Li_2O \cdot 2SiO_2$		33 mole % Li_2O 50 hrs. at 500°C		33 mole % Li_2O 100 hrs. at 500°C	
d(A)	Intensity	d(A)	Intensity	d(A)	Intensity
7.4	2			7.4	2
5.45	10	5.45	1	5.45	3
4.18	1				
3.75	10	3.75	2	3.75	10
3.65	10	3.65	2	3.65	10
3.58	10	3.58	2	3.58	1
2.95	1				
2.90	5			2.90	3
2.39	5			2.39	3
2.35	5	2.35	1	2.35	3
2.27	2				

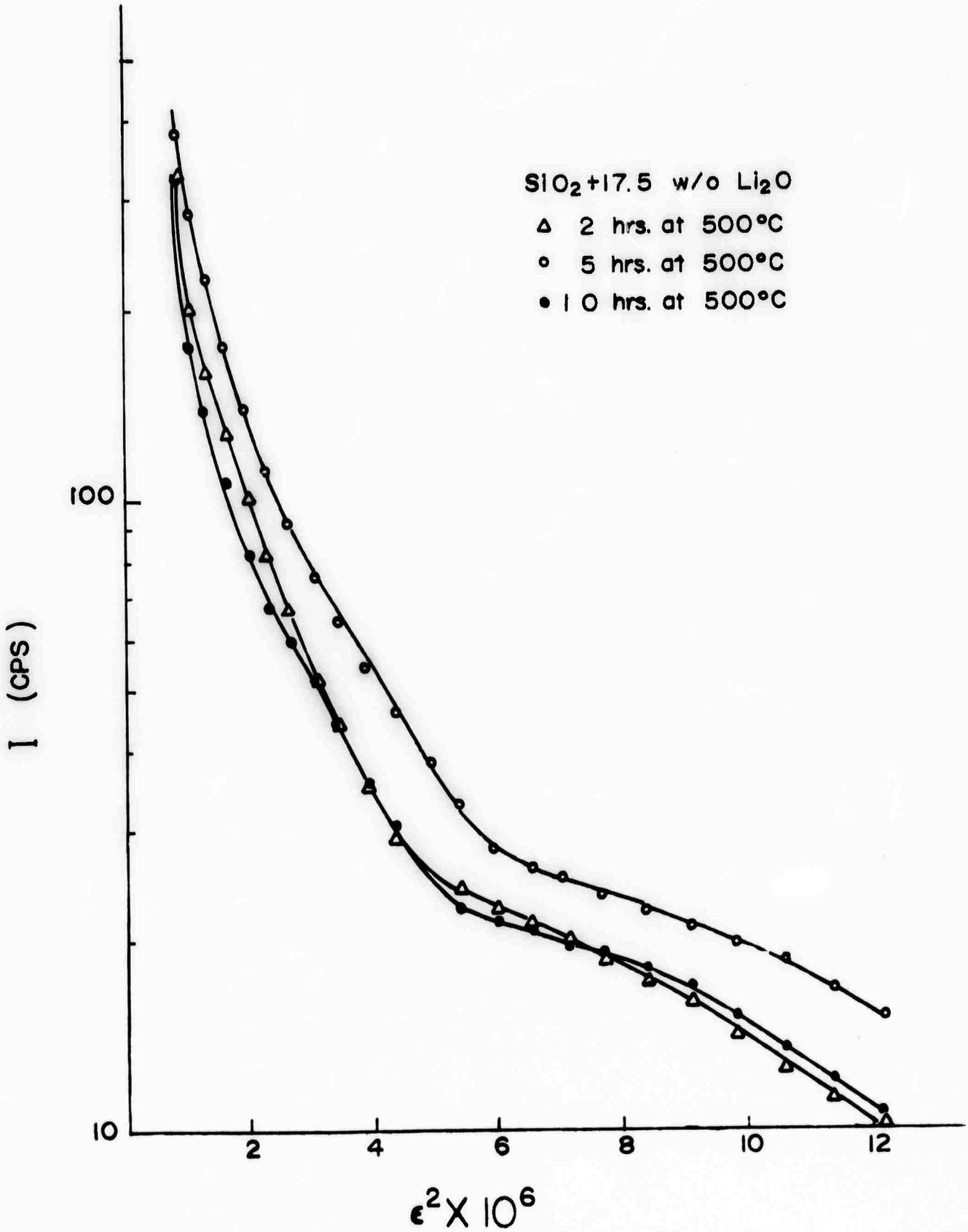


Fig. 1 X-ray small angle scattering curves from a 30 mole % (17.5 w/o) Li_2O-SiO_2 glass as a function of heat treatment.

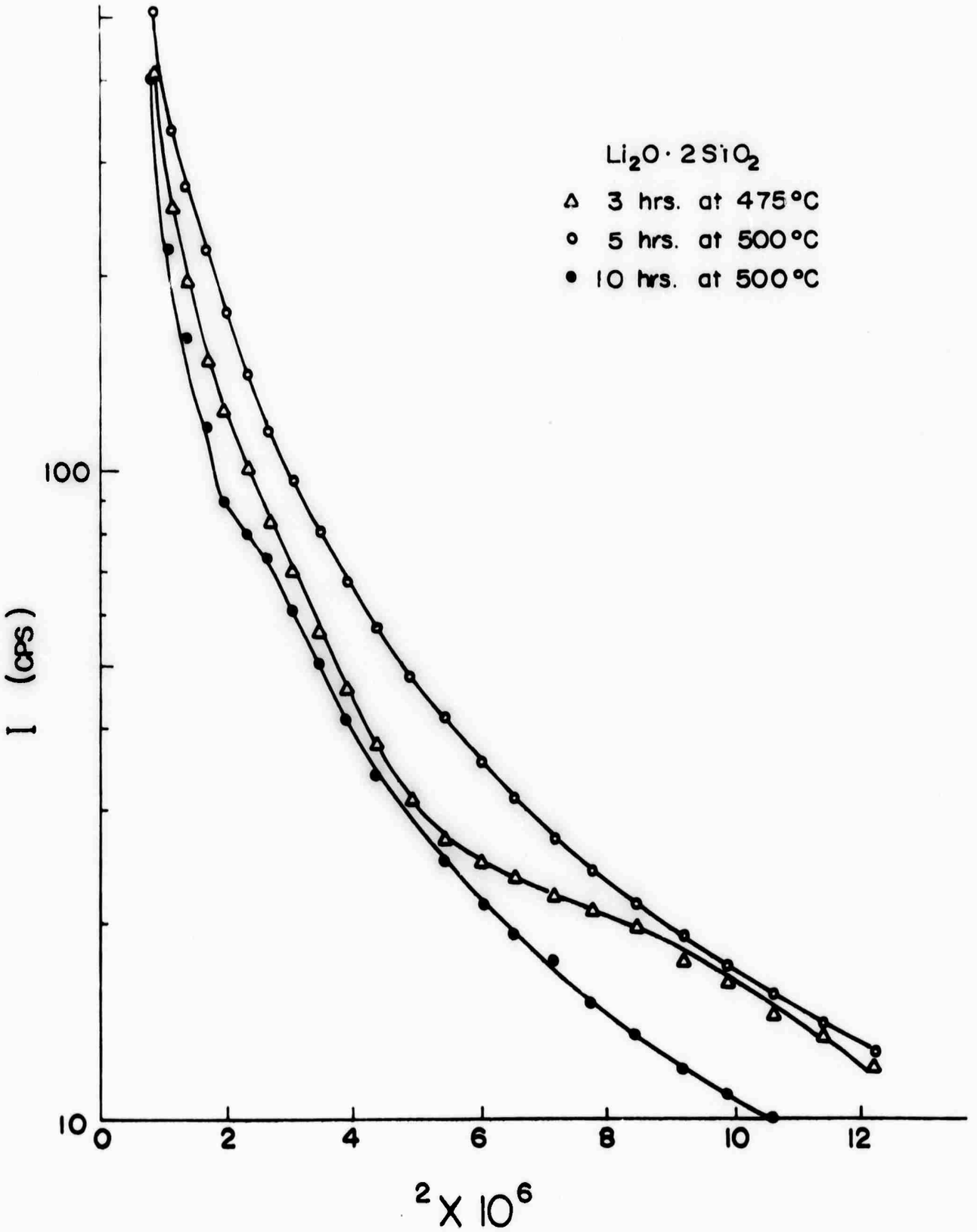


Fig. 2 X-ray small angle scattering curves from a 33 mole % Li_2-SiO_2 glass as a function of heat treatment.

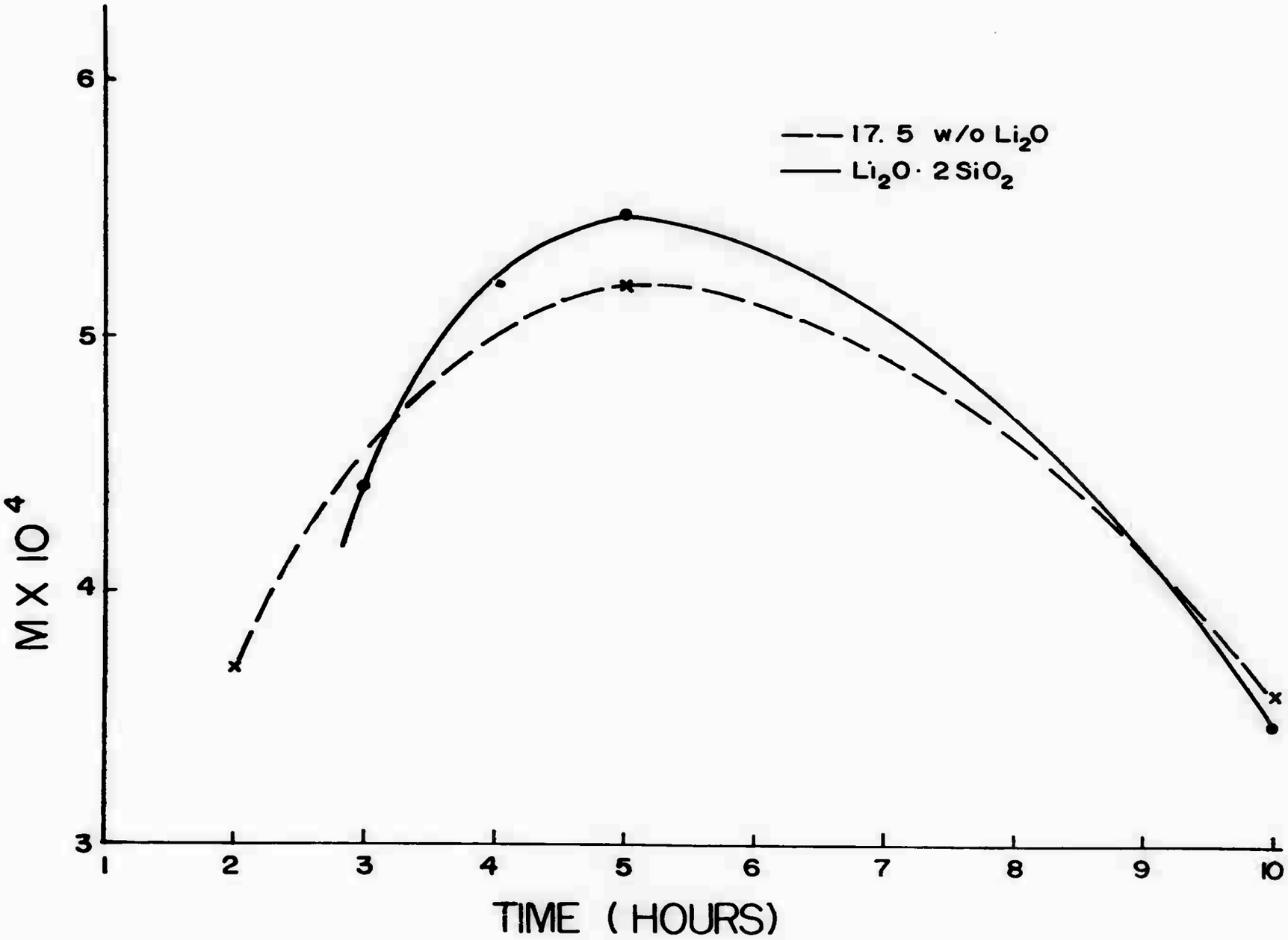


Fig. 3 Variation in integrated small angle scattered intensity with heat treatment time at 500°C for a 30 mole %(17.5 w/o) Li_2O-SiO_2 glass and a 33 mole % Li_2O-SiO_2 glass (Li_2O-$2SiO_2$).

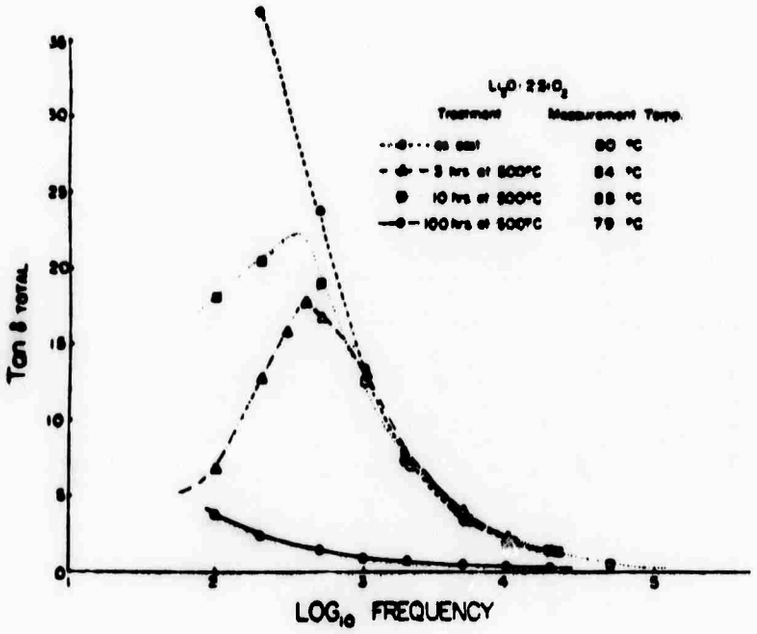


Fig. 4 Tan δ versus $\log_{10}$ frequency for the 33 mole % lithia-silica glass for various thermal treatments.

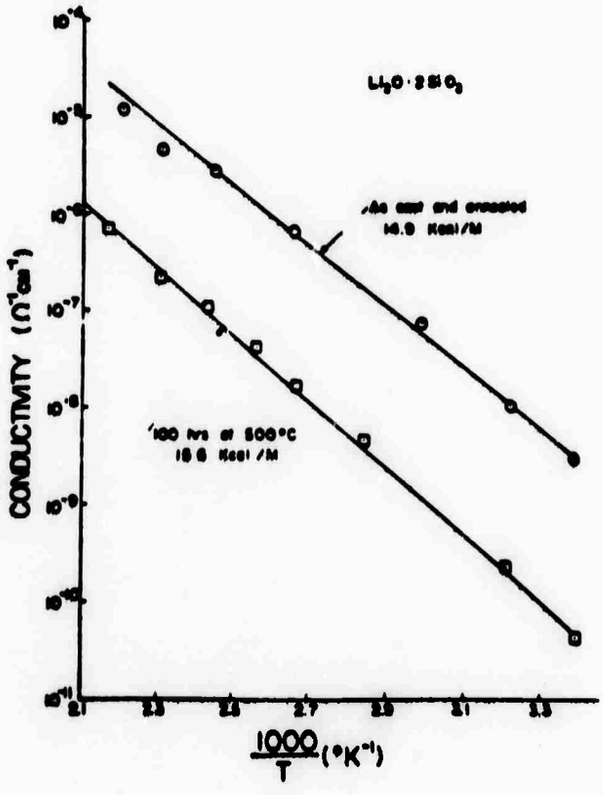


Fig. 5 Log D.C. conductivity versus reciprocal temperature for the 33 mole % lithia-silica glass.

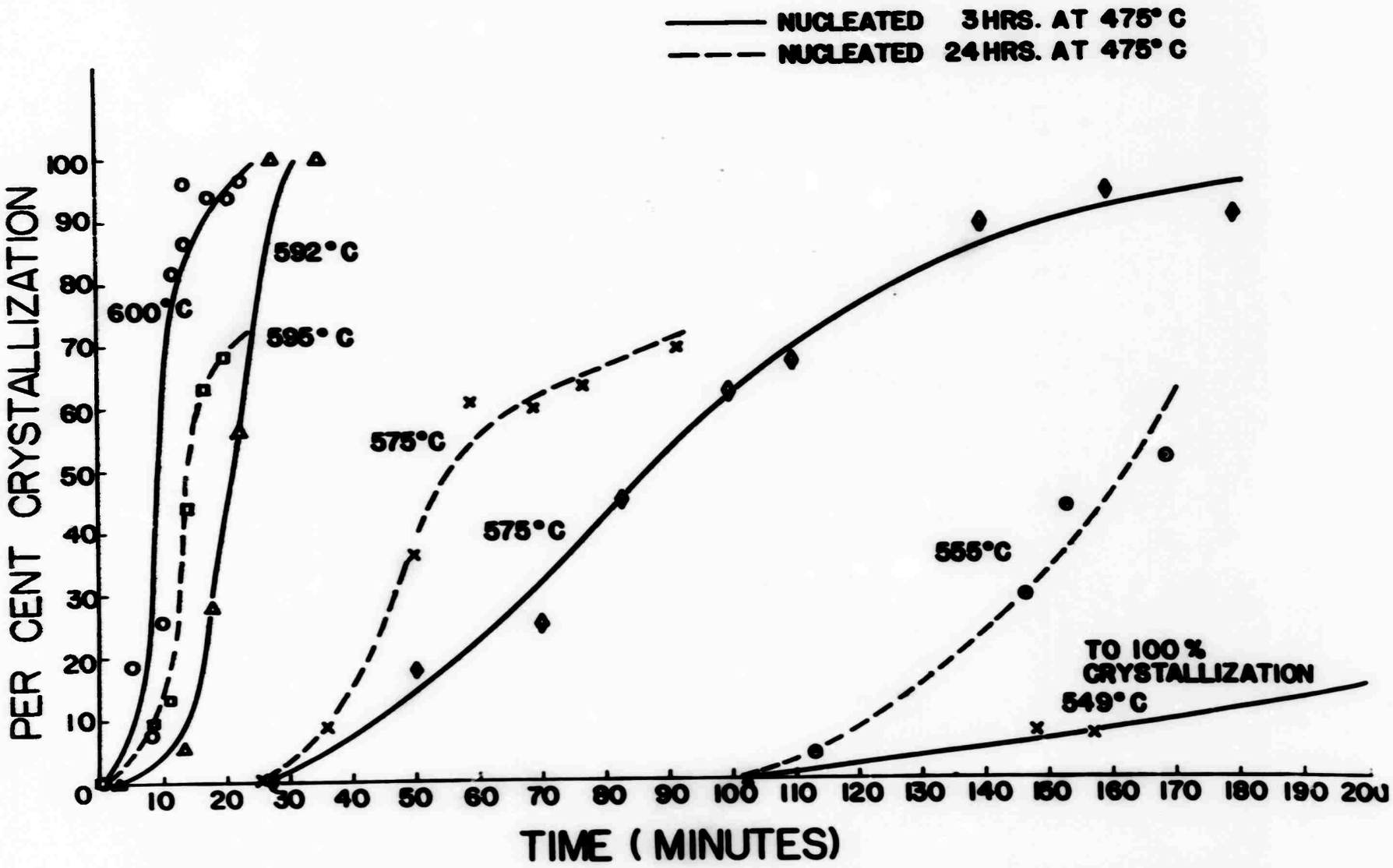


Fig. 6 Effect of increased nucleation treatment and greater concentration of metastable Li_2SiO_3 crystallites on the rate of bulk crystallization in a 33 mole % Li_2O-SiO_2 glass.

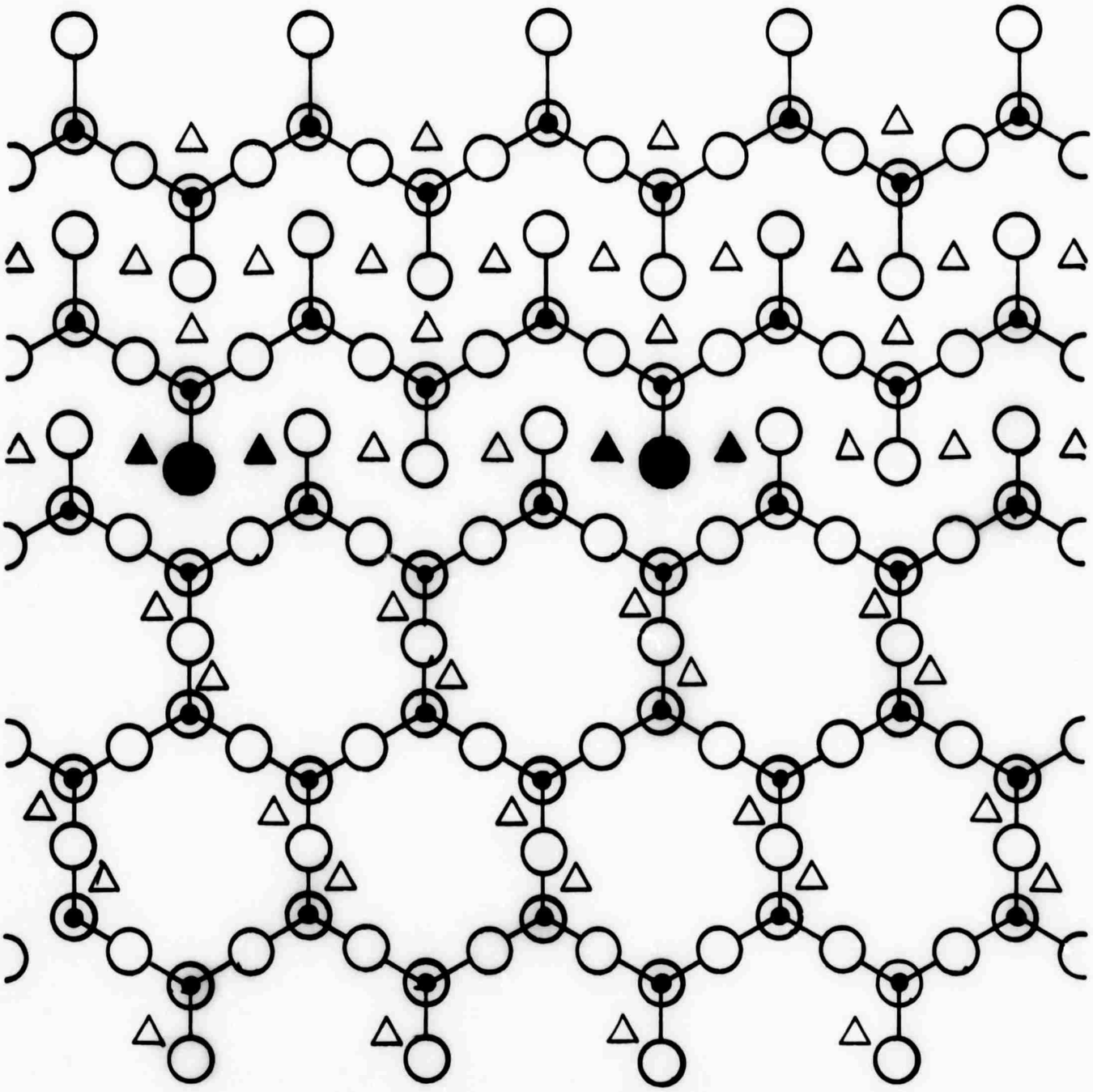


Fig. 7 Structural schematic of the interface between growing $Li_2Si_2O_5$ crystals (bottom of figure) and Li_2SiO_3 nuclei (top of figure). Diffusion of shaded Li_2O groups from the interface produces the equilibrium $Li_2Si_2O_5$ phase.

C. ELECTRICAL PROPERTIES OF GLASS (L. L. Hench and H. F. Schaake)

Introduction

The purpose of this review paper on the electrical properties of glass is threefold. First, the theoretical basis for understanding the electrical behavior of glass will be reviewed. Second, a synthesis of the understanding of composition and structure effects will be presented; and third, important problems in this area will be discussed. The theoretical review first considers a general theory of conductivity, followed by more specific discussions of ionic conductivity and electronic conductivity. Dielectric properties of glass will also be considered, including theory, polarization mechanisms, and the current status of understanding composition and structural effects. Switching phenomena will also be briefly considered.

Because of the immense scope of the subject of electrical behavior of glass, a thorough historical survey of this field will not be attempted. The reader is referred to Owens,[1] Mackenzie,[2] Pearson,[3] Morey,[4] Volf,[5] Doremus[6] and Weyl[7] for authoratative historical reviews in this area. A detailed bibliography on this subject has also been compiled for the period 1957-1968 by the International Commission on Glass.[8]

Introductory Theory of Electrical Conductivity

In order for a current to flow in a glass, charge must somehow be transported. Electrons may move through the lattice in one of two ways: either by moving with nuclei, in which case ionic conductivity (σ_i) occurs, or by becoming detached from one atom, and moving to another, in which case electronic conductivity (σ_e) is present. It is also possible, of course, that both electronic and ionic conduction processes can occur simultaneously. The transfer number, $t_x = \sigma_x/\sigma_T$, expresses the fraction of total conductivity contributed by each type of charge carrier.

The conductivity of any material is defined as the ratio of the current density to the electric field required to produce the density, i.e.,

$$\sigma = \frac{j}{E} \tag{1}$$

While in the most general case σ is a second rank tensor since both j and E are vectors, because a glass possesses spherical macroscopic symmetry, we need only consider it as a scalar. The current density is the charge crossing a unit area per second. Thus if the charge on a carrier, q, is the product of its valence times unit charge e, $q = Ze$, its velocity is v, and there are n identical carriers per unit volume, the current density will be:

$$j = nqv \tag{2}$$

We now inquire as to how the velocity, v, is related to the electric field, E. For electrons in a conduction band, the classical free electron model can be used with an argument that the electron will be accelerated by the field, and therefore its velocity at any time t will be changed from its initial random value, v_o, by an amount,

$$\Delta v(t) = (\text{acceleration} \times t) = \frac{qE}{m} t \qquad (3)$$

where m is the mass of the electron (or more properly, the effective mass, in which factors introduced by quantum mechanics are taken into account). But the electron will undergo this acceleration only for an increment of time before it collides with an atom and its random velocity is restored. If the mean time between collisions is τ, the average velocity of an electron is independent of time and is given by

$$\langle v \rangle = \langle v_o + \Delta v(t) \rangle = 0 + \frac{qE\tau}{2m} \qquad (4)$$

Combining equations (1), (2) and (4) we find that

$$\sigma = \frac{nq^2\tau}{2m} \qquad (5)$$

Equation (5) is generally written as

$$\sigma = nq\mu \qquad (6)$$

where μ is the mobility defined by:

$$\mu = \frac{q\tau}{2m} \qquad (7)$$

While equation (7) is based on an extremely naive model, a full-scale quantum statistical mechanical treatment yields similar results. Another case of interest is that of ions which hop from site to site. A relation between the ordinary diffusion constant, D_i, and the mobility can be derived from the Boltzmann transport equation,[9] yielding the so-called Einstein relation:

$$\mu = \frac{e}{kT} D \qquad (8)$$

The essential argument in deriving equation (8) is that the electric field only slightly perturbs the diffusion process

leading to a slightly larger number of ions moving in the direction of the field than against it (or vice versa, depending on the sign of the charge of the ion). The 1/T dependence of equation (8) is very slight, and is usually swamped out by the expenential temperature dependence of the diffusion coefficient, $D = D_o \exp[-Q_D/RT]$. Consequently, a general expression for ionic conduction will be of the form,

$$\sigma = nq \frac{e}{RT} D_o \exp-[Q_D/RT] = \sigma_o \exp-[Q_D/RT] \qquad (9)$$

For many years since the classical experiments of Warburg[10] which established that ions moved in glass under an electric field it was tacitly assumed that all glasses were ionic conductors. During the intervening years, details of ion transport mechanisms that led to relations such as equation (9) have become well accepted.[11] Series of glasses involving compositional variations have been investigated and have led to an understanding of the effect of network modifying ions on σ_o and Q_D. Relatively little has been done to correlate electrical properties with the recent information of glass microstructures involving both liquid phase separation and nucleation and crystal growth.

Within the last two decades the phenomena of electronic transport on oxide glasses containing transition metal ions and glasses based on the chalcogenide series of elements has also been well established.[1-3,6] In recent years extensive controvery has developed concerning the mechanisms of electronic motion in amorphous solids.[6] The theoretical understanding has been extensively hampered by a lack of a variety of conductivity data on glasses with well documented microstructures. A variety of compositional ranges in both oxide and chalcogenide systems have been studied, but accompanying phase equilibria, metastable phase separation, or glass transition data has usually not been available.

It is the author's hope that the following sections will reflect the current state of understanding in both the ionic

and electronic transport areas of glass. Examples of applications of the theoretical treatments presented will indicate the direction that theory-properties-structure correlations are headed.

Ionic Conductivity - Theory

Ionic conductivity involves the long range migration of ionic charge carriers through the glass under the driving force of an applied electric field. The charge carriers will be the most mobile ions in the glass and for silicate glasses, therefore, are usually +1 valence cations moving in an immobile SiO_2 based matrix. The electric force on the cation serves to perturb its random thermal motion by increasing the probability of a transition in the direction of the applied field. This behavior can be described easily by considering the two schematic potential energy wells shown in Figure 1. The potential barrier W represents the maximum barrier in the path of least resistance in the glass.

Consider that the ions move in one dimension parallel to the x axis jumping over the potential barrier W. The probability that an ion will move to either the right or the left is,

$$p = \alpha \frac{kT}{h} \exp[-W/kT] \qquad (10)$$

where α is an accommodation coefficient related to the irreversibility of the jump; kT/h is the vibrational frequency of the ion in the well, k is Boltzmann's constant, and T is absolute temperature. When a field E is applied, the ion coordinations are slightly distorted within the glass structure and the potential barrier to the ion motion is slightly shifted as shown in Figure 1. The field will lower the potential barrier on the side in the direction of the field and raise it on the other side by an equivalent amount. If the distance between the two wells is (b) the potential on the right side will now be smaller by an amount, 1/2 zeEb = 1/2 Fb, where F is the force on the ion

and Z is the valence of the ion. Therefore, the probability of motion to the right is:

$$p^+ = \frac{1}{2}\,\alpha\,\frac{kT}{h}\;\exp-[W - \tfrac{1}{2}\,Fb]/kT \qquad (11)$$

or $\quad p^- = p\,\exp[+Fb/2kT]$

The probability of motion to the left is thus,

$$p^- = \frac{1}{2}\,\alpha\,\frac{kT}{h}\;\exp-[W + \tfrac{1}{2}\,Fb]/kT \qquad (12)$$

or $\quad p^- = \frac{1}{2}\,p\,\exp[-Fb/2kT]$

Consequently, the positive transitions will be more frequent than the negative and therefore there will be an average drift velocity to the right in the direction of the field. The mean velocity of the drift motion is,

$$\bar{v} = b(p^+ - p^-) = \frac{1}{2}\,bp\left[\exp + \frac{Fb}{2kT} - \exp - \frac{Fb}{2kT}\right]$$

$$\bar{v} = bp\,\sinh\frac{Fb}{2kT} \qquad (13)$$

It is worthwhile to repeat that this result is based on the electric field merely directing the random diffusion of ions that occur spontaneously at any temperature, T_i. The field does not pull the ions of the wells. In fact, as long as the field strength is much smaller than kT, i.e., 1/2 Fb << kT, then sinh can be replaced by the argument and the drift velocity can be expressed as,

$$\bar{v} \cong \frac{b^2pF}{2kT} \qquad (14)$$

In the case where very large field strengths exist, the first term in sinh will dominate and,

$$\bar{v} \cong \text{Const.}\,\exp[bF/2kT] \qquad (15)$$

At room temperature bF is small compared with kT up to field strengths of 10V/cm. Therefore, the current will be proportional to the field strength in agreement with Ohm's law.

$$j = nze\bar{v}; \qquad n = \#\ \text{of ions/cc} \tag{16}$$

$$\text{so} \qquad j = \frac{nze\ b^2\ pF}{2kT} = \frac{nz^2e^2b^2pE}{2kT} \tag{17}$$

and substituting,

$$p = \frac{\alpha kT}{h}\ \exp-[W/kT] \text{ and } W = \frac{\Delta F_{DC}}{N} \tag{18}$$

where ΔF_{DC} is tne change in free energy for dc conduction in units of kilocalories/mole and N is Avagadros number = 6.023×10^{23} atoms/mole, yields

$$j = \frac{n\alpha z^2e^2b^2E}{2h}\ \exp-[\frac{\Delta F_{DC}}{RT}] \tag{19}$$

Therefore, the electrical resistance of glass

$$R_{(ohms)} = E/j = \frac{2h}{n\alpha z^2e^2b^2}\ \exp+[\Delta F_{DC}/RT] \tag{20}$$

$$\text{or} \qquad \log R = \log \frac{2h}{n\alpha z^2e^2b^2} + \Delta F_{DC}/RT$$

or conversely,

$$\log G_{(mohs)} = \log \frac{n\alpha z^2e^2b^2}{2h} - \Delta F_{DC}/RT \tag{21}$$

which is similar to the Rasch-Hinrichsen law of electrical resistivity of glass discovered empirically a number of years ago,[12] $\log \rho = A+B/T$. That a variety of glasses do behave according to equation (21) is shown in Figure 2*.

