corresponding physically to a longer relaxation time within the Debye-Hückel cloud than in the medium.

In conclusion, it should be remarked that no effort has been made to take account of interactions between neighboring clouds, undoubtedly an appreciable factor at high temperatures where the density of imperfections is large.

In retrospect, the suggestion of Lidiard[2] that Debye-Hückel effects must also play an important role in affecting the activation energies (particularly at high temperatures where the dielectric constant of the medium is increasing so rapidly) seems eminently reasonable.

Clearly, although the models discussed earlier seem plausible, more experiments, particularly as a function of frequency, are needed to establish their validity.

ACKNOWLEDGMENTS

The authors are grateful to the Office of Naval Research and the Industrial Sponsors of the University of Chicago for partial support of their experiments.

PHYSICAL REVIEW VOLUME 123, NUMBER 5 SEPTEMBER 1, 1961

Electronic Processes and Excess Currents in Gold-Doped Narrow Silicon Junctions

C. T. SAH
Research and Development Department, Fairchild Semiconductor Corporation, Palo Alto, California
(Received March 30, 1961; revised manuscript received May 3, 1961)

Large amounts of excess current in gold-doped silicon tunnel junctions are observed and interpreted as due to transition processes with the two gold energy levels in the forbidden gap of silicon as intermediate states. Eight of the ten possible processes are two-step processes. These two steps may be both of the Hall-Shockley-Read type or of the type involving electron tunneling between a trap state and a band state within the space charge region of the junction. The two steps may also consist of a Hall-Shockley-Read process as one step, and the tunneling from or to the trap state as the other step. One of the remaining two possible processes is a three-step process involving two Hall-Shockley-Read steps and one tunneling step between two trap states within the space-charge region. The last process is the usual carrier injection process. Eight of the ten processes in the gold-doped tunnel diodes have appreciable transition rates. Five of the eight processes have onset structures which appear at voltages in reasonable agreement with the predicted values. Approximate theoretical current-voltage expressions are compared with experimental data of the gold-induced excess current at 4.2°K, giving an average value of $W^2m_t/m=1.2\times10^{-23}$ volt^2-cm^3, where W is Price's matrix element of the trap potential energy in excess of the crystal potential taken between the unnormalized trap-state wave function and the band edge Bloch wave function, normalized to unit volume, and m_t/m is the transverse electron mass normalized to the free electron mass. It is also experimentally determined that the rate of tunneling from or to trap state is smaller than the rate of filling or emptying the trap in the Hall-Shockley-Read process.

I. INTRODUCTION

SINCE the discovery of the tunneling phenomenon in very narrow germanium p-n junctions by Esaki,[1] there have been many fundamental experiments designed to explore the detailed characteristics of the Esaki, or tunnel current. The negative resistance observed at the low forward bias voltage region (forward bias voltage means positive on the p-type side) between approximately 0.1 to 0.4 v, have been established by Esaki as due to the decrease of the density of states common to the two bands on the opposite side of the p-n junction.[1] In this model, the constant energy electronic transition would cease when the applied forward voltage is high enough so that the conduction and the valence band edges are uncrossed if sharp band edges can be defined. The turn-off voltage, of the order of 0.4 v, is generally so low that the diffusion current would not be important. Consequently, one would expect a region of nearly zero current flow under the forward bias condition between about 0.4 v to a voltage approximately equal to the energy gap of the material. However, the region of nearly zero current has not been observed experimentally for all known tunnel diode materials. In some semiconductors, such as silicon, the excess current is rather high with a magnitude of the order of 1/4 times the peak current, while in other materials such as germanium and gallium arsenide, the excess current may be as low as 1/10 to 1/100 times the peak current. The excess current and some associated structures have been noted by Esaki[2–4] who attributed the observations to band-to-band transition via deep impurity states in the energy gap. The first direct evidence of the excess current induced by deep impurity states was observed by Longo,[5] who created a sufficient amount of deep traps by subjecting germanium tunnel

[1] L. Esaki, Phys. Rev. **109**, 603 (1958).

[2] T. Yajima and L. Esaki, J. Phys. Soc. Japan **13**, 1281 (1958).
[3] L. Esaki, *Solid-State Physics in Electronics and Telecommunications* (Academic Press, Inc., New York, 1960), Vol. 1, pp. 514–523.
[4] L. Esaki and Y. Miyahara, Solid State Electronics **1**, 13 (1960).
[5] T. A. Longo, Bull. Am. Phys. Soc. **5**, 160 (1960).

diodes to electron irradiation. Subsequently, excess current induced by electron irradiation have been studied in silicon junctions by Logan and Chynoweth[6] and by Classen in germanium, silicon, and gallium arsenide.[7] The evidence of excess current associated with the impurity states or bands in the energy gap of GaAs and InP has recently been presented by Holonyak.[8] The investigation of the excess current, using impurity metals which have deep energy levels in the band gap and preliminary experimental findings were discussed by Sah and Tremere.[9]

A detailed model of the electron bombardment induced excess current is rather difficult to construct, since the natures of the deep traps created are not well established. In the case of excess currents associated with impurity metals, it is possible to interpret the results by means of the known energy levels and the charge conditions of the traps. Sah[9] has discussed a model of excess current in gold-doped silicon tunnel diodes based on field ionization of the electron in the trap which is subsequently refilled by an electron in the conduction band. Chynoweth[10] has considered a similar model for the electron irradiation induced traps, and more recently Price[11] has calculated the transition probability of this process and has predicted significantly short lifetime. The importance of the band-edge tail states in the very heavily doped and disordered materials on the excess current have been pointed out.[12,13] These states are presumably more or less continuously distributed and extended into the energy gap from the band edge. A detailed theoretical analysis of the excess current taking into account of the band edge tails has not been considered until recently.[14] However, it is not established if the heavily doped tunnel diode materials do have a random distribution of donor and acceptor impurities which give rise to the band edge tails.

This paper will be concerned with a detailed presentation and interpretation of the experimentally observed excess currents in gold-doped silicon tunnel diodes. A number of interesting electronic transition phenomena are discussed. Some matrix elements associated with these transitions are estimated by comparing theoretical formulas with experimental data, using the well-known energy levels and the solid solubility of gold in silicon.

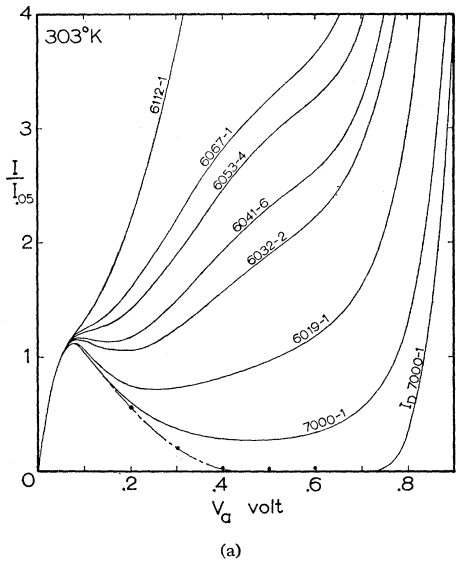
II. EXPERIMENTAL RESULTS

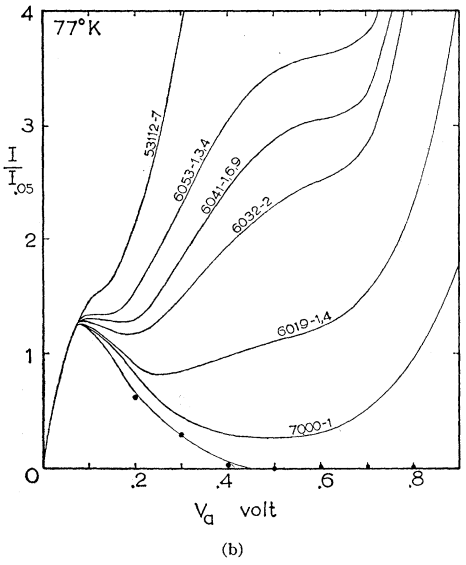
The tunnel diodes used in this study are made on phosphorus-doped and gold-diffused silicon of a few milliohm-cm resistivity materials. The gold concentration is controlled by the maximum solid solubility at the diffusion temperature. Narrow p-n junctions are made by alloying aluminum-boron alloy into the n-type silicon wafers. The boron concentration in the regrowth layer is around 6×10^{20} atoms/cc, and the phosphorus concentration of the majority of diodes for which the data are presented in this report is 6×10^{19} atoms/cc.

A. Dependence of Excess Current on Gold Concentration

Figure 1(a) shows current-voltage characteristics taken at 300°K for diodes doped with various gold concentrations which are indicated by the second, third, and the fourth figures in the numerals above each curve. For example, 6067–1 is a diode made with 6.7×10^{16} gold atoms/cc diffused into the n-type wafer. The gold concentration is assumed to be equal to the maximum solid solubility at the diffusion temperature.[15] The effect of heavy doping on the solid solubility of gold in silicon is relatively unimportant in the gold diffusion temperature range of 1000° to 1200°C,[16] and is not taken into account. The current-voltage data are taken at temperatures from 4.2° to 550°K, and some of the data are shown in Figs. 1(b) and (c). In these figures, it is clear that a large amount of excess current is produced due to gold. The results are extrapolated back to zero gold concentration as shown in Figs. 2(a) and (b) and are shown as dots in Figs. 1(a)–(c). By this extrapolation, it is seen from Figs. 1(a)–(c) that almost zero excess current is expected if the gold concentration is zero. However, due to the large experimental inaccuracy in the extrapolation, it cannot be certain if all the excess currents are due to gold. For example, the unit 7000–1 has no intentionally added gold in the n-type wafer, although there is considerable excess current as shown in Fig. 1(a) where the extrapolated diffusion current I_D for this unit is also shown. The originally present excess current may be due to the presence of some unintentionally added gold during the crystal growth or the alloying process. Other impurities or traps and band-edge tail or gap states would also give excess current. Since we can subtract this excess current

[6] R. A. Logan and A. G. Chynoweth, Bull. Am. Phys. Soc. **5**, 374 (1960).
[7] R. S. Classen, Bull. Am. Phys. Soc. **5**, 406 (1960).
[8] Nick Holonyak, Jr., J. Appl. Phys. **32**, 130 (1961); also presented at the 1960 IRE-AIEE Solid State Device Research Conference, June 13, 1960 (unpublished).
[9] C. T. Sah and D. A. Tremere, AIEE-IRE Solid State Device Research Conference, June 13, 1960 (unpublished); C. T. Sah, Bull. Am. Phys. Soc. **5**, 507 (1960); Symposium on Electronic Tunneling Phenomenon in Solids, January 29–30, 1961 (unpublished); Bull. Am. Phys. Soc. **6**, 105 (1961).
[10] A. G. Chynoweth, Fourth International Conference on Semiconductors, Prague, August, 1960 (unpublished); A. G. Chynoweth, W. L. Feldman, and R. A. Logan, Phys. Rev. **121**, 684 (1961). A review, up to 1958, was given by A. G. Chynoweth in *Progress in Semiconductors* (John Wiley & Sons, Inc., New York, 1960), Vol. 4.
[11] P. J. Price, Bull. Am. Phys. Soc. **5**, 406 (1960).
[12] R. N. Hall, General Electric Research Laboratory Report No. 60-RL-2509*G*, New York, September, 1960 (unpublished); also presented at the Fourth International Conference on Semiconductors, Prague, August, 1960.
[13] E. O. Kane, Bull. Am. Phys. Soc. **5**, 160 (1960).
[14] The band edge tails have been discussed recently by J. J. Klauder, H. Schlosser, T. Brody, and P. Aigrain at the Symposium on Electronic Tunneling in Solids, January 30–31, 1961 (unpublished).
[15] F. A. Trimbore, Bell System Tech. J. **39**, 205 (1960).
[16] C. T. Sah (unpublished).





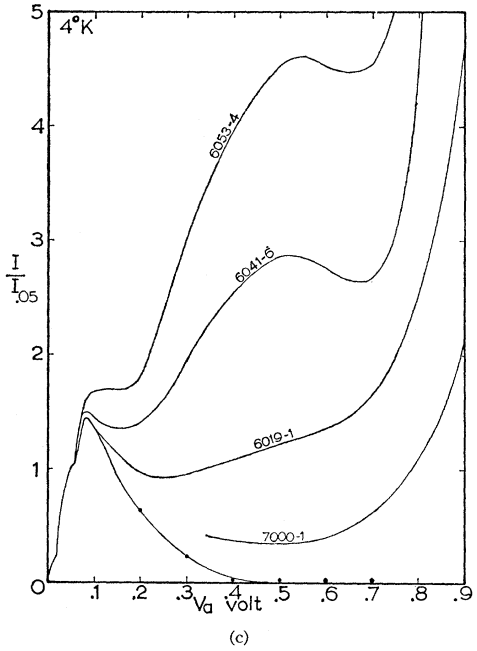


FIG. 1. Experimental current-voltage characteristics of gold-doped tunnel diodes taken from x–y recorder. Junction temperature (a) 303°K, (b) 77°K, and (c) 4.2°K. The last three numerals before the dash above each curve give the gold concentration in 10^{15} gold atoms/cc. For example: 6041–6 has 4.1×10^{16} gold atoms/cc.

component out of our experimental data, we will not be concerned with this current in our discussion.

