John A. Swanson*

Notes on Cumulative Photovoltages

Abstract: It is reasoned that the large photovoltages observed by Cheroff and Keller on ZnS crystals are theoretically plausible and can result from a wide class of different mechanisms. Conversely, the conditions under which a photoconductor with periodic inhomogeneities does not show a cumulative photovoltage are shown to be very restrictive and improbable. General theorems concerning the magnitude of the photovoltage are proved, and remarks are made on its directionality.

Introduction

It has been suggested that the large photovoltages observed recently[1] in some insulators are due to numerous internal electrostatic barriers which under illumination act as *p-n* junctions connected in series. The crystal structure of these materials does not exhibit inversion symmetry, and it is conceivable that the directionality of the crystal induces directionality in the structure of the internal barriers. Such an intrinsic directionality is necessary, since a structure of randomly alternating conductivity type could provide no basis for a preferred direction in which the voltage accumulates.

The theory of such cumulative photovoltages can be reduced almost to a theory of single-junction photovoltages by the assumption of periodicity in the structure of the internal barriers. While such periodicity is not necessarily to be expected in the actual material, the assumption does not restrict a qualitative discussion. The reason that the theory must be somewhat broader than the single-junction theory is that at least two junctions must occur in a single period. Since the photovoltages to be obtained from these junctions are according to the usual theory opposed to one another, the theory must explain why the effect of one junction preponderates over that of the other. The reduced theory in fact contains the single-junction theory as a special case, in view of the fact that ohmic contacts may be replaced by zero-lifetime regions regarded as part of a periodic structure.

The generality of the discussion is limited by the assumption throughout much of the analysis of approximate thermal equilibrium in each band. However, results to the effect that cumulative photovoltages are almost always to be expected, for example, are not diminished in interest or generality by a restriction in the cases considered. The main purposes of the discussion are in fact (1) to make the preceding statement plausible, (2) to enumerate those special cases in which photovoltages are intrinsically impossible, (3) to show the decisive importance of recombination and generation mechanisms, (4) to deduce some simple inequalities governing the possible length of the fundamental period in actual cases. Sections V, VI, and IX, have related purposes. Section V among other things supports the third consideration just mentioned by showing that a cumulative voltage may be introduced in principle into a uniform material without internal barriers merely by a (directional) periodicity in the spatial distribution of the intensity of illumination. Section VI supplements a theorem of Section IV by tending to show that the condition of the theorem (one of the two carrier mobilities is zero) is not so pathological that it cannot be relaxed in continuous fashion. Section IX is intended as a complement to the discussion, in that some of the important ways in which the assumption of approximate thermodynamic equilibrium may break down are mentioned.

I. Analysis: directionality criterion

We have as our fundamental set of equations

$$j_p = -\sigma_p \nabla \phi_p, \ (\text{or} \nabla \phi_p = -\rho_p j_p), \tag{1}$$

$$j_n = -\sigma_n \nabla \phi_n \ (\text{or} \nabla \phi_n = -\rho_n j_n), \tag{2}$$

$$\nabla \cdot j_p = G, \tag{3}$$

$$j = j_p + j_n = 0, \tag{4}$$

where G is the net generation rate (generation g minus recombination r), j's are current densities, ϕ's are electrochemical potentials, σ's are (nonequilibrium) conductivities, ρ's are resistivities, and p and n refer to holes and electrons, respectively. The last equation implies one-dimensional geometry and open-circuit conditions. Equation (3) implies a steady state. Poisson's equation will not be needed. (There may be abrupt junctions in the material; at such junctions j_p and j_n must be continuous unless there exists a recombination mechanism

*Deceased.

This work has been taken in entirety, with only minor editorial corrections, from a document prepared by John Swanson for internal laboratory circulation in June 1959. Revisions by D. C. Mattis.

at the junction itself. Whether ϕ_p and ϕ_n are taken to be continuous or not is inessential to the discussion. One may in fact assume an infinite value for ρ_n, say, at a junction, but $\int_-^+ \rho_n dx$ across the junction must be finite, and $\delta\phi_n = \int_-^+ \phi_n dx = j_n \int_-^+ \rho_n dx$.)