---

*Dimensional contribution to the conductivity of the glasses is eliminated by multiplying G by thickness and dividing by area, i.e., $\sigma(ohm^{-1}cm^{-1}) = G(ohm^{-1}) \times t(cm) \times 1/A(cm^{-2})$.

The free energy of activation for D.C. conduction is also given in Table I for several glasses, together with diffusion activation energies. The similarity between the two activation energies gives additional credence to the formulation of conductivity theory based on field directed diffusion of charged ions. It is also possible to distinguish between enthalpy and entropy changes involved in the D.C. conduction by using the Second Law of Thermodynamics, as

$$\Delta F_{DC} = \Delta H_{DC} - T\Delta S_{DC} \tag{22}$$

which yields a final expression for the conductance,

$$\log G = \log \left[\frac{n\alpha z^2 e^2 b^2}{2h} \exp \frac{\Delta S}{R}\right] - \frac{\Delta H_{DC}}{RT} \tag{23}$$

Therefore, the enthalpy and entropy can be calculated from the slope of the log σ versus 1/T plot, respectively. Some interpretations of ΔS values have been given, although they are largely speculative in nature.[1,13]

Ionic Conductivity - Electrode Polarization

Ion transport within the glass under a D.C. potential eventually leads to a build-up of charge at the glass-electrode interface if the ions are not replenished at the electrode. This phenomena is termed electrode polarization and it can drastically affect the magnitude of D.C. conduction measured. Three regions of ionic conductivity can be seen on Figure 3, which is an experimental measurement of the conductivity or current after a voltage of 1 volt had been applied for the times indicated.[14] The conductivity at very short times increases due to various dielectric contributions discussed in a following section. At long time intervals the conductivity decreases indefinitely and only at intermediate intervals is the conductivity time independent. It is within this time independent region that conductivity can be correctly termed σ_{DC}.

The reason for the continued decrease in σ at long times is that the mobile cations pile up at the electrodes which results in an induced back field. The back field concentrates the potential drop in the electrode region as shown by Proctor and Sutton[15] in Figure 4. Since the potential across the bulk of the sample is decreased to a fraction of the applied potential, the force, F, on the mobile ions in the bulk is decreased proportionally and the conductivity decreases as predicted from equation (17). The back field builds up indefinitely as time progresses, resulting in a continuous decrease in glass conductivity.

It is thus evident that D.C. conductivity is dependent on the time interval at which the measurement is made and care must be taken to conduct measurements in the time-independent region of behavior. This may be especially difficult at higher temperatures when the conductivity is in the range of $>10^{-5}$ $ohm^{-1}cm^{-1}$ because electrode polarization occurs so rapidly, <10 sec.

There are three possible solutions to the problem of accurate experimental measurement of ionic σ_{DC} posed by electrode polarization. The first is to conduct σ vs. t curves such as Figure 3 for each temperature to ensure that measurements are made in the time-independent regime. Secondly, it is possible to use non-blocking electrodes which provide a source and sink for the mobile ions, e.g., sodium amalgam electrodes for Na containing glasses.[16] However, it should also be pointed out that a brief investigation of Li-Cd alloy electrodes for Li_2-SiO_2 glasses[17] indicated that the alloy composition must be adjusted such that the chemical potential across the interface is equal before the electrode is completely non-blocking.

The third approach is to calculate the current flow required to produce the back potential necessary to decrease the conductivity of glass to a given fraction of its true non-blocked value. Kinser and Hench[18] have shown that the back potential, ϕ_b, due to the polarization charge, q, is simply:

$$\phi_b = \frac{qd}{2\varepsilon\varepsilon_o} \qquad (24)$$

where ϵ is the permittivity of the glass, d is the thickness, and ϵ_o is the permittivity of free space. The charge, q_b, to form the back potential can be related to the current flow as:

$$q_b = \int_o^t j\,dt = jt \tag{25}$$

The current, j, required to form the back charge is directly related to the applied potential, ϕ_a, and the conductivity of the glass, as $j = \phi_a \sigma$, and therefore on substitution into equation (25) yields,

$$q_b = \phi_a\, \sigma t \tag{26}$$

So, if there is a critical response time for the measuring equipment, t_x, and an acceptable level of error in the conductivity measurement of 0.1, for example, then the ratio ϕ_b/ϕ_a is equal to 0.9 and the limit of conductivity that can be accurately measured within t_x for the glass is:

$$\phi_b/\phi_a = \frac{0.9(2\epsilon\epsilon_o)}{dt_x} = \sigma_x \tag{27}$$

Since σ increases rapidly with temperature, equation (27) implies that there will be a critical temperature, T_c, where the conductivity error becomes greater than 10 % when T_c is exceeded. Such a behavior is shown in Figure 5. By using equation (27) and an estimated value of ΔF_{DC}, this problem can be eliminated.

Ionic Conductivity - Composition Effects

A number of investigators[1,19-21] have investigated the change in electrical conductivity of alkali-silicate glasses associated with the replacement of silica by alkali ions. In nearly all cases, a rapid increase in conductivity is observed until a range of 20-30 mole % modifier is reached, after which the effect becomes progressively less. The range of conductivity is from 10^{-16} $ohm^{-1}cm^{-1}$ to 10^{-7} $ohm^{-1}cm^{-1}$ at room temperature.

The relative effect of alkali valence on conductivity is $Na^+ > Li^+ > K^+$. The addition of only parts per million of Na^+ ions is sufficient to drastically reduce the conductivity of fused silica and the conductivity is directly related to the Na^+ content.[22]

Additions of divalent modifiers to a soda silica glass have consistently been observed to decrease the electrical conductivity[1,4,19] with the relative effect being Ba^{++} Pb^{++} Sr^{++} Ca^{++} Mg^{++} Zn^{++} Be^{++}. Additions of a 20 weight % Ca to an 18 % Na_2O .82 SiO_2 glass will decrease the room temperature conductivity to 10^{-15} $ohm^{-1}cm^{-1}$, where Zn^{++} additions up to 40 % will be of no influence. The effectiveness of the divalent oxides in decreasing the conductivity has been found[23] to increase as the radius of the metal ion is increased. This correlation is shown in Figure 6. The interpretation of the effect is that the divalent ions block the conduction paths of the alkali ions and as the divalent ion size increases, the blocking becomes more effective. It is also possible that the divalent and monovalent ions coordinate with mutual oxygen ions, thus maximizing the charge distribution in the glass.

Partial substitution of one alkali metal oxide for another in silicate, borosilicate and borate glasses produce such large variations in electrical conductivity that this behavior is termed "the mixed alkali effect." This effect is illustrated for three systems in Figure 7 taken from Charles.[24] Early work[25] interpreted the mixed alkali effect largely in terms of the motion of "coupled ions." The evidence suggesting the formation of a cation complex was that the activation energy for conduction also increases by almost a factor of two as the substitution occurs. Charles has recently shown[24] that the substitutions induce phase separation in the Li-Na-Cs cation glasses investigated. His conclusion, therefore, is that the smaller, more mobile ions are isolated in the dispersed phase with the larger ions in the matrix phase, thus decreasing the total D.C. conductivity. However, in order for the conductivity to be suppressed to the magnitude demonstrated, the larger cations must also be largely coordinated within the dispersed phases yielding

a nearly alkali-free silica-rich matrix which determines the D.C. conductivity.

Substitution of Al_2O_3 for SiO_2 in sodium silicate glasses is especially interesting.[26] As the Al_2O_3 content increases, the electrical conductivity is also increased and the activation energy for conduction is reduced. This behavior appears to be associated with the aluminum ion entering the network in four-fold coordination reducing the number of non-bridging oxygen ions. The non-bridging oxygen coordination shell about the sodium ions must be decreased, which could serve to increase the sodium ion mobility.

In light of Charles' observations on the mixed alkali effect it would indeed be interesting to compare the structures of the soda-silica glasses with and without the Al_2O_3 substitutions.

Ionic Conductivity - Structure Effects

With the exception of Charles' studies[24,27] there have been relatively few attempts to obtain correlations between composition-structure and the ionic conductivity of glass. It has also been recognized for many years that thermal treatments in the annealing-transformation range significantly alter the D.C. conductivity of glasses.[28,29] However, there has been relatively little work done to establish structural interpretations of thermal history variations in electrical properties either.

Charles[27] examined the electrical properties and structural differences in lithia-silica glasses resulting from different quenching rates. His results indicated that the changes in electrical properties could be related to the morphology of the phase separation. A rapid quench of a low, 7.5 mole %, lithia-silicate glass changed the phase separated structure from that of a dispersed phase to a tubular, interconnected phase, which was perhaps spinodal in origin. The D.C. conductivity was increased by nearly an order of magnitude by the rapid quenching due to the presence of the interconnected Li_2O-rich phase.

A series of investigations concerned with structural interpretations of electrical conductivity changes accompanying

thermal treatments on Li_2O-SiO_2 and Na_2O-SiO_2 glasses has recently been completed by Kinser and Hench.[17,18,30-32] These studies have shown that the D.C. conductivity of both types of alkali silicate glasses over the compositional ranges investigated decreases by over two orders of magnitude concurrent with the appearance of the equilibrium crystal phases in the glass. This behavior is illustrated in Figure 8. The major change in the conductivity is associated with the presence of only a few percent of crystals in the glass. Consequently, it appears that the more energetic cations, the ones that usually participate in the thermally activated D.C. conduction, are the first removed from the glass during the crystallization process. Since crystallization is diffusion controlled in these glasses[34] it is reasonable that it be the same ions that respond first to the free energy gradient for crystallization as would respond most easily to the D.C. potential.

Ionic Conductivity - Molten Silicates

The determination of the electrical conductance of molten glasses has been largely prompted by an interest in the structure of glasses and the manner in which the structure changes upon melting. However, the recent interest in the production of glass by means of electric furnaces provides a commercial impetus for understanding the nature of the electrical conductance of the molten glasses. In particular, it is of interest to understand the effect of composition on the conductance of glasses in the molten state. Obviously, from a commercial standpoint, the effect of temperature on electrical conductivity is of extreme importance also.

There have been several investigations of the electrical conductivity of molten silicates which contain alkali oxides.[35-39] Different workers disagree about which of the three systems, Li_2O-SiO_2, Na_2O-SiO_2, or K_2O-SiO_2 is the most conducting. The recent authoritative work of R. E. Tickle[35] has shown that the relative conductivities of the alkali oxide silicate glasses depend

upon the level of concentration of the alkali ions. For concentrations of R_2O greater than 22 %, the order of equivalent conductances was found to be $Li_2O > Na_2O > K_2O$, but below about 22 % R_2O, the order was $Na_2O > Li_2O > K_2O$. These results appear to clarify the controversy existing between the studies conducted by Endell and Hellbrügge,[36] Brockris et al.,[37] and Ur...s.[38].

A comparison of resistivity isotherms at 1400°C is made for various concentrations of Li_2O-SiO_2 glasses measured by different workers,[35-38] and is shown in Figure 9.

The resistivity of a series of soda silica glasses[35] is shown as a function of reciprocal temperature in Figure 10 over a temperature range of 900°C to 1600°C. The numbers on the curves denote the mole % of Na_2O in the binary glasses. The curvature of the plots is clearly seen, and the departure from the simple exponential temperature dependence of the Rasch-Hinrichsen law is pronounced. It is significant that the low temperature and high temperature data match well, and can be represented by a single curved line for each of the compositions studied. The resistivity of soda silica glasses decreased markedly with an increase in the percent of Na_2O present.

The conductivity of the alkaline earth silicate glasses is, in general, about an order of magnitude lower than that of alkali containing glasses of equivalent concentration. The general order of equivalent conductances for the various alkaline oxides is CaO >MgO>SiO_2>BaO. The differences between the resistivities of the high and low members of this series is very small with the resistivity of calcium containing glasses being approximately one-half that of barium containing glasses of equivalent concentration.

The change in conductivity of calcium oxide silicate glasses[36] is shown as a function of reciprocal temperature in Figure 11. The numbers on the curves represent the following percentages of CaO in the binary glasses:

Glass No.	1	2	3	4	5	6	7	8
% CaO	55	50	30	45	40	35	20	30

Binary liquid silicates containing iron and manganese modifier ions possess electrical conductivities nearly equivalent to that of the alkali silicate systems. Binary liquid silicates containing a second network former such as Al_2O_3 or TiO_2 possess electrical conductivities that are as much as three to four orders of magnitude lower than those of the alkali silicate liquids.

The electrical conductivity of glasses and liquid melts containing two alkali modifier ions has received considerable attention in the literature because of the so-called mixed alkali effect mentioned above. Tickle has established that the mixed alkali effect increases as the temperature is decreased. The conductivity of glass systems containing alkali modifier ions at high temperatures is nearly equivalent to linear additions of the conductivities of the two end members, while at lower temperatures, the conductivity is decidedly less. The mixed alkali effect may decrease the conductivity by as much as a factor of 4 at a temperature of 800°C for soda potassium silicate glasses. At 1200°C the decrease in conductivity in the mixed soda potassium glass is only a factor of 2.

Understanding the effects of composition and temperature on the electrical conductivity of the glasses of the alkali-alkaline earth-silicate systems is of the greatest importance from a commercial viewpoint. Most commercial glasses are based upon combinations of alkali and alkaline earth modifiers with silica. It is indeed unfortunate that there has been only meager systematic investigations of the electrical conductivity of glasses in these systems. Recent publications by a team of Russian investigators have been concerned with the electrical conductivity of glasses in the system Na_2O-RO-SiO_2.[40] The RO modifiers that have been studied in the ternary systems include CaO, MgO, PbO, ZnO, BaO, BeO, and CdO. The temperature range covered has been from 900°C to 1400°C. The investigations show that in the range of 1100°C to 1400°C, the conductivity of these glasses obey the well-known exponential relation, $\log \sigma = (A-B)/T$, where A and B are constants, and T is the absolute temperature.

Perhaps the most significant result of this research has been the finding that the conductivity is proportional to the concentration of sodium oxide present in the ternary systems. This result was found for all of the ternary glasses investigated. A plot of the variation of resistivity of glasses of the systems, Na_2O-RO-SiO_2, with the alkali ion concentration at 1200°C is shown in Figure 12. It can be seen from this figure that the resistivity is directly a function of concentration of sodium present in the glasses regardless of the nature and concentration of the bivalent cation present. The temperature dependence and sodium ion concentration dependence of ternary glasses in the Na_2O-RO-SiO_2 system can be expressed by the following equation:

$$\log \sigma = 1.508 - 0.0204C - \frac{4836-128C}{T}$$

where C = Na_2O content in weight percent
T = °K

The above equation has been found by the Russians, Kostanyan et al.,[41] to be valid over a temperature range of 1100°C to 1450°C for the following composition ranges: 12-20 Na_2O, 0-6 MgO, 0-11 CaO, 0-11 Al_2O_3 and 68-78 SiO_2.

It must be pointed out that these results apply to glasses with high sodium concentration (Na_2O > 10 mole %) and relatively low bivalent cation concentration. It has been stated by Kostanyan et al., that when the sodium ion concentration decreases to less than 10% or the bivalent cation concentration increases to greater than 10%, the relationship of conductivity to concentration becomes more complex.

It is tempting to extend the results of Kostanyan et al. to predict that the concentration of alkali modifier cations will be the controlling factor in the electrical conductivity of ternary glasses containing potassium or lithium ions as well as the soda containing glasses. This conclusion would seem to be reasonable considering the fact that the relative

conductivities of the alkali glasses are nearly comparable and are nearly an order of magnitude greater than the conductivities of the alkaline containing binary silicate glasses. In other words, the bivalent cations appear to have very little influence on the mobility of the sodium ions in the liquid melt and, consequently, one could conclude that they would have relatively little influence on the monovalent lithium and potassium ions as well. However, it must be cautioned that the effect of bivalent cations on the mixed alkali effect is at present unknown.

It is also of interest to note that the electrical conductivity of sodium borate glasses is also a function of sodium ion concentration. Sodium borate glasses possess approximately 50% higher values of conductivity that sodium silicate glasses of equivalent alkali concentration.

Martin and Derge[42] have conducted an extensive investigation of the effects of temperature and composition of calcia alumina silicate glags from alumina percentages of 5%-20%, and calcia percentages from 35%-50% over a temperature range of 1450°C to 1600°C. Only compositions above the liquidus surface were investigated. The results show a maximum in conductivity occurring in the region, 38% to 43% SiO_2, 44% to 50% CaO, and 10% to 15% Al_2O_3. This peak is close to the composition of the ternary eutectic in the $CaO-Al_2O_3-SiO_2$ system which is located at approximately 41% SiO_2, 47.3% CaO, and 11.7% Al_2O_3. The conductivity was observed to be a minimum at concentrations of 46% SiO_2, 37% CaO, and 17% Al_2O_3. Other than these regions of maximum and minimum conductivity, there appears to be a general trend for the conductivity to increase with increasing basicity of the melt.

Electronic Conductivity

This present section is a review of the phenomena of electronic conduction in amorphous semiconductors. An attempt has been made to unify the discussion around Anderson's theorem and hopping discussions based on Holstein's small polaron model.

Several omissions in this area are rather glaring; however, to have done justice to such topics as Hall conductivity, optical properties, and thermopower would have necessitated a great expansion into fields where rigorous theoretical work on amorphous solids is lacking, and for which there is incomplete understanding.

Band Structure

The present state of knowledge of the electronic structure of amorphous materials may best be summarized as primitive. The only calculations which have been performed are those on one-dimensional models, principally a disordered Kronig-Penney model (Figure 13). Disorder is invariably assumed to be of a strictly random nature, which is appropriate to a very high density gas but not to a glass.

This limited development is a result of two other conditions: 1) the position of the atoms in a real amorphous material has not been mathematically described, and 2) as a result, no proven methods exist for the calculation of a band structure in a glass. It is worthwhile exploring these problems a little further.

The most reliable and detailed information on the atomic structure of glasses which is currently available is the radial distribution function obtained by macroscopic diffraction methods. Accurate experimental techniques for obtaining the RDF have been developed only in the past ten years.[43] It is impossible, however, to reconstruct the three-dimensional distribution of atoms from the RDF; any given RDF can result from an essentially infinite number of three-dimensional structures. Furthermore, when more than one type of atom is present, the assignment of peaks in the RDF to specific atomic correlations is impossible without additional information. Qualitative information on the three-dimensional structure can be obtained from the RDF by regarding the atoms as nearly hard spheres, and then using geometrical arguments to construct the lattice.

For a particularly useful example, let us consider the work done on amorphous germanium.[44] It is found from the area under the first peak of the RDF which corresponds to nearest neighbor correlations that the crystalline coordination and spacing are preserved in the amorphous state. This suggests that the fundamental building block in amorphous Ge is a Ge_4 tetrahedron. These tetrahedra may be connected by having common corners. It has been suggested that the fourfold coordination of Ge may be satisfied on an extended scale if the tetrahedra are oriented in a staggered or an eclipsed configuration. Repeating only the staggered configuration leads to the diamond crystal lattice, while repeating only the eclipsed configuration (with some small distortion) leads to a regular dodecahedron, each of whose sides consist of a ring of five atoms. The lack of a periodic structure is attributed to the appearance of these elements with fivefold symmetry. By mixing the staggered and eclipsed configuration, it is found that a structure may be obtained which produces the actual RDF fairly well.[45,46] More recent work suggests that rather than the eclipsed configuration, the tetrahedra assume a configuration midway between the staggered and eclipsed configurations.[47]

We are left, therefore, with only the calculations performed on one-dimensional models. We shall review only the results of these calculations.[48]

Starting with a perfectly periodic structure, we find that all states are the normal band states (Figure 13a,b). As a small amount of disorder is introduced, the number of band states decreases and localized states are formed having energies at the top and bottom of the band. As the amount of disorder increases, more states are split off from the band and form localized states. A definite energy, E_c (Figure 14), is found below (or above) which all states are localized. Finally, if the disorder is sufficient, all states will be localized (Figure 13c,d). This latter result was first proved by Anderson in 1958.[49]

Anderson considered the tight-binding model, that is, one in which the electronic wave functions may be approximated reasonably well as a sum of atomic or molecular wavefunctions. He found that if the quantity

$$\frac{J_{ij}}{\Delta E} \tag{28}$$

is sufficiently small (of the order 1/5), then all states are localized. Here J_{ij} is defined as:

$$J_{ij} = \sum_{n \neq i} \int \psi_i^* V_n \psi_j \, dt \tag{29}$$

with ψ_i the wavefunction of the i^{th} atom or molecule,
V_n the potential of the n^{th} atom or molecule,
and ΔE is the mean spread of energy levels of the electron on each of the atoms or molecules isolated from each other.

The essential reason for this result is that due to the scattering of the wave from the various centers in which the phase undergoes a random shift, there will be at best few sites in the lattice where the scattered waves interfere constructively. If the energy of such a site is just right, then the phases can be adjusted and the wavefunction can blow up. The probability of finding such a site increases much less rapidly than the amplitude of the wavefunction decreases, provided that Anderson's criterion is met. The wavefunction therefore goes to zero at infinite distances.

Now let us consider a partially disordered solid. Considering first the sites which have smallest energy, we find that these are the least numerous (at least for random disorder), since they fall at the extremity of the energy distribution. Hence electrons localized on these sites must be the most localized (i.e., the radius of the electron cloud is the smallest), since the probability of finding another site of anywhere near the proper energy is smallest because of the low density of

these sites. This condition is analogous to the impurity problem in a semiconductor. If we have a semiconductor with a single impurity (a donor for simplicity), the energy of the donor state decreases as the charge increases, and the radius decreases.

Now let us suppose that we are in an energy region ΔE near E_c, where band states begin. Again we can identify critical sites. That is, the probability for finding an electron is large in a certain region, and we can identify a certain site which causes this modulation of the wavefunction. By this, we mean that if the energy of the site is changed appreciably, the state disappears from our range of energy interest.

We now proceed to calculate the wavefunction of an electron on each site, with all other sites so altered as to be outside the range of interest. We can use the wavefunctions so generated as a new set of "molecular" functions and apply in Anderson's theory. We will find that if $\langle J/\Delta E \rangle$ is sufficiently small, our states will be localized. Here J is the overlap integral defined in equation (29), using our "molecular" wavefunctions. By considering sites above and below E_c (the Anderson parameter of those above E_c will be exceeded), we can pin down E_c.

We need to consider one additional topic before leaving the structure problem, that of the effect of short range order on the electronic structure. As we have previously mentioned, disorder in any real glass is not completely random. If we again limit our attention to the conduction band of an n-type semiconductor, it may be possible by limiting the type of disorder to split off a relatively high density of states to below the conduction band, and hence form fairly delocalized wavefunctions, or possibly even band states. As an extreme example of this, we can imagine starting with a monatomic crystal, and then fractionally increasing every other nuclear charge, and decreasing the remaining charges by the same amount. The mean deviation of such a process would be large; however, the result will be to eventually split the conduction band in two. It is

not inconceivable that some related structure will arise in a glass. We have seen that in silicon and germanium, there are probably eclipsed and staggered configurations of tetrahedra, with comparatively small distortions of these. If the electronic energies of electrons in these configurations differ moderately, there may result a splitting of the energies into two groups about which the deviation is small. We shall consider such a possibility in a future section.

Transport

We next will consider the problem of transport in a disordered lattice. Three regimes have been recognized: band conduction, as in crystalline materials, for electrons with energy above E_c, thermally activated hopping among localized states for electrons with energy below E_c, and a diffusion mechanism for electrons with energy just slightly greater than E_c.[50]

The hopping mechanism may be understood by considering a localized electron on one site, and a second site nearby having a slightly different energy. (In order for the electron to hop it will need to gain or lose energy from thermal vibrations.) Due to its localized nature, however, the electron will polarize the lattice. Movement of the lattice distortion at high temperature (in the classical region) also requires energy.

Let us first consider the so-called non-adiabatic hopping. The classical activation energy for a hop when the energy of both sites is the same has been derived from electrostatics by Mott.[51] First, we compute the binding energy of the localized electron; that is, the energy lost by the system when the electron polarizes the lattice.

The energy of the electron in polarizing the lattice will be lowered by

$$E_e = - \frac{e^2}{r_o} \left(\frac{1}{k_\infty} - \frac{1}{k_s}\right) \qquad (30)$$

where k_s is the static dielectric constant (more precisely, to avoid many polaron effects, the dielectric constant measured at 10^{10} Hz or so), k_∞ the infinite frequency dielectric constant, and r_o the radius of the electronic state. The energy required to polarize the lattice is

$$E_\ell = \frac{1}{2} \frac{e^2}{r_o} \left(\frac{1}{k_\infty} - \frac{1}{k_s}\right) \tag{31}$$

so that the binding energy, the sum of these two terms will be

$$-E_b = E_e + E_\ell = -\frac{1}{2} \frac{e^2}{r_o} \left(\frac{1}{k_\infty} - \frac{1}{k_s}\right) \tag{32}$$

Note that since the binding energy is inversely proportional to its radius, the localized states further from the conduction band will have a greater binding energy.

In order for the electron to move classically, the electron on each of the two sites must have the same energy. The minimum energy required to do this is depicted in Figure 16.

Let the electronic energy of the electron on site 1 be 0 and on site 2 be ΔE (Figure 15a). Then when we allow the electron to polarize the lattice on site 1, the electronic energy is reduced by $2E_b$, while the lattice stores an elastic energy E_b, which means that the total energy of the system is reduced by E_b (Figure 15b). In order for the electron to jump to site 2, the electronic energies must be made equal. The minimum energy required to do this is shown in Figure 15c. If we calculated the difference in energy between this configuration, and the ground state in Figure 15b, we find that it is

$$E_H = \frac{\Delta E}{2} + \frac{E_b}{2} + \frac{(\Delta E)^2}{8E_b} \tag{33}$$

This will be the classical activation energy for hopping.

With the energy on the two sites equal, the electron will be able to tunnel from site 1 to site 2. The probability for tunneling will be proportional to

$$\exp(-2\alpha d) \tag{34}$$

where α is the tunneling constant and d the distance between the two sites. Once the electron has moved, the system will return to the configuration shown in Figure 15d, and the transition will have been accomplished.

The overall transition rate can be written as:

$$W = \nu \exp(-2\alpha d) \exp[-(\Delta E + E_b + \Delta E^2/4E_b)/2kT] \tag{35}$$

where ν is a phonon frequency. For a quantum mechanical treatment, see Schaake and Hench.[52,53]

The third term in the activation energy will be significant if $\Delta E > 2E_b$. It merely reflects the fact that if the electron is loosely bound to the lattice the probability of its gaining large amounts of energy is small.

Equation (35) is valid only over a limited range. In particular, consider the case when two sites are close together (adiabatic hopping).[54,55] Under these conditions the electron may tunnel back and forth between the two sites when their energies are similar under these conditions; the energy of the electron and hence the half-way configuration (Figure 15c) is lowered by a resonance process, yielding for the transition rate:

$$W = \nu \exp[-(\frac{\Delta E}{2} + \frac{E_b}{2} + \frac{\Delta E^2}{8E_b} - \frac{1}{2} J)/kT] \tag{36}$$

The criteria for this equation is that $J > h\omega_o$ where ω_o is the optical phonon frequency. J is as defined in equation (29).

When the temperatures are sufficiently low, neither of the above equations is valid. This results from the fact that a single phonon process may be favored over a multiphonon process such as considered above. A calculation of the transition rate for single acoustical phonons in the non-adiabatic case has been made by Schaake,[52] who finds

$$W = \frac{E_1^2 \Delta E J^2}{\pi \rho h^4 c^5} \exp[-2S(T)] \coth(-\Delta E/kT) \qquad (37)$$

where E_1 is the deformation potential, ρ the density, c the velocity of sound, and exp[-2S(T)] is a complex function of temperature which decreases more rapidly with increasing T as the binding energy is increased.

Physically, this process would correspond to the electron gaining or losing the necessary energy from a single phonon, with the electron and its distortion then tunneling to the new site.

Where the binding energy is very small, such as impurity states in silicon, the exponential term goes to unity, and the results become identical to Miller and Abrahams.[56] Of interest also is the fact that equation (35) is valid only if

$$\frac{E_b}{2h\omega_o} \coth \frac{h\omega_o}{2kT} >> 1 \qquad (38)$$

If the binding energy is very small (again as in Si), equation (37) will form a good approximation through an extensive temperature range, although interactions with single optical phonons will need to be considered also.

A consideration of the single phonon process for the adiabatic case has not been made; it appears, however, that a calculation should make the dependence on the distance between the two sites, which is implied by the overlap integral $J \propto \exp[-2\alpha d]$, disappear.

A third potential region has not been explored. If the sites are extremely close, then the electron may desire to localize itself on two sites, with the lattice polarization then surrounding both sites. Under these conditions, transitions would take place between bonding and anti-bonding orbitals, and the transition rate energy should be given by an expression in the high temperature region such as

$$E_H = \nu \exp[(-\frac{\Delta E}{2} - \frac{\Delta E^2}{8E_b})/kT] \quad (39)$$

where ΔE is the energy difference between the bonding and antibonding orbitals. The range of validity of this expression has not been determined.

Let us briefly summarize this previous discussion. For hopping of an electron between two sites, we find that the transition rate will increase as the sites come closer together, other factors being the same, and at some critical distance level off. The transition rate increases exponentially with temperature, with the activation energy also increasing. If the binding energy is small, this increase may not come until quite high temperatures. Finally, in the high temperature region, the activation energy will be smaller for closer sites, than for more distant sites.

Turning now to transport in the band states, we find comparatively little theoretical work done on them. It would appear that electrons in them would obey laws similar to conventional semiconductors with, however, a much lower mobility due to the large amount of scattering caused by the disorder.[57]

The band states just above the critical energy for localization, E_c, has been investigated by several authors.[50,52,57] In this region, the electronic wavefunction will be highly modulated, with comparatively large probabilities for finding the electron in some regions and small probabilities in the regions connecting them. In this diffusion region a mobility equation such as

$$\mu = \frac{ea^2}{kT} \nu_{el} \quad (40)$$

where a is the distance between the maxima in the probability, ν_{el} is an electronic frequency (i.e., the frequency of the electron in a maxima region).

An additional factor must be taken into consideration in these diffusion states; if the dielectron constant of the

material is high, and localization small, then there may be appreciable interaction between the lattice and the electron, leading essentially to polaron states, which we shall discuss shortly. From this, we will be able to conclude that as the temperature increases, E_c will increase, i.e., more states become localized.

Up to now, our discussion has been based on the assumption that conduction in the crystalline phase is according to conventional semiconductor theory. At least in the transition metal oxides, there is no agreement as to this.[58,59] A brief review of polaron theory is therefore in order.

Consider a narrow band semiconductor. A localized wavefunction can be constructed by summing over a number of band states. The expectation energy of such a packet will be greater than that of the states at the bottom of the band, but not by more than the width of the energy band. If the localized packet is now allowed to polarize the lattice, and if the energy decrease in polarizing the lattice is greater than the energy difference between the packet and the bottom of the energy band, then the electron will become localized, forming a polaron. Such localization occurs to some degree even in silicon; for example, see Kittel.[60] If the radius of the electron cloud is of the order of an interatomic distance, a small polaron is formed, with which we shall primarily be concerned with here.

