THEORY OF ELECTRONIC SWITCHING EFFECT AS A COOPERATIVE PHENOMENON*

Daniel C. Mattis
Belfer Graduate School of Science, Yeshiva University, New York, New York 10033
(Received 20 February 1969)

A modified form of Zener tunneling theory, in which we assume that the tunneling barrier is a decreasing collective function of electronic excitation, is shown to imply insulator-to-metal switching characteristics similar to those reported by Ovshinsky for glass film devices. In addition, our *I-V* characteristics display a region of negative resistance. Some criteria for the existence of polyconductivity are discussed, and are shown to be met by several classes of material.

A phenomenon recently reported by Ovshinsky[1] concerning the reversible switching of glassy thin films from the insulating to the conducting state, which we here denote "polyconductivity," is sufficiently startling to warrant speculative theoretical investigation. We wish to report here the results of one such speculation: Zener tunneling theory seems capable of explaining the unusual insulating-state characteristics of Ovshinsky's devices provided it is assumed that the tunneling barrier E_g is a decreasing function of the degree of electronic excitation. It is possible then that the switching properties are a new manifestation of the Mott transition.[2]

The set of equations which is discussed below was solved numerically by the present author in close collaboration with L. Landovitz and M. Plischke. Some interesting results obtained by the three of us are displayed graphically herewith. If our basic hypotheses are correct, these results imply that the glass devices should be characterized by a figure of merit which we have denoted by b, which should normally be in the range $-10 < b < +10$. If b is larger than +10 the device is essentially always metallic, whereas if it lies below −10 the device closely simulates a conventional insulator, although we predict that at sufficiently high temperature all devices become metallic. The usual variables, temperature, electric field, and energy gap (and eventually, magnetic field—which we have not yet studied) are all expressible in terms of reduced variables, so that all devices belonging to the same parameter b share a common set of characteristic curves.

Let us start with the assumption that the electrons within an impurity band are in a collective insulating state with an energy gap E_g which is a decreasing function of the degree of electronic excitation. Define n to be the fraction of carriers excited out of the collective state, and assume a linear relation[3]

$$E_g = E_0(1-2n), \text{ or } e_g = 1-2n, \tag{1}$$

which defines the "reduced-gap" parameter e_g ($e_g = 1$ for $n = 0$ and $e_g = 0$ for $n \geq \frac{1}{2}$). The fraction of carriers which are thermally excited out of the condensed state is given by

$$n_{\text{th}} = [1+\exp(E_g/2kT)]^{-1} = [1+\exp(2e_g/t)]^{-1} \tag{2}$$

in which a "reduced temperature" t is defined in terms of the critical temperature T_c as

$$t = T/T_c \text{ and } kT_c = \tfrac{1}{4}E_0. \tag{3}$$

In zero electric field $n_{\text{th}} = n$, and simultaneous solution of the above leads to[3]

$$e_g = \tanh(e_g/t), \tag{4}$$

which has a nontrivial solution for all $t < 1$, i.e., all $T < T_c$, and only the trivial solution $e_g = 0$ at higher temperatures. Since this result turns out to be independent of the applied voltage, one of our predictions is that each device has a T_c above which it is always metallic, although this temperature might not be physically attainable if it exceeds the melting temperature of the material. In units of numbers of particles/unit time the Zener current[4] is

$$j_{\text{tun}} = I_0(1-n)\exp(-4maLE_g^{\,2}/\hbar^2 eV) = I_0(1-n)\exp(-e_g^{\,2}/u). \tag{5}$$

I_0 is assumed a constant parameter, L is the film thickness, and u is the dimensionless electric field parameter

$$u = \hbar^2 eV/4maLE_0^{\,2}, \tag{6}$$

where a^{-3} = density of the impurity-band electrons (upwards of $10^{19}/\text{cm}^3$ in the experiments[1]). After a lifetime $\mathcal{T}$ an excited particle recombines, so that the net fraction n_{tun} of electrons excited out of the collective state by tunneling is

$$n_{\text{tun}} = (I_0\mathcal{T})(1-n)\exp(-e_g^{\,2}/u) = (1-n)\exp(b-e_g^{\,2}/u). \tag{7}$$