We are assuming periodic conductivity variations. Let N be the number of periods contained in a sample. Then the quantity ΔV is defined as

$$\Delta V = \lim_{N\to\infty} \frac{V}{N},$$

where V is the measured voltage across the sample. In order that ΔV be different from zero, there must clearly be a directionality in the crystal and a recombination mechanism such that the establishment of equilibrium does not depend on end effects. In the ideal case of an infinite bar, conditions are strictly periodic except for an average nonvanishing inclination of the bands, provided (as we assume) the generation rate $g(x)$ is also periodic. In other words all quantities defined so far except ϕ_p and ϕ_n are periodic. Periodicity applies also to the concentration of carriers p and n as well as their mobilities μ_p and μ_n. These facts imply the periodicity of

$$\phi_p - \frac{\Delta V}{W}x,\ \phi_n - \frac{\Delta V}{W}x,\ \psi - \frac{\Delta V}{W}x,$$

where ψ is the electrostatic potential, and W is the length of a period. We have in fact

$$\Delta V = \int_0^W \nabla\phi_p dx = \int_0^W \nabla\phi_n dx = \int_0^W \nabla\psi dx, \tag{5}$$

in which the point $x=0$ is arbitrary. Our problem thus can be reduced to one within the interval $0 \leqq X < W$ with periodic boundary conditions, i.e., the point $X=W$ is congruent to the point $x=0$.

In view of Eq. (5) we may, upon integrating Eqs. (1) and (2) and applying Eq. (4), write

$$-(a+b)\Delta V = \int_0^W (a\rho_p - b\rho_n) j_p dx, \tag{6}$$

where a and b are arbitrary constants. We may derive various interesting relationships by choosing a and b properly. First note that by taking $b=-a$, we show that

$$\int_0^W (\rho_p + \rho_n) j_p dx = 0. \tag{7}$$

Since $(\rho_p+\rho_n)$ is positive definite, this condition implies that j_p and j_n must be oscillating functions, i.e., they change sign at least once in the fundamental interval.

We next define the average of a function f as $\langle f\rangle = \frac{1}{W}\int_0^W f dx$. Taking $a = \langle\rho_p\rangle^{-1}$, $b = \langle\rho_n\rangle^{-1}$, we have from Eq. (6)

$$-C\frac{\Delta V}{W} = \left\langle j_p \left(\frac{\rho_p}{\langle\rho_p\rangle} - \frac{\rho_n}{\langle\rho_n\rangle} \right) \right\rangle, \tag{8}$$

where $C=a+b$ is positive definite, and $\Delta V/W$ is the voltage per unit length of material.

Equation (8) becomes more interesting after suitable reinterpretation of j_p. We have in fact from (3) that

$$j_p(x) = \int_0^x G dx' + \text{constant}.$$

We define

$$F(x) = \int_0^x G dx'. \tag{9}$$

Then $F(x)$ is equal to $j_p(x)$ apart from a constant, and this constant is determined by the choice of origin. We may correctly write

$$-C\frac{\Delta V}{W} = \left\langle F \left(\frac{\rho_p}{\langle\rho_p\rangle} - \frac{\rho_n}{\langle\rho_n\rangle} \right) \right\rangle, \tag{10}$$

since this equation is unaffected by a constant term in F (the average of the term in parentheses is zero).

In view of the steady state condition $\langle G\rangle = 0$, F is a periodic function. In general it is also an oscillating function, and for a suitable choice of origin $\langle F\rangle = 0$.

Let $P \equiv \left(\frac{\rho_p}{\langle\rho_p\rangle} - \frac{\rho_n}{\langle\rho_n\rangle} \right)$. Then P is also an oscillating function. The sign of ΔV depends on the correlation between F and P. Naturally Eq. (10) does not constitute a solution of the problem. However, its form renders great plausibility to the idea that the vanishing of ΔV is not to be expected in general, since despite the constraint imposed by the fundamental equations, the functions F and P should be susceptible to variations that are to a large degree independent.