The plot of the current as a function of gold concentration in Figs. 2(a) and (b) shows that the excess current induced by gold is a linear function of the gold concentration at a given bias voltage. It is particularly interesting to note that at 4.2°K a second negative resistance region appears in the current-voltage characteristics at around 0.7 v, as shown in Fig. 1(c). A replot of the difference of the two curves with different gold concentration is shown in Fig. 3. It is seen that a small negative resistance region again appears at around 0.7 v even at 300°K, although it may be fictitious due to the large error involved in subtracting the two diode characteristics. It will be shown in the next section that this negative resistance region may be associated with the turning-off of one of the transition processes through one of the gold levels at low temperatures.

In addition to the 4°K data, the slopes of the straight lines of current versus gold concentration obtained from Figs. 2(a) and (b) are also shown in Fig. 4. The 77°K and the 4°K curves are almost identical and show negative resistance at around 0.7 v. The negative resistance region at 300°K, which is quite evident in Fig. 3, is not visible in Fig. 4 due to the fairly large experimental spread of the data shown in Fig. 2(a). The normalization of these characteristics to a current at 0.05 v ($I_{0.05}$) is necessary to eliminate the uncertainty

of the active junction area from diode to diode. This normalization procedure is quite legitimate as can be seen from Figs. 5 and 6. The temperature dependence of the gold-induced excess current and the phonon-assisted indirect band-to-band tunnel current are expected to differ; however, Fig. 5 shows that the shape of the characteristics of current versus temperature at 0.05-v forward bias is very insensitive to the gold concentration. Thus it may be assumed that at this bias level the gold-induced excess current is negligible compared with the phonon-assisted current and that $I_{0.05}$ can be used for normalization to eliminate the junction area. This procedure is evidently invalid if a current at a higher bias voltage is used. For example, at 0.4 v the current-temperature characteristics cannot be normalized to one single characteristic, as shown in Fig. 6. This is due to the fact that the gold-induced current at 0.4 v is comparable to and probably greater than the phonon-assisted band-to-band transition current.

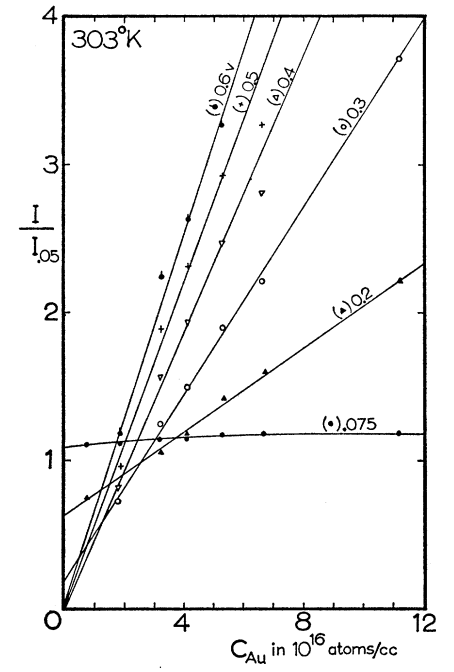


FIG. 2. The gold-induced excess current normalized to the current at 0.05 v, $I_{0.05}$, plotted as a function of gold concentration at constant voltages. (a) Junction temperature 303°K, (b) 77°K.

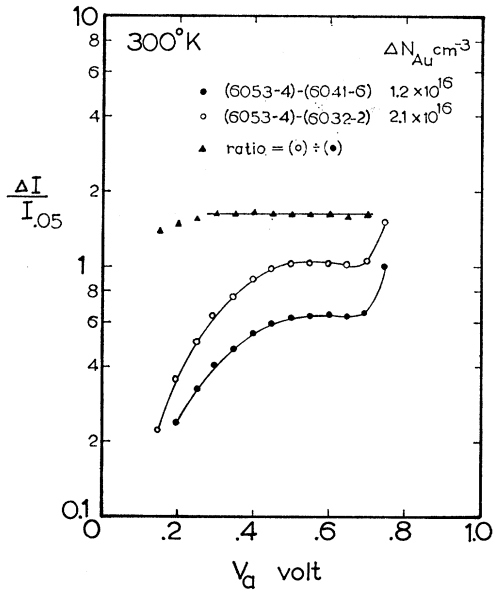


FIG. 3. The difference of current of two diodes doped with different gold concentration versus voltage at 300°K.

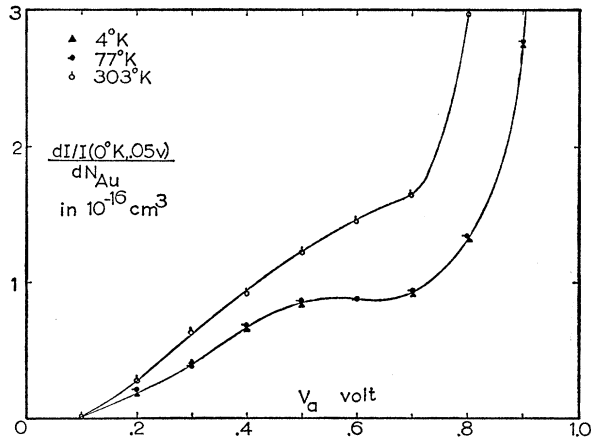


FIG. 4. The slope of the excess current-gold concentration curves versus voltage.

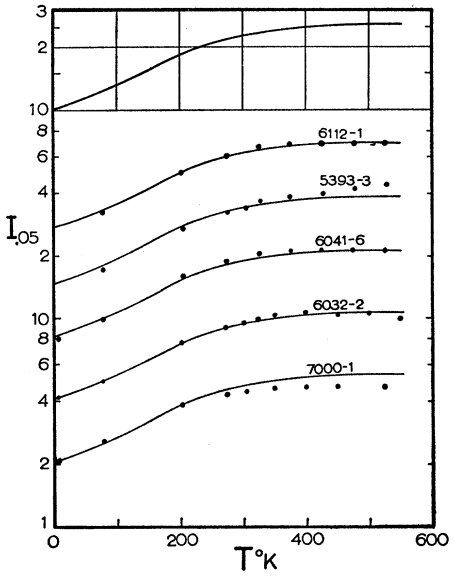


FIG. 5. The variation of the current at 0.05 v, $I_{0.05}$, with temperature for junctions doped with various gold concentrations.

To get an estimate of the active junction area of these gold-doped tunnel diodes, independent experiments have been run in which the geometric junction areas are measured visually. The experimental results of the peak current density at around 0.08 v versus the impurity concentration are shown in Fig. 7. These data, in addition to providing a means of checking the junction area, may also be used to determine an effective mass for the tunneling electrons and holes. The numerical value of the effective mass derived from the slope of the plot shown in Fig. 7 would depend on which particular tunnel formula and what kind of an average electric field is used, since in alloyed junctions the electric field is not a constant, but is a linear function of position in the space-charge region of the junction. The numerical values can range from $0.03m$ to $0.25m$, where m is the free electron mass. The range covers the light-hole mass value, $0.16m$, and the transverse electron mass, $0.19m$,[17,18] in silicon.

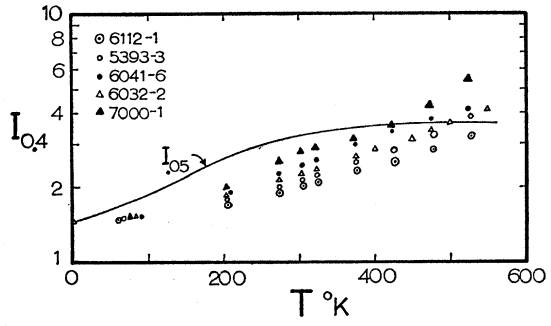


FIG. 6. The variation of the current at 0.40 v, $I_{0.40}$, with temperature for junctions doped with various gold concentrations.

[17] E. M. Conwell, Proc. Inst. Radio Engrs. **46**, 1281 (1958).
[18] F. Herman, Proc. Inst. Radio Engrs. **43**, 1703 (1955).

B. Structures or Abrupt Changes of Conductance

Very rapid change of conductance has been observed at voltages corresponding to the onset of the phonon-assisted band-to-band tunneling transitions in silicon.[12,19] These structures are evident in Figs. 8(a) and (b) for a gold-doped tunnel diode at 4.2°K, and are labeled 1, 2, 3, 4, and 4f. These phonon-assisted structures are independent of the presence of gold and are more clearly observed in lightly doped materials as shown in Fig. 9. The observed stair-case shape of the current-voltage characteristics in the low-voltage region is in accord with the theory of Keldysh.[20]

In addition to these phonon-assisted onset structures, structures associated with or without the presence of

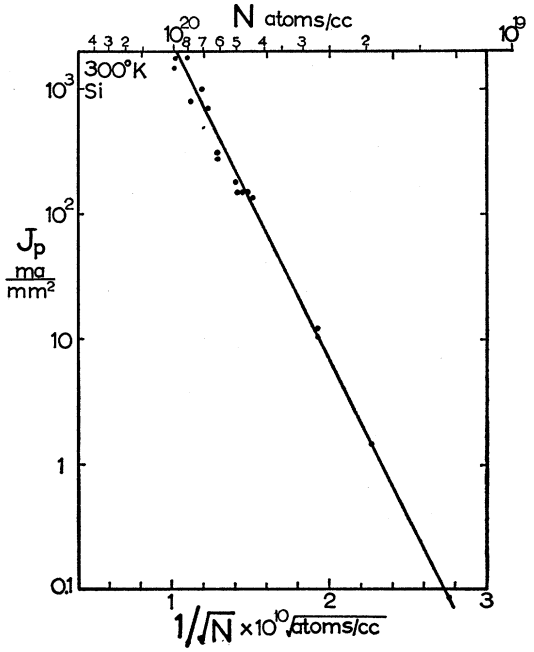


FIG. 7. The peak current density at approximately 0.075 v versus the bulk impurity concentration (phosphorus) N at 300°K for tunnel junctions with no gold doping.

gold are also observed at higher voltages. In Fig. 8(b) these structures are labeled 6^-(400 mv), 6^+(730 mv), $7^{\pm}$(930 mv), and 8(1080 mv). These structures become more evident at low temperatures and low phosphorus concentrations, and are shown for a diode with 1.1×10^{19} phosphorus/cc, and 5.3×10^{16} gold/cc in Fig. 10. The conductance curve in Fig. 10 clearly shows another minimum labeled 5^- at 0.5 v, although it is not clear whether it corresponds to a true onset voltage, or whether it appears due to the combination of several transition processes contributing to excess currents.

[19] N. Holonyak, Jr., I. A. Lesk, R. N. Hall, J. J. Teimann, and H. Ehrenreich, Phys. Rev. Letters **3**, 167 (1959).
[20] L. V. Keldysh, J. Exptl. Theoret. Phys. (U.S.S.R.) **34**, 962 (1958) [translation: Soviet Phys.—JETP **34(7)**, 665–669 (1958)].

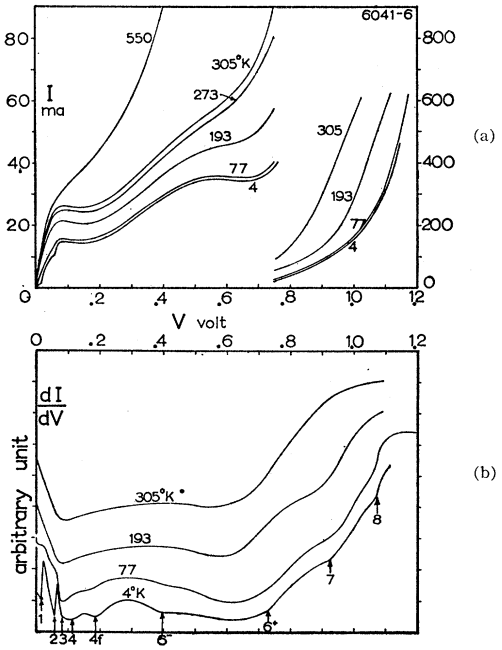


FIG. 8. (a) The current-voltage characteristics of a gold-doped tunnel junction at various temperatures; (b) the conductance-voltage characteristics of a gold-doped tunnel junction at various temperatures.