The small polaron will be transported by band diffusion at low temperatures, but at temperatures in excess of about half the Debye temperature, the mechanism will be the hopping mechanism previously discussed.[61] All previous high temperature equations are applicable with $\Delta E = 0$.

If conduction in the crystalline phase is by the small polaron, then conduction in the glassy phase should also be. Taking into consideration the lattice interaction, we find that the Anderson criterion will be met if

$$\frac{\langle J \exp(-S(T))\rangle}{\langle \Delta E\rangle}$$

is sufficiently small, and all states will be localized.[52,62] The exponential term is as described previously. Of importance is the fact that small polaron binding energies should be of the order of at least a few tenths of an electron-volt, and hence, at temperatures above half the Debye temperature (co 150°K for most solids), all states will be localized.

Conductivity

The calculation of A.C. and D.C. conductivities in the glass are separate problems in the hopping region. Let us first consider the D.C. conductivity. In the random model, the glass is considered to consist of sites on which the electron can be localized, randomly distributed and with random energy levels. If we consider hopping between pairs of sites only, each pair of sites will be connected by an element whose conductance is:

$$\sigma_{mn} = \frac{e^2 a^2_{mn}}{kT} W_{mn} \qquad (41)$$

where a_{mn} is the site separation distance, and W_{mn} is the transition rate. Because of the exponential dependence of W_{mn} on distance due to the overlap integral, σ_{mn} will rapidly decrease with increasing distance. The conductivity of such a network in the low temperature regime has been calculated by Miller and Abrahams.[56] They find that

$$\sigma_{tot} \propto \exp(-\Delta/kT) \qquad (42)$$

where Δ is the mean spacing of the energy levels at the Fermi level; that is to say, only hops in the vicinity of the Fermi level with energy differences kT or less, contribute to the conductivity.

Several additional points need to be considered. Because of the exponential dependence of the overlap integral, the conductivity will also be proportional to $\exp(-2\alpha d)$, where d is

the mean spacing of the sites in the region of the Fermi level. As the temperature is increased we eventually enter the high temperature region, and the conductivity should become proportional to:

$$\sigma \propto \exp[-(\Delta + E_b + \frac{\Delta^2}{4E_b})/2kT] \qquad (43)$$

In other words, the activation energy will increase as multiphonon processes become important. Schnakenburg[62] has estimated that the transition between these two regions will come at about half the Debye temperature. This is only true for the case of tight binding, the more general rules stated previously will always apply. At very low temperature, $\Delta \to 0$ as pointed out by Mott,[51] since hops between distant neighbors with very small energy differences will dominate the expression.

Finally, if the semiconductor is not degenerate, that is, if the Fermi level is located in a band gap, then an additional factor of exp(-Eg/2kT) will enter into the conductivity expression, reflecting the population of carriers at the bottom of the conduction band. The meaning of Δ for a rapidly changing density of localized states with increasing energy, however, is complicated.

The A.C. conductivity of this model has been considered by Pollack and Geballe.[63] Each hop will contribute to the conductivity through the factor

$$\frac{\omega^2\tau}{1 + \omega^2\tau^2} \qquad \text{with } \tau = \frac{1}{W} \qquad (44)$$

Hops can occur between distant neighbors with a large τ, and between near neighbors with a small τ. On summing all possible hops it is found that

$$\sigma_{AC}(\omega) \propto \omega^{0.8} \qquad (45)$$

The above conductivity due to non-adiabatic hopping is additive to the D.C. conductivity. An additional contribution may come

from adiabatic hopping. Since all such hops occur with approximately the same transition rate, if $\omega < W$, we will find a contribution

$$\sigma_{AC}(\omega) \propto \omega^2 \tag{46}$$

The temperature dependence occurs through τ, and through a site population factor, $\exp[-(E-E_F)/kT]$.

If the glass is heterogeneous, that is, consists of two or more phases, then additional mechanisms are possible. These have been considered recently by Schaake.[64] The D.C. conductivity will be given by a complex mixture equation. If the conductivity of the isolated (precipitated) phase is greater than the matrix phase at high temperatures and smaller at low temperatures, then the activation energy of the composite is found to obey a simple mixture formula when the conductivities are equal, and to go to the activation energy of the matrix phase at high and low temperatures.

Two new A.C. effects may be introduced by the heterogeneities. Maxwell-Sillars losses* occur when the ratio of the dielectric constants and conductivities of the two phases differ. Since the distribution of the electric field at high frequencies is determined primarily by the conductivities, if the ratios of these constants differ, a relaxation region going from one regime to the other will be observed. If the concentration of the precipitate is very small, the real A.C. conductivity will obey an equation such as equation (44).

The second mechanism occurs if the precipitated phase has greater conductivity than the matrix phase, and if it is surrounded by a barrier through which the carriers must tunnel. If these conditions are met then the D.C. path conductivity will be dependent on frequency, since above a critical frequency, the carriers will be less and less able to complete their tunneling through the barrier. An appropriate expression is

---

*See Dielectrics section.

$$\sigma_{\text{DC path}}(\omega) = \sigma_{\text{composite}} + (\sigma_{\text{matrix}} - \sigma_{\text{composite}})(1+\omega^2\tau_o^2)^{-1} \quad (47)$$

where τ_o is the tunneling time. This equation predicts a decreasing conductivity with increasing frequency when $\omega \gtrsim 1/\tau_o$

The complex part of the conductivity has not been considered, since for each of these mechanisms it is implied to follow from the Kramers-Kronig relations.

Materials

The problem of applying the previous theory to the experimental results will concern us in this section. There is no agreement as to the conduction mechanisms, in any material at the present time; indeed, different experiments appear to give conflicting results. One aspect common to most amorphous materials is that the conductivity is insensitive to small amounts of most impurities (several per cent), in marked contrast with the situation in crystalline semiconductors, where impurities in the parts per million range cause orders of magnitude change in the conductivity. This is frequently attributed to the ability of the glass to satisfy the bond requirements of the impurities.[48] It may also be a result of the effects of the impurities being swamped by defects produced by the amorphous state.

The most frequently studied amorphous materials which exhibit semiconducting properties fall into one of three categories: elemental glasses (Ge, Si, As, Te, C, B, Sb, Se),[57] chalcogenide glasses (compounds containing S, Se, and/or Te),[65] and transition metal oxide glasses (V, Ti, and Fe oxides in particular).[66]

Some recently reported conductivity data on several amorphous semiconductors is shown in Figures 16 and 17. The high temperature D.C. activation energy in the elemental and chalcogenide glasses is attributed to be due mainly to excitation

of carriers across the band gap, as these activation energies correspond to half the optically measured gap energy. In amorphous Ge and Si, the activation energy drops with decreasing temperature; this drop is attributed to a change in the conduction mechanism to a hopping of carriers in localized states near the center of the gap. These states arise from an overlap in the localized tails from the top of the valence band and the bottom of the conduction band.

The A.C. conductivity in the elemental and chalcogenide glasses, where it has been fully explored, frequency is found to have two regions; one proportional to ω and the other, at higher frequencies to ω^2 (see data for As_2S_3, Figure 17). The linear region is evidence of non-adiabatic hopping. The ω^2 region is probably due to adiabatic hopping to nearby sites.

In the transition metal oxides there is little agreement as to whether the conduction mechanism in the crystalline material is by normal band conduction or by motion of a polaron.[58] We shall limit our discussion primarily to V_2O_5-P_2O_5 glasses.

Optical measurements indicate a band gap width of 2-2.5 ev. The high temperature D.C. activation energy is found to be from 0.2 to 0.6 ev, indicating extrinsic conduction, i.e., conduction by carriers from impurities.

Mott argues that the activation energy for conduction in the glass is due to the binding energy of the polaron.[51] This seems questionable in that according to the previous sections, all hops would have this activation energy, and an A.C. relaxation peak with this temperature dependence should be observed. In the V_2O_5-P_2O_5 glasses the A.C. conductivity shows linear and quadratic regions, but it is very slightly dependent on temperature at most.[52] This is indicative of a very small binding energy (of the order of a few hundredths ev at most) of the adiabatic hopping of the polaron for most of the localized states. If we go to a heterogeneous structure, where we have a matrix in which the binding energy is high, and controls overall conduction, we would need to introduce considerable disorder in the matrix to accomplish such localization. This

in turn would cause an appreciable amount of the activation energy to be due to the term Δ (equation (43)).

Since the activation energy also decreases to a few hundredths ev at low temperatures, this latter explanation is unsatisfactory. Furthermore, since the A.C. conduction is only slightly dependent on temperature, the majority of the activation energy cannot be attributed to the thermal population of localized states. Therefore, it appears that the semiconductor is degenerate, and that the activation energy is due to thermal activation over barriers, or due to regions in the glass having band conduction states a few tenths of an electron-volt above the Fermi level. At low temperatures, hopping in the neighborhood of the Fermi level would dominate in these regions. These two cases would be particularly attractive if a large number of states are split off of the conduction band as previously described. Optical absorption evidence at this time is inconclusive on this.[67]

The thermopower results favor a hopping mechanism.[51] The effect of a heterogeneous electronic structure as postulated above on the thermopower has not been considered theoretically, however.

Heat treatment of a quenched 30% KPO_3 - 70% V_2O_5 glass has been found to produce a marked phase separation.[68] The electrical properties of these glasses are found to obey the properties discussed under the conduction mechanisms of heterogeneous glasses.

The low frequency A.C. conductivity (Figure 17) of the heat treated glass is found to fit a Debye relaxation process with a single well defined relaxation time. This relaxation time is found to have an activation energy equal to the activation energy of D.C. conduction. This is strong evidence of Maxwell-Wagner losses (see the Dielectrics section of this paper). Presumably, these losses mask the non-adiabatic hopping which is probably also present.

At high temperatures, the D.C. activation energy has a well defined break. At temperatures above this break, the D.C.

path conductivity is found to decrease with increasing frequency, decreasing to a value at 4 MHz equal to that expected by a continuation of the low temperature activation energy. This is in agreement with equation (47), where the break in the D.C. activation energy is attributed to the onset of the conductivity in the isolated phase exceeding the conductivity in the matrix phase, while the decreasing conductivity with increasing frequency is due to a tunneling process through a barrier surrounding the isolated phase.

Dielectric Properties

A. Importance

Dielectric properties are of special importance when glass is used either as a capacitive element in electronic applications or as insulation. The dielectric constant, dielectric loss factor, and dielectric strength usually determine the suitability of a particular glass for such applications. Variation of dielectric properties with frequency, field strength, and other circuit variables influence the performance of glass. Environmental effects, such as temperature, humidity, and radiation also influence the use of glass in dielectric applications. Consequently, it is necessary to examine in this section the dielectric theory of glass in terms of materials response, circuit response, and environmental response.

Glass as an insulating material has some definite advantages over plastics which are major competitors. Glass can also be molded to a variety of shapes, although flexibility still remains a problem in most instances where this feature of insulation is required. However, glass possesses superior electrical properties, is absent from creep or deformation under stresses at room temperature, and resists environmental changes particularly at high temperatures where plastics oxidize, gassify, or decompose. Glass also can be used to form gas-tight seals with metals and other ceramic components, and therefore becomes an integral part of the electronic device.

B. Theory

Dielectric properties comprise the non-long range conducting electrical characteristics of glass. Dielectric responses result from the short range motion of charge carriers under the influence of an applied electric field. The motion of the charges leads to the storage of electrical energy and the capacitance of the dielectric. Consequently, as a first step in examining the dielectric behavior of glass, let us briefly consider the definition of capacitance. Capacitance is a measure of the ability of any two conductors in proximity to store charge when a potential difference is applied across them.

$$C = Q/V = \frac{\text{Coul}}{\text{Volt}} = \text{Farad (F)} \tag{48}$$

where Q = Coulombs and V = Volts of potential difference.

The capacitance of a vacuum capacitor is determined purely by the geometry. It can be shown from elementary electrostatics[69] that the charge density on the plates, Q, is proportional to the area (A) in meters and the electric intensity applied (E = V/d), where d is the distance between the plates (in meters). The proportionality constant is defined as $_o$, the permittivity of free space, and is equal to 8.854×10^{-12} Coulombs^2/m^2 or F/m. Thus, the capacitance of a parallel plate capacitor shown in Figure 18a will be equal to:

$$Q = qA = \pm\varepsilon_o EA = \varepsilon_o(V/d)A \tag{49}$$

$$C = Q/V = \frac{\varepsilon_o(V/d)A}{V} = \varepsilon_o A/D \tag{50}$$

where A = m^2, d = m, and C = F.

When a material is inserted between the plates the capacitance is increased.

The dielectric constant (k) is defined as the ratio of the capacitance of a condenser with a dielectric between the plates to that with a vacuum between the plates.

$$k = C/C_o = \frac{\varepsilon\ A/d}{\varepsilon_o\ A/d} = \varepsilon/\varepsilon_o \qquad (51)$$

where ε is the permittivity of the material, also in units of C^2/m^2 or F/m.

Thus the dielectric constant of a material is the ratio of the permittivity of the material to the permittivity of free space. It can be seen in Table I that typical values of the dielectric constant of insulating glasses are between 5 and 10.

The charge stored in a glass is on the order of micro-coulombs or micro-micro-coulombs. So, the permittivity of glasses is ordinarily in the range of microfarads/m or picofarads/m (micro-microfarad/m).

Equation (51) shows that the presence of the material between the plates of a condenser increases the ability of the plates to store charge. The reason for this is a result of the material containing charged species, which can be displaced in response to the field applied across the material. The displaced charges within the material comprise dipoles with a moment, $\mu = Qd$ (Coul-m), where d is the separation distance. The electric dipoles, shown in Figure 18b, are oriented with respect to the applied field. The effect of the orientation is to "tie up" charges on the plates of the condenser and thus neutralize part of the applied field. The charge which is not neutralized by dipoles within the material, called "free charge", equal to Q/k, produces an electric field and voltage towards the outside; $V = \frac{Q/k}{C_o}$. Therefore, a smaller external field is required to maintain the same surface charge because some of the charge is held by the polarization in the dielectric.

There are four primary mechanisms of polarization in glasses. Each mechanism involves a short range motion of charge and contributes to the total polarization of the material. The polarization mechsnisms include: electronic polarization (P_e), atomic polarization (P_a), orientational polarization (P_o), and interfacial polarization (P_i). A schematic of the mechanism of operation of each of these major types of polarization as well as two special cases is given in Figure 19.

Electronic polarization is due to the shift of the valence electron cloud of the ions within the material, with respect to the positive nucleus. This mechanism of polarization occurs at very high frequencies (10^{15} Hz*) which are in the ultraviolet optical range. The mechanism of polarization gives rise to a resonance absorption peak in the optical range, as shown in Figure 20. The index of refraction of the glass will depend on the electronic polarization occurring within the glass.

At frequencies in the infrared range (10^{12} - 10^{13} Hz), Figure 20, atomic or ionic polarization is occurring. Atomic polarization is the displacement of positive and negative ions in a material with respect to each other, such as is shown in Figure 19b for a SiO_4^4-tetrahedron. A resonance absorption occurs at a frequency characteristic of the bond strength between the ions. If there are several types of ions in a glass or a distribution in bond strength, the infrared absorption will be quite broad.

In the sub-infrared range of frequencies, orientational polarization contributes to the dielectric properties of glass. Orientational polarization, also referred to in some texts as dipolar polarization, involves the perturbation of the thermal motion of ionic or molecular dipoles, producing a net dipolar orientation in the direction of the applied field. Mechanisms of orientational polarization can be generally divided into two categories: molecules containing a permanent dipole moment may be rotated against an elastic restoring force about an equilibrium position. This effect is especially important for a variety of liquids and gases and polar solids, such as ice and many plastics. Two models of this type of polarization in glass are shown in Figure 19c. The left hand side of the figure depicts the oscillation of a Si-O-Si bond about an equilibrium position under a sinusoidal A.C. field. Since such a bond will possess a dipole moment when it is asymmetrical, the oscillation

---

*Hz = Hertz = cps = sec^{-1}

produces an orientational polarization. The frequency of relaxation of such a mechanism is very high, 10^{11} Hz, at room temperature. A similar mode of orientational polarization due to the oscillation of an OH^- group about an equilibrium position is shown on the right hand side of Figure 19c. The dipole moment is again due to the asymmetry of the Si-OH configurations in the random glass network. Oscillation of these moments also occurs in the range of 10^{11} - 10^{12} Hz. In several sources,[1,70] these mechanisms are termed "Stevels deformation polarization".

The second mechanism of orientational polarization is an especially important contribution to the room temperature dielectric behavior of glass. It involves the rotation of dipoles between two equivalent equilibrium positions. It is the spontaneous alignment of dipoles in one of the equilibrium positions which gives rise to the non-linear polarization behavior of the ferroelectric materials that is responsible for dielectric constant values of 10^4 or more in such materials. In linear glass dielectrics, which we are concerned with in this section, orientational polarization occurs largely as a result of motion of charged ions between the interstitial positions within the ionic structure of a glass. Figure 19d shows a schematic of orientational polarization involving the oscillation of a sodium ion between two equivalent positions. Such oscillations occur continuously, the applied field makes the jumps in a direction parallel to the field more probable. Since an appreciable distance is involved in such an ionic transition, the polarization occurs at a frequency range of 10^3 - 10^6 Hz, at room temperature. Because this mechanism involves the same mobile cations that contribute to the D.C. conductivity, several authors refer to it as "migration losses".[1,70]

The last polarization mechanism, interfacial or space charge polarization, occurs when mobile charge carriers are impeded by a physical barrier that inhibits charge migration. The charges pile up at the barrier, producing a localized polarization of the material. When an A.C. field is of sufficiently low frequency, less than 10^{-3} Hz, a net oscillation of charge

can be produced between barriers as far apart as 1 cm (Figure 19e), producing a very large capacitance and dielectric constant (Figure 20). If the barriers are an internal structural feature (Figure 19f), or the density of charges contributing to the interfacial polarization is sufficiently large, the frequency range of sensitivity for interfacial polarization may extend into the kilocycle range (Figure 20). In such a case it may be impossible to distinguish the frequency response of an orientational mechanism, such as P_{o_1}, and an interfacial mechanism such as P_{i_2}.

Of the four polarization mechanisms contributing to the dielectric properties of materials, this paper will discuss only orientational and interfacial polarization since they influence the circuit characteristics of glass. Infrared and electronic polarization in glass are more appropriately discussed in an optical properties paper.

Now that we have a qualitative understanding of the physical basis of polarization, let us consider in detail: 1) the effects of polarization on the circuit behavior of glass, and 2) a quantitative description of orientational and interfacial polarization models.

Circuit Description of a Glass Dielectric

If a sinusoidal potential $V = V_o \exp(i\omega t)$ is applied to the dielectric, the charge must vary with time, as shown in equation (52), which constitutes a charging current $Q = CV$; so

$$I_c = dQ/dt = C\, dV/dt = i\omega CV = \omega CV_o \exp[i(\omega t+\pi/2)] \quad (52)$$

The charging current in an ideal dielectric thus leads the applied voltage by $\pi/2$ radians (90°). This relationship is seen vectorially in Figure 21.

In addition to the charging current, associated with storage of electric charge by the dipoles, a loss current must also be considered for real dielectrics. The loss current arises

from two sources: 1) the long range migration of charges, D.C. ohmic conduction, and 2) the dissipation of energy associated with rotation or oscillation of dipoles. The latter contribution to the dielectric losses is a consequence of the charged particles having a specific mass and, therefore, an inertial resistance to being moved. Electrical energy from the field is lost in the overcoming of this inertia during polarization. The A.C. conduction from the inertial resistance and the D.C. conduction both are in phase with the applied voltage and thus a loss current can be written as:

$$I_\ell = (\sigma_{DC} + \sigma_{AC})\ V \tag{53}$$

where σ is the conductivity in units of mho/cm or $(\text{ohm-cm})^{-1}$.

The total current for a real material is thus:

$$I_T = I_c + I_\ell = (i\omega C + \sigma_{DC} + \sigma_{AC})\ V \tag{54}$$

As shown in Figure 21 the total current in a real dielectric is a complex quantity which leads the voltage by an angle $(90 - \delta)$ where δ is called the loss angle.

An alternative way of expressing the concept of a real dielectric possessing both charging and loss processes is to use a complex permittivity to describe the material:

$$\varepsilon^* = \varepsilon' - i\varepsilon'' \quad \text{and} \quad k^* = \varepsilon^*/\varepsilon_o = k' - ik'' \tag{55}$$

Thus, the total current in the dielectric can now be expressed in terms of the single material parameter, k*, since:

$$C = k^*C_o \quad \text{so} \quad Q = CV = k^*C_oV \tag{56}$$

and $I = dQ/dt = C\ dV/dt = k^*C_oi\omega V = (k'-ik'')C_oi\omega V_o \exp i\omega t$

and thus:

$$I_T = i\omega k'\ C_oV + \omega k''C_oV \tag{57}$$

The first term on the right hand side of equation (57) describes charge storage in the dielectric and k' is thus called the charging constant or often just dielectric constant. The second term represents the dielectric losses and ε" and k" are referred to as the dielectric loss factor and relative loss factor, respectively. The loss tangent or loss angle, or dissipation factor

$$\tan \delta = \varepsilon''/\varepsilon' = k''/k'$$

represents the relative expenditure of energy to obtain a given amount of charge storage. It is the "interest rate" so to speak. The product, k' tan δ, is sometimes termed the total loss factor and provides the primary criterion for evaluating the usefulness of a dielectric as an insulator. To minimize k", the losses in the insulator, it is desirable to have a small dielectric constant and, most important, a very small loss angle. The inverse of the loss tangent, Q = 1/tan δ, is used as a figure of merit in high frequency insulation applications. In dielectric heating the critical parameters are the dielectric constant, k', and the dielectric conductivity, $\sigma_T = \omega k''$.

Relation of Dielectric Constant to Polarization

In order to obtain a quantitative understanding of the dielectric properties of glasses it is necessary to establish a relationship between the complex dielectric constant, k*, and the polarization in the glass. This can be done by considering the total electric displacement field, D, in the glass. As shown in Figure 18b, D will be the sum of the electric field established if there was not a dielectric in the condenser plus the polarization field within the material, i.e.,

$$D = \varepsilon_o E + P = \varepsilon^* E \qquad (58)$$

Thus the total electric displacement in the material is related to the external field E by the complex permittivity of the material, ε^*. Consequently, the polarization can be expressed as:

$$P = E(\varepsilon^* - \varepsilon_o) \tag{59}$$

since: $k^* = \varepsilon^*/\varepsilon_o$ then $P = E(\varepsilon_o k^* - \varepsilon_o) = \varepsilon_o(k^*-1)$

and rearranging: $k^*-1 = P/\varepsilon_o E$ or $k^* = 1 + P/\varepsilon_o E$ (60)

and defining: $P/\varepsilon_o E \equiv \chi$, the electric susceptibility (61)

we obtain: $k^* = 1 + \chi$ (62)

Equation (60) provides the relationship desired between the dielectric constant and the total polarization in the glass. However, it would be even more useful to have a relationship between k* and the fundamental polarizability of charge mechanisms contributing to the total polarization, P.

Such a relationship can be obtained through the following arguments, P is equal to the total dipole moment induced in the material by the electric field. Thus:

$$P = N_i \bar{\mu}_i \tag{63}$$

where N_i is the number of dipoles of type i and $\bar{\mu}_i$ is their average dipole moment.

Now, the average dipole moment of the charged particles is proportional to the local electric field (E') which acts on the particle:

$$\bar{\mu}_i = \alpha_i E' \tag{64}$$

where α_i is the polarizability or average dipole moment per unit local field strength and has units of $C^2 sec^2/kg$.

Thus the total polarization is:

$$P = N_i \alpha_i E' \tag{65}$$

For gases with little molecular interaction the locally acting field E' is the same as the external field, E. However, for solids, including glasses, polarization of the surrounding medium affects the magnitude of the local field. Mosotti was the first to derive the local field contribution by the integration of the normal component of the polarization vector over the surface of a spherical cavity in the material.[71] The result obtained is:

$$E' = E_{(applied)} + P/3\varepsilon_o \tag{66}$$

so since $N_i \alpha_i = P/E'$ from equation (65), then:

$$N_i \alpha_i = \frac{P}{E_a + P/3\varepsilon_o} \tag{67}$$

and by substituting equations (61) and (62) in that order we obtain:

$$N_i \alpha_i = \frac{1}{\frac{1}{(k^*-1)\varepsilon_o} + \frac{1}{3\varepsilon_o}} \tag{68}$$

Further rearrangement yields:

$$N_i \alpha_i = \frac{3\varepsilon_o{}^2 (k^*-1)}{\varepsilon_o (k^*+2)} \tag{69}$$

so: $$\frac{k^*-1}{k^*+2} = \frac{1}{3\varepsilon_o} N_i \alpha_i \tag{70}$$

This result is the famous Clausius-Mosotti equation which describes the relation between the complex dielectric constant of a material and the number of polarizable species N_i, and the

polarizability of the species α_i. As we saw earlier, there are four major classes of polarizable species in glass, α_e, α_a, α_d and α_i. Thus:

$$\frac{k^*-1}{k^*+2} = \frac{1}{3\varepsilon_o}\left[N_e\alpha_e + N_a\alpha_a + N_d\alpha_d + N_i\alpha_i\right] \qquad (71)$$

Orientational Polarization Theory

As was mentioned earlier, the physical theory of resonance absorption which explains the contribution of α_e, and α_a to k^* in the optical and infrared region, will not be considered in this discussion.

It is important to discuss orientational polarization, however, since it markedly influences the insulation and capacitive applications of glass and is strongly affected by the composition, structure, and thermal history of glass. For several other theoretical treatments the reader is referred to Frohlich[72] and Daniels.[73]

Let's consider a bistable dipole model such as shown in Figure 19d. A change in the coordination of the Na^+ ion from the position at left to that at right involves a change in energy as represented in Figure 1. There is a random oscillation of Na^+ ions between these positions at any temperature above 0°K with the probability of a jump, ω, being exponentially related to the temperature and the energy barrier W:

$$\omega = A\, e^{-W/kT} \qquad (72)$$

where $k = 1.37 \times 10^{-16}$ erg/°K, and T = °K.

However, with an electric field applied the potential energy of the two sites will become unequal by an amount:

$$\phi_1 - \phi_2 = e(bE) = ebE\cos\theta \qquad (73)$$

where b is distance separating the potential wells and θ is the angle between the field vector and the jump vector. Thus, this model is equivalent to a turn of 180° of a dipole of the moment:

$$\mu_d = 1/2\ Zeb \tag{74}$$

where Z is the valence of the ion.

Let us assume that there are N bistable dipoles per unit volume with N being small enough that there will be no dipolar interaction. Also assume that cos θ = 1 for all dipoles, $\omega_1 = \omega_2$ without a field applied, and W >> kT.

The probability of jumps from 1 to 2 can thus be written as:

$$\omega_{12} = A \exp - \left[\frac{W+ E}{kT}\right] \tag{75}$$

or $$\omega_{12} = A \exp - [W/kT] \exp - [\mu E/kT] \tag{76}$$

Since $\mu \cong 10^{-18}$ esu, then as long as E is less than 10^5 esu where dielectric breakdown begins to occur, $\mu E/kT$ will be much less than unity. So if the last term in equation (32) is expanded as

$e^{-x} = 1 - x + x^2/2 + \ldots$, we obtain:

$$e^{-\mu E/kT} = 1 - \mu E/kT \tag{77}$$

which on substitution into equation (76) yields:

$$\omega_{12} = A \exp - [W/kT](1-\mu E/kT) = \omega(1-\mu E/kT) \tag{78}$$

The probability of a jump of the ion in the opposite direction will be:

$$\omega_{21} = A \exp - [W-\mu E/kT] \tag{79}$$

which yields the following expression from the above argument equation (77):

$$\omega_{21} = \omega(1+\mu E/kT) \tag{80}$$

Under equilibrium conditions the average population of charges in wells 1 and 2 will not change with time. Consequently, those going into well 1 must come out of well 2 and vice versa. Thus:

$$N_1\omega_{12} = N_2\omega_{21} \tag{81}$$

Thus: $$N_1\omega(1-\mu E/kT) = N_2\omega(1+\mu E/kT) \tag{82}$$

which upon rearranging yields:

$$N_1 - N_2 = (N_1+N_2)\ \mu E/kT \tag{83}$$

And since the number of wells occupied per unit volume is constant: $$N_1 + N_2 = N \tag{84}$$
the polarization per unit volume is:

$$P = (N_1 - N_2) \tag{85}$$

Therefore: $$P = \frac{N\mu^2 E}{kT} \tag{86}$$

and since $\mu = 1/2\ Zeb$, $$P = \frac{Z^2Ne^2}{4}\left(\frac{b^2E}{kT}\right) \tag{87}$$

Consequently, the dielectric constant of a glass containing neighboring positions of such bistable dipoles depends on the number of dipoles, oscillation length, and temperature as:

$$k = 1 + \frac{P}{\varepsilon_o E} = 1 + \frac{Z^2Ne^2b^2}{4kT} \tag{88}$$

Time Dependence of Orientational Polarization

In order to describe the A.C. behavior of the dielectric properties of glass, it is necessary to discuss the time dependent response of the above bistable dipole model. The change in number of dipoles in site 1 is equal to the outflow to site 2

minus the inflow from site 2; thus:

$$\frac{dN_1}{dt} = -N_1\ 12 + N_2\ 21 \tag{89}$$

and since $N_1+N_2 = N$(constant)

$$\frac{dN_2}{dt} = -dN_1/dt \tag{90}$$

and thus $$\frac{d(N_1-N_2)}{dt} = 2\ dN_o/dt \tag{91}$$

By using expansions for $E/k << 1$, as in equation (77), we can obtain for the rate of change of occupancy:

$$\frac{1}{2}\left(\frac{d(N_1-N_2)}{dt}\right) = N_1\omega_{12} + N_2\omega_{21} = -N_1\omega(1-\mu E/kT) + N_2\omega(1+\mu E/kT) \tag{92}$$

By rearranging this result is obtained:

$$\frac{1}{2}\left(\frac{d(N_1-N_2)}{dt}\right) = -\omega(N_1-N_2) + \omega(N_1+N_2)\ \mu E/kT \tag{93}$$

which is a differential equation for N_1-N_2. Since $P = (N_1-N_2)$, equation (93), is also a differential equation describing the time dependence of P:

$$\frac{1}{2}\left(\frac{dP}{dt}\right)\frac{1}{\mu} = -\omega/\mu(P) + \omega N\mu E/kT \tag{94}$$

which upon rearranging yields:

$$\frac{1}{2}\omega\left(\frac{dP}{dt}\right) + P = N\mu^2 E/kT \tag{95}$$

Equation (95) is a relaxation equation with a relaxation time $\tau = \frac{1}{2}\omega$ characteristic of the rate of relaxation.

So, $$\tau \frac{dP_D}{dt} + P_D = N\alpha_D E = P_S \qquad (96)$$

where α_D is the dipolar polarizability, μ^2/kT, and the subscript D has been assigned to designate that it is the dipolar polarization that is time dependent. In this sense P_S is the static or zero frequency value of the polarization. Equation (96) shows that when E changes with time, P, at a given moment, will generally differ from P_S. As the oscillation time increases there will be a trend towards an equilibrium value of P_S at a rate of change of dP_D/dt.

Integration of equation (96) will lead to a solution of the time or frequency dependence of the dipolar contribution to the dielectric constant. However, several steps of simplification are desirable. For example, since α_e and α_a occur very rapidly; i.e., when $\tau > 10^{-11}$ sec, a high frequency polarization contribution can be defined as:

$$P_\infty = P_e + P_a \qquad (97)$$

and at frequencies of 10^{11} sec^{-1} or greater the dielectric constant will depend only on P_∞ as:

$$k_\infty \equiv 1 + \frac{P_\infty}{\epsilon_o E} \cong n^2 \qquad (98)$$

where n is the index of refraction of the glass.