The important parameter in our theory is $b = \ln(I_0\mathcal{T})$. Positive b implies long lifetimes and easy excitation by applied fields, i.e., easy breakdown of the insulating state, whereas negative b implies the diametric opposite. Equations (1), (2), (5), (6), and (7) together with

$$n = n_{\text{th}} + n_{\text{tun}} \tag{8}$$

are a set of self-consistent nonlinear equations with rather unexpected structure, as we shall soon see. At constant t it is possible to solve for u as a function of e_g, viz.,

$$u = e_g^{\,2}\left[b + \ln\frac{1+e_g}{\tanh(e_g/t)-e_g}\right]^{-1} \tag{9}$$

and by inverting the solution, i.e., considering e_g to be a function of u and t, to obtain the tunneling current as a function of u and t:

$$j_{\text{tun}} = \tfrac{1}{2}I_0[\tanh(e_g/t)-e_g]. \tag{10}$$

At fixed u, we may solve for t given e_g:

$$t = 2e_g[\ln(1+F)/(1-F)]^{-1}, \tag{11}$$

where $F = e_g + (1+e_g)\exp(b-e_g^{\,2}/u)$, and invert the solution to find $e_g(t)$ at fixed u. Figure 1 shows the solutions $e_g(u)$ at various values of b and t. The resulting tunneling current, Eq. (10), is plotted in Fig. 2 for various values of b and t.

The Ohmic conductivity $ne\mu$ yields a current which has the same qualitative features as the tunneling current, as can be seen by use of Fig. 1 and the relation $n = \frac{1}{2}(1-e_g)$ of Eq. (1). It is however essential that the mobility μ be extremely small in the insulating state, else the Ohmic current will shunt out the applied field and eliminate the effect we are interested in observing.

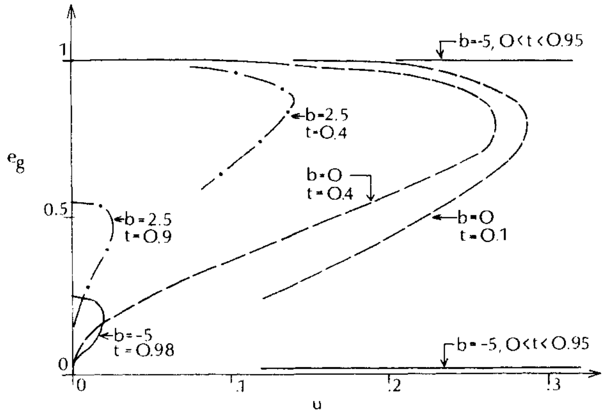

FIG. 1. Reduced-gap function versus applied electric field for various reduced temperatures t and figure-of-merit parameters b. For negative b at low temperature the curve is not continuous but has two branches —one near $e_g = 1$ and the other near $e_g = 0$. In the example shown, the $b = -5$ curve does not become continuous until t exceeds 0.95, whereupon it rapidly shrinks to zero (cf. plot shown for $t = 0.98$).

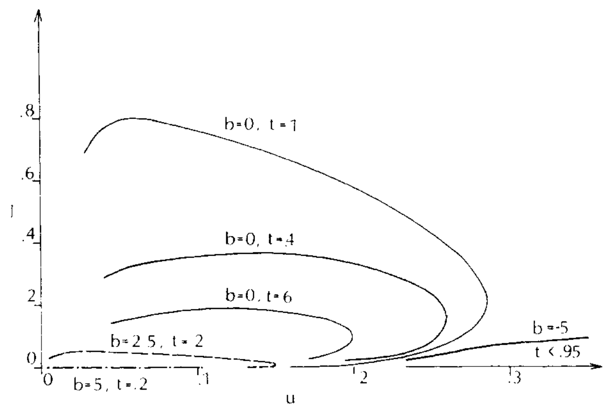

FIG. 2. Tunneling current in the insulating phase at various b and t. Note that the characteristics are double valued and for $b \geqslant 0$ have a maximum field, which depends on b and t. All curves shrink to a point at the origin at temperatures above $t = 1$. We take $\frac{1}{2}I_0 = 1$.