More extensive data, involving several hundred gold-diffused, phosphorus- or arsenic-doped silicon tunnel diodes of majority impurity concentration covering the range, 1 to 14×10^{19} atoms/cc, of which Figs. 8(a), 8(b), and 10 are samples, show that the structures which occur at 5^-, 6^-, 6^+, and $7^\pm$ are present only in gold-doped tunnel diodes, while structure 8 occurs for practically all diodes made with or without gold or with other impurity metals. In many of these diodes, doped with impurities, a structure near the onset voltage of $7^\pm$ also appears, although structures at 5^-, 6^- and 6^+ are not detected. Additional breaks in the current voltage characteristics of gold-doped diodes appear to be present, but are much less clearly brought out. In Fig. 11, the current is plotted on a semilog scale showing a low-voltage structure labeled 5^+ at around 0.25 v, and several other doubtful ones near and below the structure at 6^+.

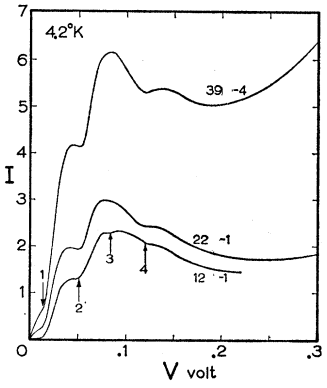


FIG. 9. The current-voltage characteristics of tunnel junctions with no gold doping at 4.2°K and low voltage.

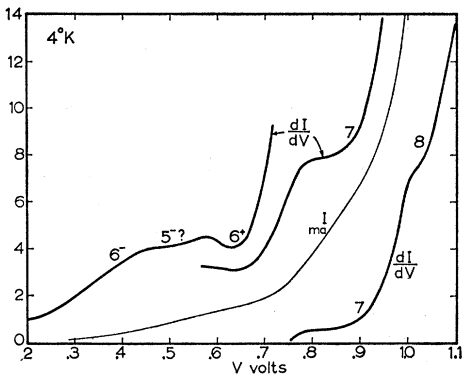


FIG. 10. The current and conductance-voltage characteristics of a tunnel junction of 1.1×10^{19} phosphorus atoms/cc with 5.3×10^{16} gold atoms/cc in the bulk at 4.2°K.

III. INTERPRETATION OF DATA

The experimentally observed structures and the current-voltage characteristics will be qualitatively and quantitatively explained by considering transition processes involving Hall-Shockley-Read processes[21,22] in combination with tunneling transitions or field ionization processes from one of the three possible gold charge states.

A. Energy Band Scheme

The energy band in silicon at 0°K under thermal equilibrium may be constructed from published data[18,23,24] and is shown for an idealized and constant electric field case in Fig. 12 near and including a *p-n*

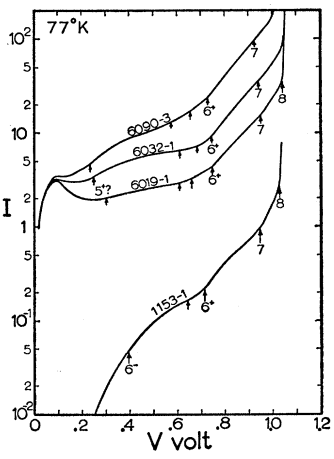


FIG. 11. The semilog plot of the current-voltage characteristics of gold-doped tunnel junctions at 77°K.

[21] R. N. Hall, Phys. Rev. **83**, 228 (1951); **87**, 387 (1952).
[22] W. Shockley and W. T. Read, Jr., Phys. Rev. **87**, 835 (1952).
[23] G. G. Macfarlane, T. P. McLean, J. E. Quarrington, and V. Roberts, Phys. Rev. **111**, 1245 (1958).
[24] H. Brooks, *Advances in Electronics and Electron Physics*, edited by L. Marton (Academic Press, Inc., New York, 1955).

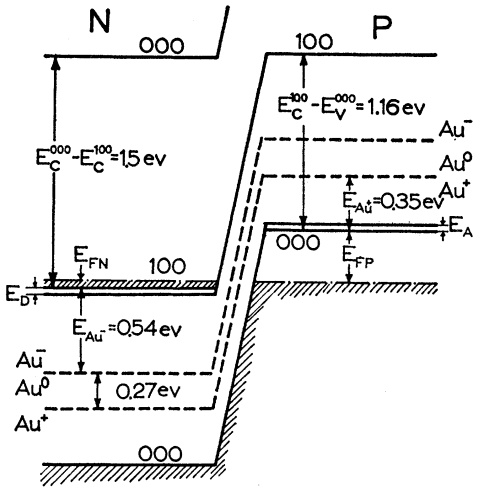


FIG. 12. The energy band diagram of a gold-doped silicon tunnel junction at 0°K.

junction. The cross-hatched regions represent filled electron states, while the empty regions outside of the band gap represent empty electronic states. The Fermi levels in the n and the p sides are inside the bands due to heavy impurity doping. These Fermi energies can be approximately calculated using degenerate Fermi statistics. In Fig. 13, the Fermi energies are plotted for several approximations. The variations of the major impurity level with its concentration included in the calculation for the Fermi levels are estimated from the data of Pearson and Bardeen[25] and from a theoretical analysis by Castellan and Seitz.[26] Experimental data on ϵ_v, the Fermi level in the p side, are also plotted and are obtained from the valley voltages of a large number of silicon tunnel diodes of different and known boron concentration in the aluminum-boron-silicon regrowth

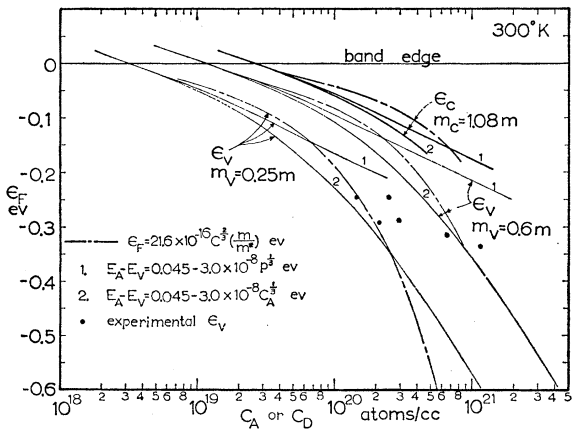


FIG. 13. The Fermi energies versus acceptor or donor concentrations at 300°K.

[25] G. L. Pearson and J. Bardeen, Phys. Rev. **75**, 865 (1949).
[26] G. W. Castellan and F. Seitz, *Semiconducting Materials* (Butterworths Scientific Publications, Ltd., London, 1951).

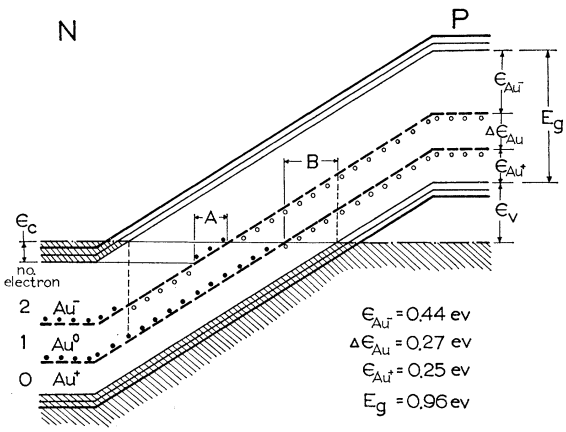


FIG. 14. The modified energy band diagram of a gold-doped silicon tunnel junction at 0°K, taking into account impurity banding.

layer from spectrographic analysis of the boron-aluminum alloy and by assuming unity distribution coefficient of boron in silicon. These experimental data appear to give the right trend compared with that predicted for the valley voltage based on the Esaki model.

The variation of the major impurity level with its concentration comes from impurity banding at high concentrations. An estimate made from published analysis[27,28] shows that for the silicon materials used here, the impurity levels, E_D and E_A are widened and merged with the conduction and the valence band. These conclusions can also be obtained from the experimental data of Pearson and Bardeen.[25] Furthermore, the analytical results of Baltensperger,[27] which assume uniform or non-random impurity distribution, show that the impurity bands also extend away from the respective band edges approximately symmetrically with respect to the impurity energy level at low doping with a sharp band edge. From this discussion, we may construct a slightly revised band picture shown in Fig. 14, in which the effective energy band gap is now given by $E_g=E_C^{100}-E_V^{000}-2E_D-2E_A=1.16-0.1-0.1=0.96$ ev. Although this is a rather rough estimate, it appears to be sufficient to explain the experimental data of gold-induced excess currents. In particular, there are probably less densely packed states extending far into the band gap from the position of the sharp band edges.[14]

The gold energy levels are experimentally determined and measured from the band edges.[17] In essence, gold atoms can exist in silicon in one and only one of the three charge conditions: +, 0, or −. The energy associated with the capture of a conduction electron by a neutral gold atom is 0.54 ev measured from the conduction band edge. Similarly, the emission of an electron from a neutral gold atom to the valence band involves 0.35 ev. Taking into account the banding of the

[27] W. Baltensperger, Phil. Mag. **44**, 1355 (1953).
[28] E. M. Conwell, Phys. Rev. **103**, 51 (1956).

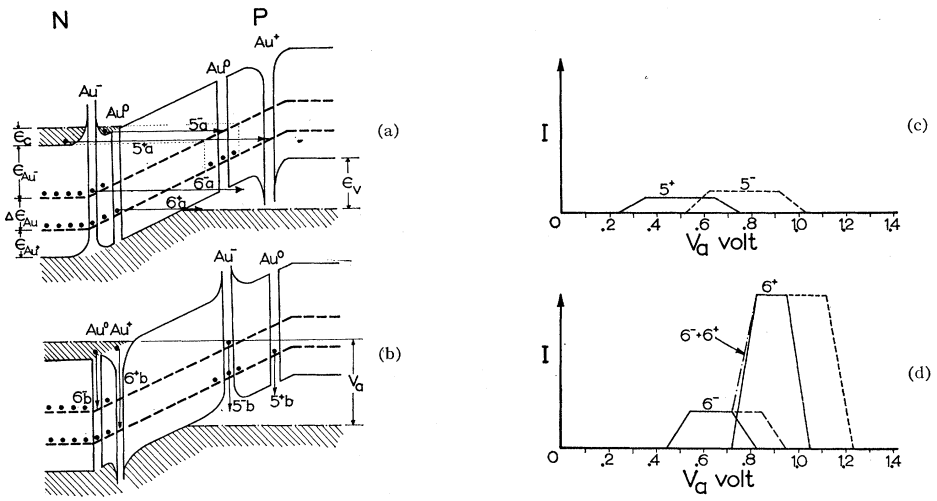


FIG. 15. The transition processes with gold impurity levels as intermediate step, processes 5^-, 5^+, 6^-, and 6^+. (a) The first step involving electrons tunneling to or from the gold impurity levels or trap states; (b) the second step involving refilling or emptying of the trap states or gold impurity levels according to the Hall-Shockley-Read process; the idealized trapezoidal current-voltage characteristics assuming that the Hall-Shockley-Read process is not rate limiting: (c) processes 5^+ and 5^-; (d) processes 6^+ and 6^-.

majority impurity levels, these figures become, respectively, 0.44 ev for the acceptor level and 0.25 ev for the donor level. The energy spacing between the two gold levels is assumed unchanged and is 0.27 ev. These energy levels are drawn to scale in Fig. 14.

The electron occupation conditions on the gold atoms at 0°K in and near the space charge region of the *p-n* junction is shown in Fig. 14. Ouside of the two vertical dashed lines, the gold atoms are either filled with two electrons (*n*-type side) or filled with two holes (*p*-type side). Circles are drawn below the levels for the unfilled gold levels in this diagram to indicate that one may also consider these empty levels as filled with holes. In some of the subsequent energy band diagrams, these circles are not shown for simplicity, following the more customary procedures. However, it should be emphasized that the second gold level, the acceptor level of E_{Au^-}, is not an excited state of the gold donor level in the usual sense, in that the acceptor level represents the energy involved for emission or capture of a second electron and is meaningful or in existence only when there is already one electron trapped by the gold atom.

The electron occupation conditions on the gold levels in the region between the two vertical dashed lines in the space charge layer shown in Fig. 14 are more difficult to determine. At equilibrium, the situation shown in Fig. 14 would finally be reached by electron tunneling horizontally from the conduction band to the trap states in region *A* and by electron tunneling from the trap states in region *B* horizontally to the valence band.