II. The sensitivity of ΔV to $g(x)$ and $r(p,n,x)$

As an example of the use of Eq. (10) we shall show in this section that given any initial band edge configuration (i.e., the equilibrium electrostatic potential $\psi_0(x)$ in the absence of illumination) extreme choices of either the generation function $g(x)$ or of recombination function $r(p,n,x)$ can make ΔV negative or positive, as desired. This fact means that the directionality of the initial electrostatic potential configuration has in a sense less to do with the type of photovoltage to be expected than the generation and recombination variations which may exist, and (especially) the correlation of these variations with the initial band edge configuration.

Something of this sort is rather obvious from the following consideration. Suppose that the initial configuration is a simple alternation of p-n and n-p junctions. Suppose further that the excess carrier lifetime in the vicinity of the p-n junctions is zero, while it is infinite in the neighborhood of the n-p junctions. Then clearly we obtain a photovoltage appropriate to n-p junction photocells connected in series. On the other hand a reversal of these conditions would produce the situation analogous to p-n junction photocells connected in series, i.e., the effect would be the same, except for a sign reversal. It should be noted that in this example no directionality of $\psi_0(x)$ is required.

We may similarly determine the sign of ΔV by proper choice of $g(x)$, given that the behavior of $r(p,n,x)$ does not conspire against us. Taking for $g(x)$ the delta-function $g(x)=\delta(x-x_0)$ we observe that $F(x)$ can

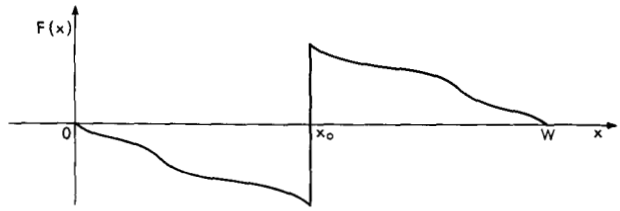

Figure 1

have only one sign change in the interval (see Fig. 1). The equilibrium resistivity function $P_0(x)$ in the case of simple alternating junctions can also be made to have only one sign change in the interval, but depending on the choice of interval $P_0(x)$ and $F(x)$ will correlate or anticorrelate. We may thus determine the sign of ΔV by proper choice of the point X_0. In more physical terms, by shining light exclusively in the neighborhood of the *p-n* or the *n-p* junctions we can obtain either sign of photovoltage. This intuitively evident fact may be of interest in connection with reversals in the sign of the photovoltage as a function of light frequency observed by Keller and Cheroff and others.[1,2] Different frequencies may be absorbed at different points in the fundamental period.

III. Intrinsic vanishing of ΔV

The arguments of this section are designed to show that one should expect a nonvanishing photovoltage *except* in very special cases. These cases fall into three classes; the vanishing of ΔV is respectively (a) due to symmetry, (b) accidental, (c) intrinsic. Class (a) refers to structures having no directionality[3] exposed to uniform illumination. We are naturally not concerned with such situations. Class (b) refers to structures possessing directionality, but such that any variation in the initial electrostatic potential $\psi_0(x)$ would with unit probability cause ΔV to "reappear." Class (c) refers to situations exhibiting a vanishing of ΔV independently of the variation of initial band-edge configurations.

