F. Fang
S. Triebwasser

Carrier Surface Scattering in Silicon Inversion Layers

Abstract: The field effect surface-channel conductance and transconductance of both *p*-type and *n*-type Si inversion layers were measured as a function of external field. In the small signal region, the channel conductance was found to vary logarithmically with the transverse field. The results are interpreted in terms of reduction of carrier mobility that is due to surface scattering. A model which consists of a uniformly distributed charge layer and self-consistent field is proposed to explain the observed results. It was found that in most samples measured a combination of specular and diffuse scattering is involved. Examples of completely diffuse scattering as well as the diffuse and specular combination are given. The temperature dependence of the surface mobility between 77° and 300°K is presented.

Introduction

The electric conductivity of a semiconductor surface is usually different from that of the bulk material because of the effect of band bending associated with the surface potential that is due to the surface states or external fields. The differences are manifested as effects on charge distributions and on the mobility of the carriers. If the transverse field (i.e., one normal to the surface) near the surface is sufficiently large, the surface scattering frequency may be comparable to or larger than that associated with bulk scattering. If in addition the surface is a diffuse scatterer, i.e., the carriers are randomized thermally upon impinging at the surface, then one would expect a reduction of the carrier mobility relative to the bulk value.

The effect of surface scattering in semiconductors was first discussed by Schrieffer,[1,2] following similar analyses by Fuchs[3] and by Sondheimer[4] in the calculation of size effect of conductivity in metallic thin films. At constant high transverse field, Schrieffer was able to show that the carrier mobility due to diffuse surface scattering is inversely proportional to the field strength. Many more elaborate calculations[5-7] arrive at essentially the same conclusion. Recent experiments on conductivity in inversion layers in *p*-type Si[8] indicate that the field dependence of the carrier surface mobility is somewhat different from Schrieffer's result. A quadratic potential well near the surface was suggested to account for the observed field dependence. In this paper, we shall present some experimental results which show the effect of the combination of diffuse and specular surface scattering for both electrons and holes in the appropriate inversion layers.

Experiment

The samples used for the present work are typical insulated-gate field effect transistors. The surface contacts, i.e., the source and drain, are diffused areas of a different conductivity type from that of the bulk. Thus, ohmic contact is made only to the surface inversion layer on the surface, if it exists. The configuration of a sample with an inversion layer on a *p*-type substrate is shown schematically in Fig. 1a, and the corresponding energy band diagram is shown in Fig. 1b. Complementary structures were used for the study of the inversion layers on *n*-type samples.

The *p*-type substrates were boron-doped to about $1.5 \times 10^{15}/\text{cm}^3$. The *n*-type substrates were phosphorus-doped to about $4 \times 10^{14}/\text{cm}^3$. For both, the background impurities were less than $10^{13}/\text{cm}^3$. Since no intentional alteration of the impurity concentration of the Si near the $Si\text{-}SiO_2$ interface was made to change the surface properties, and the breakdown voltages between the substrates and the diffused areas correspond to the respective substrate impurity concentrations, it is believed that the band bending near the interface at zero gate voltage is

due to the interface states and/or the work function differences of the metal-oxide-silicon gate system.

The distance between source and drain is of the order of 10 microns, the thickness of the oxides is about 5600 Å, and the width of the channel is about 500 microns. The channels are usually in a closed pattern to eliminate possible leakage. Surface conductance was measured with a small ac signal while a dc voltage was applied between the gate and the source. Transconductance was measured by superimposing a small ac signal on the dc bias between gate and source while the source-drain voltage was smaller than 0.5 volt. The ac signals used in all measurements were smaller than 8 mV (rms). The ac channel currents were measured across a small resistance ($\sim 1\Omega$) in series with the channel. A narrow band detection system was used throughout the measurements. Care was taken that for all samples measured, the series resistances of the source or drain contacts would be negligible compared to the channel conductance. The small dc bias in the source-drain circuit ensures that the transverse field in the channel has negligible longitudinal variation.