When the gold atom has trapped one electron, it is in the neutral charge condition and has a rather localized and nearly square potential well as shown under Au^0 in Fig. 15.[24,29] However, when the gold atom is filled with two electrons or is completely empty, there is long-range Coulomb interaction giving potential barriers which are labeled Au^- or Au^+ in Figs. 15(a) and (b). There are probably a large number of excited states associated with the Coulomb well of the charged gold atoms and these are not shown in Figs. 15(a) and (b) for the sake of simplicity.

B. Transition Processes

In addition to the band-to-band phonon-assisted tunneling,[12,19,20] there are ten possible transition processes associated with gold levels in gold-doped tunnel diodes in silicon. Of the ten processes, nine are associated with the presence of gold within the space charge region, while the tenth one is the usual diffusion or injection current. Two of these have no onset voltage and will give excess current starting from zero bias. In the silicon tunnel diodes studied in this paper, one of the transition processes cannot occur since the Fermi energy on the *n* side is smaller than the energy spacing of the gold levels. These processes are labeled numerically from 5 to 10 with superscript − or +, indicating whether it is associated with the gold acceptor level or the gold donor level. These processes are considered in detail in the following discussions.

(*1*). *Process* 5^-, 5^+, 6^-, *and* 6^+

These processes are similar and are two-step processes. In the first step, step *a*, a trapped electron or hole in a gold atom located in the space charge region makes a tunneling transition horizontally to one of the bands. In the second step, step *b*, the gold trap is refilled by an electron or a hole at the same location from the other band. These two steps may occur in the reverse order. The four possible processes involving these two steps

[29] Melvin Lax, J. Phys. Chem. Solids **8**, 66 (1959); Phys. Rev. **119**, 1502 (1960).

are shown in Figs. 15(a) and (b). The energy band scheme corresponds to the condition of a fairly large forward applied voltage, V_a. Let us take 6^+ shown in Fig. 15 as an example. The first step of process 6^+, labeled 6^+a, involves the field ionization or tunneling of the trapped electron from a neutral gold atom located in the space charge region to the valence band horizontally across the junction barrier. In the second step, 6^+b, shown in Fig. 15(b), the gold atom which is now positively charged is rapidly refilled by an electron at the same location from the valence band.

The transition rate (or transition probability per unit time) of 6^+a has been calculated by Price.[11] (See Appendix I.) The exact rate expression of 6^+b, which is the Hall-Shockley-Read mechanism, is not available. The rate expression used by Shockley involves some arbitrary assumptions[29] and applies only to the low electric field region. The use of this rate expression to our case is made but not justified. The Hall-Shockley-Read transition is also calculated using an effective matrix element in the usual first order formula of the transition probability per unit time. The refilling rate of the traps by the majority carriers is rather short, since the majority carrier concentration is rather high in these heavily doped tunnel diode materials. Although the more exact rate expression for the Hall-Shockley-Read process is not available, it is still possible to perform a detailed balance calculation in order to determine an approximate shape for the current voltage characteristics if the Hall-Shockley-Read process is not the rate limiting process. In Appendix II, detailed balance calculations have been carried out for processes 5^+, 6^-, and 6^+, and the result will be summarized.

In the gold-doped tunnel diodes studied in this paper, process 5^- does not occur theoretically. By inspection of Fig. 15(a), which represents nearly to scale the energy diagram of the tunnel diodes used in this study, it is clear that the absence of process 5^- is due to the condition that $\epsilon_c < \Delta\epsilon_{Au}$. Because of this situation, the gold level on the p side can be filled by only one electron which comes from the filled conduction band states on the n side by tunneling. This is the transition which is labeled 5^+a.

To obtain the steady-state transition rate of process 5^+, a detailed balance calculation of the two transition steps involved is required. Let the rate of the first step, 5^+a, which involves the tunneling of a conduction electron, be $w_t(5^+a)$, and the rate of emission of the trapped electron to the valence band be $w_e(5^+b)$, then the steady-state transition rate for process 5^+ is (Appendix II):

$$w = w_t w_e/(w_t + w_e). \tag{1}$$

The general relation for the current density at or near 0°K may be obtained from

$$J = \int qwN_t dx, \tag{2}$$

where N_t is the total gold concentration per unit volume, and q is the magnitude of the electronic charge. The integral covers a range common to both the filled states in the conduction band and the empty states in the valence band through a 90° reflection at the acceptor gold level. In the constant electric field case, it is more convenient to transform the integration over distance to that over energy by using

$$dx = d\epsilon_1/F,$$

where F is the electric field times the electronic charge.

The expression (2) can be written into a product of two factors[30–32] (see Appendix III):

$$J = J_0 D, \tag{3}$$

where J_0 includes a tunneling attenuation factor which is exponentially dependent on F^{-1}, and D is the density-of-state integral which comes from normalization of the integral of (2). For the constant-field case, the main voltage dependence of the tunneling current given in (3) comes from the density of state integral, D. At low field, when the perpendicular energies are smaller than the Fermi energies, D has a trapezoidal shape as shown in Fig. 15(c) labeled 5^+, if the Hall-Shockley-Read process is not rate limiting. The flat portion of the trapezoid has a length $\epsilon_v - \epsilon_c$, the current decreasing or increasing regions have a width ϵ_c each, and the total base width of the trapezoid is $\epsilon_v + \epsilon_c$. The onset at $\epsilon_{Au}{}^+$, in general, has a discontinuity of slope or conductance at 0°K, while the other three corners of the trapezoid at voltages $\epsilon_{Au}{}^+ + \epsilon_c$, $\epsilon_{Au}{}^+ + \epsilon_v$, and $\epsilon_{Au}{}^+ + \epsilon_c + \epsilon_v$ are rounded off when the finite perpendicular energies are taken into account.

In processes 6^- and 6^+, due to the condition that $\Delta\epsilon_{Au} < \epsilon_v$, the two gold levels at a given location may simultaneously see empty states across the barrier in the valence band. Thus, at the forward bias voltage when both 6^- and 6^+ can proceed on the gold atoms at a given x plane in the junction, these gold atoms will in general be distributed in all three charge conditions. At a given location on the x axis, the fraction of the gold atoms which is in one of the three charge conditions is determined by the relative magnitude of the four transition rates, $w_t(6^-a)$, $w_e(6^-b)$, $w_t(6^+a)$, and $w_e(6^+b)$ through the coupling of processes 6^-b and 6^+b by the gold atoms. The steady-state transition rates for 6^- and 6^+ are more complicated than that of 5^+ given by (1) due to the coupling, and have different forms in different ranges of applied voltage. The dependence on the applied voltage comes about since 6^- would start

[30] P. J. Price and J. M. Radcliff, IBM J. Research Develop. **3**, 364 (1959).

[31] E. O. Kane, AIEE Tunnel Diode Symposium Record, February 17, 1960 (unpublished), pp. 1–15; also published in J. Appl. Phys. **32**, 83 (1961).

[32] R. N. Hall, unpublished notes of lectures given at the Electrical Engineering Department, Stanford University, February, 1960.

and turn off at lower applied voltages than 6^+, so that there are applied voltage ranges where only one of 6^- and 6^+ can occur. In this range, the relation given by (1) for the steady-state rate is applicable. In the range where both 6^- and 6^+ can occur simultaneously, the steady-state rates are given by (see Appendix III):

$$w(6^-)=w_e(6^-b)/\{1+[w_t(6^+a)/w_e(6^+b)]+[w_e(6^-b)/w_t(6^-a)]\}, \quad (4)$$

$$w(6^+)=w_t(6^+a)/\{1+[w_t(6^+a)/w_e(6^+b)]+[w_e(6^-b)/w_t(6^-a)]\}. \quad (5)$$

The rate expressions given by (4) and (5) may be substituted into (2) for w to get the current density. The results can again be written in the form given by (3). The density-of-state integral, D, has the low-field forms shown in Fig. 15(d) which are labeled 6^- and 6^+ and are constructed by assuming

$$w_t(6^-a)=w_e(6^-b)\ll w_t(6^+a)=w_e(6^-b),$$

although this assumption, which assumes the equality of the rate expression of the Hall-Shockley-Read process and the tunneling process from traps, is generally not valid. The dashed portion of the trapezoid for process 6^- shown in Fig. 15(d) represents the I–V plot of the process 6^- if 6^+ does not exist or if the rate of 6^+ were considerably smaller than 6^-. It is interesting to note that under the assumptions stated, the combined current-voltage plot of 6^- and 6^+ would not show negative resistance for process 6^-, although the onset of 6^+ at a voltage $E_g=\epsilon_{Au}{}^+$ would still give a distinctive structure or abrupt change of conductance.

It is possible to obtain a qualitative estimate of the relative magnitude of the processes 5^+, 6^-, and 6^+ by considering the energy band diagram shown in Figs. 15(a) and (b). It is evident from these two figures that process 6^+ has the highest probability rate since it involves both the shortest tunneling forbidden path [Fig. 15(a), 6^+a], the light hole mass, and the attractive Coulomb barrier [Fig. 15(b), 6^+b] for electron capture by the positively charged trap. Similarly, process 5^+ has the smallest probability rate of the three processes since it involves the longest tunneling forbidden path [Fig. 15(a), 5^+a] and an approximate square barrier for hole capture by the neutral trap [Fig. 15(b), 5^+b]. The Coulomb interaction in step 5^+a and the parabolic potential in the true step junction would reduce the barrier height somewhat and increase the tunneling rate of 5^+a. An approximate correction to the tunneling rate from traps due to Coulomb barrier is given in Appendix I. Furthermore, electrons from the conduction band may also tunnel to the excited levels of the Coulomb well in step 5^+a and would considerably increase the total transition rate of 5^+a at the higher voltage end of the trapezoid. A similar situation applies to process 6^-a, which also involves Coulomb barrier. These qualitative conclusions are kept in mind in constructing the current-voltage characteristics in Figs. 15(c) and (d), except that the modification due to the excited states of the Coulomb well is not taken into account.

In addition to the decrease of the voltage span of process 6^- due to the interaction with process 6^+ on the same gold atoms, there are other processes which may occur with appreciable transition probabilities in a portion of the voltage range where 5^+, 6^-, or 6^+ transition is proceeding. Considerable coupling would be expected. The modification of the current-voltage characteristics due to these other processes will be discussed in the following sections.

(2). *Processes 7^- and 7^+*

These two processes are the Hall-Shockley-Read processes in which both of the two transitions, $7^{\pm}a$ and $7^{\pm}b$ are trap filling processes. These are the recombination processes in the space charge region for which the expression of the current density has been derived previously based on the Hall-Shockley-Read statistics at high temperatures.[33] At low temperatures and in heavily doped materials such as that for tunnel diodes, the Fermi levels are inside the energy bands and these recombination processes would have a definite onset voltage which should be independent of the type of the deep traps or their energy levels if these traps do not affect the energy band structure. The two steps of process 7^- are shown in Fig. 16. In 7^-a, an electron from the conduction band drops into a neutral gold trap. In 7^-b, a hole from the valence band drops into the negatively charged gold trap and recombines with the trapped electrons. In view of the Coulomb barrier in step 7^-b, 7^-a should be the limiting step. It is evident from Fig. 16 that near 0°K both 7^- and 7^+ would start when the two vertical dashed lines coincide, or when the applied voltage is greater than or equal to the energy of the forbidden band gap. It is possible to get an explicit expression for the current density if it is assumed that the Hall-Shockley-Read statistics are applicable or an effective first order matrix element may be assumed in the "golden rule" expression for the transition probability per unit time. Detailed calculations are given in Appendix IV. Under the simplified assumption that only one of the gold energy levels is

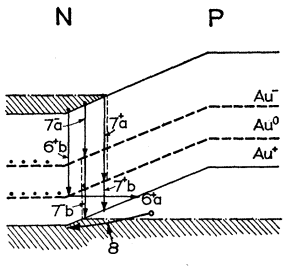


Fig. 16. The transition processes 7^- and 7^+ in gold-doped silicon tunnel junctions.

[33] C. T. Sah, R. N. Noyce, and W. Shockley, Proc. Inst. Radio Engrs. 45, 1228 (1957).

important in providing the recombination traffic and the electron and hole capture rates are equal, a simple result can be obtained near the onset voltage. The current density may again be split into two factors in the form given by (3). The density-of-state factor near the onset is given by (H-S-R approximation)

$$D=\frac{\epsilon_v^{\frac{3}{2}}}{\epsilon_v^{\frac{3}{2}}+\epsilon_c^{\frac{3}{2}}}\frac{V_c-E_g}{\epsilon_c} \tag{6}$$

for $V_a \geq E_g$. An estimate of the cross sections, the matrix elements, and the transition probability rate may be obtained for the Hall-Shockley-Read and the first order matrix approximation by evaluating the exact result of D at $V_a=E_g+\epsilon_c$ from the experimentally observed current density at this applied voltage.