To express this more precisely, ΔV is a function

$$\Delta V=\Delta V[\psi_0(x), \mu_p(p,n,x), \mu_n(p,n,x), r(p,n,x), T(x), g(x)],$$

where $T(x)$ is some distribution of traps. (If a variation of gap width is allowed we must explicitly add a gap-width function $E_g(x)$ and define $\psi_0(x)$ uniquely. The addition of such a function would not invalidate our discussion—indeed it presents us with a further and unnecessary degree of freedom.) Let $\delta\psi_0$ be a variation of $\psi_0(x)$. Then an intrinsic vanishing of ΔV is one such that both

$$\Delta V=0, \text{ and } \delta\Delta V/\delta\psi_0=0, \tag{11}$$

where the variational derivative is taken keeping all other functions the same. We further require as a subsidiary condition that conditions (11) continue to hold under replacement of $g(x)$ by $\alpha g(x)$, where α is a constant; i.e., an intrinsic vanishing of ΔV is by definition not dependent on the over-all integrated light intensity.

The general condition (11) for an intrinsic vanishing will be satisfied by a certain ordered set of functions μ_p, μ_n, r, T, and g. However we are not interested in this complete set, but only in the subset having some physical likelihood. Thus, for example, the condition

$$r(p,n,x)=r(p_0,n_0,x)=g(x),$$

is sufficient to make $\Delta V=0$, but has little probability of occurrence. We have seen above that ΔV is extremely sensitive to $r(p,n,x)$ and to $g(x)$. Although it may often be physically reasonable to assume $g(x)=$const., the function $r(p,n,x)$ will in general not depend on x vacuously. We search first therefore for a condition yielding an intrinsic disappearance of ΔV independent of g and of r.

- *The "strong" intrinsic case:* $\delta\Delta V/\delta r=\delta\Delta V/\delta g=0$

In view of the discussion of Section II, it is clear that we must have $F=$const. in Eq. (10) to insure that $\Delta V=0$ independent of variations of r and g. But $F=$ const. implies $j_p=$ const., and therefore (because of Eq. (7)) $j_p=0$. But j_p vanishes in general only if

$$\mu_p=0, \text{ or } \mu_n=0. \tag{12}$$

Eq. (12) constitutes the desired condition: both carriers must be mobile in order to make a photovoltage possible (with the fundamental assumption of thermal equilibrium in each band).

- *The "weak" intrinsic case:* $g=$const., $r=r(p,n)$

This is an interesting case, since uniform illumination may in many cases give rise to uniform generation and since recombination mechanisms independent of position exist. We argue again that although F and P are connected by the equations of the system, the average of their product is not fixed independently of the initial band edge configuration. Let us first assume $F=j_p=-j_n=0$. From Eqs. (1) and (2) it follows that

$$\phi_p=\text{const.},$$

$$\phi_n=\text{const.},$$

and therefore $\phi_p-\phi_n=$const. But it is easily checked that $\phi_p-\phi_n=$const. if and only if $p\cdot n=$const. Since in general neither p nor n will be constant individually, and since the dependence of r upon p and n is to be meaningful, we must have

$$r(p,n)=f(pn), \tag{13}$$

for the initial assumption $F=j_p=j_n=0$ to be self-consistent. In particular bimolecular recombination (direct recombination across the gap) satisfies this condition, namely

$$r(p,n)=kpn,$$

where k is a constant independent of position in the usual case of constant band gap. Deviations from the law (13) will allow for weak photovoltages, depending on the initial band edge configuration.

It is not likely that there exist any other intrinsic cases. Although we cannot exclude the possibility that "matched" functions $r(p,n,x)$ and $g(x)$, say, might lead to $\Delta V=0$ independently of $\psi_0(x)$, this is unlikely, the more so in view of our subsidiary condition. Furthermore, such cases, should they exist, would appear to be artificial from the physical viewpoint.

IV. Theorems involving $\Delta V=0$

We can generalize some of the results of the preceding section to nonperiodic structures. In the following two theorems we are to imagine a piece of material of no special shape upon which two ohmic contacts A and B are placed. Theorem II requires a single semiconductor of constant band gap—otherwise there are no restrictions, except approximate thermal equilibrium in each band and steady-state conditions. (Note that Eq. (4) does not necessarily hold; we have instead $\nabla \cdot j=0$.)

- *Theorem I*

The open-circuit voltage A between contacts A and B is zero under illumination unless both carriers are mobile.