Consequently, at frequencies of 10^2 sec^{-1} to 10^{11} sec^{-1}, the low frequency or static value of the dielectric constant (k_S) can be expressed as:

$$k_S - 1 = \frac{P_D + P_\infty}{\epsilon_o E} \qquad (99)$$

By inserting equation (98) into (99) and simplifying, the dipolar polarization can be written in terms of the relaxed static dielectric constant, k_S, and the unrelaxed high frequency constant, k_∞.

$$k_S - 1 = \frac{P_D + (k_\infty - 1)\varepsilon_o E}{\varepsilon_o E} \tag{100}$$

$$(k_S - k_\infty)\varepsilon_o E = P_D = P_S - P_\infty \tag{101}$$

Upon inserting equation (101) into equation (96), one obtains the differential equation describing the change of the dielectric constant from k_S to k_∞ resulting from the dependence of the dipolar polarization:

$$\tau \frac{dP_D}{dt} + P_D = P_S - P_\infty = (k_S - k_\infty)\varepsilon_o E \tag{102}$$

The general solution of the above equation is most easily obtained by using complex variables; i.e., let $E^* = E_o \exp(i\omega t) = E_o (\cos \omega t + i \sin \omega t)$. Therefore, the real part of $E^* = E = E_o \cos \omega t$. If we assume that relevant physical laws hold for E^* as for E, we can express the time dependent orientation polarization as:

$$P^*_D = \varepsilon_o k^* E^* \tag{103}$$

Substitution of equation (99) into equation (97) and solving for P^*_D yields:

$$P^*_D = k_e^{-t/\tau} + \frac{(k_S - k_\infty)_o}{1 + i\omega\tau} E_o \exp(i\omega t) \tag{104}$$

The first term on the right hand side of the equation describes the time dependent decay of the D.C. charge on the capacitor. The second term describes the A.C. behavior of the polarization when a field of magnitude E_o and frequency ω is applied.

Since the electronic and atomic polarization are frequency independent in the range of interest, they can be separated from the time dependent expression of k^* by the following definition:

$$k^* - k_\infty = \frac{P^*_D}{\varepsilon_o E^*} \tag{105}$$

Now by substitution of equation (101) into the above definition:

$$k^* - k_\infty = \frac{k_S - k_\infty \varepsilon_o E^*}{1 + i\omega\tau \varepsilon_o E^*} \tag{106}$$

or $$k^* = k_\infty + \frac{k_S - k_\infty}{1 + i\omega\tau}$$

And, recalling that:

$$k^* = k' - ik''$$

we can separate equation (106) into real and imaginary parts as:

$$k' = k + \frac{k_S - k_\infty}{1 + \omega^2\tau^2} \tag{107}$$

and $$k'' = (k_S - k_\infty) \frac{\omega\tau}{1 + \omega^2\tau^2} \tag{108}$$

and $$\tan \delta = \frac{k''}{k'} = \frac{(k_S - k_\infty)\omega\tau}{k_S + k_\infty\omega^2\tau^2} \tag{109}$$

Equations (107) to (109) are the desired frequency dependent relationships of the charging and loss constants and the loss tangent. These equations are known as the Debye equations and they yield the graphical relationships shown in Figure 22.

At low frequencies the charging constant is frequency independent at a value which reflects the contribution of $P_e+P_a+P_d$. As the applied frequency $\omega = 2\pi f$ approaches a value equal to $1/\tau$, k' passes through an inflection and at a higher frequency becomes asymptotic to k_∞ which depends only on P_e+P_a When $\omega = 1/\tau$, the oscillating charges are coupled directly with the oscillating field and absorb a maximum in electrical energy and k'' goes through a maxima as a result. The magnitude of the loss peak will be $(k_S-k_\infty)/2$ and thus is directly dependent on the number of oscillating charges and their distance of motion

as given in equation (87). Since P_D is inversely proportional to temperature, equation (87), the magnitude of the peak height should vary with the reciprocal of the absolute temperature of the glass.

Figure 22 and equation (109) also show that the loss tangent goes through a maximum when plotted as a function of log frequency. However, the loss tangent peak is displaced to a higher frequency by the quantity $\sqrt{k_s/k_\infty}$. Consequently, it is often possible to observe a maxima in tan δ where the k" maxima is at a frequency too low to be measured.

Temperature Dependence of Orientational Polarization

The development of the bistable dipole model also provides a basis for understanding the temperature dependence of the dielectric properties of glass. Equation (77) expressed the probability of a transition of the oscillating charge in terms of Boltzmann statistics as $\omega = A \exp[-Q/kT]$. It was also shown in equation (96) that ω is directly related to the relaxation time of the orientational process, $\tau = 1/2\ \omega$. Consequently, the relaxation time is exponentially temperature dependent:

$$\tau = \frac{1}{2A} \exp[+Q/RT] \qquad (110)$$

where Q is an activation energy for the relaxation process in units of kilocalories per mole. It is also possible to define an intrinsic relaxation time for the relaxation process as:

$$\frac{1}{2A} \equiv \tau_o \qquad (111)$$

where τ_o will be proportional to a term $(\frac{\alpha kT}{h})^{-1}$, which includes a frequency factor, $\frac{kT}{h}$, describing the average frequency of oscillation of an harmonic oscillator in an assembly of temperature, T. τ_o also includes the term α which is an accommodation coefficient that is related to the irreversibility of the oscillations.

By combining equation (110) and the Debye equations it is possible to describe the temperature variation in the location of the tan δ loss peak and the ε'' loss peak. Since tan δ_{max} occurs when $\omega = \sqrt{k_s/k_\infty}/\tau$ and $\omega = 2\pi f$

$$\text{then} \qquad f_{max}(\tan \delta) = \frac{\sqrt{k_s/k_\infty}}{2\pi\tau_o \exp[+Q/RT]} \tag{112}$$

$$\text{or} \qquad f_m = f_o \sqrt{k_s/k_\infty} \exp[-Q/RT] \tag{113}$$

$$\text{and similarly:} \quad f_m(\varepsilon'') = \frac{1}{2\pi\tau}$$

$$\text{so:} \qquad f_m(\varepsilon'') = f_o \exp[-Q/RT] \tag{114}$$

As shown in Figures 23 and 24, tan δ_{max} will vary with temperature in the manner predicted by equation (113). The slope of the log f_m versus $1/T\ ^\circ K^{-1}$ plot in Figure 24 will be equal to $-Q/R$ as given in equation (113). Thus, an experimental measurement of the dielectric loss peaks as a function of temperature enables calculations of relaxation activation energies to be made.

There are several other important parameters which can also be obtained from dielectric data. τ_o can be calculated from data such as shown in Figure 24. Values of 10^{-13} sec to 10^{-14} sec are indicative that the relaxation process is ionic in origin. However, very low values of τ_o are suggestive that the relaxation process is electronic space charge polarization.

Since the quantity, $k_s - k_\infty$, is directly related to the dipolar polarization, P_D, in the glass, the magnitude of tan δ_{max} or ε''_{max} can be used to calculate the number of oscillating charges and their distance of oscillation by employing equation (87). The inverse temperature dependence of P in equation (87) also provides somewhat of a test for the presence of dipolar polarization in the glass. When the measuring temperature increases, the polarizability of the dipoles decreases and the loss peak height should decrease. However, the rapid increase

in σ_{DC} with increasing temperature in glass usually results in the A.C. loss peak being swamped in magnitude by D.C. losses at temperatures of one-fourth the softening point. Consequently, analysis of a decrease in the height of tan δ peaks with temperature is often not possible. If temperature independence can be observed, it is suggestive that an interfacial polarization may be operating.

Dielectric loss data in glasses can also be employed as a measure of the distribution of ion energies or configurations in the glass structure. Often, an appreciable broadening in the Debye peaks is observed. The broadening is due to a distribution in relaxation times which may be a consequence of a distribution of Q or a distribution in the oscillation distances between equivalent sites, b. There are a variety of techniques in the literature[73] which can be used to analyze distribution of relaxation times. Often a log normal distribution is a satisfactory description:

$$\text{Let} \qquad \frac{\varepsilon''}{\varepsilon'_s - \varepsilon'_\infty} = \int_0^\infty \frac{\omega\tau}{1 + \omega^2\tau^2}\, g(\tau)\, \ell n\tau \qquad (115)$$

where $g(\tau)$ is a weighting function which gives the relative strength of a given relaxation time, τ_i. Assuming a log-normal distribution yields:

$$\frac{\varepsilon''}{\varepsilon'_s - \varepsilon'_\infty} = \int_0^\infty \frac{\omega\tau}{1 + \omega^2\tau^2} \frac{1}{(2\pi)^{1/2}\, s} \exp - \frac{1}{2} \left[\frac{\ell n\tau/\tau_o}{s}\right] d\ell n\ \tau/\tau_o \qquad (116)$$

where s is the standard deviation of the distribution. This formidable relation comes out to be an error-function type of solution for s and has been nicely published in tabular form with instructions as to application by Nowick and Berry.[74]

Analyses of loss peak distribution functions have led to a number of authors concluding that there is a distribution in cation activation energies in glass.[1,75] Several other authors[1,76] have reached the conclusion that the broadening

response is due to variations in ion configurations. Thus, if Q is represented as $Q = \Delta F = \Delta H - T\Delta S$, there will be localized variations in the entropy within the glass due to the randomness of the structure.

Interfacial Polarization

Although the atomistic dipolar model may provide a reasonable interpretation of the dielectric behavior of many glasses, it does not usually represent a unique interpretation. The frequency response and frequency range of interfacial polarization mechanisms is equivalent to that of the Debye type of model in many cases. Consequently, the possibility of an interfacial interpretation of the dielectric behavior of a glass must always be considered as an alternative to the Debye model. Independent evidence that interfaces are or are not involved in the behavior is required to achieve an unambiguous interpretation of glass behavior.

There are two principal categories of interfacial polarization and both can be important for glasses. The first involves a variation of electrode polarization. The second consists of the dielectric response due to heterogeneities in the glass.

If a sharp compositional gradient is present at one of the electrode-glass interfaces due to electrode polarization, heat treatment, or atmospheric attack, a two layer capacitor results.* As shown in Figure 25, such a two layer capacitor involves a region described by a permittivity, ε_1, a conductivity, σ_1, and a thickness t_1, which is in series with the matrix of ε_2, σ_2, and t_2. Analysis of this model and numerous other series and parallel R-C circuit models of materials have been reported in the literature.[73,77] The frequency dependence of the dielectric parameters of the glass model in Figure 25 are:

---

*Water films and oxidized electrode layers produce especially strong contributions to interfacial polarization.

$$k'_{(\omega)} = \frac{k_s + k_\infty \omega^2\tau^2}{1 + \omega^2\tau^2} \tag{117}$$

and $$\sigma_{(\omega)} = \frac{\sigma_s + \sigma_\infty \omega^2\tau^2}{1 + \omega^2\tau^2} \tag{118}$$

and $$\tan \delta = \frac{\sigma_\infty}{\omega k_\infty} \frac{(\frac{k_\infty \sigma_s}{k_s \sigma_\infty}) + \omega^2\tau\delta^2}{1 + \omega^2\tau\delta^2} \tag{119}$$

The relaxation times and static and high frequency parameters are defined as below:

$$\tau = \frac{k_s - k_\infty}{\sigma_\infty - \sigma_s} = \frac{k_1 d_2 + k_2 d_1}{\sigma_1 d_2 + \sigma_2 d_1} \tag{120}$$

$$k_s = \frac{(d_1 + d_2)}{(\frac{d_1}{\sigma_1} + \frac{d_2}{\sigma_2})^2} \cdot \frac{d_1 k_1}{\sigma_1^2} + \frac{d_2 k_2}{\sigma_2^2} \tag{121}$$

and $$k_\infty = \frac{d_1 + d_2}{\frac{d_1}{k_1} + \frac{d_2}{k_2}} \tag{122}$$

$$\sigma_s = \frac{d_1 + d_2}{\frac{d_1}{\sigma_1} + \frac{d_2}{\sigma_2}} \tag{123}$$

$$\sigma_\infty = \frac{d_1 + d_2}{(\frac{d_1}{k_1} + \frac{d_2}{k_2})^2} (\frac{d_1 \sigma_1}{k_1^2} + \frac{d_2 \sigma_2}{k_2^2}) \tag{124}$$

The frequency response of equations (117), (118) and (119) is equivalent to that shown in Figure 22 for the Debye equations. k' goes through a relaxation decay from k_s at low frequencies to k_∞ at high frequencies described by the relaxation time τ. k" and tan δ likewise go through maxima as a function of frequency. Rearranging equation (119) to obtain the location of

the loss angle maximum on a tan δ-log frequency plot yields:

$$f_{m_{(\tan \delta)}} = \frac{\sigma_1 t_2 + \sigma_2 t_1}{2\pi[k_1 t_2 + k_2 t_1]} \left(\frac{k_s}{k_\infty}\right)^{1/2} \qquad (125)$$

Equation (125) indicates that a glass dielectric may be responding as a two layer capacitor due to a surface film, for example, and exhibit a loss peak in the same frequency range as Debye peaks if the ratios of σ_1/σ_2 and t_1/t_2 are appropriate. Since σ_1 and σ_2 will also be exponentially temperature dependent, the loss peaks of the two layer glass capacitor will also shift to higher frequencies with increasing temperature. However, the interpretation of such a shift is not straightforward, since Q_1 and Q_2 for the two conducting layers may be different. Therefore the ratio of σ_1/σ_2 may change with temperature as well, which would produce a frequency shift proportional to $\Delta(\sigma_1/\sigma_2)$.

Heterogeneities are another source of polarization in glass. Wagner[78] was the first to extend Maxwell's[79] multilayer dielectric analysis to include dispersed spherical phases. Sillars[80] later generalized heterogeneous dielectrics to include dispersed ellipsoidal phases. An important consequence of the final M-W-S theory is that the dielectric parameters of a glass may be interpreted in terms of a distribution of ellipsoidal particles of a given volume fraction and size, instead of atomistic dipoles as Debye theory requires. Since the frequency location of loss peaks and the frequency and temperature dependence of k' and k" can be equivalent in both the M-W-S and Debye theories, independent confirmation of a model for the dielectric behavior of a particular glass series must be obtained.

A schematic of the M-W-S heterogeneous dielectric model is shown in Figure 26. The diagram shows isolated ellipsoidal particles of conductivity σ_2 in a matrix of conductivity σ_1. When σ_2 is greater than σ_1, the phase boundary between the particles and the matrix will serve as a barrier to charge migration and A.C. charge oscillation and interfacial polarization will occur (Figure 19f). Sillars has shown that for a volume

fraction of spheroids, q, of conductivity σ_2 and permittivity ε_2 and axial ratio of a/b, imbedded in a perfect dielectric matrix of conductivity σ_1 and permittivity ε_1, the dielectric parameters of the composite material are:

$$\varepsilon' = \varepsilon_\infty + \frac{\varepsilon'_1 N}{1 + \omega^2\tau^2} \tag{126}$$

$$\varepsilon'' = \frac{\varepsilon'_1 N\omega\tau}{1 + \omega^2\tau^2} \tag{127}$$

$$\tan\delta = \frac{\varepsilon'_1 N\omega\tau}{\varepsilon_\infty(1+\omega^2\tau^2) + \varepsilon'_1 N} \tag{128}$$

where $$\tau = \frac{[\varepsilon'_1(\lambda-1) + \varepsilon'_2]\,\varepsilon_o}{\sigma_2} \tag{129}$$

and $$N = q\,\frac{\lambda^2\varepsilon'_1}{\varepsilon'_1(\lambda-1) + \varepsilon'_2} \tag{130}$$

and $$\varepsilon_\infty = \varepsilon'_1\left[1 + q\,\frac{\lambda(\varepsilon'_2-\varepsilon'_1)}{\varepsilon'_1(\lambda-1)+\varepsilon'_2}\right] \tag{131}$$

The relationship between λ and a/b has been evaluated by Sillars and is reproduced in part in Figure 27.

The similarity of the form of the above equations with the Debye equations is readily apparent. However, the location of the M-W-S dielectric relaxation is primarily a function of the axial ratio of the ellipsoidal particles and the conductivity of the dispersed phase. For example,

$$f_{m_{(\varepsilon'')}} = \frac{1}{2\pi\tau} = \frac{1}{2\pi}\,\frac{\sigma_2}{[k'_1(\lambda-1) + k'_2]} \tag{132}$$

If the axial ratio of particles changes during a growth process from spherical to dendritic with an equivalent ellipsoidal ratio of a/b = 7, then the loss peak location would shift downwards

by an entire order of magnitude. If the conductivity of the dispersed phase were decreasing as a result of growth, the downward shift would even be more marked. However, elongated growth accompanied by a $+\Delta\sigma_2$ would tend to be self-compensating.

The M-W-S model also shows that the magnitude of both ε' and ε'' are dependent on the volume fraction of the dispersed particles and their axial ratio. Consequently, a change in magnitude of ε' or ε'' cannot be unambiguously interpreted unless other evidence is available to independently estimate axial ratios or the volume fraction. With either information, though, it is possible to obtain the other structural features from the above equations.

Composition and Structural Effects

The next objective is to attempt to interpret the dielectric behavior of some selected glass compositions in terms of the various theories presented above. Special emphasis will be placed on the relation of structural models to the behavior observed.

A number of dielectric studies of fused SiO_2 [81,82] have demonstrated the presence of two dielectric loss peaks at cryogenic temperatures. Recent results [83] have shown that the magnitude of the loss peak at 1.3°K can be related to the H_2O content of the glass (Figure 28). The second cryogenic fused silica loss peak occurs in the temperature range of 28-60°K at audio frequencies and has an activation energy of 1.2 kilocalories per mole. Consequently, the peak shifts to a frequency range of 10^{11} - 10^{12} Hz at room temperature, beyond the range of concern for audio or radio circuit applications of fused silica, but certainly within the range of interest for microwave devices.

The water based loss peak has been analyzed as an oscillating permanent dipole and can be described by the PO_2 mechanism shown on the right side of Figure 19c. The higher temperature peak is associated with what Stevels terms "deformational losses". The dipolar mechanism, Pd_2, that has been suggested

for the peak involves the oscillation of the bridging oxygens in nearly colinear Si-O-Si bonds.[84] The model suggests that if the Si-O-Si bond is nearly 180° it will be possible for the oxygen to oscillate between two equivalent positions, as Figure 19c, with only a small energy barrier. No spatial rearrangements or bond-breaking would be involved. This interpretation would seem to be in reasonable agreement with the latest structural interpretation of fused silica advanced by Warren and Mozzi.[85] They concluded that the bond angles could be best described by a skewed Gaussian distribution about 144°. The half breadth of the distribution is nearly 35° and about 10% of the bond angles lie between 175° -180°. It is presumably the oscillation of bonds within the tail of this distribution that produces the "deformational losses".

Perhaps the most detailed atomistic analysis of a dipole relaxation mechanism in glass has been published by Charles[86] In this publication he attempts a quantitative correlation of dielectric relaxation and D.C. conductivity in an alkali-silica glass. These are two basic postulations: 1) that near a given non-bridging oxygen ion in a glass there are several equivalent sites for an alkali metal ion, and 2) that the alkali ion will respond to the field in two major types of motion. The alkali ion may oscillate between equivalent sites about the same non-bridging oxygen which leads to A.C. polarization but no D.C. conduction. Or, the alkali ion may exchange non-bridging oxygen ions which usually will involve both A.C. polarization and D.C. conduction. Charles analyzed these two types of alkali ion motion in detail using a modification of the ionic conductivity and bistable dipole models presented herein and obtained reasonable agreement between theory and observation.

Charles[27] has also examined the influence of microstructure on the dielectric behavior of several alkali-silicate glasses. In this classic experiment he found that a rapid quency of 7.5 mole % and 15 mole % Li_2O-SiO_2 glasses changed a dispersed Li_2O rich glassy phase to a tubular-interconnected phase. The dielectric losses of the rapidly quenched sample

were increased by an order of magnitude and the location of the loss peak also shifted by a factor of 10 to higher frequencies. Binary silicate glasses containing higher percentages of Li_2O, up to 40 mole %, did not show appreciable changes in dielectric loss upon changes in quench rate and the microstructure also remained that of a dispersed lithia rich phase in a SiO_2 rich matrix. His interpretation of the change in loss behavior with thermal history of the various glasses is based on Maxwell-Wagner-Sillars interfacial polarization. The conclusion that was reached was that the continuously distributed lithia-rich phases associated with spinodal decomposition gives rise to a large A.C. loss due to a very high a/b axial ratio in the M-W-S heterogeneous loss model. The nearly spherical dispersed phase, associated with a nucleation and growth phase separation, exhibits a much lower loss due to the low axial ratio in the interfacial polarization model.

An interfacial M-W-S polarization model has also been used by Kinser and Hench[30,32,33] to interpret the changes in dielectric properties associated with the thermal treatment of several Li_2O-SiO_2 and Na_2O-SiO_2 glasses. They found, as shown in Figure 23, that appreciable losses occur in 30 mole % Li_2O-SiO_2 glasses after a heat treatment of 5 hours at 500°C. Their results also show, Figure 29, that loss peaks are not present for the same samples before they are heat treated in contrast to Charles' data.[86] The source of the dielectric losses in the heat treated glasses has been attributed to the presence of a high conductivity metastable crystal phase of lithium metasilicate. Guinier x-ray diffraction, hot stage electron microscopy, and x-ray small angle scattering were also used to establish the presence of the metastable phase. Analysis of the change in frequency of the dielectric loss peak with thermal treatment employing the M-W-S theory led to the conclusion that the axial ratio of the metasilicate precipitate was decreasing from 100/1 to 5/1 during the nucleation of the equilibrium lithium disilicate phase. A summary of the changes in dielectric loss with thermal treatment is given in Figure 30. Similar

results have also been found in 26.4 % Li_2O-SiO_2 glasses as well as 33 mole % Li_2O-SiO_2 glasses, which are not phase separated, and 33 mole % Na_2O-SiO_2 glasses. The activation energy for the loss process, Figure 24, was calculated to be 15 kilocalories/mole, equivalent to the activation energy for D.C. conduction. Extension of the temperature range has shown that the appearance and disappearance of the dielectric loss peaks in the Li_2O-SiO_2 glasses occurs at least as low as 475°C, 24 hours required, and as high as 525°C. Consequently, the presence of dielectric loss peaks in some glass systems may be attributable to annealing treatments that result in precipitation of metastable crystallites which give rise to M-W-S interfacial losses.

Thus, it appears at the present time, based on the case histories cited, that the detailed analysis of dielectric behavior of glasses with well characterized water contents, complete amorphous x-ray structure analysis, transmission electron micrographic evaluation of phase separation, and controlled thermal histories are needed in order to firmly develop a generalized view of the dielectric properties of glass.

Switching Behavior

In recent years, considerable attention has been given to the ability of amorphous semiconductors (and many crystalline semiconductors) to switch from a state of low conductivity to one of high conductivity. A typical current versus voltage curve is shown in Figure 31.

An explanation of this behavior has been given by Mott.[50] Due to the presence of localized electron states, carriers are injected into the semiconductor by the electrodes forming a space charge region (Figure 32). This concentration of carriers causes the Fermi level of the semiconductor to drop with respect to the Fermi level of the electrode. Carriers from the electrodes may now be injected into the semiconductor in the high mobility region where the density of semiconductor states is

high. This leads to a decreased resistance. When the field is removed, the semiconductor returns to its initial high resistance state.

A second type of glass has the ability to remember its low resistance state when the field is removed. This effect has been shown to result from a structural change in a filamentary path between the two electrodes.[87,88] Thus, when a pulse with a slow fall time is applied to the glass, a high conductivity filament is formed due to melting or softening of the glass, and the device is "on". When a pulse with a fast fall time is subsequently applied, the filament is destroyed, the glass returning to its original state, turning the device "off". It appears at this time that the high conductivity filament results partially from crystallization of the glass, and partially due to a change in the glass structure resulting from going above the glass transition temperature.[89]

BIBLIOGRAPHY

1. A. E. Owen, "Electric Conduction and Dielectric Relaxation in Glass," Progress in Ceramic Science, Vol. 3, pp. 77-1971, J. Burke, ed., Pergamon Press (1963).

2. J. D. Mackenzie, in Modern Aspects of the Vitreous State, Vol. 3, ed., J. D. Mackenzie, pp. 126-147, Butterworths, Washington (1964).

3. A. D. Pearson, in Modern Aspects of the Vitreous State, Vol. 3, ed., J. D. Mackenzie, pp. 29-58, Butterworths, Washington (1964).

4. G. W. Morey, The Properties of Glass, 2nd Ed., pp. 465-542, Reinhold Publishing Corp. (1954).

5. M. B. Volf, Technical Glasses, Putman & Sons, Ltd., London (1961).

6. W. Doremus, "Semiconducting Effects in Amorphous Solids," reprinted from J. Non-crystalline Solids, Vol. 2, North Holland Publishing Co., Amsterdam (1970).

7. W. Weyl, The Constitution of Glass, Vol. II, Electrical Properties.

8. Electrical Properties of Glass and Glass-Ceramics: A bibliography covering the period 1957-1968, Int. Commission on Glass, Charleroi, Belgium.

9. W. Reif, Statistical Mechanics.

10. E. Warburg, Ann. Physik., 21,622 (1884).

11. W. D. Kingery, Introduction to Ceramics, pp. 647-761, J. Wiley & Sons (1960).

12. E. Rasch and F. Hinrichsen, Z. Elektrochem, 14, 41 (1908).

13. R. L. Myuller, Fizika Tverdogo Tela, 2, 1333 [Soviet Physics Solid State 2, 1213 (1960)].

14. D. L. Kinser and L. L. Hench, J. Am. Ceram. Soc., 52, 12 (1969) 638.

15. T. M. Proctor and P. M. Sutton, J. Am. Ceram. Soc., 43, 173 (1960).

16. J. O. Isard, Proc. Instn. Elect. Engrs., 109 B No. 22, p. 440, paper No. 3636, June (1961).

17. D. L. Kinser, "The Early Stages of Crystallization in Alkali-Silicate Glasses," Ph.D. Dissertation, University of Florida (1968), Diss. Abstracts, University of Michigan (1968).

18. D. L. Kinser and L. L. Hench, J. Am. Ceram. Soc., 52, 12 (1969) 638.

19. H. E. Taylor, Trans. Farad. Soc., 52, 873 (1956).

20. H. Moore and R. C. Dilva, J. Soc. Glass Tech., 36, 5T (1952).

21. E. Seddon, E. Tippet, and W. Turner, J. Soc. Glass Tech., 16, 450T (1932).

22. A. E. Owen and R. W. Douglas, J. Soc. Glass Tech., 32, 159T (1959).

23. O. V. Mazurin and E. S. Borisovskii, J. Tech. Phys., Moscow, 27, 275 (Soviet Phys. Tech. Phys., 2, 243).

24. R. J. Charles, J. Am. Ceram. Soc., 48, 8 (1965) 432.

25. B. Lengyel and Z. Boksay, Z. Phys. Chem., 204, 157 (1955).

26. J. O. Isard, J. Soc. Glass Tech., 43, 113T (1959).

27. R. J. Charles, J. Am. Ceram. Soc., 46 (5) 235-43 (1963).

28. J. T. Littleton and W. L. Wetmore, J. Am. Ceram. Soc., 19 (9) 243 (1936).

29. J. W. Rebbeck, M. J. Mulligan, and J. B. Ferguson, ibid., 8 (6) 329 (1925).

30. D. L. Kinser and L. L. Hench, "Electrical Properties of Thermally Treated Alkali-Silicate Glasses," presented 1968, Am. Ceram. Soc. Meeting, Chicago.

31. D. L. Kinser and L. L. Hench, J. Materials Science, 3 (1970).

32. L. L. Hench, D. L. Kinser, and S. W. Freiman, "The Early States of Crystallization of a Li_2O-$2SiO_2$ Glass," VIIIth Int. Glass Congress, London (1968).

33. L. L. Hench, W. D. Tuohig, and D. A. Sanders, "Kinetics of Resorbtion of a Metastable Precipitate in a Lithia Silica Glass," to be submitted to Phys. and Chem. of Glasses.

34. S. W. Freiman and L. L. Hench, J. Am. Ceram. Soc., 51,382 (1968).

35. R. E. Tickle, Phys. and Chem. of Glasses, 8 (3) 101 (1967).

36. K. Endell and H. Hellbriigge, Glastech. Ber., 277 (1942).

37. J. O. M. Bockris, J. A. Kitchener, S. Sgnatowicz, and J. W. Tomlinson, Trans. Faraday Soc., 48, 75 (1952).

38. S. Urnes, Glass Industry, 40, 237 (1959).

39. C. Babcock, J. Am. Ceram. Soc., 329 (1934).

40. K. A. Kostanyan, K. S. Saakyan, E. M. Avetisyan, "Electrical Conductivity of Molten Glasses," Electrical Properties and Structure of Glass, O. V. Mazurin, ed., New York, Consultants Bureau, pp 81-83 (1965).

41. K. A. Kostanyan, K. S. Saakyan, and O. K. Geokchyan, Izv. Akad. Nauk Armyan SSR, 17 (4) 357 (1964).

42. A. E. Martin and G. Derge, Metals Technology, 10 (5) 1 (1943).

43. B. E. Warren, X-ray Diffraction, Chap. 7, Addison-Wesley, New York (1969).

44. R. Grigorovici, J. Non-crystalline Solids, 1, 303 (1969).

45. R. Grigorovici and R. Manaila, Thin Solid Films, 1, 343 (1968).

46. M. V. Coleman and D. J. D. Thomas, Phys. Stat. Solidi, 24, K111 (1967).

47. J. Chang, "Structure of Amorphous Ge Films," M.S. Thesis, University of Florida (1970).

48. N. F. Mott, Adv. Phys., 16, 49 (1967).

49. P. W. Anderson, Phys. Rev., 109, 1492 (1958).

50. N. F. Mott, Festkorperprobleme, 9, 22 (1969).

51. N. F. Mott, J. Non-crystalline Solids, 1, 1 (1968).

52. H. F. Schaake and L. L. Hench, J. Non-crystalline Solids, 2, 292 (1970).

53. H. F. Schaake, "Electron Transport Phenomena in Transition Metal Oxide Glasses," M. S. Thesis, University of Florida (1968).

54. I. G. Austin, J. Non-crystalline Solids, 2, 474 (1970).

55. I. G. Austin and N. F. Mott, Adv. Phys., 18, 41 (1969).

56. A. Miller and E. Abrahams, Phys. Rev., 120, 745 (1960).

57. E. A. Davis and R. F. Shaw, J. Non-crystalline Solids, 2, 406 (1970).

58. D. Adler, Solid State Physics, 21, 1 (1968).

59. J. Appel, Solid State Physics, 21, 193 (1968).

60. C. Kittel, Quantum Theory of Solids, Chap. 7, J. Wiley & Sons, New York (1963).

61. T. Holstein, Ann. Phys., 8, 325 (1959).

62. J. Schnakenburg, Phys. Stat. Solidi, 28, 633 (1968).

63. M. Pollack and T. H. Geballe, Phys. Rev., 122, 1742 (1961).

64. H. F. Schaake, "Heterogeneous Amorphous Semiconductors," in Physics of Electronic Ceramics, L. L. Hench and D. B. Dove, eds., Marcel Dekker, Inc., New York (in press).

65. A. D. Pearson, Modern Aspects of the Vitreous State, J. D. Mackenzie, ed., Butterworths, Washington (1964) p. 29.

66. J. D. Mackenzie, Modern Aspects of the Vitreous State, J. D. Mackenzie, ed., Butterworths, Washington (1964) p. 126.

67. G. W. Anderson,"Structural Characterization and Optical Transmission Studies of Vanadate Glasses,"Ph.D. Dissertation, University of Illinois (1969).