It is evident from Fig. 16 that 7^- and 7^+ would be in competition with and coupled to process 6^+. If 7^- and 7^+ have considerably higher probabilities than 6^+, 6^+ would start to turn off at the onset voltage of process 7^- and 7^+, $V_a=E_g$. On the other hand, if 7^- and 7^+ have smaller probabilities than 6^+, one would expect 6^+ to extend to its normal turn-off voltage which is shown as dashed lines in Fig. 15(d), and a negative-resistance region would occur when 6^+ starts to turn off. An inspection of the experimental data, however, would not reveal a definite conclusion as to which one of the processes, 6^+ or $7^\pm$ has higher transition rate. The difficulty lies in another current, the normal diffusion or injection current which starts to flow before 6^+ starts to turn off if it turns off at the higher turnoff voltage and completely obscures the negative resistance region if it exists. The distinct and rather strong structure observed experimentally, which corresponds to the onset voltage of 7^- and 7^+, seems to favor a situation in which 7^- and 7^+ are more probable than 6^+. A numerical calculation from the experimental data in a following section shows that the rate of $7^\pm$ is probably three times higher than that of 6^+.

(3). *Process 8*

This process is also indicated in Fig. 16 and gives the usual diode injection or diffusion current. At low temperatures or near 0°K, holes start to flow into the n region from the p region when the applied voltage is high enough so that the Fermi level for holes on the p side coincide horizontally with the top of the valence band on the n side. A sharp rise or onset of current due to this effect should be observed at a forward bias voltage corresponding to $V_a=E_g+\epsilon_c$. Experimental data have shown that this is the case for all tunnel diodes made with or without gold doping and with various phosphorus concentrations in the bulk. The observed onset voltage of this current varies only slightly for units with different phosphorus concentrations, indicating that ϵ_c does not vary greatly with phosphorus concentration in the range of 1 to 10×10^{19} atoms/cc, and is in accord with the calculation shown in Fig. 13. For heavily gold-doped tunnel diodes, there seems to be a considerable shift to lower voltage of both structures 7 and 8, as indicated in the experimental data in Fig. 11. This shift is probably due to the temperature rise in the junction at the very high current density at this voltage. For low gold concentration diodes, the slope of the log of the current versus voltage plot above $V_a=E_g+\epsilon_c$ such as that shown in Fig. 11 is very nearly given by q/kT, and is in agreement with the usual diode diffusion or injection formula,

$$J=J_s\exp(qV/kT).$$

(4). *Process 9^- and 9^+*

The experimental results shown in Figs. 1(a), (b), (c), or 4 indicate that there is considerable gold-induced current in the voltage range below the onset of process 5^+ at approximately 240 mv given in Fig. 15(c). This induced current is probably associated with the transition processes 9^- and 9^+ shown in Fig. 17(a), in which electrons tunnel from the conduction band horizontally to the valence band in two steps. In the first step, $9^\pm a$, a conduction electron tunnels to a gold level at the same energy as the electron within the space charge region and in the second step, $9^\pm b$, the trapped electron tunnels out of the gold level into the valence band again horizontally or at constant energy. The transverse momentum of the electron need not be constant in this transition and may be carried away by the phonon emitted or absorbed by the trap.

The transition probability can be calculated from a relation similar to (1). In the present case, the transition probabilities of both the first and the second step are given by the general result derived by Price[11] and the steady-state rate is

$$w=w_{ct}w_{tv}/(w_{ct}+w_{tv}), \tag{7}$$

where w_{ct} and w_{tv} are the transition rate of the first and the second step. The current density due to these two

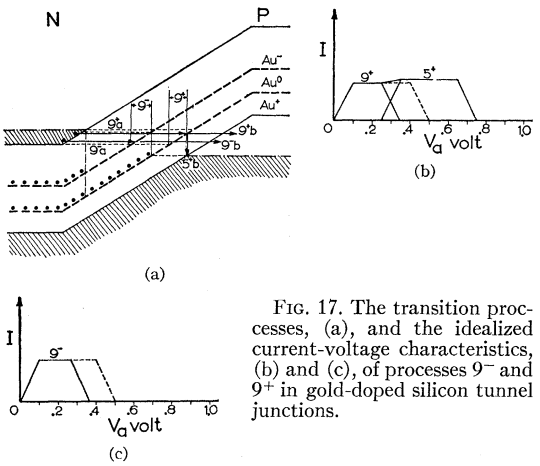


FIG. 17. The transition processes, (a), and the idealized current-voltage characteristics, (b) and (c), of processes 9^- and 9^+ in gold-doped silicon tunnel junctions.

processes can be computed from (2) and may be written in the form given by (3). The integral, D, may be calculated and put into explicit forms in the various voltage range if $\epsilon_c \ll E_{11}$ and $\epsilon_v \ll E_{12}$, e.g., in the high-field limit. (See Appendix V.)

The density-of-state integral has again the trapezoidal form in the low-field limit. It is evident from Fig. 17(a) that both 9^- and 9^+ start at zero voltage with no on-set structure. Since $\epsilon_c < \Delta\epsilon_{Au}$, the two processes through the two gold levels would not occur on gold atoms at the same x axis location or they are not coupled. However, there should be a sizeable coupling between processes 9^+ and 5^+, and between 9^- and 5^- since there are voltage ranges where both of the two processes in each of the two pairs may have non-zero density of states.

For example, at $V_a = \epsilon_{Au}{}^+$, approximately 0.25 v, process 5^+ would start as shown in Fig. 15(c) and would greatly affect the second tunneling step of process 9^+, 9^+b, since the second step of 5^+, 5^+b, shown in Fig. 15(b) may have comparable or appreciably greater transition rate than the second step of 9^+, 9^+b. The steady state transition rates may be obtained by a procedure similar to that used previously for processes 6^- and 6^+ and are given by ($V_a \geq \epsilon_{Au}{}^+$)

$$w(9^+) = w_{ct} w_{tv} / [w_{ct} + w_{tv} + w_e(5^+b)], \tag{8}$$

$$w(5^+) = w_{ct} w_e(5^+b) / [w_{ct} + w_{tv} + w_e(5^+b)]. \tag{9}$$

Equation (9), which takes into account process 9^+, is the modified transition rate for process 5^+ which was given in (1) neglecting process 9^+.

The modified trapezoidal current-voltage characteristics of 9^+ and 5^+ are shown in Fig. 17(b) for a limiting case in which $w_{tv} \ll w_e(5^+b)$, and the excited states of the Coulomb well are not taken into account.

A nonzero coupling between 9^- and 5^- through the second step of 5^-, 5^-b shown in Fig. 15(b), would be expected if $E_g - \epsilon_{Au}{}^- < \epsilon_v$. However, for the tunnel diodes studied in this paper, this condition is not satisfied and, in addition, 5^- does not exist due to $\Delta\epsilon_{Au} < \epsilon_v$. Transitions in a different route, however, would reduce the voltage span of process 9^-. Since $\Delta\epsilon_{Au} < \epsilon_v$, at $V_a = \Delta\epsilon_{Au} = 0.27$ v, the electron at the donor level on the gold atoms where 9^- is proceeding may now tunnel to the empty states in the valence band. These gold atoms, which have lost the electron on the donor level, have no acceptor level and become unavailable for process 9^-. Thus 9^- would start to turn off at $V_a = \Delta\epsilon_{Au}$.

The relative magnitude of 9^- and 9^+ may be estimated from the tunneling rate of 5^-, 5^+, 6^-, and 6^+ shown in Fig. 15(a). Process 9^- is essentially the combination of 5^-a and 6^-a at a constant energy. The forbidden path of the two steps of 9^- and the Coulomb interaction are the determining factors. Thus the tunneling rate of 9^- is limited by the first step, 9^-a, i.e., the tunneling of an electron from the conduction band to a neutral gold trap. The transition rate is given by

$$w(9^-) \approx w(9^-a). \tag{10}$$

For process 9^+, the heavier electron effective mass and the considerably longer forbidden path of the first step, 9^+a, compared with those of the second step, 9^+b, make the first step rate limiting. The transition rate of the first step is somewhat increased by the lowering of the potential barrier due to Coulomb interaction and by the parabolic and concave potential barrier in a real step junction. However the rate of the first step taking these factors into account is still considerably smaller than the rate of the second step.

(5). *Process 10*

This process is indicated in Fig. 18 and involves three steps. In the first step, $10a$, a neutral gold trap in the space charge region is filled by an electron from the conduction band and becomes negatively charged; in the second step, $10b$, the trapped electron tunnels horizontally from the negatively charged trap to an adjacent empty and positively charged gold trap; and in the third step, $10c$, the electron from the second step is emitted to the valence band from the gold trap, completing the transition from the conduction to the valence band. The steady-state transition rate is given by

$$w(10) = (w_a{}^{-1} + w_b{}^{-1} + w_c{}^{-1})^{-1}. \tag{11}$$

The probabilities of step a and c are large, due to the large concentration of electrons and holes in the conduction and valence band. The rate limiting step, b, has a rather small transition probability due to the large height of the potential barrier. The rate in step b is further reduced, since the gold concentration is small and the separation of the gold atoms is large compared with the size of the gold atom, which makes the solid angle sustained by a gold atom very small. In addition, step b involves transition between two rather sharp gold levels. The width of these gold levels is small and further reduces the transition probability. Thus, the probability rate of process 10 is expected to be very small and this process may be neglected compared with other processes just discussed.

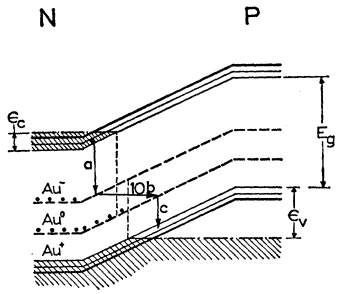


FIG. 18. The transition-process 10 in gold-doped silicon tunnel junctions.

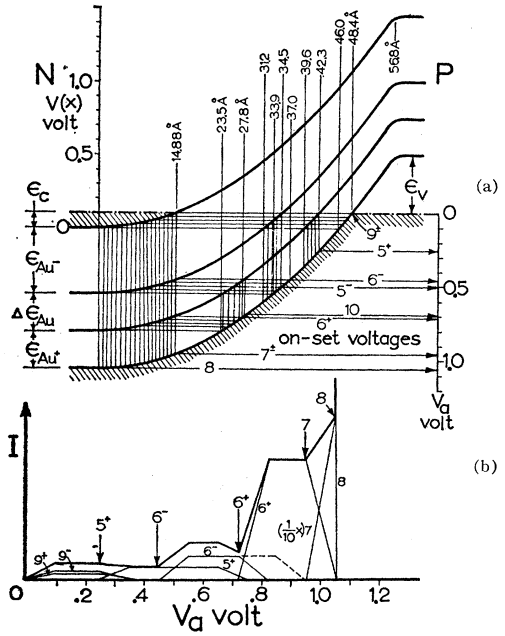


FIG. 19. (a) The parabolic potential energy diagram in gold-doped silicon tunnel junctions, and the ten transition processes with gold energy levels as intermediate states and their onset voltages, (b) the idealized theoretical current-voltage characteristics of a gold-doped tunnel junction of $N_D=6\times10^{19}$ atoms/cc, assuming $(m_\perp/m)W^2=10^{-24}$ volt^2-cm^3.

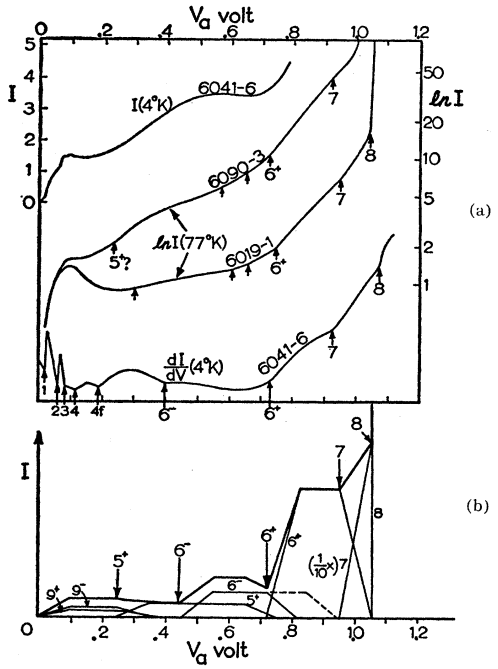


FIG. 20. A comparison of the observed and predicted onset voltages. (a) Experimental data and (b) idealized theoretical current-voltage characteristics from Fig. 19(b).