This may be the key to the success of the glass devices as opposed to the use of "good" semiconducting materials. The usual Mott insulators, transition-metal oxides and salts of all kinds, are also good candidates for the switching effect, owing to their generally low mobility.

In order to discuss the switching effect, it is also necessary to turn to the metallic phase. Let us first assume that the insulating phase becomes unstable, for reasons that will become apparent with study of Fig. 3, and that a large current is made to flow through the material. In the metallic phase internal fields and potentials are screened out, trapped carriers are liberated, and the mobility will in general be relatively high. The high-current characteristics of a semiconducting device, such as a p-n junction in reverse bias, obey the quasi-Ohmic law

$$j_{\text{cond}} = \sigma(u - u_h) \tag{12}$$

in which the parameter u_h, related to a threshold potential denoted as the "holding voltage" in Ref. 1, may be related to the avalanche breakdown of the insulating state or to some other parameter such as the excess free energy per carrier required to maintain the conducting state, assuming the insulating state to be the ground state. The metallic-state characteristic is shown in Fig. 3 as the straight line RQX, and the insulating characteristic is curve $OZQP$. Straight lines a, b, and c are load lines for three different applied voltages across a device in series with a fixed load resistance:

$$jR_L = u_{\text{batt}} - u. \tag{13}$$

Thus the intersection of the load line with the device curve yields the operating point of the circuit.

As the applied potential is increased from zero to u_b one follows the insulating solution (as indicated by point 1). Beyond u_b the solution jumps from point 4 to point 5 and then travels up the metallic characteristic as u is further increased. The jump from 4 to 5 should occur in a time not exceeding the inverse of the plasma frequency, i.e., very fast compared with the usual transit times in semiconductor phenomena. Upon decreasing the voltage, a metallic solution is retained until the load line intersects Q, whereupon a jump to the insulating characteristic at a point slightly to the right of point 1 is mandated. Thus portions of the curves RQ and QP are never observable, and $OZQX$ represents the observable characteristics.

Region ZQ has the property of negative resistance, and a properly designed circuit configuration (in Fig. 3, a bias slightly to the left of u_b) can be used for current, voltage, or power amplification. Evidently, the use of a high load resistance and a large bias voltage is most favorable since then the largest possible range of intersections with ZQ will be allowed.

In comparing the theory with experimental data on glass,[1] we find qualitative agreement, such as

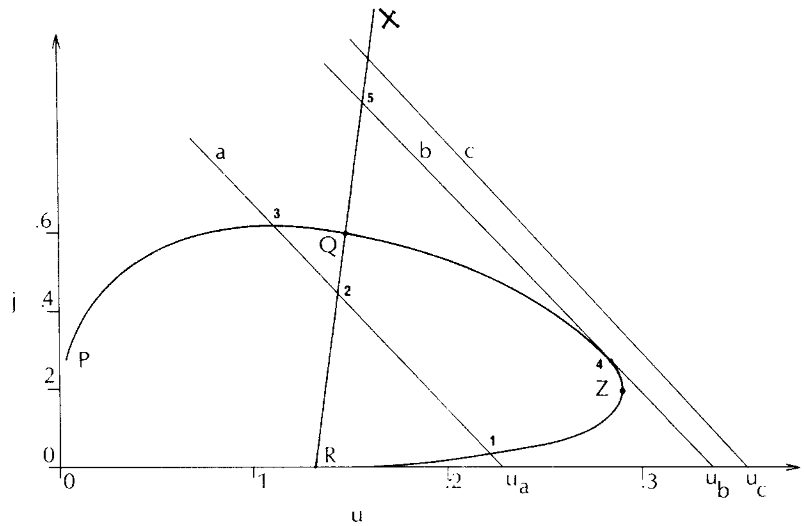

FIG. 3. I-V characteristics for $b = 0$ and $t = 0.2$. The scale of current is in units of $\frac{1}{2}I_0 = 1$. (It is assumed that the mobility in the conducting phase greatly exceeds that of the insulating phase.)