Below source-drain pinch-off, simple linear theory predicts that in the absence of trapping the transconductance G_m will be given by[9]

$$\frac{\partial I_D}{\partial V_g} \equiv G_m = \frac{W}{L} C\mu V_D, \qquad (1)$$

where C is the capacitance per unit area between the gate and the channel; V_D, the source-drain voltage; I_D, the drain current; V_g, the gate voltage; L, the channel length; and W, the channel width. It was found experimentally that the capacitance C is constant over a large range of gate voltage after an inversion layer is formed.[10,11] Figure 2 shows a typical plot of channel capacitance as a function of the gate voltage. In the voltage range where the inversion layer is formed, the channel-to-gate capacitance is essentially that of the simple parallel plate capacitance of the SiO_2 layer. When the gate voltage is increased sufficiently in the negative direction, the channel begins to deplete. As the voltage is further increased an accumulation layer is formed. The capacitance is again an equivalent of a simple parallel-plate capacitor.

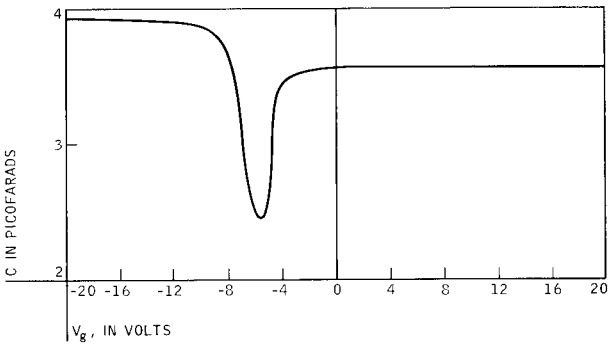


Figure 2 **The gate capacitance as a function of the gate voltage of an n-channel FET.**

Figure 1 **The schematic diagram of an n-channel field effect transistor.**

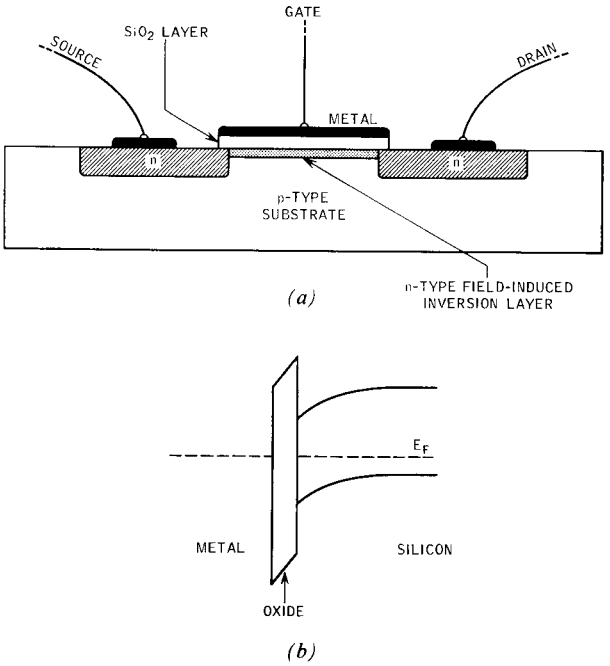
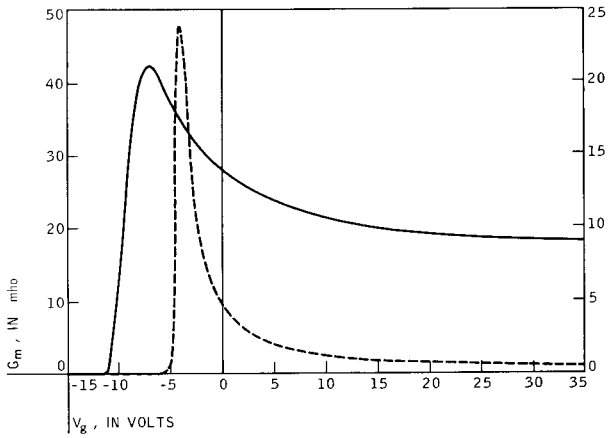