C. Analysis of Experimental Data

The transition processes associated with gold levels are computed numerically using the low-field trapezoidal approximation. The linear field variation is approximately taken into account by using the average slope of the energy-distance diagram for each of the transition processes. In Fig. 19, the energy distance diagram near and within a *p-n* junction is plotted using the parabolic approximation. The *p* side is assumed to be heavily doped so that the shape of the energy curve is fixed and independent of applied voltage. The effect of neutralization of impurities outside the depletion region would modify the shape of the energy-distance curve[34]; however, this is not taken into account.[35] In Fig. 19(b), the trapezoidal current-voltage characteristics are sketched out approximately to the right proportion in height. The onset voltages are also labeled, both in Figs. 19(a) and (b).

The comparison of the experimental onset voltages with the calculated values is made in Figs. 20(a) and (b). Reasonably good agreements are obtained for all processes from both the experimental conductance-voltage and the current-voltage curves. The onset of 5^+ appears to be very weak in Fig. 20(a) if it exists at all.

[34] R. N. Hall, General Electric Research Laboratory, Scientific Report No. 1, 1960 (unpublished).

[35] Calculations of the junction potential distribution and capacitance taking into account of de-ionization of impurity outside of the depletion region are made and planned to be published elsewhere.

By inspection of Fig. 19(a) and 17(b), this is to be expected, and in the limiting case where both 5^+ and 9^+ are rate controlled by 5^+a or 9^+a, one would expect no onset for 5^+, as indicated in Fig. 17(b).

The comparison of the observed shape of the gold-induced excess current-voltage characteristics with the trapezoidal approximation as shown in Figs. 20(a) and (b) or Figs 21(a) and (b), gives less satisfactory results. Some rather unambiguous differences appear below the onset voltage of 6^-. In particular, the experimental current-voltage characteristics at 4°K, shown in Fig. 21(a), are increasing near 0.4 v, while the idealized theoretical current-voltage curve shown in Fig. 21(b) is rather flat or slightly decreasing in this voltage range. This discrepancy may be partly removed by doing a more exact calculation taking into account the physical distribution of the donor and the acceptor in a real alloy junction. In addition, several other mechanisms may account for part of this discrepancy. One mechanism which is neglected involves tunneling to the excited levels of the Coulomb well in process 5^+. These transitions would increase the total transition rate considerably for process 5^+ above $\epsilon_c+\epsilon_{Au^+}\approx0.34$ v. One might expect to detect onset structures for these transitions to the excited states; however, they have not been experimentally observed. A second possible mechanism which has not been considered involves tunneling transition between conduction band tail or gap states on the *n* side and the positively charged gold trap states. Such a process would also increase the

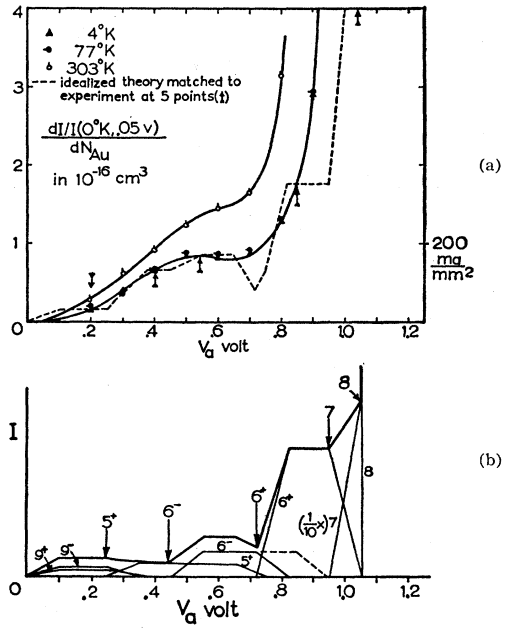


FIG. 21. A comparison of the observed gold-induced excess current-voltage characteristics with (a) the empirically matched theoretical calculation from Table II, and (b) the idealized theoretical calculation from Fig. 19 (b).

transition rate of 5^+ in the voltage region above $\epsilon_c+\epsilon_{Au}{}^+$ where the discrepancy exists. In addition to these mechanisms, the more exact calculations of the excess current-voltage characteristics must take into account the many-valley wrapped energy surfaces of the band structures in silicon.

In view of the various complications just discussed and the large number of unknown quantities involved in the theoretical formula, a detailed comparison of the calculated and observed characteristics has not been carried out. However, it may be useful to derive from the experimental data some estimate of the matrix elements and cross sections of the various tunneling and Hall-Shockley-Read processes.

In the following tables, results of numerical computations are listed for the various processes using the low-field trapezoidal approximation. The following formulas and numerical values are used in the calculations: [See also, Fig. 19(a)]

N_A (boron concentration in the regrowth layer) $=6\times10^{20}$ atoms/cc,

N_D (phosphorus concentration in the bulk) $=6\times10^{19}$ atoms/cc,

ϵ_c [Fermi energy in the n side) $=0.10$ ev,

ϵ_v (Fermi energy in the p side) $=0.40$ ev,

E_g (modified energy gap width) $=0.96$ ev,

$\epsilon_t(\mathrm{Au}^-)=0.44$ ev,

$\epsilon_t(\mathrm{Au}^+)=0.25$ ev,

$\Delta\epsilon_{Au}=0.27$ ev,

V_D (built-in voltage) $=1.46$ ev,

W_D (junction width at $V_a=0$ volt) $=56.8$ A,

F_D (maximum force on electron at $V_a=0$ v) $=2V_D/W_D=5.14\times10^6$ ev/cm,

$V(x)=V_D(x/W_D)^2$,

F_t (force on electron at trap) $=F_D(\epsilon_t/V_D)^{\frac{1}{2}}$,

x_t (length of the forbidden path to the trap) $=2\epsilon_t/F_t$ or $(46-2\epsilon_t/F_t)$ A,

$\Delta\epsilon_t=2(qF_t/4\pi K\epsilon_0)^{\frac{1}{2}}=0.219(F_t/10^6)^{\frac{1}{2}}$ ev, $K=12$ for silicon,

$$E_\perp=\hbar F_t(\epsilon_t-\Delta\epsilon_t)^{\frac{1}{2}}/2(2m_x)^{\frac{1}{2}}\epsilon_t$$
$$=[1962/x_t(\mathrm{A})][(m/m_x)(\epsilon_t-\Delta\epsilon_t)]^{\frac{1}{2}}\times10^{-3}\ \text{ev},$$
$$T=\exp\{-\tfrac{4}{3}[(2m_x)^{\frac{1}{2}}/\hbar F_t]\epsilon_t(\epsilon_t-\Delta\epsilon_t)^{\frac{1}{2}}\}$$
$$=\exp\{-0.341x_t(\mathrm{A})[(m_x/m)(\epsilon_t-\Delta\epsilon)]^{\frac{1}{2}}\},$$
$$D=\{\epsilon_c-E_\perp[1-\exp(-\epsilon_c/E_\perp)]\}/\epsilon_t,$$
$$w_0=1.59\times10^{13}(m_\perp/m)(10^{24}W^2)[T/x_t(\mathrm{A})]\ \mathrm{sec}^{-1},$$
$$J=1270(10^{-16}N_t)(m_\perp/m)(10^{24}W^2)TD\ \mathrm{ma/mm}^2.$$

The numerical computations are summarized in Table I. An example of taking into account the Coulomb interaction is given for 5^+a and 9^+a. The Coulomb correction amounts to almost an order of magnitude increase of both the transition probability rate at the

TABLE I. Numerical calculations of the tunneling probability T, the density of state integral, D, the tunneling transition probability rate at the band edge, w_0, and the current density J for the various tunneling transition processes.

Process	F_t [a] (10^6 ev/cm)	ϵ_t	$\Delta\epsilon_t$ (10^{-3} ev)	$\epsilon_t-\Delta\epsilon_t$	x_t (A)	m_x/m	$E_\perp$ (10^{-3} ev)	T	D	w_0 [c] (10^{10}/sec)	J [c] (ma/mm²)
$5^+a, 9^+a$	3.60	710	415	295	39.4	0.19	62.2	0.042	0.071	1.68	37.4
			0	710			99.5	0.007	0.051	0.29	4.8
9^+b	3.60	250	0	250	6.4	0.16	383	0.647	0.049	160	400
$5^-a, 9^-a$	2.83	440	0	440	31.2	0.19	96.2	0.047	0.086	2.40	51.2
9^-b	2.83	520	368	152	14.8	0.16	130	0.463	0.058	49.7	344
6^+a	2.13	250	0	250	23.5	0.16	104	0.202	0.142	13.6	363
6^-a	3.07	520	254 [b]	266	33.9	0.16	74.7	0.093	0.086	4.37	102

[a] Field at trap or at valence band edge at (onset voltage $+\epsilon_c$).
[b] $F_t=1.35\times10^6$ ev/cm at trap is used.
[c] $(m_\perp/m)W^2=10^{-24}$ volt²-cm³ and $N_t=10^{17}$ Au atoms/cc.

TABLE II. Evaluation of the tunneling matrix element, $(m_\perp/m)W^2$ and the tunneling transition probability at the band edge, w_0, for the various tunneling and recombination processes from experimental data.

Process	V_a (volts)	J_{exp} [a] (ma/mm²)	J_{cal} [b] (ma/mm²)	$(m_\perp/m)W^2$ [c] (10^{-24} v² cm³)	$\bar{C}$ (cm³/sec)	σ (10^{-14} cm²)	w_0 [f] (10^{10} sec^{-1})	J_{exp} [g] (ma/mm²)
9^-+9^+	0.20	28	8.2	3.4			9^- 8.2	15.2
							9^+ 5.7	12.8
5^+	0.40	130	3.74	34.8			5^+ 58.5	130
6^-+5^+	0.55	168	13.9	2.74			6^- 12.0	38
6^+	0.85	350	36.3	9.6			6^+ 130	350
7	$E_g+\epsilon_c$ =1.06	3060 [d]		0.25 [e]	9.3×10^{-8}	1.8	555 142[e]	

[a] From Fig. 21(a).
[b] Based on Fig. 21(b) and Table I with $(m_\perp/m)W^2=10^{-24}$ volt²-cm³, $N_t=10^{16}$ Au/cm³.
[c] Value required to bring J_{exp} and J_{cal} in agreement.
[d] From diode 6019 with $N_t=19\times10^{16}$ Au/cm³.
[e] First order matrix approximation.
[f] Use $(m_\perp/m)W^2$ from *c*.
[g] Use w_0 from *f*.

band edge, w_0, and the total current density J in the flat region of the trapezoid for these two processes. Similar corrections are made for processes 9^-b and 6^-a. The effective mass of the light hole in the V_2 valence band and the transverse electron mass are used in the calculation. The relative magnitude of the transition probability rate, w_0, and the current density in the flat region, J, obtained in Table I are in accord with the discussion on the various transition processes given previously in Sec. III(B).

In Table II, the theoretical results of Table I are added at various applied voltages and compared with the experimental data taken from Fig. 21(a). The column w_0 is computed from the experimental data and the calculations of Table I. An average value for the matrix element for tunneling from trap states is obtained, which has a value of

$$W^2(m_\perp/m)=1.2\times10^{-23}\ \text{volt}^2\text{-cm}^3.$$

This is quite a reasonable value in view that the square well depth of the trap in silicon is around 5 ev and the diameter of the trap well is about 10^{-8} cm.[24]

In the last row of Table II, several quantities associated with the Hall-Shockley-Read process are calculated from diode 6019–1, shown in Fig. 11. For process 7^- one would expect $\bar{C}_P\gg\bar{C}_n$, since the hole capture process involves attractive Coulomb barrier, while the electrons are captured by neutral centers which have nearly a square barrier. In the case of process 7^+, $\bar{C}_p\ll\bar{C}_n$, since the barrier for electrons is of the Coulomb type and that for hole is nearly a square well. (See Fig. 16.) Since the acceptor doping or hole concentration is much higher than the electron concentration or donor doping, the density-of-state integral given by (A.4.7) in Appendix IV has about the same value for both 7^- and 7^+. Thus, in the calculation given in the last line of Table II it is assumed that one of the gold levels is important in contributing to the recombination in the space charge region and a value of $D=0.36$ is adopted which corresponds to $(\bar{C}_nN_D/\bar{C}_pN_A)(\epsilon_v/\epsilon_c)^{\frac{3}{2}}=1$ and $\epsilon_v/\epsilon_c=4$ and was numerically computed and listed in Appendix IV following Eq. (A.4.8). It is interesting that the cross section obtained, 1.8×10^{-14} cm², is in reasonable agreement with published results.[29] The transition probability rate for the Hall-Shockley-Read process is also obtained for electrons or holes at the Fermi surface from experimental data and Eq. (A.4.2) in Appendix IV. The rate listed in Table II is $w_0=555\times10^{10}$/sec and is considerably higher than the tunneling rate from traps for the various processes, except for process 6^+, listed in Table II. The large rate for the Hall-Shockley-Read process is reasonable, since this process in tunneling diodes involves majority carriers which have very high concentrations.