in the temperature dependence and the fact that the applied voltages must scale with film thickness. Quantitative agreement with the predicted threshold voltage (point Z) in the neighborhood of $u = 0.28$ for a $b = 0$ device at low temperature results if we assume $E_0 = \frac{1}{2}$ eV, $a = 50$ Å (corresponding to 10^{19} donors/cm^3), and a field of 3×10^5 V/cm, with the first two values being rough estimates and the last being the experimental value. The comparison of theory with experimental results in iron oxide are given in a companion article.[5] At this stage, several mysteries remain in the understanding of the phenomenon of polyconductivity. But if our partial explanation is correct, the dynamic properties of the cooperative insulator—the "Wigner electron lattice" or the "Mott insulator"—should prove an exciting object of study, theoretically as well as experimentally. Optical excitation of carriers, and the effects of applied magnetic fields, are but two aspects which now bear close scrutiny.

The author thanks Dr. S. Ovshinsky for the privilege of visiting his laboratory and is grateful to Professor L. Landovitz and Mr. M. Plischke for collaboration in the numerical calculations.

---

*Research supported by Air Force Office of Scientific Research, U. S. Air Force Grant No. AFOSR AF 69-1642.

[1]S. Ovshinsky, Phys. Rev. Letters 21, 1450 (1968). An earlier paper describing this effect is A. Lebedev et al., Zh. Tekh. Fiz. 26, 2131 (1956) [translation: Soviet Phys.–Tech. Phys. 1, 2071 (1957)], and there may exist a number of other papers predating Ovshinsky's publication. Nevertheless, he is clearly responsible for developing this invention and first "making it work."

[2]N. F. Mott, Rev. Mod. Phys. 40, 677 (1968) gives a review; other papers in the same issue of the journal are also devoted to the topic of the metal-nonmetal transition, although none of the authors seems to have considered the dynamic mechanisms of applied electric field, optical excitations, etc., as possible vehicles for achieving a Mott transition.

[3]D. Adler and H. Brooks, Phys. Rev. 155, 826 (1967). This work is reviewed and Eqs. (1)-(4) of the present Letter are introduced in S. Methfessel and D. Mattis, in Handbuch der Physik, edited by S. Flügge (Springer-Verlag, Berlin, Germany, 1969), Vol. 18A.

[4]C. Zener, Proc. Roy. Soc. (London), Ser. A 145, 523 (1934). A more correct formula for one-particle tunneling has $e_g^{3/2}$ in the exponent. Extensive computations with this exponent are now under way; however, the curves shown in the present article are very insensitive to the exact choice of exponent. This is fortunate, as the precise formula for many-body tunneling, which is undoubtedly occurring, is not yet known.

[5]M. Lipsicas and D. Mattis, to be published. We have found switching and negative resistance usuable at frequencies up to 1 MHz, in iron oxide. A search of the literature indicated that this effect was also found by J. E. Christopher, R. V. Coleman, Acar Isin, and R. C. Morris, Phys. Rev. 172, 485 (1968), and has been found in other transition-metal insulators: in vanadium oxide, by K. Van Steensel et al., Philips Res. Rept. 22, 170 (1967); in titanium oxide, by F. Argall, Solid-State Electron. 11, 535 (1968); in various transition-metal oxides, by K. Chopra, J. Appl. Phys. 36, 184 (1965); in niobium oxide, by D. Geppert, Proc. I.R.E. 51, 223 (1963). I wish to thank Dr. W. Doremus, Dr. J. Hall, and Dr. R. Gayley for providing some of these references.