Figure 3 **Typical differences in the dependence of small signal transconductance on gate voltage.** *The broken curve applies to a typical sample for which $V_D = 0.1V$; ordinate values should be read from the left scale. The solid curve applies to a second typical sample with $V_D = 0.2V$; ordinate values should be read from the right scale.*



The region of interest for the work reported here is that over which the inversion layer exists and for which the capacitance is that of the SiO_2 dielectric layer. In this region G_m should remain constant with respect to the gate voltage provided that the carrier mobility is not a function of gate voltage. However, measurements of G_m show a quite strong gate voltage dependence. Figure 3 shows typical plots of G_m versus V_g for two typical samples. The decrease of G_m with increasing V_g can be due either to surface state trapping or to reduction of carrier mobility or to both. In the case of a highly inverted n-type surface, for which most of the measurements were performed, the Fermi level at the Si-SiO_2 interface is very close to the conduction band edge. The deep trapping states are filled and should have little effect on steady-state small-signal measurements. The incremental charges should all be mobile and may be determined from the differential capacitance.[12] To good approximation,

$$qN = C(V_g + V_{ox}), \tag{2}$$

where q is the electronic charge, N is the surface density of the carriers, and V_{ox} is the built-in potential in the oxide at zero bias. Hence, the observed reduction of G_m with increasing gate voltage can only be due to the reduction of carrier mobility caused by surface scattering.

The channel conductance G_{SD} is given by

$$G_{SD} = \frac{W}{L} Nq\mu. \tag{3}$$

Thus, from (2)

$$\mu = \frac{L}{W} \frac{G_{SD}}{C(V_g + V_{ox})} = \frac{L 4\pi\delta G_{SD}}{W\epsilon_{ox}(V_g + V_{ox})}, \tag{4}$$

where δ is the SiO_2 thickness and ϵ_{ox} is the dielectric constant of SiO_2. Figure 4 shows typical plots of the dependence of the channel conductance on the gate voltage.

The effective mobility calculated from Eq. (4) is plotted in Fig. 5 as a function of the transverse field $(\epsilon_{ox}/\epsilon_{Si}) \cdot (V_g + V_{ox})/\delta$ where ϵ_{Si} is the dielectric constant of Si. The distinctive difference between the two typical curves lies in the differing values for their asymptotic behavior at high fields.

Interpretation

The transverse fields involved in the measurements described above are of the order of 10^5 volts/cm, for which the mean free path for surface scattering is about 50 Å at room temperature. The bulk mean free path estimated from the bulk mobility is larger than 200 Å. Thus the effective mobility should be dominated by surface scattering if it is diffuse. In the constant field approximation at high field limit μ_{eff} is proportional to F^{-1}, where F is the transverse field in the channel. Since the carrier density N is proportional to F then, according to Eq. (3), there should