Proof: Suppose $\mu_p=0$ but $V\neq 0$. Then $j_p=0$, and $\nabla \cdot j_n=0$. Either $j_n\equiv 0$, or $j_n\not\equiv 0$ somewhere. If $j_n\equiv 0$, then according to Eq. (2) $V=-\int_A^B \nabla\phi_n \cdot ds=0$, contrary to hypothesis. If $j_n\not\equiv 0$ somewhere, then, since there are no sources of j_n within the material nor on the free surfaces, it is possible to follow "lines of j_n" from A to B, such that the sign of j_n does not change along these lines. Let us integrate $\nabla\phi_n \cdot ds$ along one of these lines. Since ρ_n is positive definite, we see from Eq. (2) that the sign of ΔV is the same as that of j_n along the line. However since the total current from contact A is zero there must exist lines having both signs of current. Thus $V=0$.

- *Theorem II*

If the recombination rate is a function of pn alone (not of position) and if the generation rate is uniform in space, $\lim_{d\to\infty} \frac{V}{d}=0$ where d is the distance between contacts A and B.

Proof: We construct a solution of the fundamental set of equations by taking $\phi_p=$const., $\phi_n=$const. (except in the proximity of the contacts). Then choose $\phi_p-\phi_n$ such that $r(p,n)=f(pn)=g$ (recall that pn must be constant if $\phi_p-\phi_n$ is constant). We have now satisfied the equation $\nabla \cdot j=0$, since in fact $j=j_p=j_n=0$. The degree of freedom left over, namely, the height of the average $\bar{\phi}=\frac{\phi_n+\phi_p}{2}$ with respect to the band edges, is just sufficient to enable us to satisfy Poisson's equation; in fact the form of Poisson's equation is not essentially different from that obtaining at equilibrium, for which condition a solution must exist.

We now prove a third theorem rather unrelated to the material thus far presented. It will provide a basis for the following section.

- *Theorem III*

Let two ohmic contacts A and B be placed upon a piece of material having no internal fields at equilibrium (no impurity gradations or interior dipole layers). There may, however, exist lifetime variations. Let this material be exposed to weak, but not necessarily uniform, radiation. Then, if V is the open-circuit voltage between the contacts and if I is the integrated intensity of radiation, $dV/dI=0$ at $I=0$, i.e., there is no photovoltage of first order in the light intensity.

Proof: To first order in the case of light intensity

$$j_p=-\sigma_{0p}\nabla\phi_p, \qquad j_n=-\sigma_{0n}\nabla\phi_n,$$

where σ_{0p} and σ_{0n} are the unilluminated conductivities.

Thus $j=-\nabla\phi$ where $\phi\equiv\sigma_{0p}\phi_p+\sigma_{0n}\phi_n$.

But since $\nabla j=0$, $\nabla^2\phi=0$. $\phi=$const. is a (unique) solution having $j=0$ on all surfaces and interfaces. But since contacts A and B are ohmic $\phi_p=\phi_n$ at A and B. Thus $\phi_p(B)-\phi_p(A)=0$. But this quantity is equal to V. We note that this theorem is without meaning when the dark conductivities σ_{0n} and σ_{0p} are very small. Therefore we shall now consider higher-order effects.

V. Materials with no internal fields

It has been seen from the preceding sections, and as is well known, advantage can be taken of pre-existing electrostatic fields in a material to separate light-created carriers, and thus to produce photovoltages. It has also been noted that there is no limit to the cumulative value of a photovoltage in an alternating structure. A natural and rather interesting question now arises; namely, is it possible to produce arbitrarily large photovoltages in a structure having no pre-existing electrostatic fields by the device of nonuniform generation or recombination? Specifically, let $0\leqq x<W$ again be a fundamental interval, with $x=W$ congruent to $x=0$. Let $(g)x$ and $r(p,n,x)$ be arbitrary, but $\psi_0(x)=0$. Is it possible that ΔV need not vanish?