68. S. R. Bates and L. L. Hench, to be published.

69. Shortley and Williams, Elements of Physics, Reinhold Press.

70. J. M. Stevels, Progress in the Theory of the Physical Properties of Glass, Elsevier, Ansterdam (1948).

71. L. Azaroff, Electronic Properties of Materials, Addison Wesley, New York (1962).

72. H. Frohlich, Theory of Dielectrics, Oxford University Press, London (1949).

73. V. V. Daniels, Dielectric Relaxation, Academic Press, London (1967).

74. Nowick and Berry, IBM Technical Review.

75. M. Gevers and F. K. DuPre, Disc. Faraday Soc. 42 N, 47 (1946).

76. A. E. Owen, "Dielectric Properties of Glass," M.S. Thesis, University of Sheffield, England (1959).

77. L. K. H. von Beek, in Progress in Dielectrics, Vol. 7, Birks, ed., p. 69, CRC Press (1966).

78. K. W. Wagner, Arch. Electrotech., 2 (9) 371 (1914).

79. J. C. Maxwell, Electricity and Magnetism, Vol. 1, p. 328, Clarendon Press, London (1892).

80. R. W. Sillars, J. Inst. Elec. Engrs., 80, 378 (1937).

81. J. Volger and J. M. Stevels, Phillips Res. Rep., 11, 452 (1956).

82. R. E. Jaeger, J. Am. Ceram. Soc., 51, 57 (1968).

83. S. H. Mable and R. D. McCammon, Phys. and Chem. of Glasses, 10 (6) 222 (1969).

84. O. L. Anderson and H. E. Bömmel, J. Am. Ceram. Soc., 38, 125 (1955).

85. B. E. Warren and R. Mozzi, "The Structure of Vitreous Silica," submitted to J. Applied Crystallography (1969).

86. R. J. Charles, J. Appl. Phys., 32, 1115 (1961).

87. A. D. Pearson, J. Non-crystalline Solids, 2, 1 (1970).

88. R. Uttecht, H. Stevenson, C. H. Sie, J. D. Greiner, and K. S. Raghavan, J. Non-crystalline Solids, 2, 358 (1970).

89. H. Fritzsche and S. R. Ovshinsky, J. Non-crystalline Solids, 2, 148 (1970).

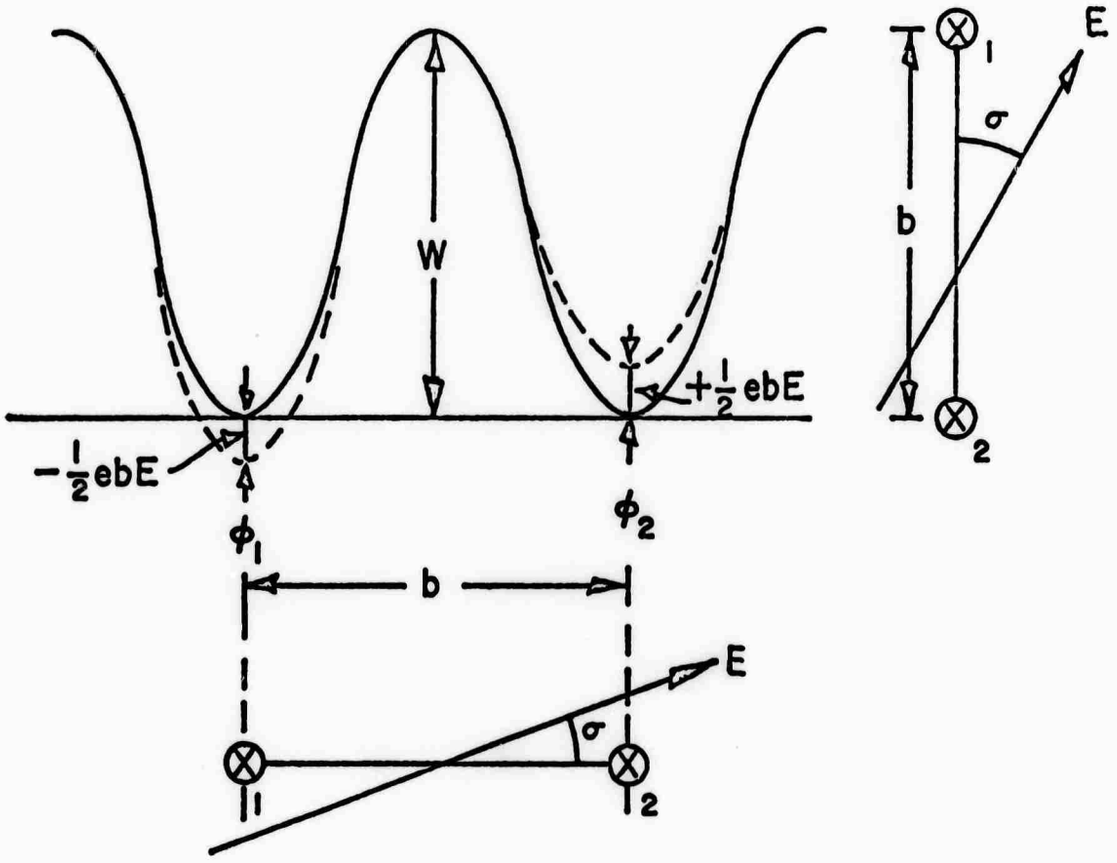


Figure 1. In Site Energy Configuration in a Glass. Solid lines indicate energy without field; dashed lines indicate energy with field E applied.

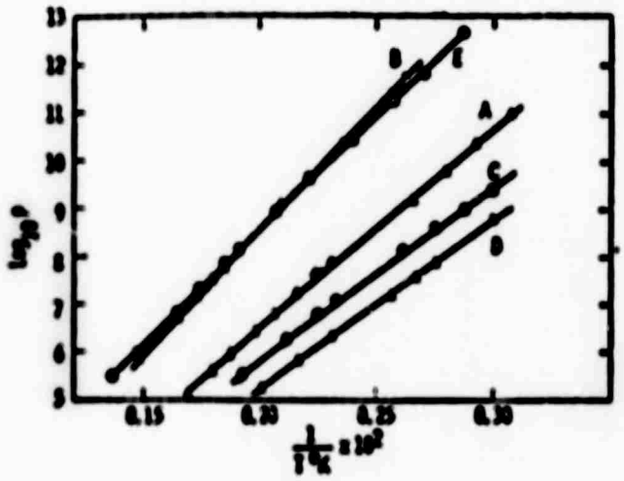


Figure 2(a).

Conductivity of Some Ionic Glasses

A) 18 Na_2O·10 CaO·72 SiO_2

B) 10 Na_2O·20 CaO·70 SiO_2

C) 12 Na_2O·88 SiO_2

D) 24 Na_2O·76 SiO_2

E) Pyrex

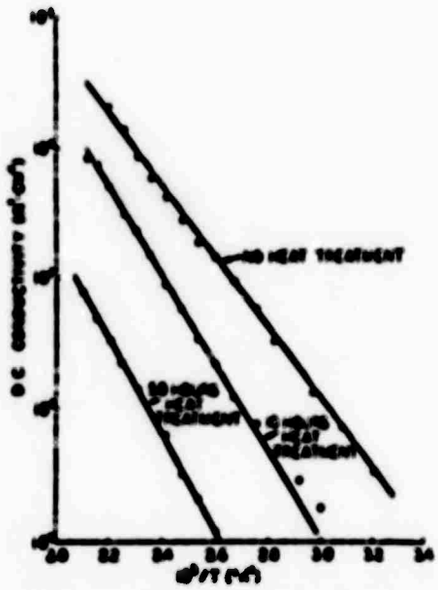


Figure 2(b).

Conductivity of a 30 Li_2O·70 SiO_2 Glass.

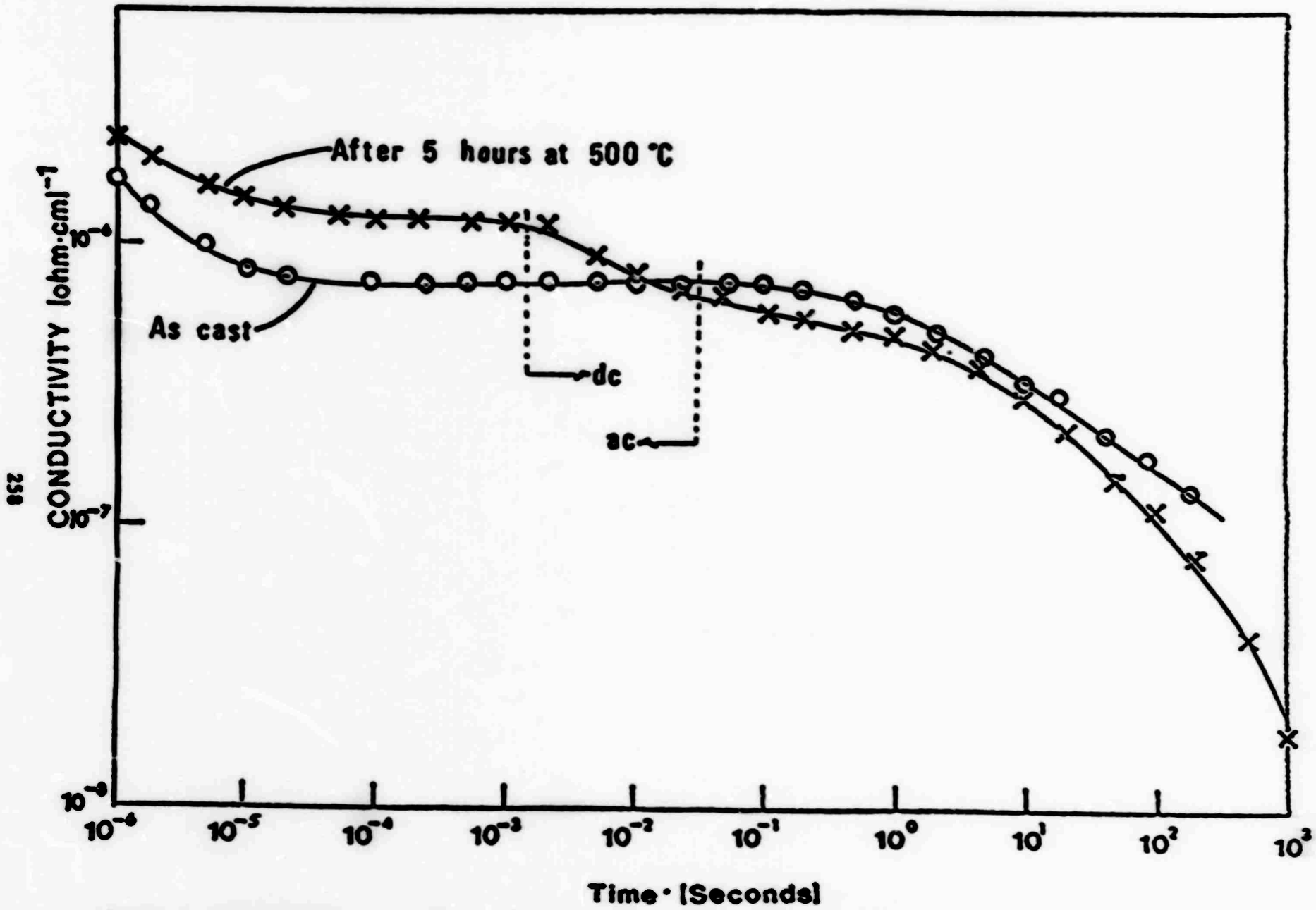


Figure 3. Conductivity as a Function of Time for a 30 Li_2O·70 SiO_2 Glass as Cast and Heat Treated.

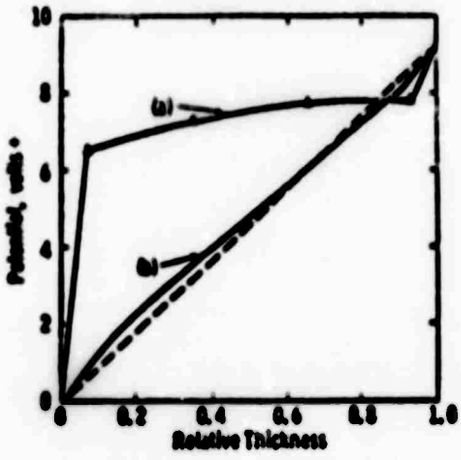


Figure 4.

Potential Distribution in Glass

(a) An alkali-lead-silicate after 10 hrs. at annealing temperature (383°C).

(b) An alkali-free glass after 18 hrs. at annealing temperature (460°C). Field strength 160-180 V/cm.

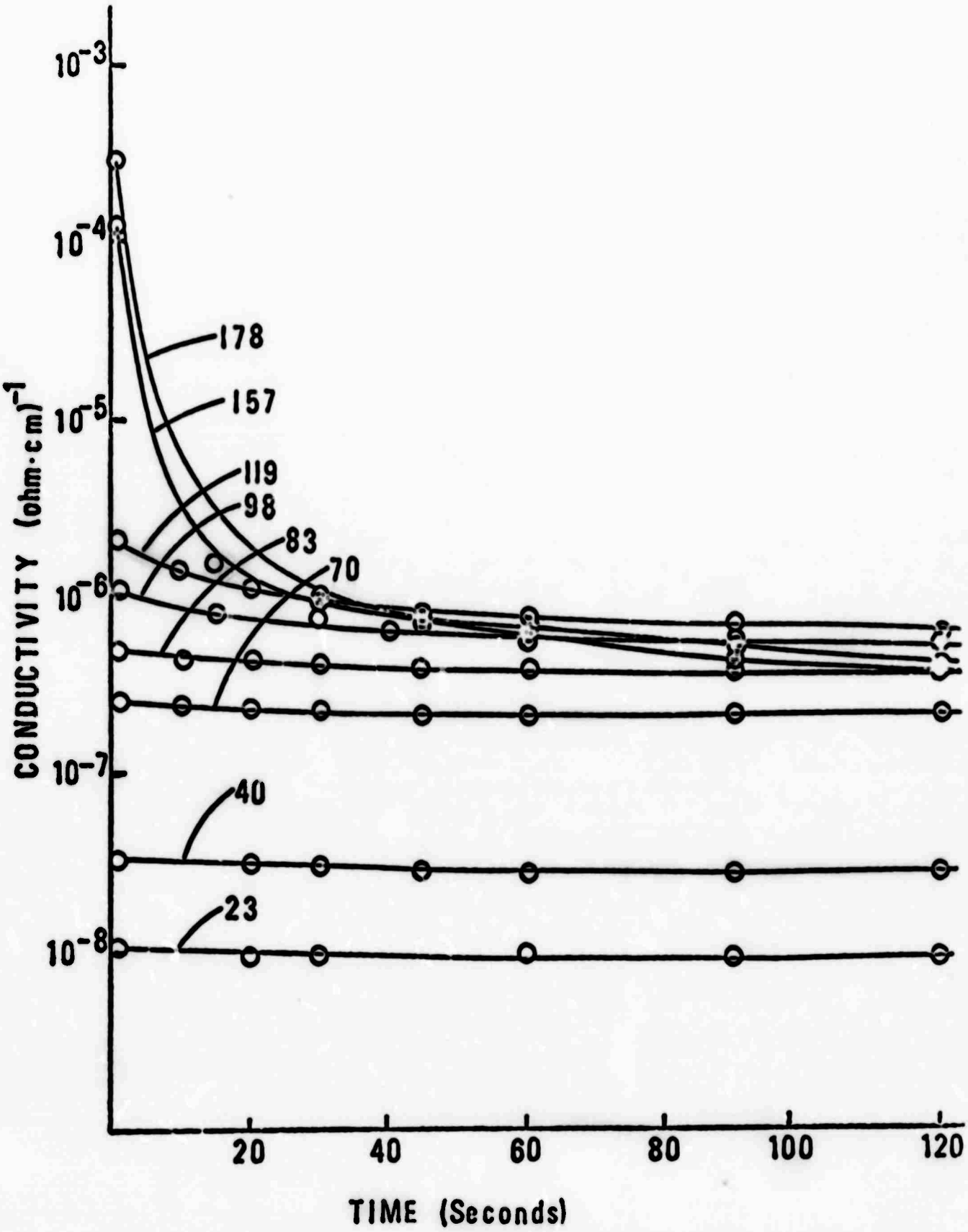


Figure 5. Conductivity as a Function of Time for a 33 $Na_2O \cdot 67\ SiO_2$ Glass at Various Temperatures (in °C).

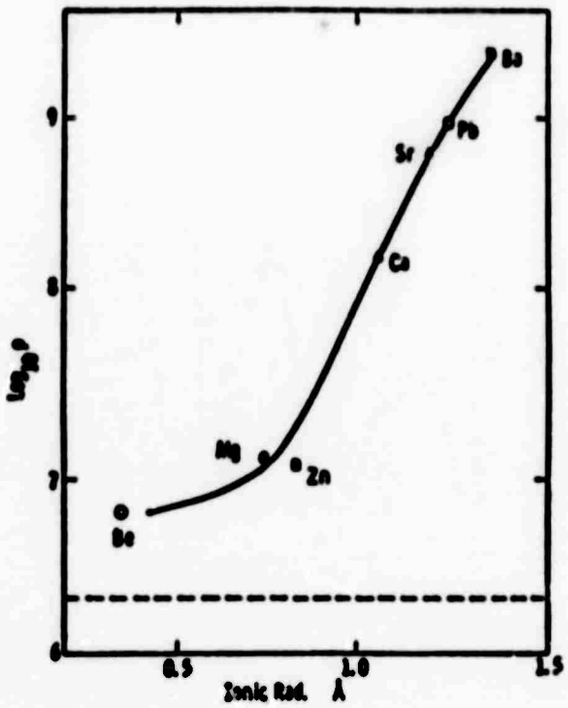


Figure 6.

The Effect of Divalent Ion Radii on Resistivity of 20 Na_2O·20 MO·60 SiO_2 Glass. (Dashed line represents resisitivity of 20 Na_2O·80 SiO_2 Glass. Measurements at 150°C)

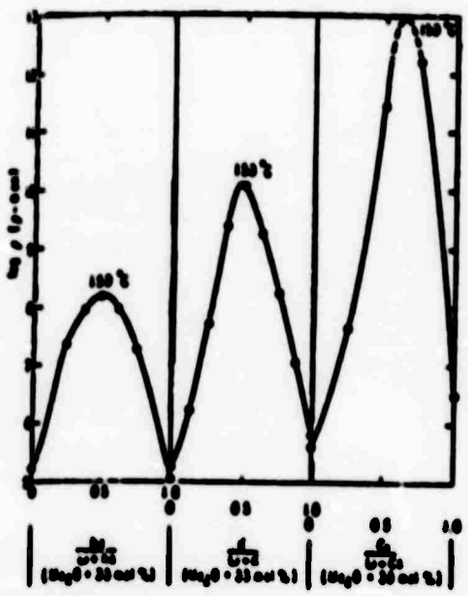


Figure 7.

Resistivity of Silicate Glasses Containing Mixed Alkali Ions.

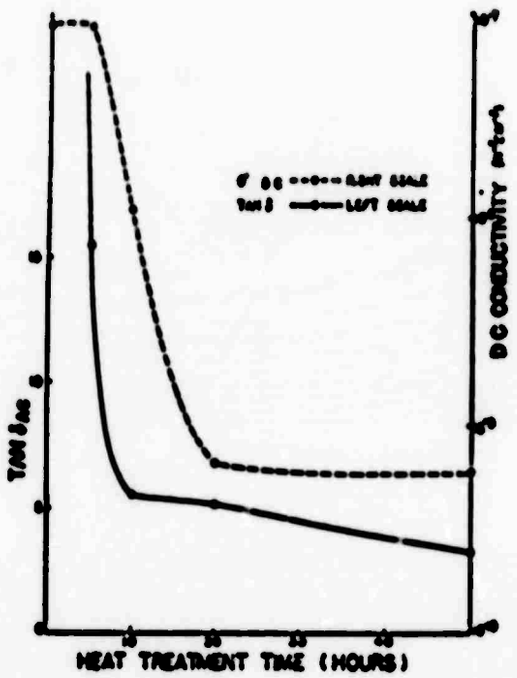


Figure 8.

Conductivity and Loss Tangent for a 30 Li_2O·70 SiO_2 Glass as a Function of Heat Treatment at 500°C.

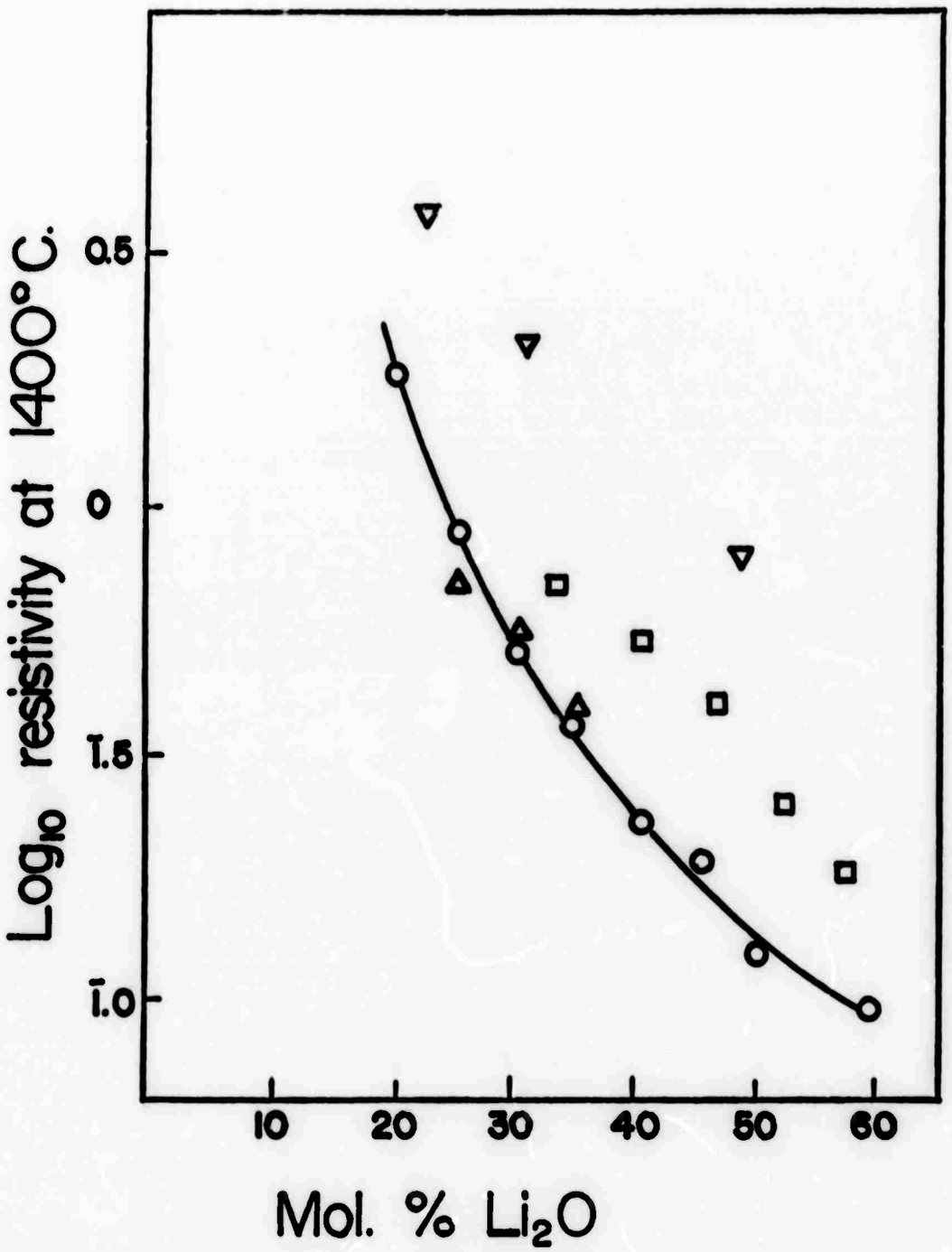


Figure 9. Resistivity of Molten Lithia-Silica Glasses as a Function of Mole % Li_2O.

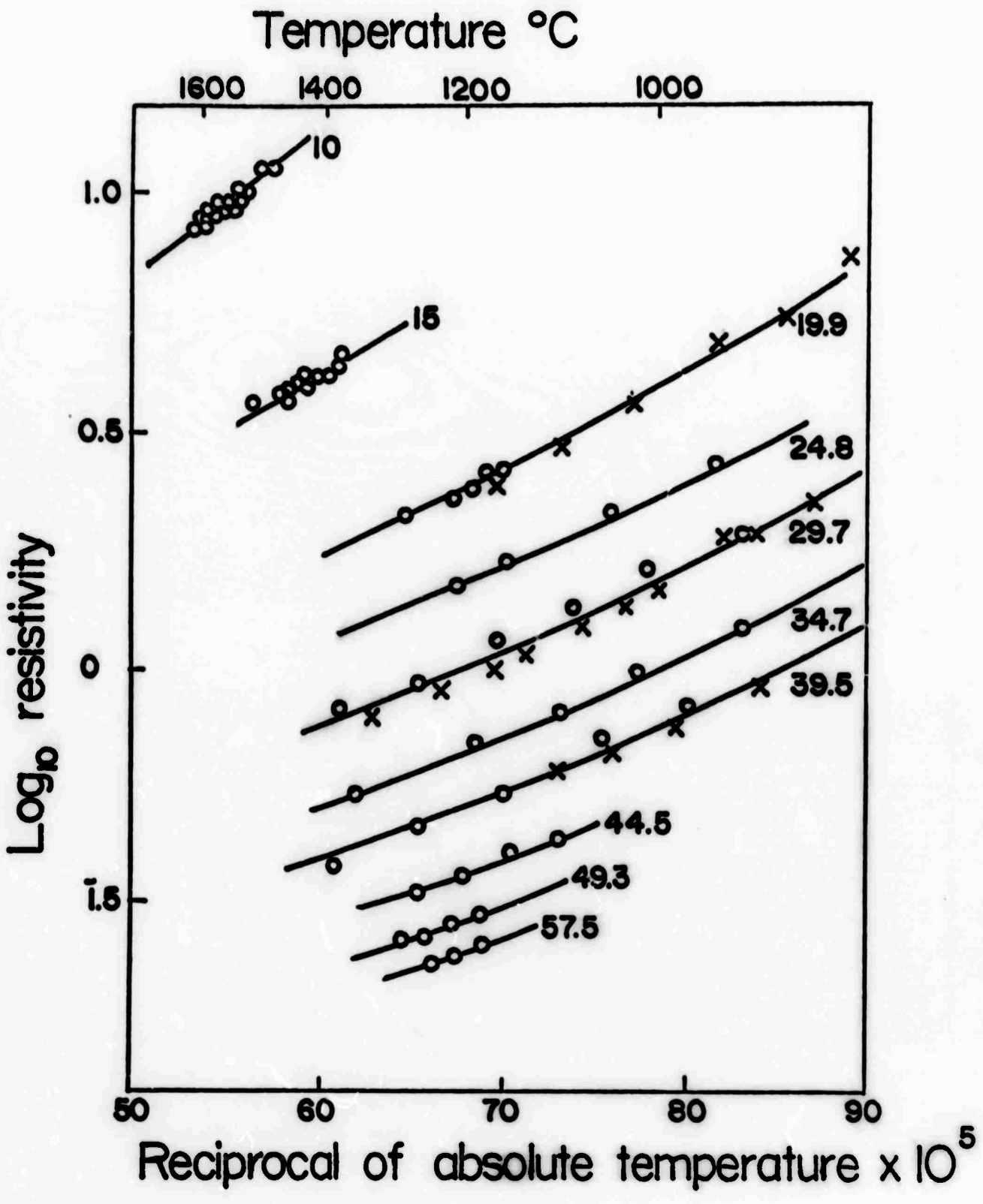


Figure 10. Resistivity of Molten X $Na_2O \cdot (100-X)\ SiO_2$ Glasses as a Function of Temperature.

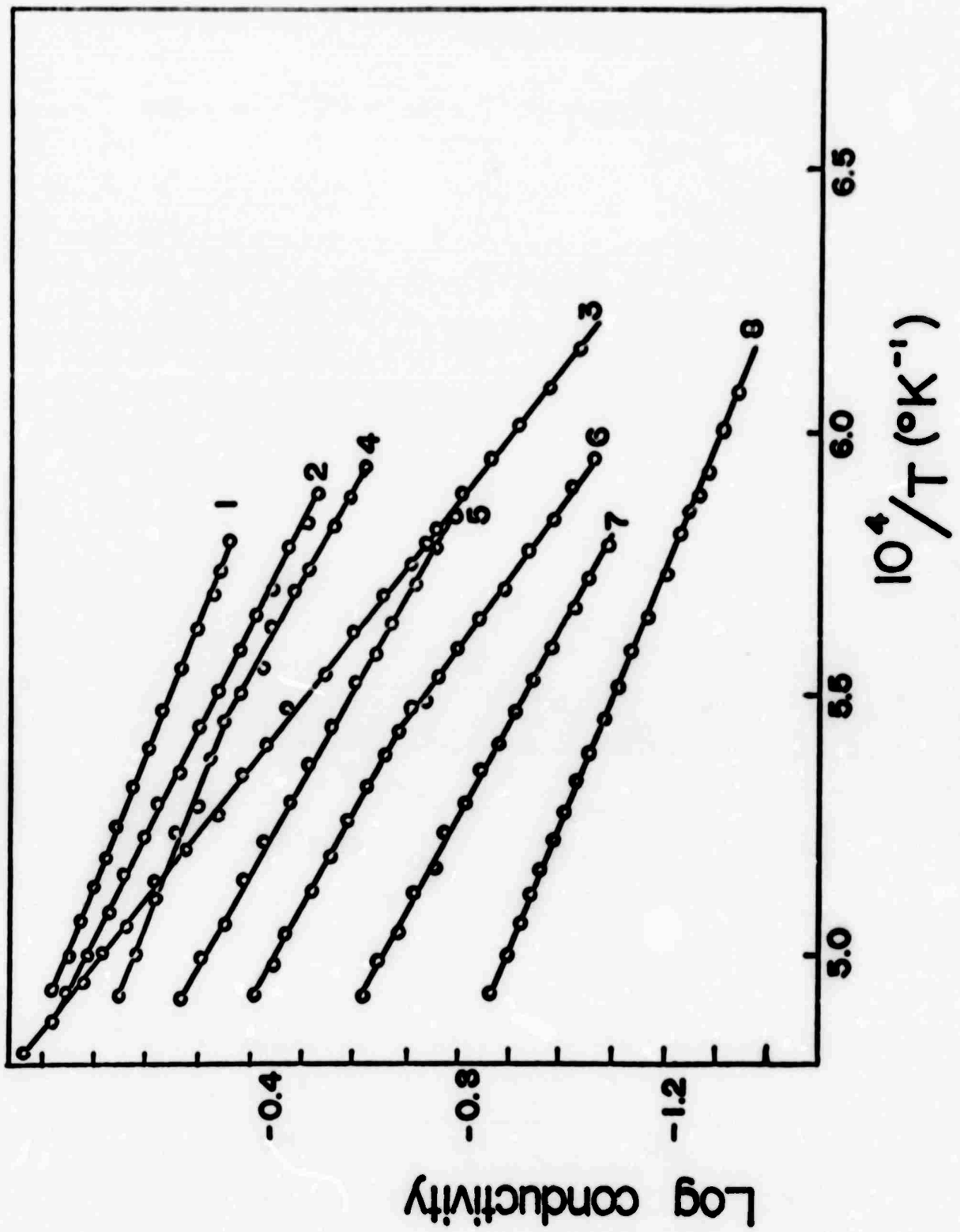


Figure 11. Conductivity of Molten $CaO-SiO_2$ Glasses as a Function of Temperature. (Glass #1 is 55% CaO, #8 is 30% CaO, remainder given in text.