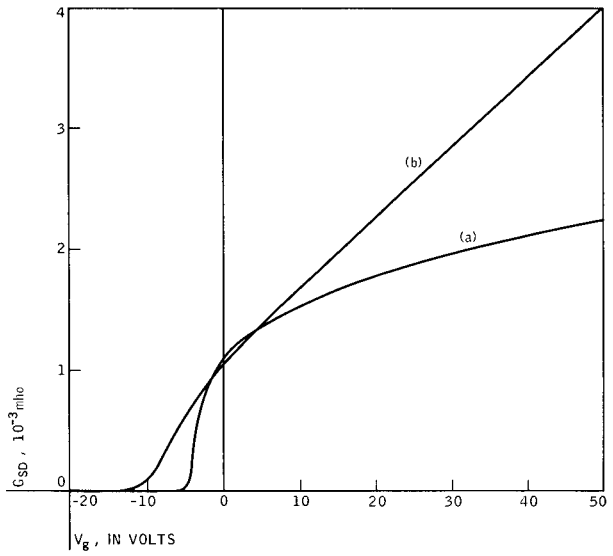
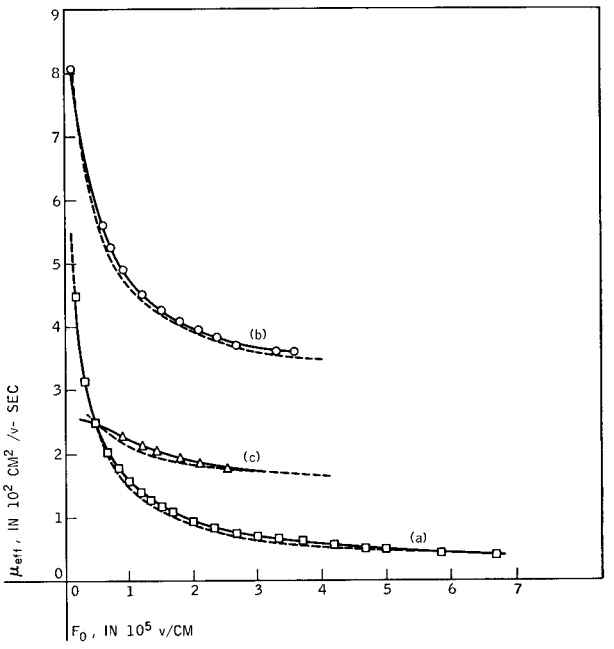


Figure 4 **Channel conductances of the two typical devices as a function of the gate voltage.**

Figure 5 **Effective mobilities of the carriers in the inversion layer: (a) An example for complete diffuse scattering of the electrons; here $\mu = 3.78 \times 10^6\ (1/F_0)\ \ln [F_0/(F_1/2)]$. (b) An example for specular and diffuse scattering of the electrons; here $\mu = 278 + 4.65 \times 10^6\ (1/F_0)\ \ln [F_0/(F_1/2)]$. (c) Mobility of the holes with combination of specular and diffuse scatterings; here $\mu = 142 + 1.8 \times 10^6\ (1/F_0)\ \ln [F_0/(F_1/2)]$.** *Dashed lines represent curves drawn from calculated values; $F_1 = 5 \times 10^3 V/cm$ for (a), (b) and (c).*



be no conductivity modulation in the channel. The fact that substantial transconductance is observed seems to suggest a weaker field dependence of the effective mobility. In this connection, consider that in the region of interest the channel width is about 25 Å (half of the surface mean free path) and that the de Broglie wavelength of the carrier with a wall energy of approximately $kT/2$ is about 100 Å. Thus, it is not appropriate to use the Thomas-Fermi model for conventional three-dimensional band states in calculating the carrier concentration in the channel. Rather, for the purpose of spatial population density calculations, it appears to be more correct to assume that the carriers encounter an average potential. For a first approximation, it is assumed that the density of the free carriers is uniform in the channel. The field is then given by

$$F(x) = F_0 - \frac{x}{l}(F_0 - F_1), \tag{5}$$

where x is the distance from the interface; l is the channel width; F_0, the field in Si at the Si-SiO_2 interface, is given by

$$F_0 = \frac{\epsilon_{ox}}{\epsilon_{Si}} \frac{(V_g + V_{ox})}{\delta} \cong \frac{4\pi q N}{\epsilon_{Si}}; \tag{6}$$

and F_1, the residual field at the edge of the free carrier sheet, is given by

$$F_1 = \frac{4\pi q X_1 N_a}{\epsilon_{Si}} \equiv \frac{\epsilon_{ox}}{\epsilon_{Si}} \frac{V_1}{\delta}. \tag{7}$$