The experimental data for processes 7^+ and 7^- are also calculated using the first order matrix or the golden rule approximation given in Appendix IV instead of the Hall-Shockley-Read approximation. The effective matrix element is about 0.25×10^{-24} volt²-cm³, as listed in Table II, and is almost 50 times smaller than the matrix element of tunneling from trap states. In the process of tunneling from trap states, the transition is horizontal or the energy is constant, while in the capture or emission process in traps, the transition involves a dissipation of about 0.5 ev energy which may be carried away by phonons in successive one-phonon processes as discussed by Lax.[29] Thus, in view of the large energy dissipation which must be carried away by a cascade process, the effective matrix element of the Hall-Shockley-Read process would be considerably smaller than the tunneling processes from traps which involves constant energy.

The value of the matrix element $(m_\perp/m)W^2$ obtained empirically at five voltages given in Table II are used to calculate the idealized theoretical curve shown in dashed lines in Fig. 21(a). Although the region near 0.4 v now shows good correlation between experiment and calculation, the negative resistance region near 0.7 v of the calculated curve has a considerably larger dip than observed. This is due to the very large empirically determined value of $(m_\perp/m)W^2$ for process 5^+ shown in Table II. However, there is no reason that the

matrix element of 5^+ should be greatly different from that of 9^+ since these two processes are rate limited by the identical first step, i.e., 5^+a or 9^+a, involving electron tunneling from the conduction band states horizontally to positively charged trap states in the space charge layer. Thus, it seems more reasonable to assume that the additional mechanisms discussed previously, particularly transition to the band edge tail states, are probably responsible for the discrepancy which exists between the experimental data of Fig. 21(a) and the calculation of Fig. 21(b), which is preferred over the empirically matched calculation of Fig. 21(a).

IV. SUMMARY

The electronic transition processes in gold-doped silicon tunnel junctions are examined. Ten processes associated with gold are considered. The observed and expected structures at the onset voltages for five of the processes are in reasonably good agreement. However, the observed excess current at around 0.4 v is not completely accounted for and may be due to tunneling from band edge tails to trap states, or due to tunneling to the excited trap states.

Matrix elements for the process of tunneling from trap states are estimated from the experimental data and are about 50 times greater than the first order effective matrix element of the Hall-Shockley-Read process. However, since the Hall-Shockley-Read process involves high concentration of majority carriers in tunnel junctions, the transition probability rate for the processes are rate controlled by the step involving tunneling from or to the trap states. At above 0.7 v, a two step process, process 6^+, which involves conduction electrons filling the positively charged and empty gold donor level and subsequently tunneling out of the trap horizontally to the empty states in the valence band, may have comparable transition probability rates for both the Hall-Shockley-Read process in the first step and the tunneling from trap states in the second step.

ACKNOWLEDGMENTS

The author wishes to thank G. E. Moore and Jean Hoerni for helpful discussions and P. J. Price for sending us a summary of his paper. He is indebted to D. A. Tremere, D. Tavares, and G. Lessard for fabricating the tunnel diodes used in this study. He is also indebted to L. Chang for help in preparing part of the manuscript.

APPENDIX I. THE TRANSITION PROBABILITY RATE OF TUNNELING TO OR FROM A TRAP

Price[11] has derived an effective first order matrix element for electron tunneling out of the trap to the valence band which is given by

$$M=W\left(\frac{v_x}{2Vu_t}\right)^{\frac{1}{2}}e^{-\theta} \qquad \text{(A.1.1)}$$

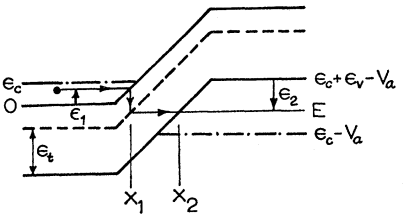


FIG. 22. The energy band diagram labeled with the notations used in deriving the transition probability rate.

where W is the matrix element $(\phi_t|\mathcal{U}_t|\psi_0)$ of the trap potential energy, $\mathcal{U}_t$, in excess of the crystal potential taken between the unnormalized trap-state wave function, ϕ_t, and the band-edge Bloch wave function, ψ_0, normalized in unit volume; $\hbar u(x)=|\partial E/\partial k_x|$, and u_t is evaluated at the trap; v_x is the component of the electron velocity at infinity in the valence band; V is the volume of the crystal; and

$$\theta=\int|k_x|dx, \qquad \text{(A.1.2)}$$

where the integration covers the forbidden path. The transition probability rate is given by the usual "golden rule" expression:

$$w=(2\pi/\hbar)|M|^2\rho(\epsilon_2), \qquad \text{(A.1.3)}$$

where $\rho(\epsilon_2)$ is the density of the final state in the valence band. An expression similar to (A.1.3) may be obtained for conduction electrons tunneling into an empty trap. In order to facilitate the calculation of (A.1.3), we shall use the labeled energy diagram shown in Fig. 22.

For an ellipsoidal energy surface, the wave-vector component, k_x, in (A.1.2) may be obtained in the forbidden region by using the effective-mass equation:

$$E\psi_0(r)=\left(\frac{\hbar^2k_{2x}^2}{2m_{2x}}+\frac{\hbar^2k_{2y}^2}{2m_{2y}}+\frac{\hbar^2k_{2z}^2}{2m_{2z}}+V\right)\psi_0(r). \qquad \text{(A.1.4.)}$$

Thus,

$$|k_{2x}|=\frac{(2m_{2x})^{\frac{1}{2}}}{\hbar}[E-V(x)-\epsilon_{21}]^{\frac{1}{2}}, \qquad \text{(A.1.5)}$$

where

$$\epsilon_{21}=\frac{\hbar^2k_{2y}^2}{2m_{2y}}+\frac{\hbar^2k_{2z}^2}{2m_{2z}}. \qquad \text{(A.1.6)}$$

Substitution of (A.1.5) into (A.1.2) gives

$$\theta=\int_{x_1}^{x_2}|k_{2x}|dx=\frac{2}{3}\frac{(2m_{2x})^{\frac{1}{2}}}{\hbar F}[(\epsilon_{21}+\epsilon_t)^{\frac{3}{2}}-\epsilon_{21}^{\frac{3}{2}}] \qquad \text{(A.1.7)}$$

in which the constant field approximation, $V(x)=Fx-\epsilon_g$ is used. It is usually assumed that appreciable tunneling occurs near $\epsilon_{21}=0$, so that (A.1.7) may be approximated by the first term of a Taylor series expansion which gives

$$\theta\approx\frac{2}{3}\frac{(2m_{2x})^{\frac{1}{2}}}{\hbar F}[\epsilon_t^{\frac{3}{2}}+\tfrac{3}{2}(\epsilon_t)^{\frac{1}{2}}\epsilon_{21}]. \qquad \text{(A.1.8)}$$

The function, u_t, may be obtained from

$$\hbar u_t=(\partial/\partial k_x)(\hbar^2k_x^2/2m_{2x})$$
$$=\hbar\{2[E-V(x)]/m_{2x}\}^{\frac{1}{2}}=(2\epsilon_t/m_{2x})^{\frac{1}{2}}. \quad (A.1.9)$$

The density of the final state may be obtained from

$$\rho(\epsilon_2)d\epsilon_{2x}=2V\frac{2k_{2x}dk_{2y}dk_{2z}}{(2\pi)^3}=\frac{V(m_{2y}m_{2z})^{\frac{1}{2}}}{\hbar^3v_x2\pi^2}d\epsilon_{2x}d\epsilon_{2\perp}. \quad (A.1.10)$$

Thus, combining (A.1.1), (A.1.8), (A.1.9), and (A.1.10) with (A.1.3) and integrating over all possible perpendicular energy from 0 to ϵ_2, we have

$$w=\frac{2\pi}{\hbar}\sum|M|^2\rho(\epsilon_2)$$
$$=\frac{\pi^2}{\hbar^3}\frac{(m_{2y}m_{2z})^{\frac{1}{2}}}{\epsilon_t}F_tW_2^2\exp\left[-\frac{4}{3}\frac{(2m_{2x})^{\frac{1}{2}}}{\hbar F}\epsilon_t^{\frac{3}{2}}\right]$$
$$\times\left\{1-\exp\left[-\frac{2(2m_{2x})^{\frac{1}{2}}}{\hbar F}(\epsilon_t)^{\frac{1}{2}}\epsilon_2\right]\right\}$$
$$=w_0\left\{1-\exp\left[-2\frac{(2m_{2x})^{\frac{1}{2}}}{\hbar F}(\epsilon_t)^{\frac{1}{2}}\epsilon_2\right]\right\}. \quad (A.1.11)$$

Coulomb Correction

The Coulomb attraction which lowers the tunnel barrier, such as in process 5^+a shown in Fig. 15(a), may be taken into account approximately. The simplest approximation is to replace the exact barrier by a triangular one with a height which is equal to the maximum height of the true barrier. Thus, in the constant-field approximation, the barrier height ϵ_t in (A.1.11) is replaced by

$$\epsilon_t=\epsilon_{t0}-2\left(\frac{qF_0}{4\pi K\epsilon_0}\right)^{\frac{1}{2}}, \quad (A.1.12)$$

where ϵ_{t0} is the barrier height or the trap energy measured from the band edge and F_0 is the constant force at zero Coulomb interaction. The field F in (A.1.11) must be replaced by

$$F=F_0\left[\epsilon_{t0}-2\left(\frac{qF_0}{4\pi K\epsilon_0}\right)^{\frac{1}{2}}\right]\Big/\epsilon_{t0}. \quad (A.1.13)$$

APPENDIX II. DETAILED BALANCE CALCULATION OF PROCESSES 5^+, 6^-, AND 6^+

Referring to Figs. 15(a) and (b), the steady-state rate of 5^+ is

$$w=U/N_t=w(5^+a)N^+/N_t=w(5^+b)N^0/N_t, \quad (A.2.1)$$

where

$$N_t=N^+ +N^0 \quad (A.2.2)$$

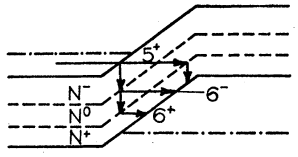


FIG. 23. The coupling of processes 6^- and 6^+.

is the total gold concentration, since no gold in the space charge region can trap two electrons (or 5^- does not occur), i.e., $N^-=0$. N^+, N^0, and N^- are the gold concentrations in the three charge conditions. Thus, from (A.2.1) and (A.2.2)

$$w=U/N_t=w(5^+a)w(5^+b)/[w(5^+a)+w(5^+b)]. \quad (A.2.3)$$

For processes 6^- and 6^+, (A.2.3) can be used if only one of the 6^- and 6^+ processes can occur at the same atoms. If 6^- and 6^+ can occur simultaneously, we must take both into account. (See Fig. 23.) Thus,

$$U(6^-)=w_e^-(6^-b)N^0=w_t^-(6^-a)N^-=w^-N_t, \quad (A.2.4)$$

$$U(6^+)=w_e^+(6^+b)N^+=w_t^-(6^+a)N^0=w^+N_t, \quad (A.2.5)$$

and

$$N_t=N^+ +N^0+N^-. \quad (A.2.6)$$

Combining (A.2.4), (A.2.5), and (A.2.6), we have the relation given by (4) and (5) in the text.