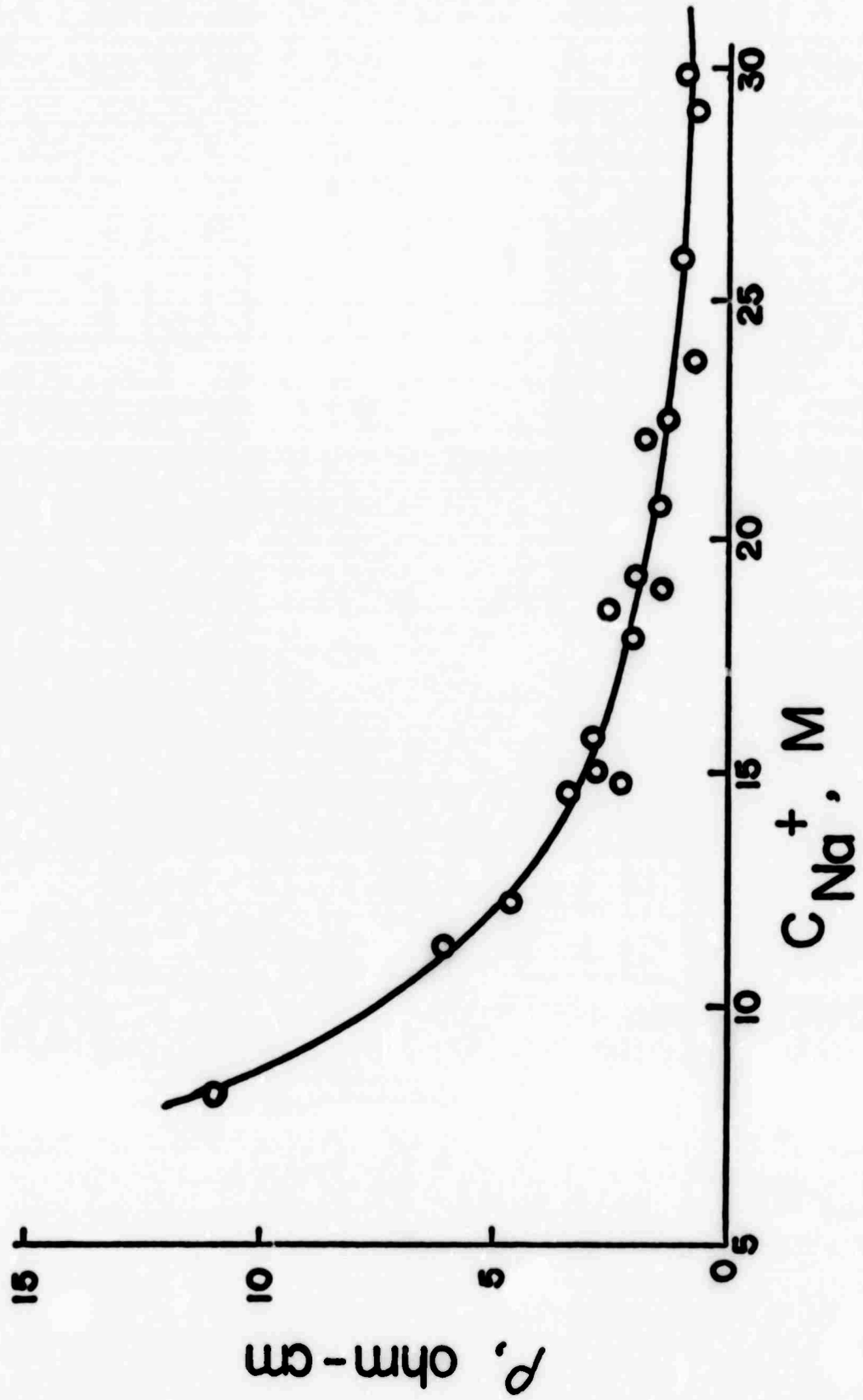


Figure 12. Dependence of Resistivity of Na_2O-RO-SiO_2 Glasses on Weight % of of Na^+ Ions at 1200°C. (Data points represent MgO, CaO, BaO, BeO, CdO containing glasses with 68-78% SiO_2).

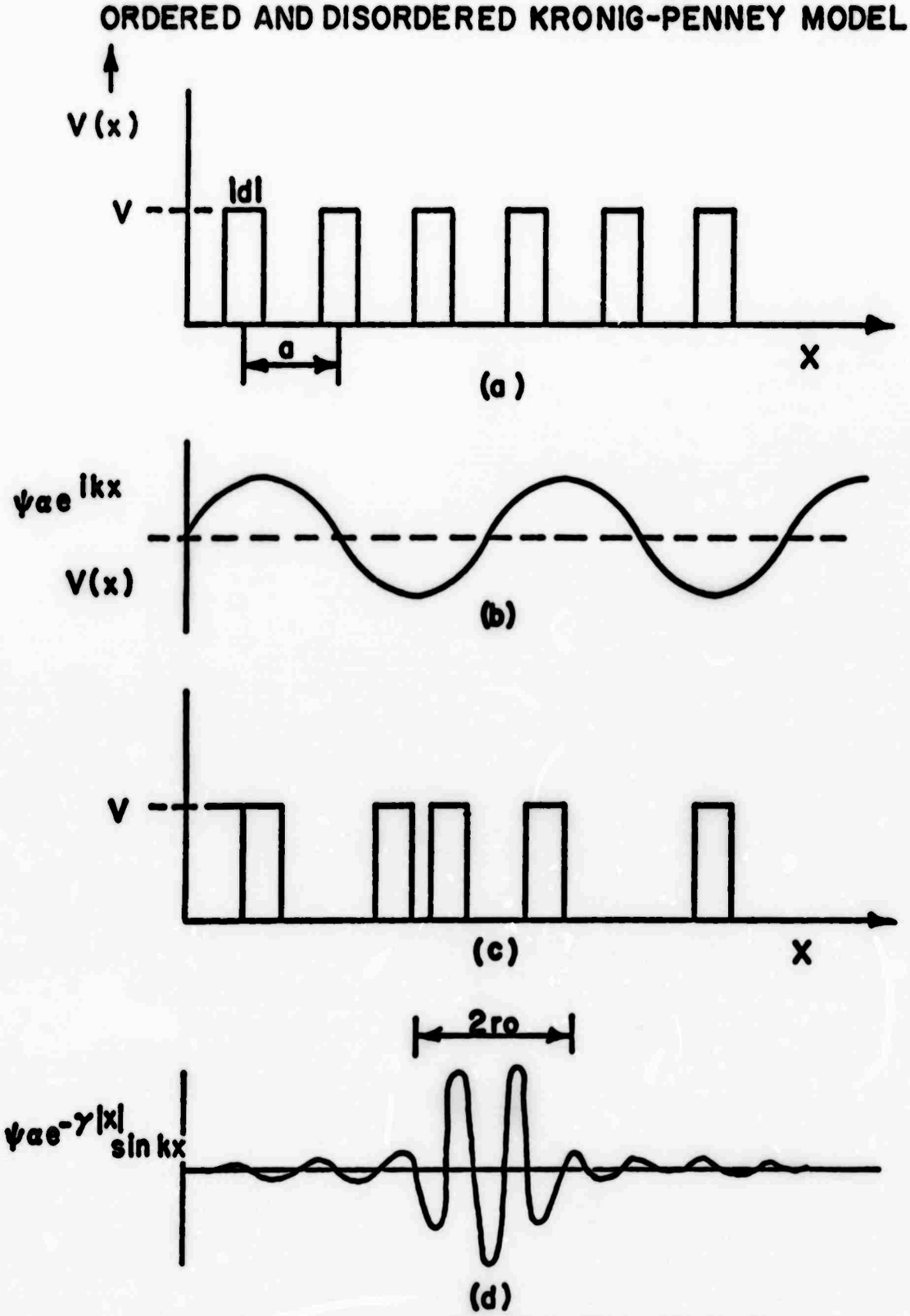


Figure 13. Ordered and Disordered Kronig Penney Model

(a) Ordered
(b) Qualitative Solution to Schrodinger Equation
(c) Disordered
(d) Qualitative Solution to Schrodinger Equation

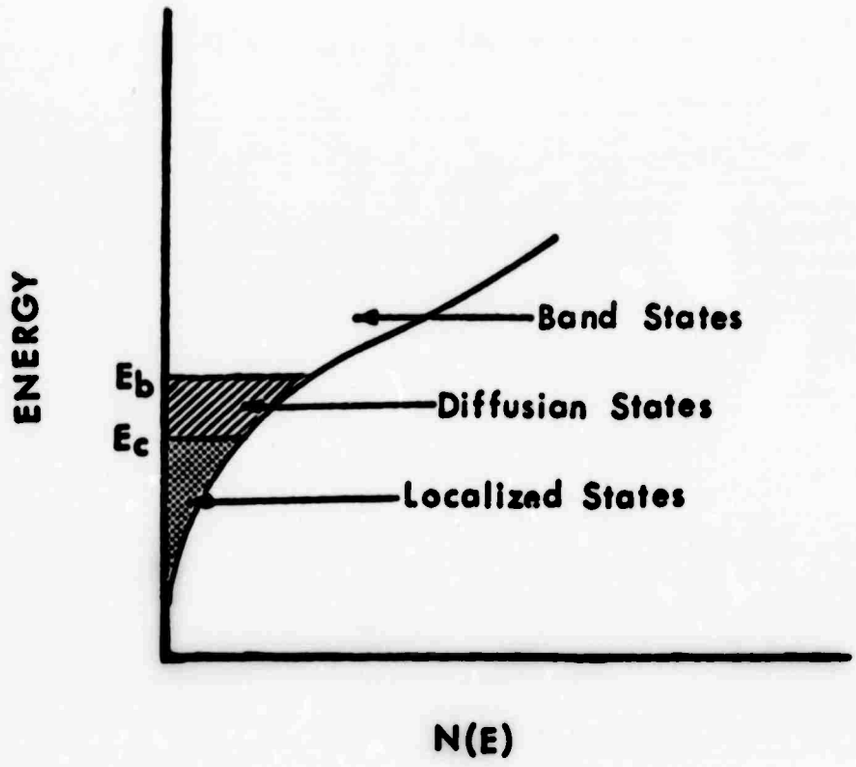


Figure 14. Density of States (Conduction Band) in a Partially Disordered Lattice.

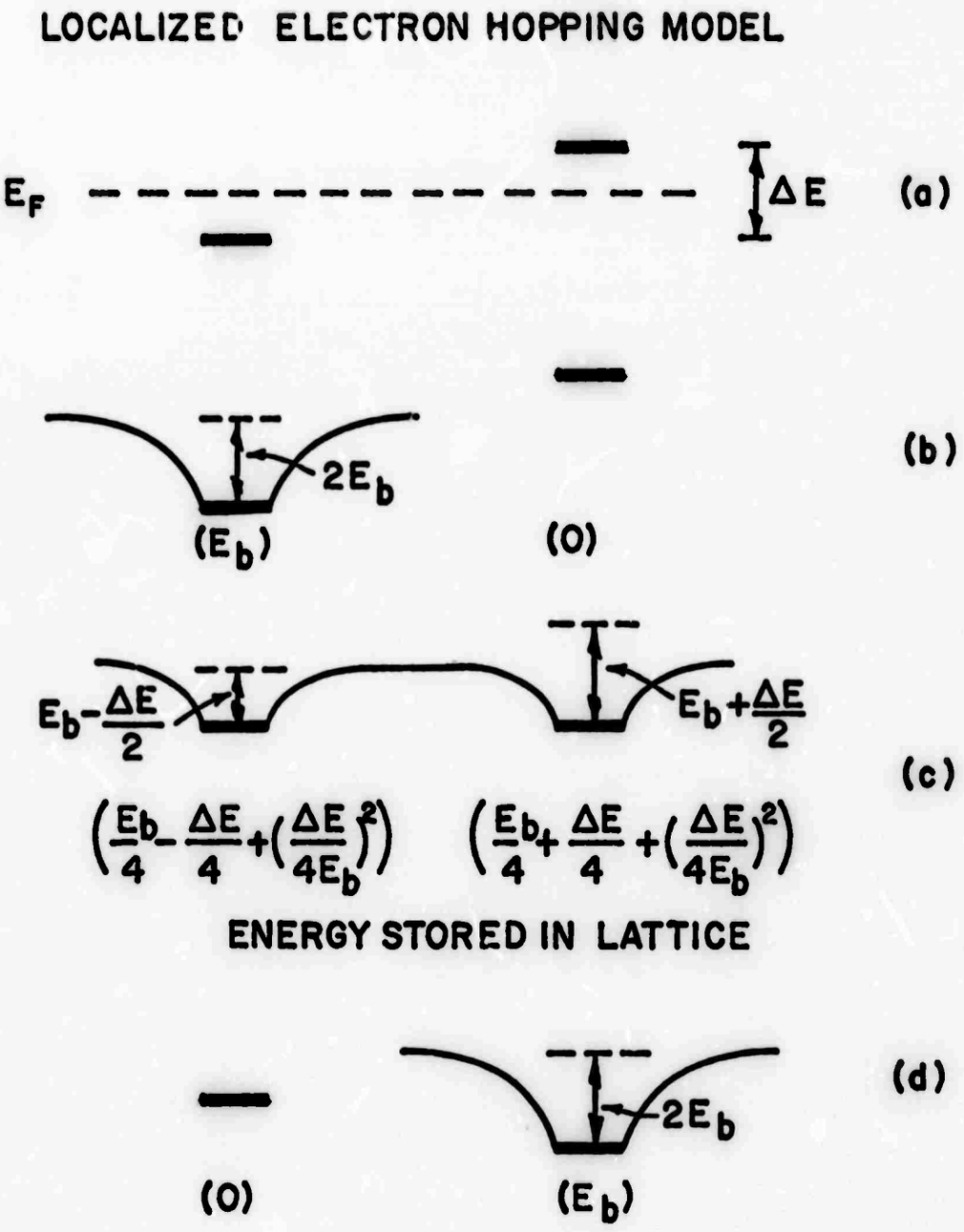


Figure 15. Localized Election Hopping Model

(a) Site Energies
(b) Lattice Polarization on Site 1.
(c) Configuration at Mid-Hop
(d) Configuration at Conclusion of Hop

(Energies in parenthesis are polarization energies stored in the lattice at each site).

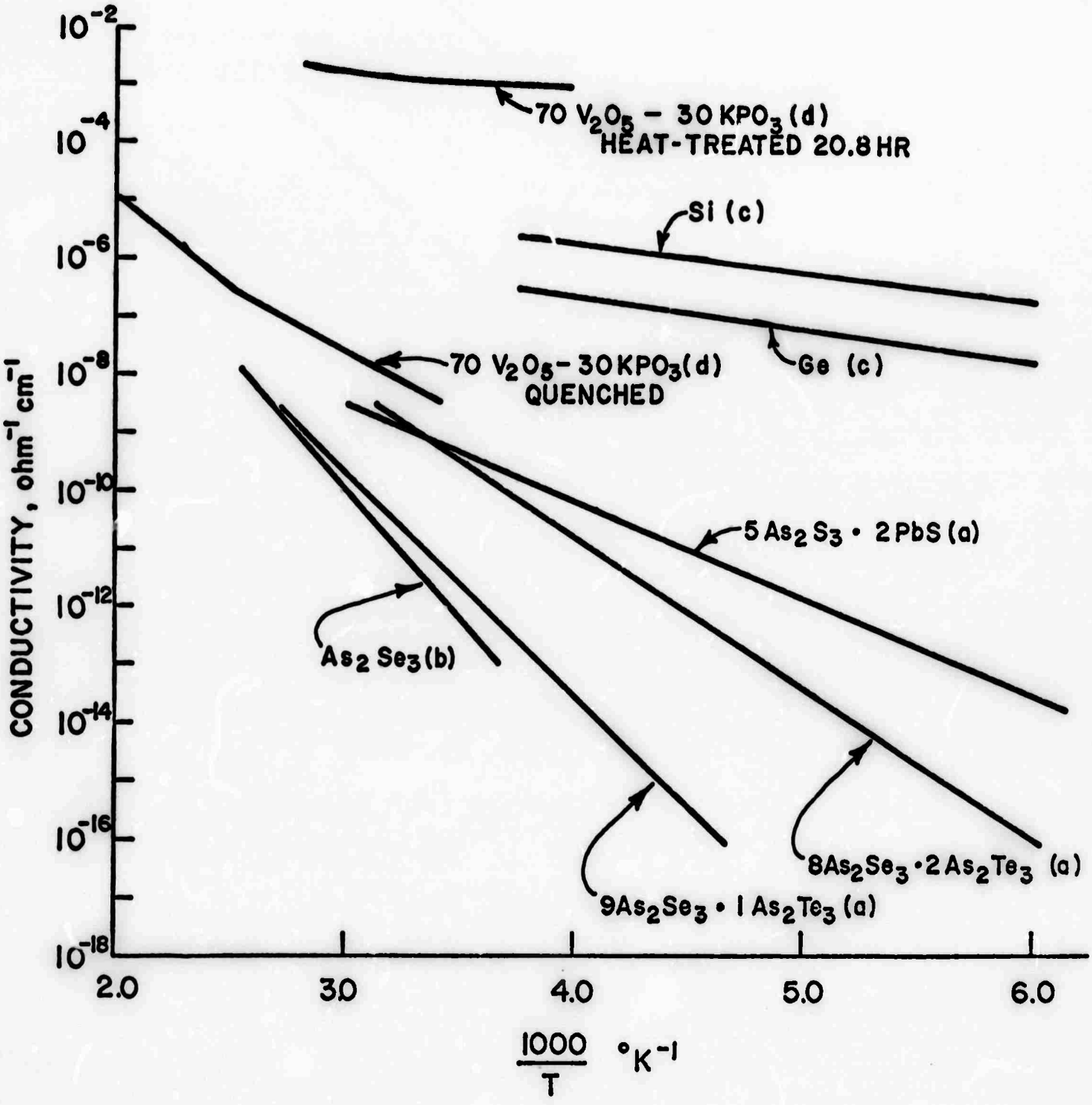


Figure 16. D.C. Conductivity of Some Amorphous Semiconductors.
a) A. E. Owen and J. M. Robertson, J. Non-crystalline Solids, 2, 40, (1970).
b) J. E. Lang, Ph.D. Thesis, Rensselaer Polytechnic Institute, July (1968).
c) P. A. Walley, Thin Film Solids, 2, 327 (1968).
d) H. F. Schooke and L. L. Hench, J. Non-crystalline Solids, 2, 292 (1970).

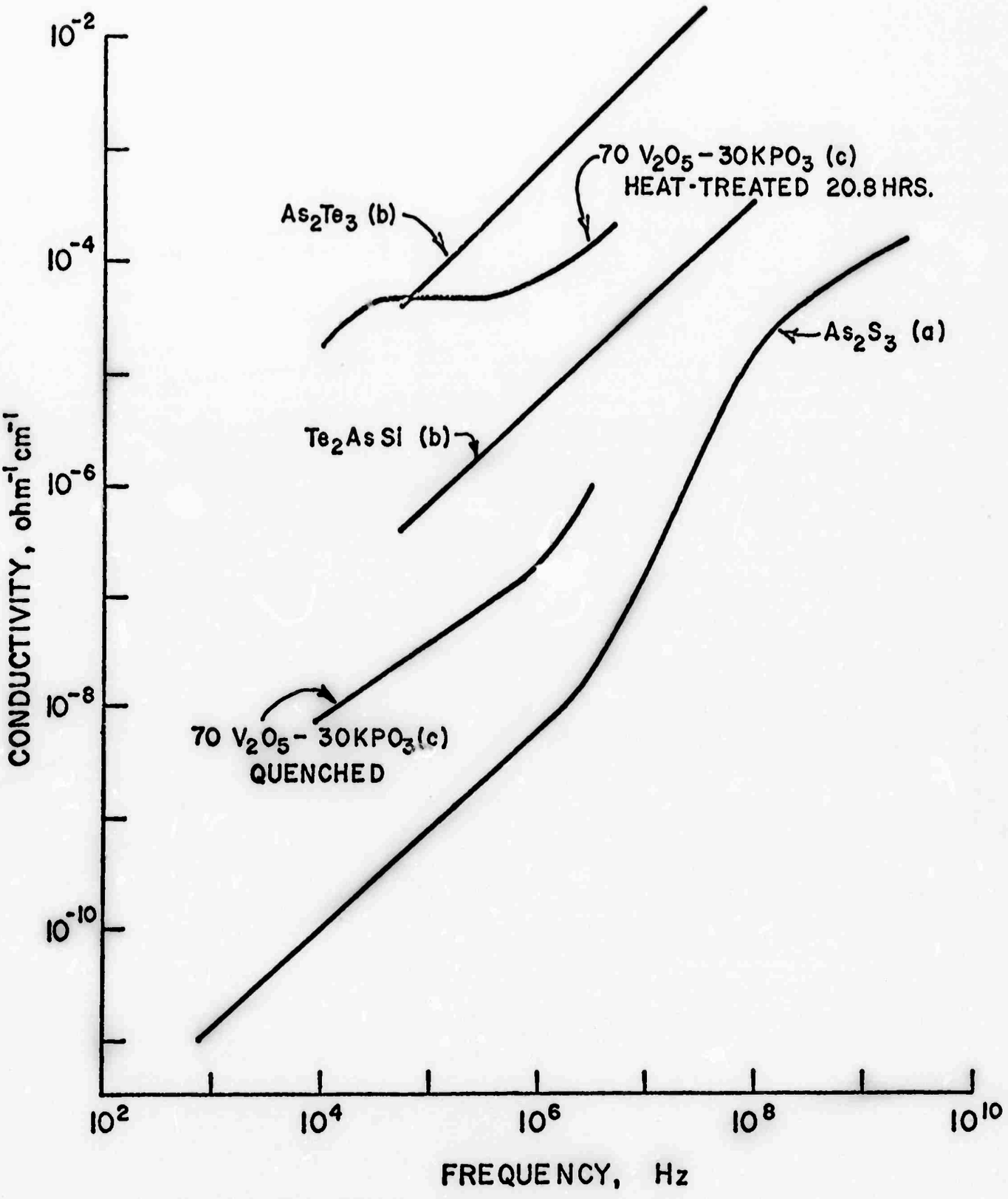
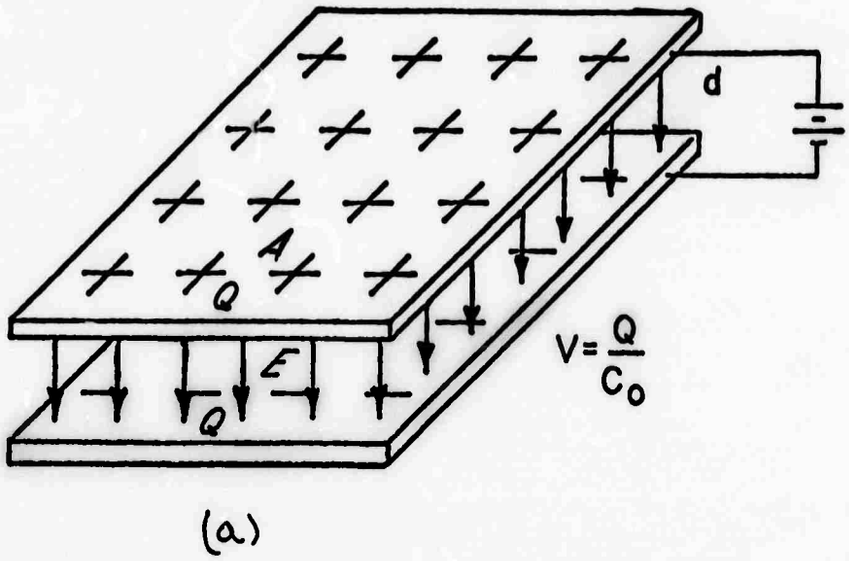
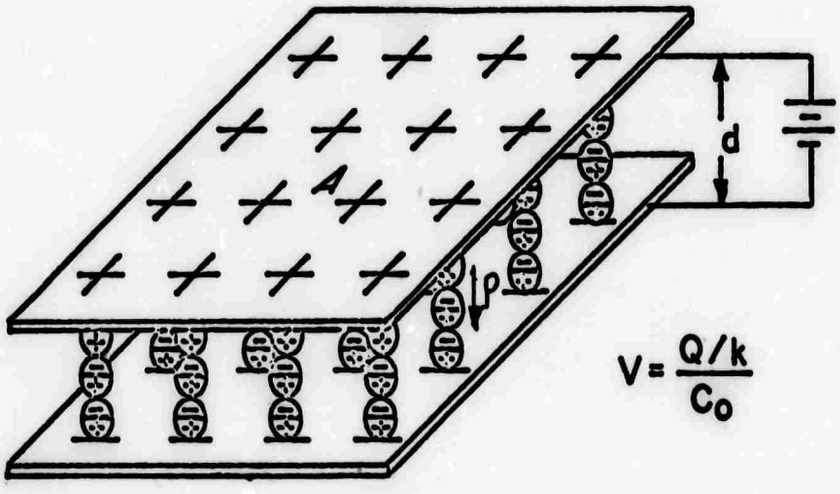


Figure 17. A.C. Room Temperature Conductivity of Some Amorphous Semiconductors
a) A. E. Owen and J. H. Robertson, J. Non-crystalline Solids, 2, 40 (1970).
b) H. K. Rockstad, J. Non-crystalline Solids, 2, 192 (1970)
c) H. F. Schooke and L. L. Hench, J. Non-crystalline Solids, 2, 292 (1970).



(a)



(b)

Figure 18. Charge on a Parallel Plate Capacitor

a) With Vacuum as Dielectric
b) With a Dielectric Material between the Plates.

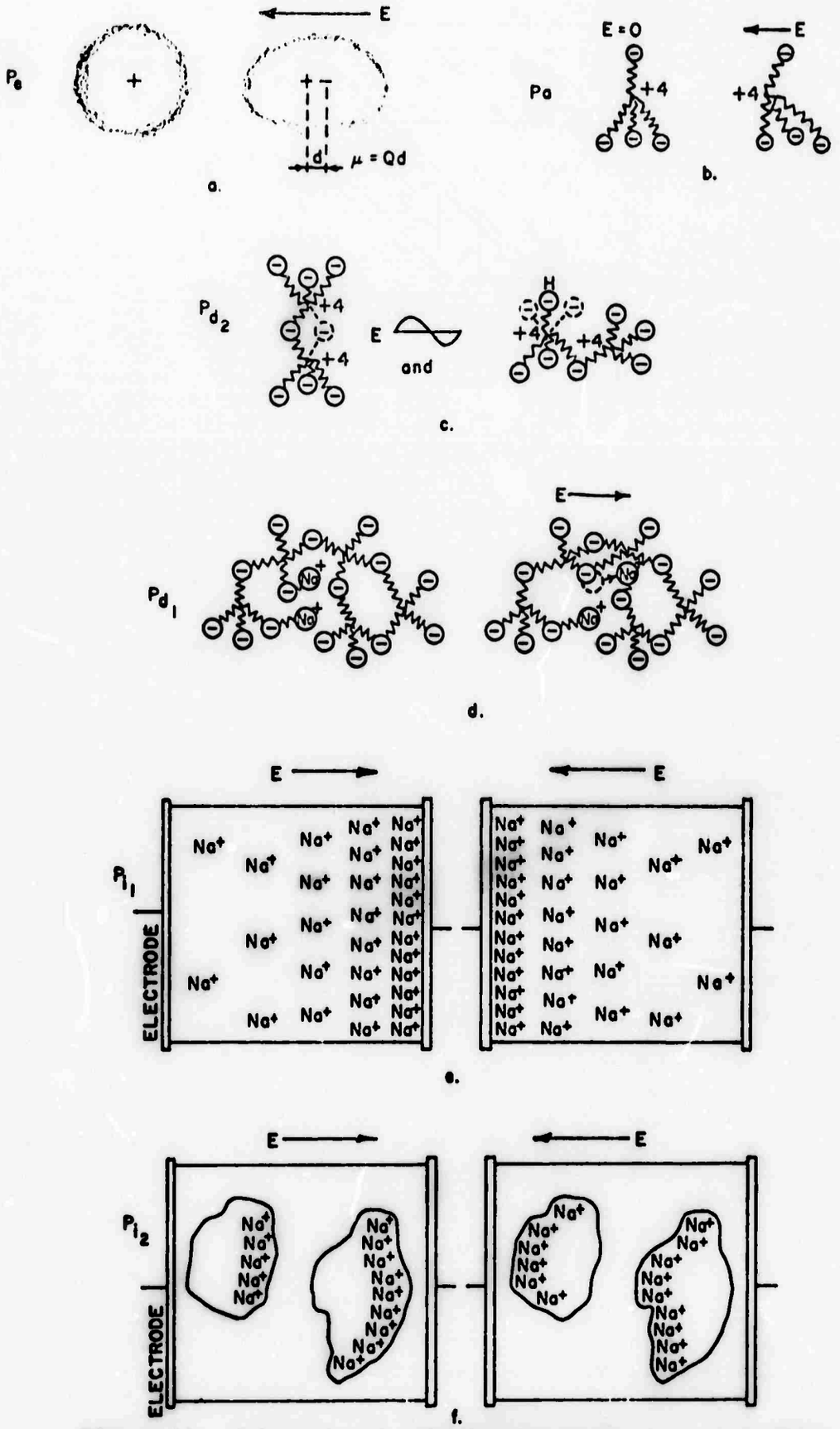


Figure 19. Schematic of Polarization Mechanisms in a Glass a) electronic; b) atomic or ionic; c) high frequency oscillatory dipoles; d) low frequency cation dipole; e) interfacial space charge polarization at electrodes; f) interfacial polarization at heterogeneities.

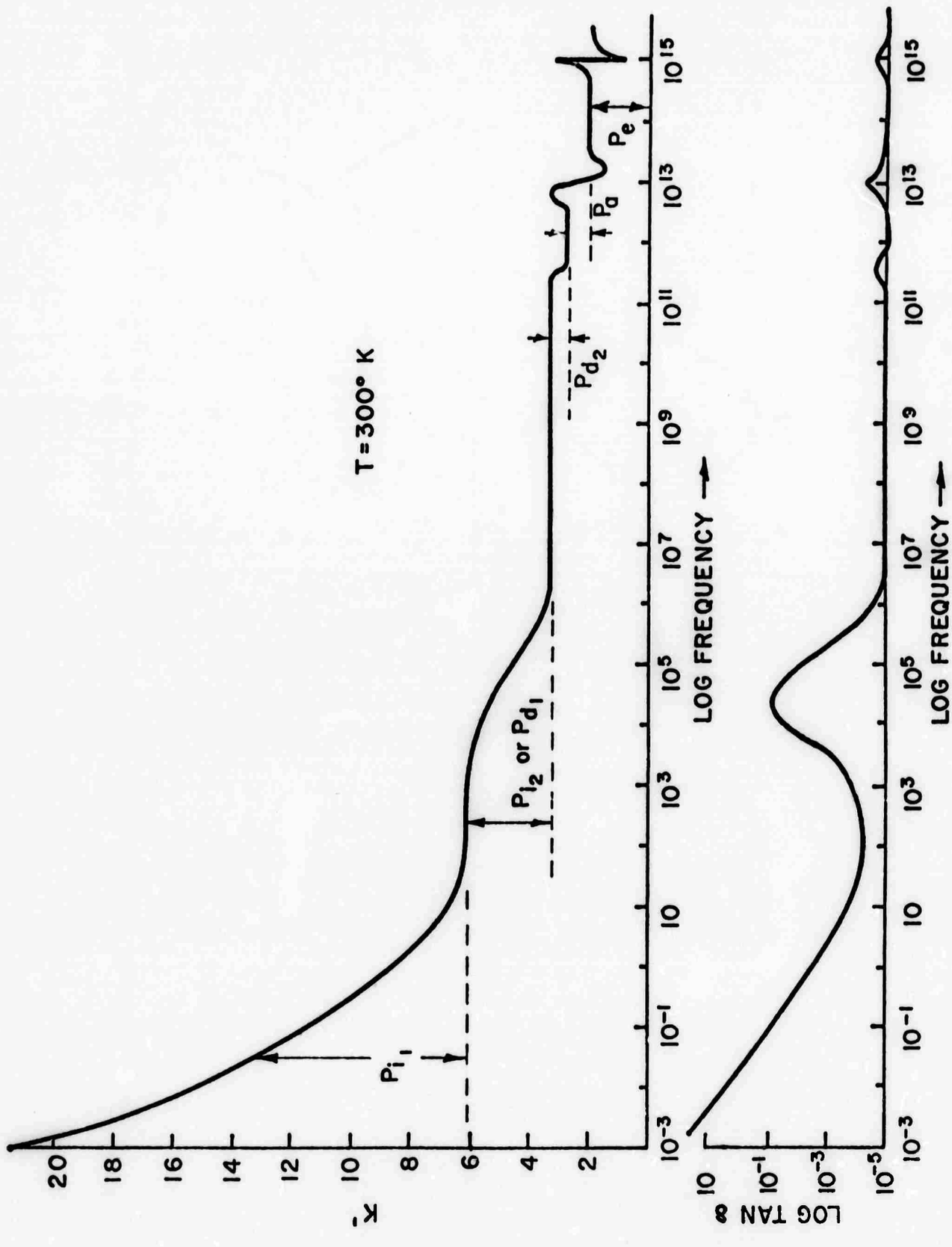


Figure 20. Frequency Dependence of the Polarization Mechanisms in Glass. Upper: contribution to the charging constant (representative values of k^1); Lower: contribution to the loss angle (representative values of tan δ).