In Eq. (7) X_1 is the width of the depletion region and V_1 is defined by the latter equality. In a one dimension case, for classical particles scattered from the interface with an energy U_0 moving in the conservative field F,

$$U_0 = \int_0^l qF\,dx. \tag{8}$$

From Eqs. (5) and (8), we obtain

$$l = \frac{2U_0}{q(F_0 + F_1)}. \tag{9}$$

From Eq. (5), the potential the carrier encounters is

$$V(x) = \frac{(F_0 - F_1)}{2l} x^2 - F_0 x. \tag{10}$$

The classical electronic motion is described by

$$\frac{d^2x}{dt^2} = \frac{q}{m^*}\left[F_0 - \frac{x}{l}(F_0 - F_1)\right], \tag{11}$$

where m^* is the carrier effective mass. The first integration of Eq. (11) can be performed immediately, yielding

$$\frac{dx}{dt} = \left\{\frac{q(F_0 - F_1)}{m^* l}\left[x^2 - \frac{2lF_0 x}{F_0 - F_1} + \frac{l^2(F_0 F_1)}{F_0 - F_1}\right]\right\}^{\frac{1}{2}}, \tag{12}$$

which includes the boundary condition

$$U_0 = \tfrac{1}{2}m^*(dx/dt)^2 = (F_0 + F_1)/2lq \quad \text{at} \quad x = 0.$$

The scattering time, τ, is given by

$$\tau = 2\int_0^l \frac{dx}{\left(\frac{dx}{dt}\right)}. \tag{13}$$

When $F_0 \gg F_1$, a condition which is satisfied in the region of interest,

$$\tau = (2U_0 m^*)^{\frac{1}{2}} \frac{2}{qF_0} \ln \frac{2F_0}{F_1}. \tag{14}$$

The effective mobility due to diffuse scattering is found from

$$\mu_d = \frac{q}{m^*}\tau = \left(\frac{2U_0}{m^*}\right)^{\frac{1}{2}} \frac{2}{F_0} \ln \frac{2F_0}{F_1}, \tag{15}$$

which except for the logarithmic factors is similar to Schrieffer's result in the high field limit.

The above considerations are valid only for the case of diffuse scattering at the surface. However, there is no fundamental reason for a perfectly plane surface to exclude the elastic scattering at the surface. Fuchs[3] introduced a parameter p which was interpreted as the probability that a carrier striking the surface will be specularly scattered. The resulting effective mobility may be written[13]

$$\mu_{eff} = p\mu_b + (1 - p)\mu_d, \tag{16}$$

where μ_b is the bulk mobility, μ_d is the diffuse surface scattering limited mobility given by Eq. (15), and p may be a function of the field. It follows from Eqs. (3), (6), (7), (15), and (16) that the channel conductance is then

$$G_{SD} = \frac{W}{L}\left[\frac{p\mu_b\epsilon_{ox}(V_{ox} + V_g)}{4\pi\delta} + \frac{(1 - p)\epsilon_{Si}}{2\pi}\left(\frac{2U_0}{m^*}\right)^{\frac{1}{2}} \ln \frac{V_{ox} + V_g}{V_1/2}\right]. \tag{17}$$

Accordingly,

$$\frac{G_m}{V_D} = \frac{W}{L}\left[\frac{p\mu_0\epsilon_{ox}}{4\pi\delta} + \frac{(1 - p)\epsilon_{Si}}{2\pi(V_{ox} + V_g)}\left(\frac{2U_0}{m^*}\right)\right]^{\frac{1}{2}}. \tag{18}$$