APPENDIX III. CALCULATION OF TUNNEL CURRENT OF PROCESSES 5^+, 6^-, AND 6^+

The tunnel current may be obtained from

$$J=\int qwN_tdx=\int qwN_td\epsilon_1/F. \quad (A.3.1)$$

If the Hall-Shockley-Read process is not rate limiting and only one transition process can occur at the gold atoms at a given x plane, we have (for example, process 5^+), $w\approx w(5^+a)$, where $w(5^+a)$ may be obtained from (A.1.11) by changing subscript 2 to 1. Thus,

$$J=J_0D, \quad (A.3.2)$$

where

$$J_0=qN_t(\pi^2/\hbar^3)(m_{1y}m_{1z})^{\frac{1}{2}}W_1^2T=qN_tw_0x_t, \quad (A.3.3)$$

$$x_t=\epsilon_t/F=\text{length of forbidden path}, \quad (A.3.4)$$

$$T=\exp\left(-\frac{4}{3}\frac{(2m_{1x})^{\frac{1}{2}}}{\hbar F}\epsilon_{\text{Au}^-}^{\frac{3}{2}}\right), \quad (A.3.5)$$

$$D=\frac{1}{\epsilon_t}\int[1-\exp(-\epsilon_1/E_{1\perp1}]d\epsilon_1, \quad (A.3.6)$$

$$E_{1\perp1}=\hbar F/2(2m_{x1}\epsilon_t)^{\frac{1}{2}}. \quad (A.3.7)$$

There is a region: $\epsilon_c+\epsilon_g-\epsilon_{\text{Au}^-}\lesssim V_a\lesssim\epsilon_v-\epsilon_c+\epsilon_g-\epsilon_{\text{Au}^-}$, where D is constant and is given by

$$D=\{\epsilon_c-E_{1\perp1}[1-\exp(-\epsilon_c/E_{1\perp1})]\}/\epsilon_{\text{Au}^-}. \quad (A.3.8)$$

If the Hall-Shockley-Read process has a transition rate comparable to that of the tunneling from traps, the density of state integral would have a different form from (A.3.6). In the Hall-Shockley-Read approximation, the transition rate for the second step, $5^{+}b$, is

$$w(5^{+}b)=\int_0^{\epsilon_2} C_p(\epsilon_2)V\rho(\epsilon_2)d\epsilon_2\cong\bar{C}_p p(\epsilon_2/\epsilon_v)^{\frac{3}{2}}, \quad \text{(A.3.8)}$$

and

$$D=\frac{1}{\epsilon_t}\int\frac{[1-\exp(-\epsilon_1/E_{\perp 1})](\epsilon_2/\epsilon_v)^{\frac{3}{2}}d\epsilon_1}{[1-\exp(-\epsilon_1/E_{\perp 1})]+a^{-3}(\epsilon_2/\epsilon_v)^{\frac{3}{2}}}, \quad \text{(A.3.9)}$$

where

$$a^{-3}=\bar{C}_p p\Big/\left[\frac{\pi^2(m_{1y}m_{1z})^{\frac{1}{2}}}{h^3\epsilon_t F}W_1^2\exp\left(-\frac{4}{3}\frac{(2m_{1x})^{\frac{1}{2}}}{\hbar F}\epsilon_t^{\frac{3}{2}}\right)\right]. \quad \text{(A.3.10)}$$

The relation between ϵ_1 and ϵ_2 can readily be obtained from Fig. 22, and is given by

$$\epsilon_1+\epsilon_2=\epsilon_c+\epsilon_v+\epsilon_g-V_a. \quad \text{(A.3.11)}$$

The range of integration for (A.3.9) is the same as that of (A.3.6) and is discussed in the text. The function D is smoothly varying with V_a, except at the onset, where it has a finite slope or discontinuity in the D versus V_a plot. In general, (A.3.9) must be evaluated numerically; however, for $E_{\perp 1}\gg\epsilon_1$, an explicit result can be obtained.

APPENDIX IV. RECOMBINATION IN THE SPACE CHARGE REGION (PROCESSES $7^{\pm}$)

For simplicity, we shall consider only a one-level recombination center at 0°K. Let's consider process 7^{-}. The steady-state rate is given by an expression similar to (A.2.3) and is

$$w=w(7^{-}a)w(7^{-}b)/[w(7^{-}a)+w(7^{-}b)]. \quad \text{(A.4.1)}$$

The individual rates $w(7^{-}a)$ and $w(7^{-}b)$ may be obtained from the result of Hall-Shockley-Read statistics or by using an effective matrix element in the "golden rule" expression of (A.1.3).

(1) *Hall-Shockley-Read Approximation*

The transition rates are:

$$w(7^{-}a)=\int_0^{\epsilon_1} C_n(\epsilon_1)V\rho(\epsilon_1)d\epsilon_1\cong\bar{C}_n n(\epsilon_1/\epsilon_c)^{\frac{3}{2}}, \quad \text{(A.4.2)}$$

$$w(7^{-}b)=\int_0^{\epsilon_2} C_p(\epsilon_2)V\rho(\epsilon_2)d\epsilon_2\cong\bar{C}_p p(\epsilon_2/\epsilon_v)^{\frac{3}{2}}, \quad \text{(A.4.3)}$$

where the electron and hole capture probabilities, $\bar{C}_n$ and $\bar{C}_p$ are assumed constant and their concentrations are

$$n=8\pi(2m_c\epsilon_c)^{\frac{3}{2}}/3h^3, \quad \text{(A.4.4)}$$

and

$$p=8\pi(2m_v\epsilon_v)^{\frac{3}{2}}/3h^3. \quad \text{(A.4.5)}$$

The current density obtained from (A.3.1) is given by

$$J=(qN_t\bar{C}_nN_D/F)\epsilon_c D, \quad \text{(A.4.6)}$$

where

$$D=\int\frac{\epsilon_2^{\frac{3}{2}}\epsilon_1^{\frac{3}{2}}d\epsilon_1}{\epsilon_2^{\frac{3}{2}}+(\bar{C}_nN_D/\bar{C}_pN_A)(\epsilon_v/\epsilon_c)^{\frac{3}{2}}\epsilon_1^{\frac{3}{2}}}\frac{1}{\epsilon_c^{\frac{3}{2}}}. \quad \text{(A.4.7)}$$

The electron and hole kinetic energies, ϵ_1 and ϵ_2, are related by $\epsilon_2+\epsilon_1=\epsilon_c+\epsilon_v+\epsilon_g-V_a$ as given by (A.3.11). The integration range of (A.4.7) covers energies vertically common to the filled conduction and empty valence band. For example, in the range, $\epsilon_g\lesssim V_a\lesssim\epsilon_g+\epsilon_c$, $\epsilon_c+\epsilon_g-V_a\lesssim\epsilon_1\lesssim\epsilon_c$, when $V_a\sim\epsilon_g$, and $(\bar{C}_nN_D/\bar{C}_pN_A)(\epsilon_v/\epsilon_c)^{\frac{3}{2}}=1$, D is given by

$$D\approx\frac{\epsilon_v^{\frac{3}{2}}}{e_v^{\frac{3}{2}}+e_c^{\frac{3}{2}}}\left(\frac{V_a-\epsilon_g}{\epsilon_c}\right). \quad \text{(A.4.8)}$$

At $V_a=\epsilon_g+\epsilon_c$, the onset voltage of the injection current D is given by the following table for $(\bar{C}_nN_D/\bar{C}_pN_A)\times(\epsilon_v/\epsilon_c)^{\frac{3}{2}}=1$:

ϵ_v/ϵ_c	1.0	2.0	3.0	4.0	∞
D	0.0983	0.278	0.338	0.362	0.4000

(2) *First-Order Matrix Approximation*

The transition rates are

$$w(7^{-}a)=\frac{2\pi}{\hbar}M_a^2\frac{1}{\pi^2\hbar^3}(2m_{1x}m_{1y}m_{1z})^{\frac{1}{2}}(\epsilon_1)^{\frac{1}{2}}, \quad \text{(A.4.9)}$$

and

$$w(7^{-}b)=\frac{2\pi}{\hbar}M_b^2\frac{1}{\pi^2\hbar^3}(2m_{2x}m_{2y}m_{2z})^{\frac{1}{2}}(\epsilon_2)^{\frac{1}{2}}. \quad \text{(A.4.10)}$$

The current density is

$$J=\frac{qN_t}{F}\left(\frac{2\pi}{\hbar}M_a^2N_D\right)D, \quad \text{(A.4.11)}$$

where

$$D=\frac{3}{e_c^{\frac{3}{2}}}\int\frac{(\epsilon_1)^{\frac{1}{2}}(\epsilon_2)^{\frac{1}{2}}}{(\epsilon_2)^{\frac{1}{2}}+(M_a/M_b)2(m_{1x}m_{1y}m_{1z}/m_{2x}m_{2y}m_{2z})^{\frac{3}{2}}(\epsilon_1)^{\frac{1}{2}}}d\epsilon_1. \quad \text{(A.4.12)}$$

The integration range of (A.4.12) covers the same energy as (A.4.7). At $V_a=\epsilon_g+\epsilon_c$,

$$(M_a/M_b)^2(m_{1x}m_{1y}m_{1z}/m_{2x}m_{2y}m_{2z})^{\frac{3}{2}}=1$$

and $V_a=\epsilon_g+\epsilon_c$; (A.4.12) can be evaluated numerically At $\epsilon_v/\epsilon_c=4$, $D=1.408$.

APPENDIX V. CALCULATION OF TWO-STEP CASCADE TUNNELING VIA TRAP (PROCESSES 9^+ AND 9^-)

Let us consider 9^- and disregard coupling to other processes. Then the rates are

$$w(9^+a)=\frac{(m_{1y}m_{1z})^{\frac{1}{2}}FW_a^2}{8\pi\hbar^3\epsilon_{Au^-}}\exp\left(-\frac{4}{3}\frac{(2m_{1x})^{\frac{1}{2}}}{\hbar F}(\epsilon_{Au^-})^{\frac{3}{2}}\right)\times\left[1-\exp\left(-\frac{2(2m_{x1})^{\frac{1}{2}}}{\hbar F}(\epsilon_{Au^-})^{\frac{1}{2}}\epsilon_1\right)\right]$$

$$=w_{10}\left[1-\exp\left(-\frac{2(2m_{2x})^{\frac{1}{2}}}{\hbar F}(\epsilon_{Au^-})^{\frac{1}{2}}\epsilon_1\right)\right] \quad \text{(A.5.1)}$$

and

$$w(9^+b)=w_{20}\left[1-\exp\left(-\frac{2(2m_{2x})^{\frac{1}{2}}}{\hbar F}(\epsilon_g-\epsilon_{Au^-})^{\frac{1}{2}}\epsilon_2\right)\right]. \quad \text{(A.5.2)}$$

The current density is (for small F)

$$J=qN_t\frac{w_{10}w_{20}}{w_{10}+w_{20}}\frac{\epsilon_c}{F}D. \quad \text{(A.5.3)}$$

Here, $D=$ trapezoid for

$$E_{\perp 1}=\frac{\hbar F}{2(2m_{1x}\epsilon_{Au^-})^{\frac{1}{2}}}\ll\epsilon_c$$

and

$$E_{\perp 2}=\frac{\hbar F}{2[2m_{2x}(\epsilon_g-\epsilon_{Au^-})]^{\frac{1}{2}}}\ll\epsilon_v.$$

For $E_{\perp 1}\gg\epsilon_c$ and $E_{\perp 2}\gg\epsilon_v$, D is given by

$$D=\frac{1}{\epsilon_c^2}\int\frac{\epsilon_1(\epsilon_c+\epsilon_v-V_a-\epsilon_1)d\epsilon_1}{\epsilon_1(1-a)+a(\epsilon_c+\epsilon_v-V_a)}$$

$$=\frac{1}{\epsilon_c^2}\left[\frac{-\epsilon_1^2}{1-a}+\frac{\epsilon_c+\epsilon_v-V_a}{(1-a)^2}\left\{\epsilon_1-\frac{a(\epsilon_c+\epsilon_v-V_a)}{1-a}\times\ln[(1-a)\epsilon_1+a(\epsilon_c+\epsilon_v-V_a)]\right\}\right]_{\epsilon_a}^{\epsilon_b}, \quad \text{(A.5.4)}$$

where

$$a=\frac{w_{20}}{w_{10}}\left(\frac{m_{2x}(\epsilon_g-\epsilon_{Au^-})}{m_{1x}\epsilon_{Au^-}}\right)^{\frac{1}{2}}. \quad \text{(A.5.5)}$$

For $a=1$, the peak voltage occurs at

$$v_a=V_a/\epsilon_c=1+r-(1+r^3)^{\frac{1}{3}}<1, \quad \text{(A.5.6)}$$

where

$$r=\epsilon_v/\epsilon_c.$$

In a symmetrical case with respect to electrons and holes, i.e., $\epsilon_{Au^-}=\epsilon_g/2$, $w_{10}=w_{20}$, and $m_{1x}=m_{2x}$, the integral can be calculated explicitly, giving

$$D=\left[\frac{1}{\epsilon_c}(\epsilon_1+\tfrac{1}{2}E_\perp G)\right]_{\epsilon_1=\epsilon_a}^{\epsilon_1=\epsilon_b}, \quad \text{(A.5.7)}$$

where

$$G=\frac{b}{(1-b)^{\frac{1}{2}}}\ln\left[\frac{2(1-b)^{\frac{1}{2}}(1-b-2u+u^2)^{\frac{1}{2}}}{u}+\frac{2(1-b)}{u}-2\right], \quad \text{(A.5.8)}$$

$$b=\exp[-(\epsilon_c+\epsilon_v-V_a)/E_\perp], \quad \text{(A.5.9)}$$

$$u=1-\exp-\left(\frac{\epsilon_c+\epsilon_v-V_a}{2E_\perp}\right)\cosh\left[\frac{\epsilon_1-\frac{1}{2}(\epsilon_c+\epsilon_v-V_a)}{E_\perp}\right]. \quad \text{(A.5.10)}$$