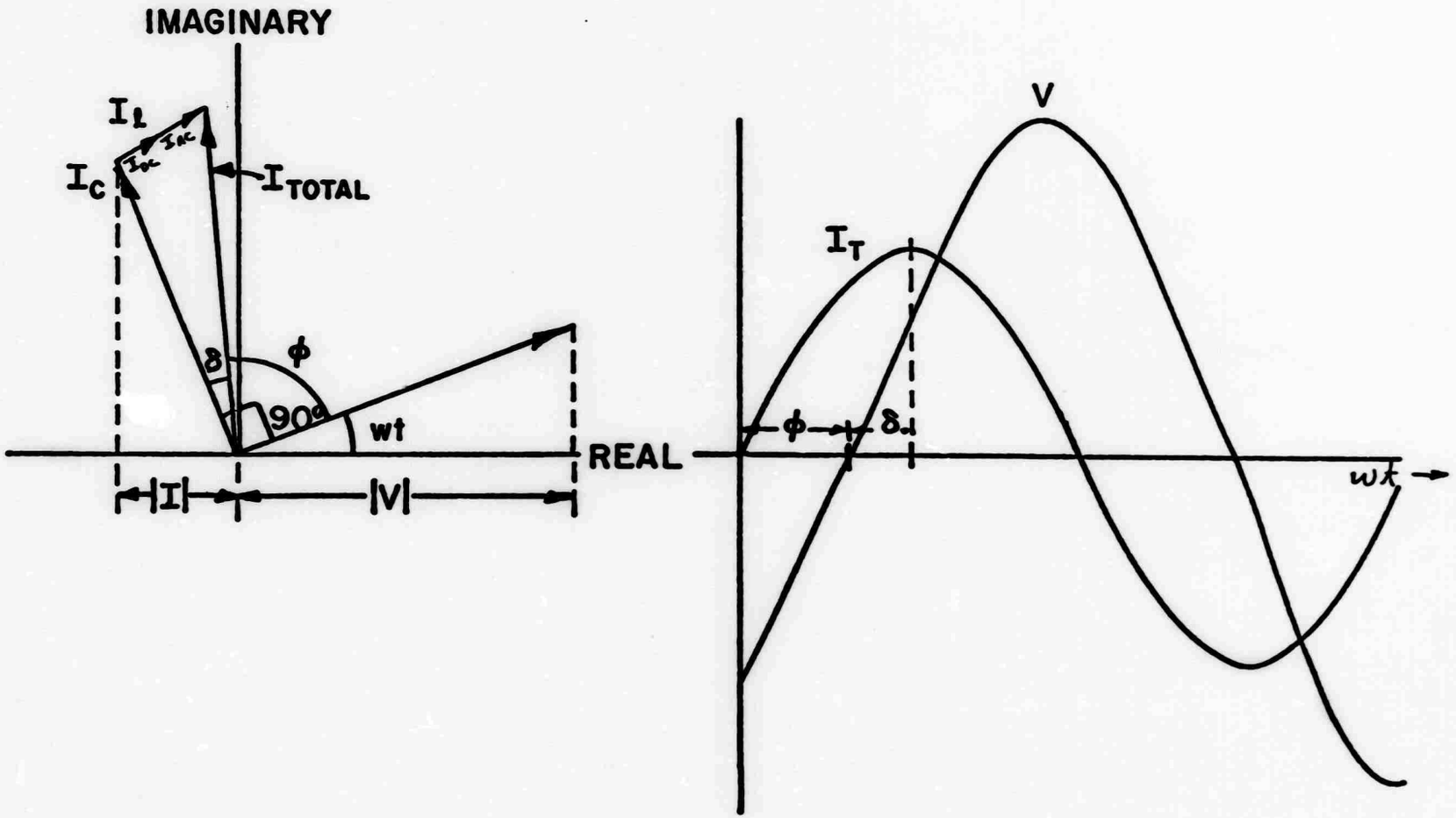


Figure 21.

a) Vector Diagram of Charging, Loss, and Total Currents in a Glass Dielectric.

b) Periodic Variation of Dielectric Current in a Glass with Sinusoidal Voltage V.

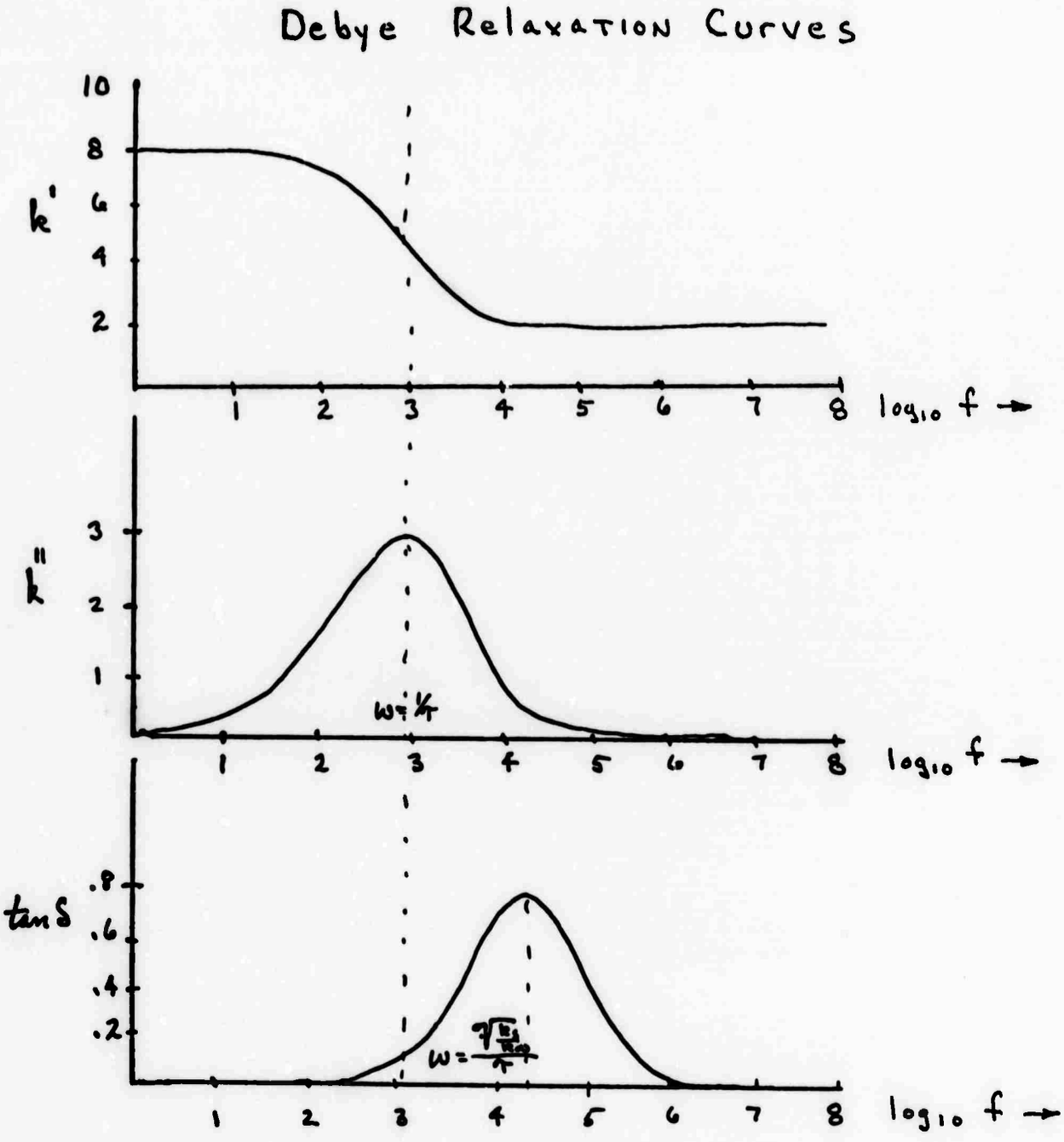


Figure 22. Frequency Variation of Dielectric Parameters due to a Debye Relaxation Process.

(a) k', (b) k'' (c) $\tan\delta$

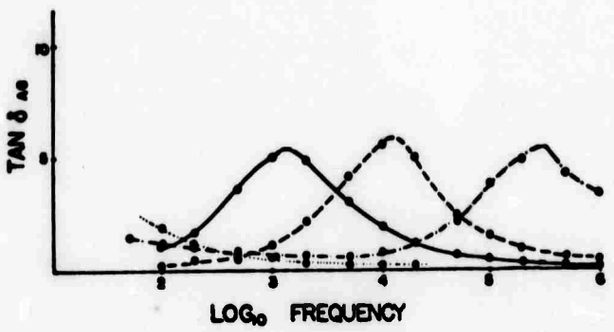


Figure 23. Shift of the Dielectric Loss Peak due to Increasing Temperature.

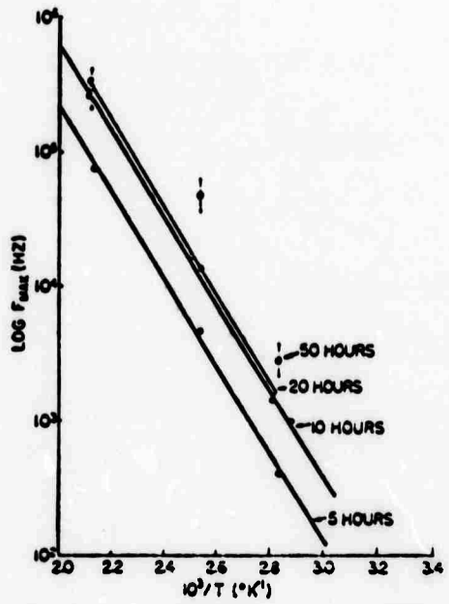


Fig. 3. Log of frequency maximum vs. reciprocal temperature for various thermal treatments at 500°C.

Figure 24. Logarithmic Temperature Dependence of the Frequency Location of the Dielectric Loss Peak.

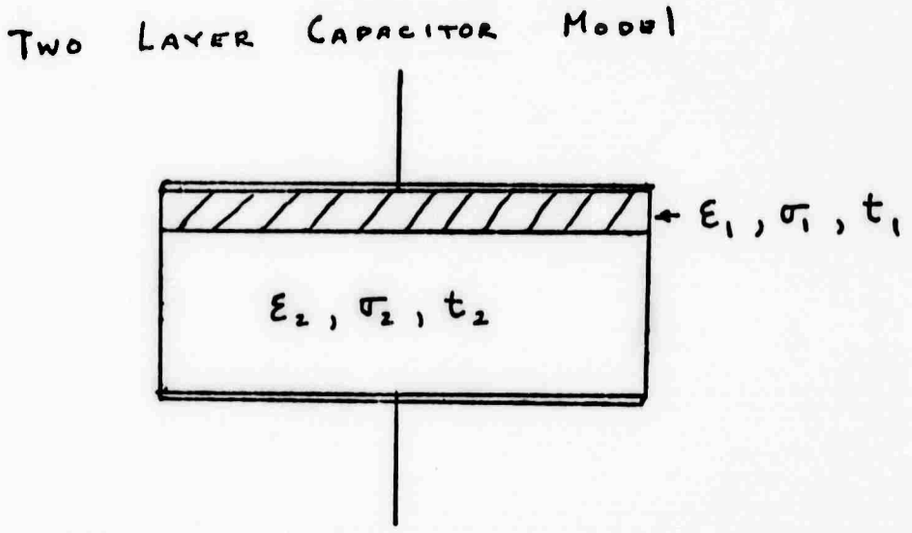


Figure 25. Schematic Two-Layer Model of a Glass Dielectric.

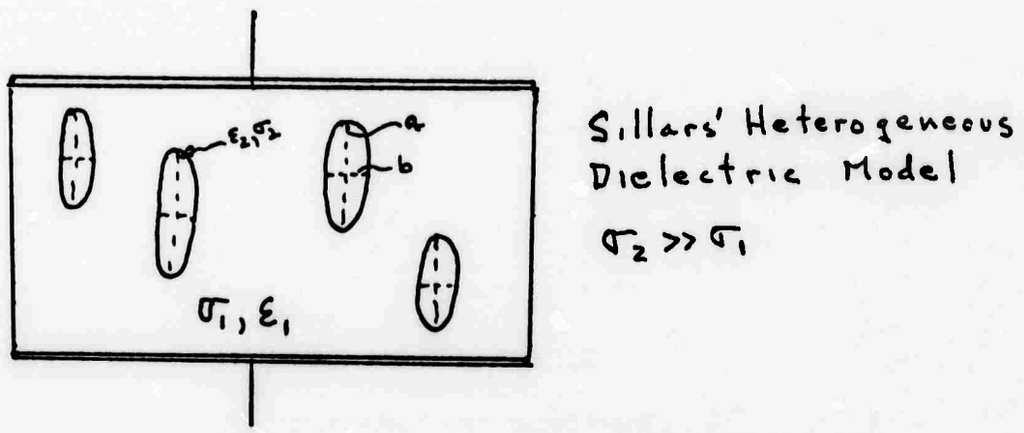


Figure 26. Schematic of a Glass Dielectric Containing High Conductivity Ellipsoidal Heterogeneities of Axial Ratio a/b.

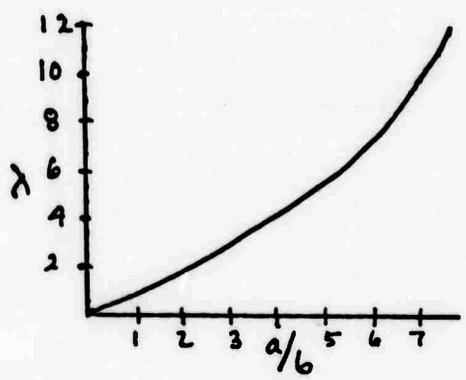


Figure 27. Variation of Sillars' Dimensional Parameter λ with Axial Ratio of Ellipsoids, a/b.

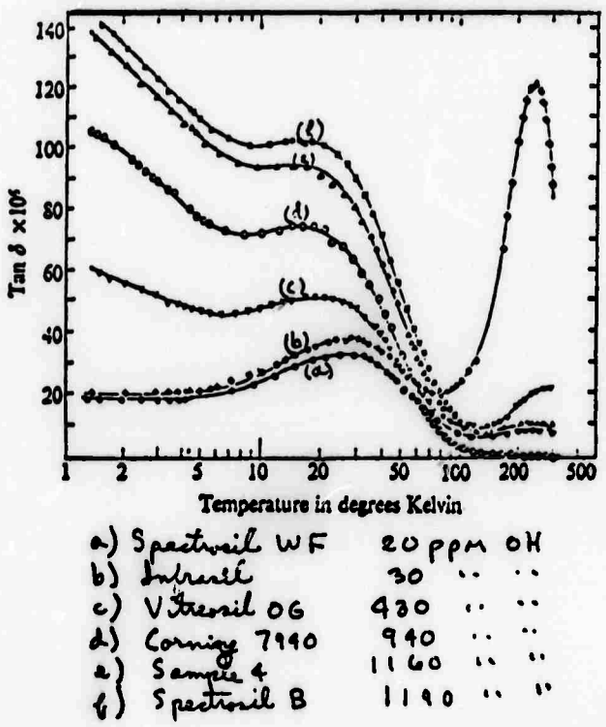


Figure 28. Dielectric Loss Behavior of a Variety of Fused Silica Glasses as a Function of Temperature.

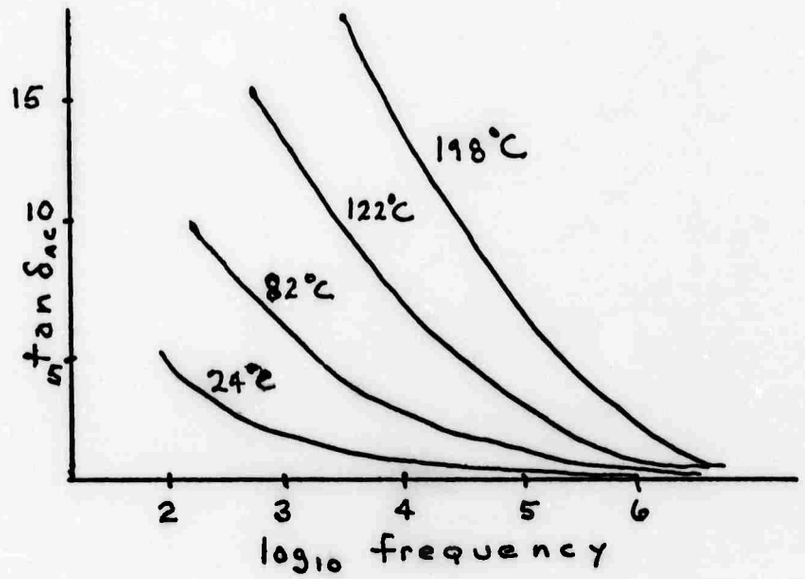


Figure 29. Frequency Dependence of the Dielectric Loss Angle of a 30 Mole % Li_2O-SiO_2 Glass as Quenched.

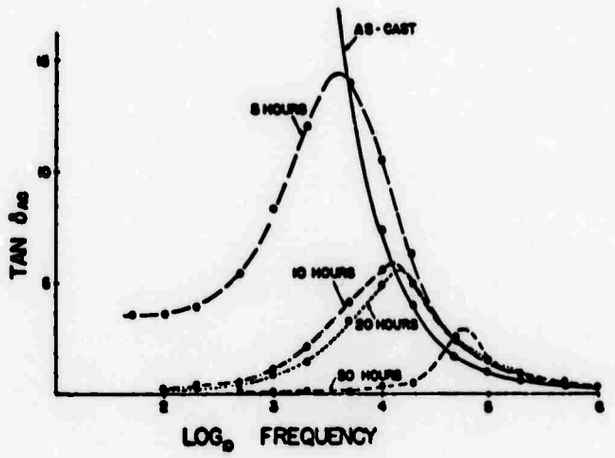


Figure 30. Variation is Dielectric Loss Angle of a 30 Mole % Li_2O-SiO_2 Glass with Thermal Treatment at 500°C.

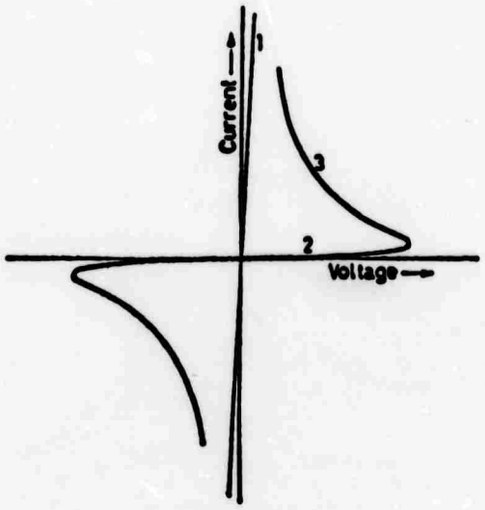


Figure 31. Typical Current-Voltage Characteristics of a Amorphous Semiconductor Switch.

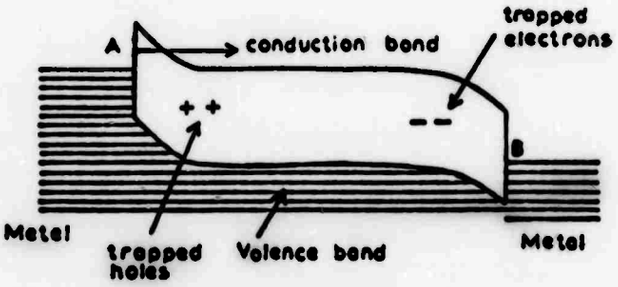


Figure 32. Mott's Electrode Injection Model of Reversible Switching of an Amorphous Semiconductor.

IV. Measurement Techniques (R. W. Gould, L. L. Hench, and J. J. Hren)

A. SOLID SOLUTION DECOMPOSITION OF ALUMINUM-RICH ALUMINUM-ZINC-SILVER ALLOYS (S. R. Bates and R. W. Gould)

Introduction

Precipitation from the solid solution has been a subject of great importance for many years. Ever since Wilm[1] published results showing that the mechanical properties of certain aluminum alloys changed at room temperature, a tremendous amount of energy has been expended in understanding the phenomena.

The early endeavors that followed Wilm's work demonstrated that precipitation processes could radically change the properties of materials. Experimental work was greatly expanded following the First World War. These investigations demonstrated that the so-called "age-hardening" changes were due to atomic rearrangements accompanying the decomposition of supersaturated solid solutions. Subsequent investigations showed that the atomic rearrangements were often submicroscopic and often very complex.

A major breakthrough occurred when Guinier[2] and Preston[3] demonstrated the clustering of solute atoms could give rise to diffuse x-ray scattering. These solute atom clusters are now known as Guinier-Preston or G. P. zones. Thus, it then

became possible to follow changes that were occurring in the solid solution from the very earliest stages.

A great deal of research has been directed toward understanding the solid solution decomposition of the aluminum-zinc and aluminum-silver alloys. The results of these investigations indicate that the property changes of the solid solution in the preprecipitation stage are caused by atomic distribution of heterogeneities in the supersaturated solid solution. The excess dissolved atoms have a tendency to form clusters or G.P. zones which are coherent with the matrix.

With the level of understanding that has been achieved in the aluminum-zinc and aluminum-silver systems, it seems very reasonable to extend this type of research into the ternary aluminum-zinc-silver system. As the decomposition sequences are quite different in the aluminum-zinc and aluminum-silver systems, one can examine how the decomposition of solid solution changes as the silver and zinc concentrations are varied. The amount of experimental work carried out in this system has been very limited. Shashov and Buynov[4] examined the effects caused by the addition of silver to aluminum-zinc alloys and concluded that the majority of silver atoms should be present in the G.P. zones, since this would decrease the strain between the G.P. zones and the matrix. They also concluded that there was only one type of G.P. zone present. Koster et al.[5] determined the 300°C isotherm for the ternary aluminum-zinc-silver phase diagrams but directed little attention to the solid solution decomposition. Murakami et al.[6] recently studied the addition of 0.1 percent silver alloy and concluded that two types of G.P. zones were present.

The present research investigates the solid solution decomposition of the aluminum-zinc-silver solid solutions. The principal tools used in this investigation were x-ray small angle scattering and transmission electron microscopy.

Experiment Equipment and Procedures

The alloys used in this investigation were prepared by melting 99.999 aluminum, 99.999 zinc and 99.999 silver. The aluminum was melted in a reactor grade graphite crucible, and a layer of graphite powder was spread over the molten aluminum prior to the addition of the alloying constituents. The material was cast into pre-heated graphite molds and subsequently heated approximately three weeks at 500°C. The alloys were then swaged and rolled to 3 mm thick sheets. Several flash anneals were necessary in order to roll some of the sheets to approximately 0.1 mm foils.

Portions of the 3 mm sheets were used for growing of alloy single crystals. The growth of the single crystals was accomplished by means of an adaptation of the molten zone technique developed by Harkness et al.[7] The primary modification to the previous experimental arrangement involved a new specimen holder. In this new arrangement, the focused light passes through an optical flat which covers the specimen chamber in order to allow a controlled atmosphere to surround the heated specimen.

After growth, the single crystals were given a homogenization treatment for two weeks at 500°C. Electron microprobe scans along the length of the crystal revealed that the composition fluctuations in the crystals were less than 0.1 percent of the mean alloy composition.

Before any of the samples were given the prequenching heat treatment, they were annealed at 500°C for one hour to relieve any remaining cold work acquired during forming. All samples were quenched from a vertical tube furnace fitted with an Inconel block. The samples were quenched into a water bath maintained at room temperature (24°C), and subsequently aged in a Lauda constant temperature bath. The constant temperature bath could be controlled within ±0.1°C.

In the small angle scattering investigations a Siemens-Kratky camera was used. Balanced nickel and cobalt filters

used in conjunction with a pulse-height analyzer allowed nearly monochromatic copper Kα radiation to be measured. A Lupolen standard provided by Kratky was used in determination of the intensity of the primary x-ray beam.

Discussion and Results

The basic aim of this research was to examine the decomposition of the solid solution for alloys located in the aluminum-rich corner of the ternary aluminum-zinc-silver system. In all cases, the concentration of both zinc and silver was always less than or equal to, 15 atomic percent. Extensive investigations have shown that preprecipitation states are found in binary aluminum-zinc and aluminum-silver alloys. It is also known that in the binary alloys the decomposition of the supersaturated solid solution is governed by the existence of a metastable miscibility gap. With this type of information available, one can summarize some of the basic questions that can be considered when examining the preprecipitation phenomena and precipitation sequence in the aluminum-zinc-silver system. Some of the general areas that were considered are:

1. Do G.P. zones exist in the ternary alloys?
2. If G.P. zones exist, what is their morphology?
3. Is there more than one type of G.P. zone?
4. Do the G.P. zones have an ordered internal structure?
5. Is the G.P. zone formation governed by a metastable miscibility gap?
6. Do the G.P. zones act as nuclei for precipitates as in aluminum-zinc, or are defects, such as dislocations, sites for precipitates as in aluminum-silver?
7. Is there a transition precipitate phase as in aluminum-silver and aluminum-zinc?
8. What is the crystal structure of the precipitate?
9. What is the morphology of the precipitate?

10. What is the orientation of the precipitate relative to the matrix?

The existence of G.P. zones in the ternary alloys was established by x-ray small angle scattering and transmission electron microscopy. The G.P. zones are observed to be present immediately after the quench, and are assumed to form during the quench as is the case with the aluminum-silver alloys. This fact was established by quenching the alloy to room temperature and observing the small angle scattering within one minute after the quench. Since the G.P. zones did not grow at room temperature, it must be concluded that they were formed during the quench. Alloys which were examined in the electron microscope were always polished in a solution below 0°C to inhibit G.P. zone growth during the polishing procedure. In the quenched conditions the zones have a Guinier sphere radius of 10-20 Å, thus making it difficult to establish their shape from visual observations. In addition, the large volume fraction of G.P. zones produces overlap problems, which also makes visual shape interpretation difficult.

Streaks were observed in the electron diffraction patterns of alloys in the quenched state. As the microscope used in this investigation was equipped with a goniometer stage, it was possible to examine a given area of the foil in different orientations. By taking the diffraction patterns and examining the streaking in two different orientations, one can determine in the foil the orientation of centers that are giving rise to the streaking.

Since the streaking in reciprocal space is in the <111> direction, it can be concluded that the scattering centers in the foil lie in the {111} planes. This suggests the change in shape of the G.P. zone that Gerold[8] observed for aluminum-zinc alloys.

Gerold considered the G.P. zone shape change and its effect on the scattering curve. As the change of shape is related to the change in total surface area, shape is more closely related to the intensity in the "Porod tail" than to the rest of the scattering curve.

Gerold states that a practical measure of the changes in the "tail intensity" is the ratio at two scattering angles S_1 and S_2, where the intensities have a distinct ratio J_1/J_2. J_M is the maximum scattered intensity, $J_1 = 0.75\ J_M$ and $J_2 = 0.125\ J_M$. Gerold's results can be summarized by observing that the ratio S_2/S_1 is equal to or greater than 1.9 when the zinc-rich aluminum-zinc G.P. zones have undergone a spherical to ellipsoidal shape change. Aluminum-silver zones are known to remain spherical throughout their growth, and the ratio of S_2/S_1 remains less than 1.8. For the ternary alloys the ratio of S_2/S_1 is always greater than 1.9, indicating that the G.P. zones are nonspherical. The occurrence of nonspherical G.P. zones is somewhat surprising as the G.P. zones in aluminum-silver are always spherical, and in aluminum-zinc they remain spherical until the Guinier radius exceeds 25-20 Å.

The occurrence of nonspherical G.P. zones suggests a distortion between the G.P. zones and the matrix. As has been stated previously, the shape of the (000) node is unaffected by strain, whereas strain leads to a broadening of the Bragg peaks. It should be emphasized that the small angle scattering represents only part of the total picture in the study of atomic configurations. The scattering due to heterogeneities in a solid solution is a combination of pure disorder scattering, observed at low angles, and the combined disorder and size contributions which are observed at larger diffraction angles.

For alloys exhibiting clustering, the diffuse scattering can be modified by a size effect term as has been shown by Warren and Averbach,[9] Guinier[10] and Sparks and Borie.[11] This modification is shown schematically in Figure 1. Examining Figure 1, it can be seen that when the scattering factor, f_A, of the A atoms in the cluster is greater than the scattering factor, f_B, of the B atoms in the matrix, and the A atoms are smaller than the B atoms, then the diffuse scattering is not symmetric about the node.

All samples used in this investigation were examined in a Guinier-DeWolff camera. The initial objective was to

ascertain whether or not precipitation had occurred during the quenching or aging sequence. The Guinier-DeWolff camera can be operated in a vacuum, which permits very long exposure times without appreciable "fogging" of the film. This has the decided advantage that very low volume fractions of second-phase particles can be detected.

Examination of the diffraction patterns obtained from the Guinier-DeWolff camera reveals a diffuse band with a maximum located at the high-angle side of the Bragg lines. A typical microdensitometer trace of one of the films is shown in Figure 2. One can readily see the similarity between Figure 1 and Figure 2. Analogous effects have been observed in this laboratory for aluminum-zinc and aluminum-silver alloys. Gerold[12] reported such an effect for aluminum-silver, but the magnitude of the size effect is much smaller due to the close similarity in size of the aluminum and silver atoms when in solid solution. It should be noted that the size effects have not been commonly observed in systems displaying clustering phenomena, as most workers have used either a Debye-Scherrer camera or a diffractometer in their investigations. The diffuse scattering intensity is much less than that of a Bragg line and can be easily masked by line broadening due to strain or $K\alpha_1$ - $K\alpha_2$ splitting. The focusing characteristic of the monochromator used in the Guinier-DeWolff camera is such that there is an achromatization of the diffraction line. The line profile is therefore independent of the band width, $\Delta\lambda$, of the radiation incident on the specimen.

If one now defines the average scattering factor of atoms in the G.P. zones as f_z, and the average scattering factor of atoms in the matrix as f_M, one can conclude from electron microscope observations that f_z is greater than f_M. Inasmuch as f_z is greater than f_M and the diffuse intensity maxima is on the high-angle side of the Bragg Reflection, it can be concluded that the lattice parameter of the G.P. zone is smaller than the lattice parameter of the matrix.

With the existence of G.P. zones established, it was necessary to consider the possibility of more than one type of G.P. zone being present. Murakami et al.[6] have proposed that two types of G.P. zones exist in the ternary aluminum-zinc-silver alloys, but their proposal is not supported by this research. If, as they suggest, there are two types of G.P. zones present; i.e., one zinc-rich and the other silver-rich, then it follows that there would likely be two types of precipitation taking place. Judging from past experiences with aluminum-zinc alloys, it is reasonable to assume that the zinc-rich G.P. zones would act as nuclei for the precipitation of the hexagonal close packed β-zinc. A series of experiments were carried out, whereby the alloys were quenched from various temperatures ranging from 568 to 350°C. The time, at temperature, became progressively longer as the temperature was decreased. After each treatment a diffraction pattern was taken, using the Guinier-DeWolff camera. In all cases only one precipitate type was observed to form. This fact is not really surprising, as Koster et al. showed that, for the alloys used in this research, the equilibrium precipitate should be the hexagonal close packed ε-phase. If the quenched alloys had contained two types of zones, then two types of precipitates would have been observed, and the G.P. zones should have disappeared at different times and different temperatures. If zinc-rich G.P. zones were present, it is likely that the kinetics and transition sequence would be similar to those observed for aluminum-zinc alloys. That is, they would grow rapidly at temperatures above 25°C and give rise to transition phases.

In addition to the above experiment, a series of alloys were examined in the electron microscope. Using the hot stage, the growth and dissolution of the G.P. zones was monitored and the formation of precipitates was observed. In every experiment only one type of G.P. zone could be observed visually, and only one precipitate type observed to form.

The G.P. zones in the aluminum-silver system have been shown by Auer and Gerold[13] to have an ordered internal

structure. Thus it was necessary to consider the possibility of an ordered zone state existing in the ternary alloys. Within the experimental sensitivity of the research carried out, there is no indication that the G.P. zones have an ordered internal structure. No superstructure reflections were observed in either the single crystal oscillation photographs or the electron diffraction patterns.

Of course it must be realized that in order to establish conclusively the presence or absence of internal G.P. zone order, a more sensitive experimental arrangement would be necessary. Ideally, this would consist of a primary beam monochromator utilizing a crystal that would not give rise to any $\lambda/2$ diffraction effects. The crystal should also have a radius of curvature large enough to provide an adequate working distance. An adequate working distance for this type of experiment would be one in which the beam would come to a focus at the film and still have sufficient space for experimental accessories.