In Fig. 6 the data of Fig. 3 is replotted with $G_m(L/WV_D)$ as ordinate values and $(V_{ox} + V_g)^{-1}$ as abscissa values in which V_{ox} was chosen so that $(V_{ox} + V_g)$ is zero at the onset of channel conductance. Curve (a) exemplifies the condition where $p = 0$, and indicates that the carriers in this device undergo completely diffuse scattering. Curve (b) exemplifies the condition of partial specular scattering. The deviation from linearity in this plot at large $(V_g + V_{ox})^{-1}$ is thought to be due to a decrease in the specular portion of the surface scattering at low field or to a departure from the approximation of Eq. (2) caused

by a variation of the capacitance associated with the onset of depletion of the space charge region in silicon. In addition, the surface scattering time may not be sufficiently small to allow neglect of the bulk scattering time. We note that in Fig. 6 the slopes of the curves are 5.4×10^{-6} mho and 6.3×10^{-6} mho for curve (a) and (b), respectively. In a nondegenerate electron gas, the one-dimensional thermal energy is $kT/2$. Taking $m^* = 0.25\ m_0$, the theoretical value for slope would be 25.6×10^{-6} mho. As obtained from the plot of ln $(V_{ox} + V_g)$ vs G_{SD}, V_1 is 0.84 volts compared with a theoretical value of 1.22 volts calculated from Eq. (7).

Figure 5 shows μ_{eff} effective mobility as a function of transverse field reduced from G_{SD} data presented in Fig. 4. In Fig. 5, the dashed curves are calculated from Eqs. (15) and (16) which are fitted at two points for each curve. F_1 was calculated from the experimental V_1 and Eq. (7). For the sample for which the data is shown in curve (b) the specular portion of the effective mobility of electrons is 287 cm^2/V sec. If the drift mobility in the bulk is 1400 cm^2/V sec,[14] one finds $p = 0.205$; about 20% of the surface scattering is specular. Curve (c) shows the effective mobility for holes in the inversion layer on an n-type sample. The specular portion in this case is 142 cm^2/V sec. Taking the bulk value to be 500 cm^2/V sec (Ref. 13), then the value of p is found to be 0.28.

Discussion

The effective mobility of the carriers in an inversion layer has been shown in general to be dominated by a combination of specular and diffuse surface scattering. Most of the n-p-n samples examined indicate that the electrons undergo about 15–25% specular scattering. This percentage may vary from device to device depending on the surface conditions. In addition, there are samples in which the mobility data cannot be fitted with a single elastic factor p; that is, p is field dependent. The roughness of the surface compared to the electron wavelength[15] and density of the ionized centers[16] at the surface combine to play an important role in this regard. The fact that values of p for holes are greater than for electrons in this work seems in accord with Coovert's results,[16] in which bulk behavior of the holes in an inversion layer on n-type Si was observed.

The temperature dependence of the effective mobility can be examined in terms of the temperature dependence of the transconductance. For example, for a sample which shows complete diffuse scattering, the slope of curve (a) in Fig. 6 gives $(\epsilon_{Si}/2\pi)(2U_0/m^*)^{\frac{1}{2}}$ which should have the same temperature dependence as the mobility. G_m of such a sample was measured at temperatures from 77°K to 300°K. For each temperature the field dependence of G_m follows Eq. (18). The temperature coefficient $dG_m/d[(V_{ox} + V_g)^{-1}]$ is plotted in Fig. 7. The temperature dependence of the mobility is seen to be quite different from the bulk behavior.[16] The temperature dependence of the surface mobility arises from the energy of the particle U_0 in the above analysis. For a nondegenerate inversion layer $U_0 \approx kT/2$. One would expect a $T^{\frac{1}{2}}$ dependence for the mobility. There is some question as to whether the carriers in the inversion layer are degenerate. For a degenerate layer, U_0 is essentially independent of the temperature. The present simple model does not seem to be adequate to interpret the complex temperature dependence of the surface mobility.

Finally, it should be remarked that although charged trapping states could have a role in the scattering mechanism, no direct evidence of trapping effects on charge kinetics was observed.[8] Measurements of capacitance and transconductance from 20 cps to 10 Mc/sec show an essentially flat response.

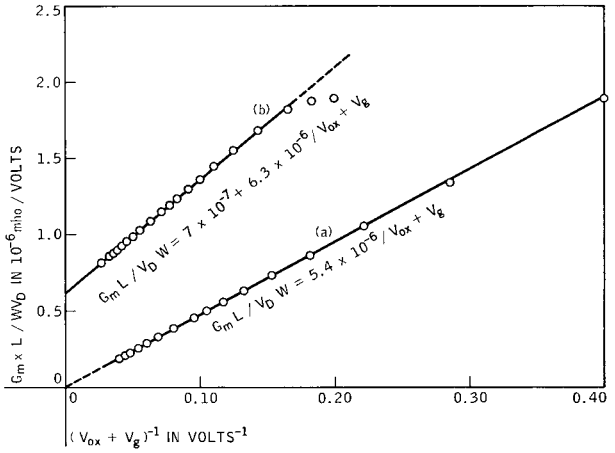
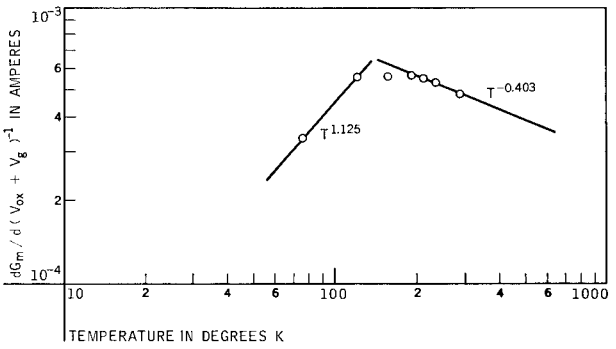


Figure 6 **Curves for the data of Fig. 3 replotted to show the functional dependence of G_m with $(V_{ox} + V_g)$.**

Figure 7 **Temperature dependence of $dG_m/d[(V_{ox} + V_g)^{-1}]$ for a sample with total diffuse scattering.**



References

1. J. R. Schrieffer, *Phys. Rev.* **97**, 641 (1955).
2. J. R. Schrieffer, *Semiconductor Surface Physics*, edited by R. H. Kingston, Univ. of Penn. Press, Philadelphia, 1956, p. 55.
3. K. Fuchs, *Proc. Cambridge Phil Soc.* **34**, 100 (1938).
4. E. H. Sondheimer, *Phys. Rev.* **80**, 401 (1950).
5. R. L. Petritz, *Phys. Rev.* **110**, 1254 (1958).
6. R. F. Greene, D. R. Frankl, and J. Zemel, *Phys. Rev.* **118**, 967 (1960).
7. F. S. Ham and D. C. Mattis, *IBM Journal* **4**, 143 (1960).
8. F. Fang and S. Triebwasser, *Applied Physics Letters* **4**, 145 (1964).
9. See, for example, P. K. Weimer, *Proc. IRE* **50**, 1462 (1962).
10. R. Lindner, *Bell System Tech. J.* **41**, 803 (1962).
11. P. M. Marcus, *IBM Journal.* To be published.
12. A. B. Fowler, F. Fang, and F. Hochberg, *IBM Journal* **8**, 427 (1964).
13. Strictly, Eq. (16) is correct only for the case of spherical energy surfaces; however, it is adequately accurate for the admittedly simple model assumed here and for the small value of p observed.
14. E. M. Conwell, *Proc. IRE* **46**, 1281 (1958).
15. See, for example, J. M. Ziman, *Electrons and Phonons*, Oxford Press, London, 1960, p. 451.
16. R. E. Coovert, *J. Phys. Chem. Solids* **21**, 87 (1961).
17. See, for example, F. J. Blatt, *Solid State Physics* **4**, edited by F. Seitz and P. Turnbull, Academic Press, New York, (1957).

Received June 17, 1964.