Gerold[14] has shown that the concentration of solute atoms inside the binary G.P. zones in both aluminum-zinc and aluminum-silver alloys is a function of aging temperature only, and not of the initial concentration of the solid solution. This implies the existence of a metastable miscibility gap. The limits of the miscibility gap for a binary alloy may be determined from the quantitative measurement of the total integrated intensity, Q. Gerold has also shown that Q is independent of time and G.P. zone radius.

It can be concluded that formation of G.P. zones in the ternary alloys are also governed by the existence of a metastable miscibility gap. This can be supported by the fact that the relative integrated intensity, M, is independent of time and G.P. zone radius as can be seen in Figures 3 and 4. For the alloys considered in this research, one can speculate that the miscibility gap does not have the unusual shape that is observed in the binary aluminum-silver system. This can be determined by examining the relative integrated intensity, M, as a

function of aging temperature, T. Bauer and Gerold[15] have carried out this experiment for aluminum-silver and their results are shown in Figure 5. The integrated intensity as a function of aging temperature for the aluminum-zinc system is quite similar to that of aluminum-silver below 125°C. Figure 6 is a plot of volume fraction of G.P. zones for several aluminum-silver alloys as a function of aging temperature. As can be seen in Figure 6, the volume fraction of G.P. zones remains quite constant up to the temperature where the abrupt change in the shape of the metastable miscibility gap occurs. Since the volume fraction of G.P. zones changes only slightly, the variation in integrated intensity must be due to changes in the composition of the matrix and G.P. zones. A plot of the relative integrated intensity, M, versus aging temperature, T, is given for the ternary alloys in Figure 7. Note that these curves do not exhibit sharp changes in integrated intensity such as is seen for the aluminum-silver alloys.

The metastable miscibility gap should exist within the two-phase region ($\alpha + \varepsilon$) of the stable ternary diagram. Koster et al. have experimentally determined the 300°C isotherm for this system and it can be seen in Figure 8. Koster et al. also gave the lattice parameter of the equilibrium ε-phase as a function of composition. The lattice parameter of the ε-phase was experimentally determined for the alloys used in this research and the results are given in Table 1. By comparing these results with the data of Koster et al., the composition of the precipitate was determined.

Table 1

Lattice Parameters of ε-Phase at 300°C

Alloy	a_o	c_o
Al - 5 a/o Zn - 5 a/o Ag	2.8240	4.4107
Al - 8 a/o Zn - 2 a/o Ag	2.8398	4.4562
Al - 10 a/o Zn - 3 a/o Ag	2.8308	4.4366
Al - 3 a/o Zn - 12 a/o Ag	2.8810	4.5685

Whereas the metastable miscibility gap should exist within the two-phase region ($\alpha + \varepsilon$) of the ternary diagram, the highest possible concentration of silver, zinc and aluminum atoms in the G.P. zones is thus given by the phase boundary of the stable phase. If one uses the composition of the equilibrium precipitate as the composition of the G.P. zones, the electron density of the G.P. zones can be calculated. By comparing the relative integrated intensity, M, which has been shown to be proportional to $(\rho-\rho_o)^2$, with the calculated electron density, it can be seen that the order of increasing M proceeds in the right direction.

The complete precipitation sequence for the ternary alloys was observed in the electron microscope equipped with a commercial hot stage. The use of the hot stage yielded meaningful information, since the surface precipitation problem that is encountered with aluminum-zinc alloys was not found in the ternary alloys. Thus, it was possible to establish the precipitation sequence: G.P. zones → ε' → ε directly in the thin foils. Hren and Thomas,[16] while studying binary aluminum-silver alloys, concluded that the majority of vacancies are trapped in the silver-rich G.P. zones; thus, surface precipitates are not observed in the transmission electron microscopy of thin foils of aluminum-silver alloys.

G.P. zone formation was observed as the first stage in the metastable phase sequence that was progressing toward the equilibrium state. The G.P. zones in aluminum-zinc-silver alloys are quite stable during the aging sequence and the growth kinetics are primarily the result of a redistribution process. It was not possible to revert to the G.P. zones prior to the appearance of the ε' precipitates. Panseri and Federighi[17] reported the binding energy between a zinc atom and a vacancy to be 0.06 eV. Hren and Thomas reported the binding energy between a silver atom and a vacancy as approximately 0.4 eV. This means that if the G.P. zones in the ternary alloys contain silver, then the binding energy between the vacancies and G.P. zones must also be quite large. This can also help to explain why

the ternary alloys behave in a manner similar to aluminum-silver alloys, even when the concentration of zinc atoms is much greater than the concentration of silver atoms.

Westmacott and Barnes[18] considered the problem of vacancy trapping in quenched aluminum alloys. They assume that immediately after quenching a new equilibrium between the concentration of free vacancies, c_v, and the concentration of vacancies trapped by an impurity, c_1, is established. Thus the total number of vacancies $c_t = c_v + c_1$. They further state that the concentration of impurity atoms having one trapped vacancy, c_1, is

$$c_1 = 12c_a \exp - \frac{E_f - E_b}{kT_q} \exp \frac{S_1 - S_o}{k} \qquad (1)$$

where c_a = concentration of impurity atoms
E_b = binding energy between a vacancy and solute atom
E_f = activation energy for the formation of vacancies
S_o = entropy term associated with placing a single impurity atom into the lattice
S_1 = entropy term associated with placing an impurity atom with its vacancy into the lattice.

If the assumption is made that equation (1) can be approximated by equation (2), since the activation energy for the formation of vacancies in aluminum-silver alloys is approximately the same as that for an aluminum-zinc alloy, then

$$c_1 = 12A\, c_A \exp \frac{E_b}{kT_q} \qquad (2)$$

and if we assume that $c_{A\text{-}Zn}$ = 12 a/o Zn and $C_{A\text{-}Ag}$ = 3 a/o Ag, then

$$c_{1\text{-}Zn} = 12A(0.12) \exp \frac{0.06}{kT_q}$$

$$c_{1\text{-}Ag} = 12A(0.03) \exp \frac{0.4}{kT_q}$$

$$\frac{c_{1\text{-}Ag}}{c_{1\text{-}Zn}} = 0.36$$

which means that even though $(c_{A\text{-}Ag})/(c_{A\text{-}Zn}) = 0.25$, 36 percent of the vacancies are associated with the silver atoms.

As a natural consequence of extended aging treatment, the ternary alloys undergo a transition from the G.P. zone state to the occurrence of a true precipitate. The ε' precipitates were observed to nucleate on defects in the matrix and lie on {111} planes of the matrix. That is, the {0001} planes of ε' are parallel to the {111} planes and coherent with the matrix. The ε'-phase has a close packed hexagonal structure, and the atomic arrangement within the cell is quite similar to that observed for $CuZn_3$. The precipitate does not appear to have an ordered structure as indicated by the absence of superstructure reflections. In the early stages of precipitation, as can be seen in Figure 9, G.P. zones and precipitates can co-exist. The precipitates are platelet in shape with their length being very much greater than their thickness.

The final step in the precipitation sequences is the appearance of the equilibrium ε precipitate. The ε-phase also has a close packed hexagonal structure but is totally incoherent with respect to the matrix. It is a discontinuous precipitation process, and the ε-phase is much less regular in its spatial distribution, even though differences in the lattice parameter between ε and ε' could not be detected. They are assumed to exist due to the loss of coherence of the {111} planes of the matrix and {0001} planes of the precipitate.

Conclusions

Thus, in summary it can be concluded:

1. G.P. zones do exist in the ternary aluminum-zinc alloys.
2. The G.P. zones are nonspherical and are most likely ellipsoidal in shape.

3. Only one type of G.P. zone exists and is of ternary composition, and the lattice parameter of the G.P. zone is smaller than the matrix.
4. The G.P. zones do not appear to have an ordered internal structure.
5. G.P. zone formation is governed by the existence of a metastable miscibility gap.
6. There is a transition, hexagon, close packed precipitate, ε'.
7. ε' is nucleated on defects in the matrix such as dislocation.
8. ε' is a platelet precipitate with {0001} planes of the precipitate coherent with the {111} planes of the matrix.

BIBLIOGRAPHY

1. Wilm, A., Metallurgie, 8, 225 (1911).

2. Guinier, A., Nature, 142, 569 (1938).

3. Preston, G. D., Nature, 142, 570 (1938).

4. Shashkov, D. and Buynov, N. N., Fiz. Metal Metalloved, 17, 278 (1964).

5. Koster, W., Müller, R. and Seelhorst, J., Z. Metall., 55, 589 (1964).

6. Murakami, M., Kawano, O. and Murakami, Y., Acta Met., 17, 29 (1969).

7. Harkness, S. D., Gould, R. W., Hren, J. J. and Sheble, A., Rev. Sci. Inst., 38, 1676 (1967).

8. Gerold, V., Conference on Small Angle X-Ray Scattering (Syracuse University, 1965).

9. Warren, B. E. and Averbach, B. L., Modern Research Techniques in Physical Metallurgy (ASM, Cleveland, 1953).

10. Guinier, A., X-Ray Diffraction in Crystals, Imperfect Crystals and Amorphous Bodies (W. H. Freeman and Co., San Francisco, 1963).

11. Sparks, C. J. and Borie, B., Local Atomic Arrangements Studied by X-Ray Diffraction (Gordon and Breach, New York, 1965).

12. Gerold, V., Physica Status Solidi, 1, 37 (1961).

13. Auer, H. and Gerold, V., Z. Metall., 56, 240 (1965).

14. Gerold, V., Physica Status Solidi, 1, 37 (1961).

15. Baur, R. and Gerold, V., Acta Met., 10, 637 (1962).

16. Hren, J. J. and Thomas, G., TAIME, 227, 208 (1963).

17. Panseri, C. and Federighi, T., Acta Met., 8, 217 (1960).

18. Westmacott, K. H. and Barnes, R. S., Phil. Mag., 51,929 (1961).

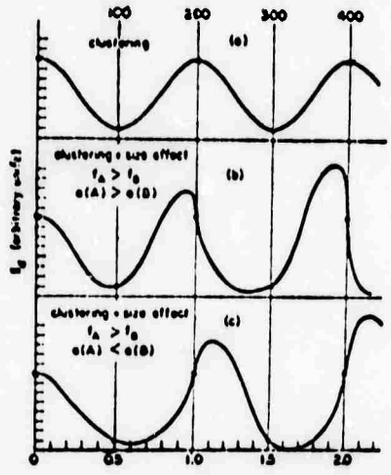


Figure 1. Idealized diffuse scattering along [100] direction for a binary single crystal with clustering and size effects.

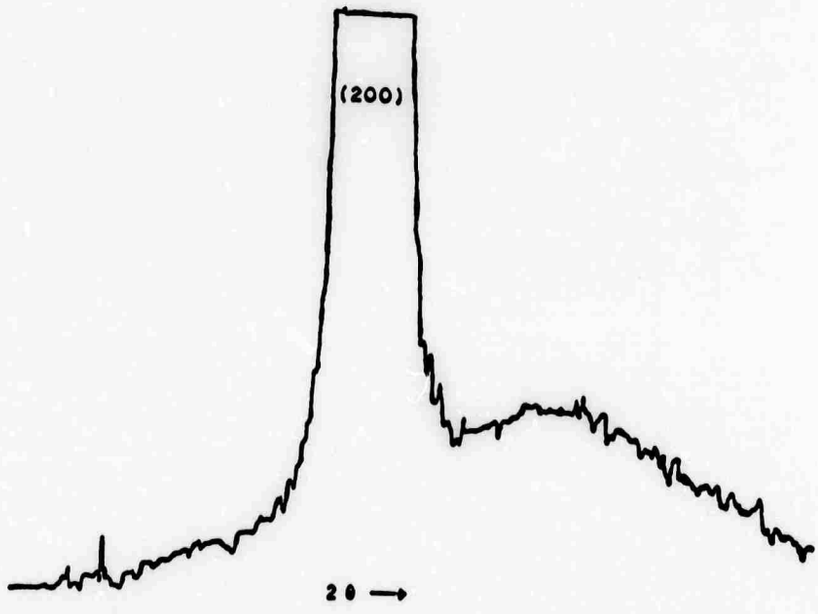


Figure 2. Microdensitometer trace of the diffuse side band on the high-angle side of the (200) Bragg reflection for Al - 10 a/o - 3 a/o Ag.

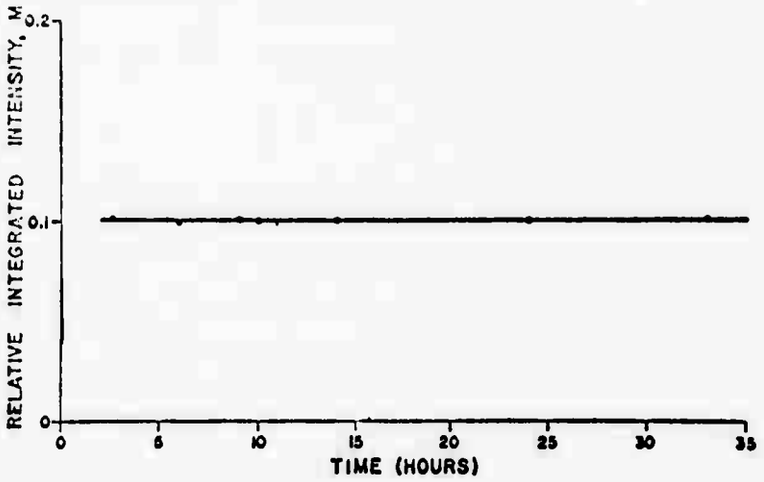


Figure 3. Relative integrated intensity vs. time for aluminum-5 a/o zinc-5 a/o silver aged at 175°C.

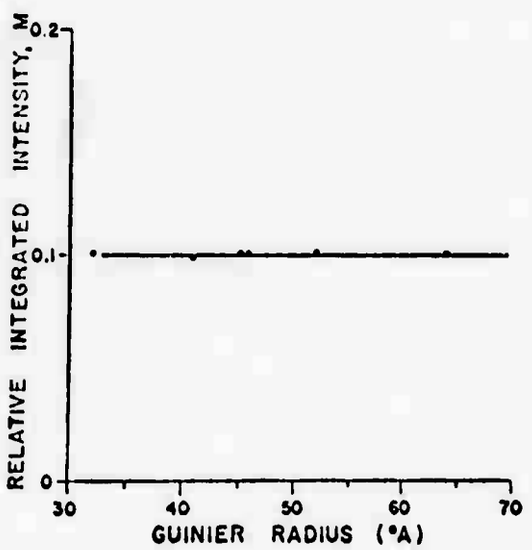


Figure 4. Relative integrated intensity vs. Guinier radius for aluminum-5 a/o zinc-5 a/o silver.

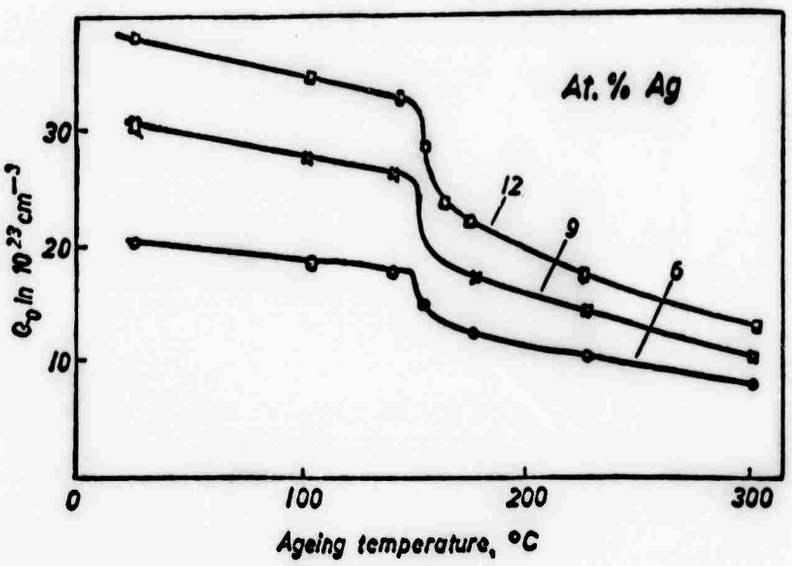


Figure 5. Integrated intensity vs. temperature for aluminum-6 a/o silver, aluminum-9 a/o silver and aluminum-12 a/o silver.

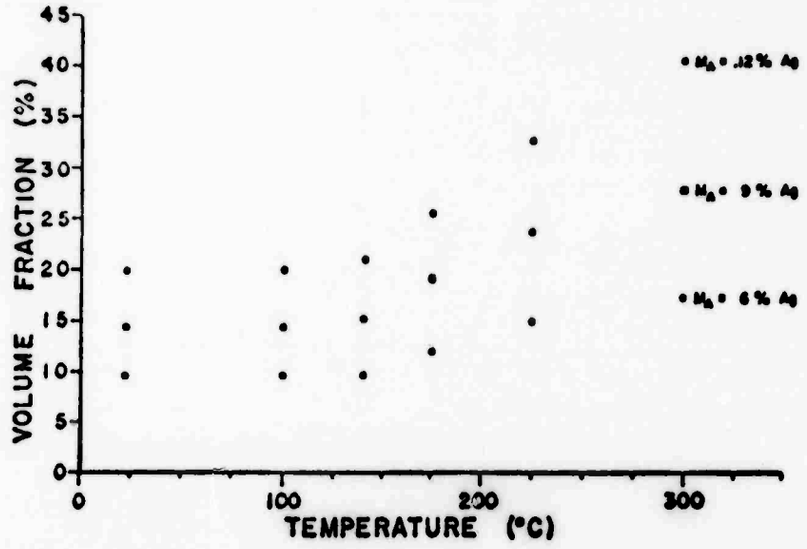


Figure 6. Volume fraction of G.P. zones vs. temperature for aluminum-6 a/o silver, aluminum-9 a/o silver and aluminum-12 a/o silver.

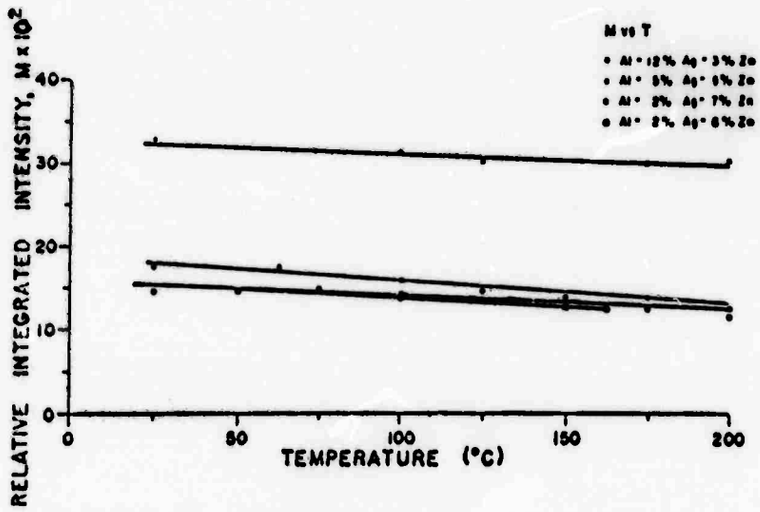


Figure 7. Relative integrated intensity vs. temperature for ternary aluminum-zinc-silver alloys containing Al-12 a/o Ag-3 a/o Zn (Δ), Al-5 a/o Ag-5 a/o Zn (o), Al-3 a/o Ag-7 a/o Zn (o), Al-2 a/o Ag-8 a/o Zn.

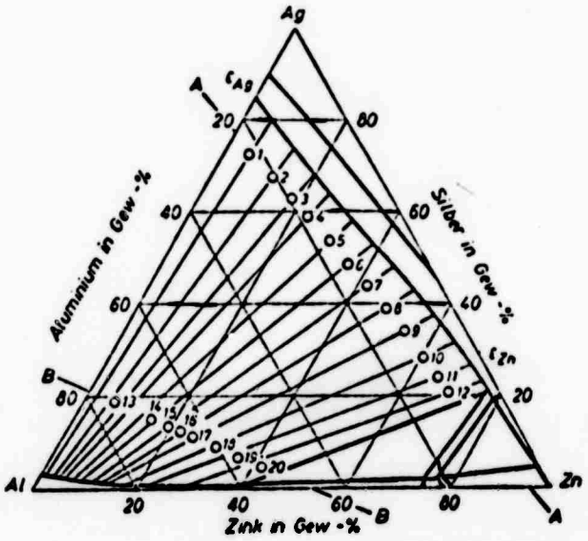


Figure 8. Aluminum-zinc-silver 300°C isotherm of the equilibrium phase diagram.

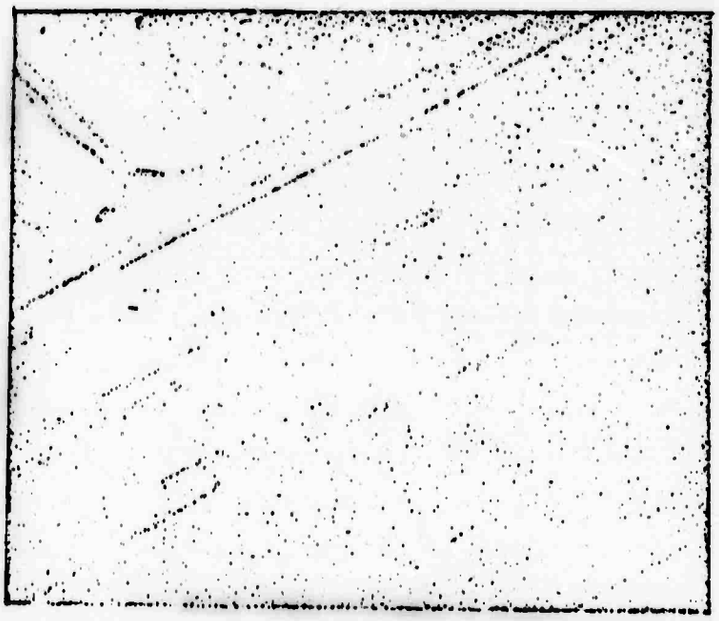


Figure 9. Transmission electron micrograph for aluminum-7 a/o Zn-3 a/o Ag quenched from 500°C to 25°C and aged in the microscope 40 minutes at 125°C. (92,600X)

B. METAL TIP HOLDER FOR FIELD ION MICROSCOPE (D. A. Jenkins and J. J. Hren)

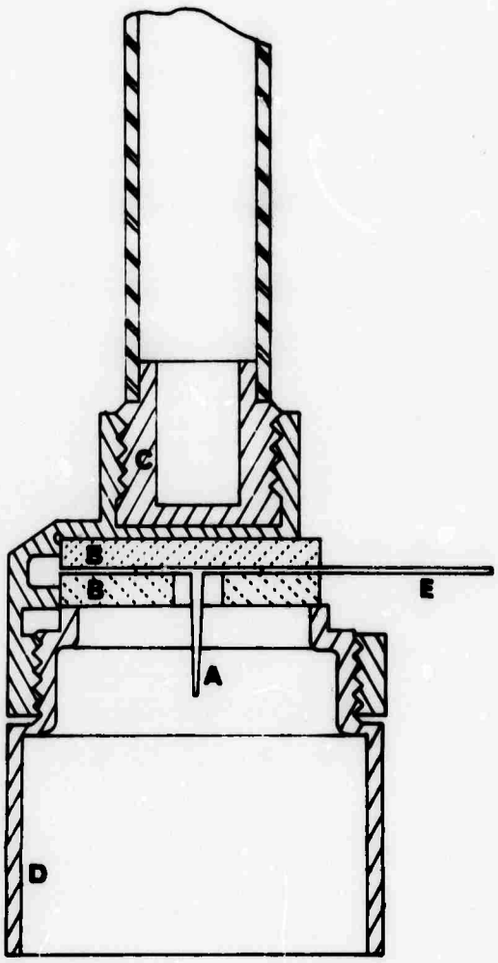
Tip holders for field-ion microscopes have traditionally been made using glass-to-metal seals in various configurations. Unfortunately, such assemblies are subject to occasional catastrophic failures. Thus a need arises for a durable tip holder with electrical and thermal properties equal to or better than the glass-metal types.

A useful design using no glass in its construction is shown in the figure. The body of the tip holder is made of OFHC copper and the tip mounting studs (A) are insulated by means of two sapphire disks (B) in a sandwich arrangement. The whole assembly is attached, by means of a threaded cap (C), to the end of a sealed stainless steel tube which is supplied with a cryogenic liquid from the outside of the microscope body. The threaded cap on the end of the tube is made of copper and is fairly thin on the bottom to allow efficient heat transfer. The sapphire disks combine good electrical insulation characteristics with good thermal conductance . The lower portion of the copper cylinder (D) serves as a grounded shield and a thermal radiation shield. Electrical connection is made through a small stainless steel rod (E) which protrudes through a clearance hole in the side of the body. This rod is attached directly to a copper disk into which the mounting studs are fastened. This disk can be cut in half and a lead and a mounting stud can be attached in each side, making it possible to heat the wire loop attached to the tip. All of the metal parts are gold-plated to facilitate maintenance of clean surfaces.

To load the tip holder, it is taken off the cooling tube after disconnecting the high voltage lead. The specimen can then be mounted in the holder at a place and from an angle that is convenient for the operator. The assembly is then reattached to the cooling tube, the high voltage lead is attached, and the system can be pumped down.

Catastrophic failures which can destroy costly screens and image intensifiers are much less probable with this design. For example, failure of either of the sapphire disks by cracking has no effect on the integrity of the structure. Accidentally ruining the tip by brushing it with the hand is made far less likely by the shape of protective shield.

The assembly has been found to be cryogenically effective, but noticeably slower to cool down than glass-metal types because of its relatively large mass. After an appropriate cooling down period (about 15 minutes), very sharp images can be obtained, indicating that the tip has reached a temperature reasonably close to that of the cryogenic liquid (e.g., hydrogen). From time to time, a buildup is observed to form on the sapphire surfaces leading to premature breakdowns. The problem is readily solved by thoroughly cleaning the sapphire disks with suitable organic solvents. Further refinements to the design will be aimed at reducing the mass of the metal parts in order to shorten the cool-down period, and at improving the ultimate insulation capability (currently about 30 KV), but the device has proven to be a useful addition to the field-ion microscope in its present form.


C
B
B
E
A
D

V. Discussion

From the research of this sixth semiannual period of contract support have come findings of technical significance upon which our future efforts will build. Moreover, the findings relate directly to fields of interest to the D. O. D.

Our noise research has identified sources of burst noise at the surface as well as in bulk material, thus providing a full characterization of this phenomenon. Burst noise is a critical source of low-frequency noise in bipolar transistors and has plagued the manufacturers of semiconductor integrated circuit amplifiers. By disclosing the physical origin of burst noise, this research furnishes an understanding that enables the control of its occurrence.

Additionally, the noise research has yielded a full understanding of phototransistor noise performance and has provided, as a by-product, means for obtaining information important to device design that was previously unattainable. Because the noise inherent in the operation of a device limits its sensitivity as a detector and its stability as a switch, these results bear on the design of detectors and amplifiers, with attendant applications in many fields of military importance: detection surveillance, telemetry, bio-medicine, communications, etc.

To achieve superior performance, the military semiconductor industry has shown a direction recently that makes essential the role that degenerate regions take in determining device performance. Gaps that exist in understanding this role preclude the optimization of device design. The research reported here has filled some of these gaps. We have calculated the density of states and the Fermi level in degenerate silicon, taking account of the dependence on energy and on impurity concentration. Based on the understanding thus gained, we are exploring the feasibility of a technique for extracting, by approximation, the essence of the complicated impurity concentration dependence of the density of states. If successful, this procedure will yield a tractable theory suitable for the study of some hitherto incompletely explained aspects of pn junction device behavior; among these is the temperature dependence of the common-emitter current gain.

We have reported research relevant to the design of electronic equipment for operation at low temperatures. Our study of steady-state recombination

and trapping processes of injected carriers in gold- and phosphorus-doped silicon has provided insight concerning the relationship between parameters useful in device design, such as lifetime, and parameters determined by fabrication, such as the density of gold. Building upon the results of this study, we are now exploring the use of silicon overcompensated with gold as a detector or a laser-beam modulator.

Two studies reported here have described techniques enabling better design of components of military systems relevant to such diverse fields as detection, navigation, surveillance, telemetry, bio-medicine, control and communications. The first of these studies deals with a unifying approach to the derivation of models for field-effect devices. These devices find application in many military systems. The resulting models are especially well-suited for use in the computer-aided design of semiconductor integrated circuits. The second study has described a means for automated simultaneous ac and dc design of linear semiconductor integrated-circuit amplifiers. The design procedure takes into account both fabrication parameters and mask geometry.

From the research concerning insulating and semiconducting glasses has come understanding directly applicable to such fields as bio-medicine, deep-sea technology, and the radiation-insensitive design of many different types of electronic systems. Our research has identified the sources of strength-enhancement and strength-reduction in certain species of glass ceramics. The detailed study of the early stages of crystallization of glasses has provided knowledge basic to the design of mechanical and dielectric properties. Additionally, we have described research relating composition and structural parameters with the parameters that describe electric and dielectric behavior.

The plan of the research program remains unchanged. Emphasis will continue to focus on the relationship between the structural properties of materials, on the one hand, and the consequent electrical, quantum, chemical, mechanical and thermal properties, on the other. Further correlation will be sought with radiation effects. As the research reported here indicates, device and circuit applications made possible by the understanding gained from materials studies will receive increasing attention. The materials research will both suggest collateral device studies and will be guided by them.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
University of Florida
Engineering and Industrial Experiment Station
Gainesville, Florida 32601

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
A CENTER OF COMPETENCE IN SOLID STATE MATERIALS AND DEVICES

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Scientific Interim

5. AUTHOR(S) *(First name, middle initial, last name)*
Fred A. Lindholm, Stephen W. Director, Sheng S. Li
Arthur J. Brodersen, Robert W. Gould, James K. Watson
Eugene R. Chenette, Larry L. Hench, Aldert van der Ziel
John L. Hren

6. REPORT DATE
10 October 1970

7a. TOTAL NO. OF PAGES
319

7b. NO. OF REFS
246

8a. CONTRACT OR GRANT NO.
F 19628-68-C-0058 ARPA Order No. 1060

b. PROJECT, TASK, WORK UNIT NOS.
8687 n/a n/a

c. DOD ELEMENT
61101D

d. DOD SUBELEMENT
n/a

9a. ORIGINATOR'S REPORT NUMBER(S)
Scientific Report No. 6

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
AFCRL-71-0030

10. DISTRIBUTION STATEMENT
1- This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES
This research was supported by Advanced Research Projects Agency.

12. SPONSORING MILITARY ACTIVITY
Air Force Cambridge Research Laboratories (LQ)
Laurence G. Hanscom Field
Bedford, Massachusetts 01730

13. ABSTRACT This report describes technical findings in three main subject areas: semiconductors and semiconductor devices, insulating and semiconducting glasses, and measurement techniques. In the research concerning semiconductors and semiconductor devices, a model is presented that identifies and fully characterizes the physical origins of burst noise in bipolar transistors. A new characterization of phototransistor noise performance is reported. The density of states and Fermi level for silicon is derived over the whole range of impurity concentrations of technological interest, taking account that the density of states depends on impurity concentration. Steady-state recombination and trapping processes in gold- and phosphorus-doped silicon are studied for impurity concentrations and temperatures such that the impurity density greatly exceeds the equilibrium carrier density. A unifying approach is described for deriving models for field-effect devices. Algorithms for the automated design of semiconductor integrated-circuit amplifiers are developed that grow transistors, in the computational sense, so as to attain specified design goals. In the research concerning insulating and semiconducting glasses, variations of thermal expansion coefficient, elastic modulus, and fracture strength of Li_2O-SiO_2 glass-ceramics are determined as functions of the nucleation treatment and volume fraction of crystals present. X-ray small-angle scattering results indicate that a metastable reaction precedes equilibrium crystallization in certain Li_2O-SiO_2 glasses. The results of our continuing research effort concerning the electrical properties of glass are summarized and related to results obtained elsewhere. In the research concerning measurement techniques the design for a more reliable tip-holder for field-ion microscopes is proposed. X-ray scattering, x-ray diffraction and transmission electron microscopy are employed to study the solid solution decomposition of aluminum-rich, aluminum-zinc-silver alloys.

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
transistors						
semiconductors						
noise						
glass						
field-ion microscope						