That there can be no first-order effect is clear from Theorem III. However two kinds of higher-order effects may exist, as we shall show.

Consider first a nonuniform distribution of electron traps such that the trapping levels are above the Fermi level initially. Irradiation causes these traps to become filled, and the space charge these created gives rise to electric fields. These electric fields may separate mobile charges in the usual way, causing photovoltages.

The second effect we shall discuss is bound to exist if there is any directionality in $g(x)$ or $r(p,n,x)$ or both (excluding the accidental cancelling of the two nonuniformities). It is due to the space charge of mobile carriers required to compensate for the mobility difference of holes and electrons, and though small in one interval, may of course be significant through many intervals. However we are not concerned to calculate possible magnitudes of the effect, but merely to show its existence.

For the same reason we specialize to an intrinsic sample, i.e., one for which the equilibrium concentrations n_0 and p_0 are equal. We assume no other change of charge density than that due to the mobile carriers. Poisson's equation then takes the simple form

$$\varepsilon\psi''=n-p,$$

where (14)

$$\psi''=d^2\psi/dx^2.$$

We shall expand solutions in powers of ε. To zeroth order in ε

$$n=p, \tag{15}$$

and we find on setting $j_n=-j_p$ that

$$E=V_t\frac{1-b}{1+b}\frac{p'}{p}, \tag{16}$$

where V_t is the "thermal" voltage kT/e which enters into the Einstein relation, $b=\mu_n/\mu_p$, $p'=dp/dx$, $E=-d\psi/dx=-\psi'$. This a well known formula in the semiconductor field. In deriving it we have used the more familiar variants of Eqs. (1) and (2),

$$j_p=e\mu_p(pE-V_Tp'), \tag{1'}$$

$$j_n=e\mu_n(nE+V_Tn'). \tag{2'}$$

The average $<E>$ is clearly equal to zero, as one might expect, so that the first possible contribution to ΔV comes from the first-order solution. Since we are seeking merely to show the existence of an effect and not to obtain a solution, we may without loss of generality set the currents j_p and j_n to first order in ε equal to zero (the *zeroth* order currents of course do not vanish). This is equivalent to regarding the currents as *prescribed*, and taking them equal to the zeroth order currents. This is a consistent procedure, since there exists *some* correct current distribution in any problem. The problem then becomes: given $j_p(x)$ what is the solution $\psi(x)$ expanded in powers of ε? If we can show that $\psi(W)\neq\psi(0)$ for a $j_p(x)$ possessing "directionality," we will have proven what we set out to prove since certainly a directional j_p can be produced by a direction g or r.

Proceeding then, on substituting $j_{p'}=j_{n'}=0$ in equations (1') and (2') and using the condition that $p-n=0$ to zeroth order in ε, we find

$$0=p_1E+pE_1-V_Tp'_1, \tag{17}$$

$$0=n_1E+pE_1+V_Tn'_1, \tag{18}$$

where the subscript "1" refers to terms of first order in ε, and lack of subscript indicates a zeroth order term. Subtracting these equations we have

$$V_T(p_1+n_1)'=(n_1-p_1)E=0,$$

or

$$V_T(p_1+n_1)'+\varepsilon E'E=0,$$

and

$$V_T(p_1+n_1)=-\frac{\varepsilon}{2}E^2+C. \tag{19}$$

Adding Eqs. (17) and (18), we have

$$(p_1+n_1)E+2pE_1+V_T(n_1-p_1)'=0,$$

or

$$(p_1+n_1)E+2pE_1+V_T\varepsilon E''=0,$$

and

$$E_1=-\frac{(p_1+n_1)E+\varepsilon V_TE''}{2p}, \tag{20}$$

$$E_1=\varepsilon\frac{E^3-2EC-2V_T^2E''}{4V_Tp}.$$

Now

$$\left\langle\frac{E}{p}\right\rangle=\left\langle\frac{p'}{p^2}\right\rangle=0.$$

Thus

$$<E_1>=\frac{\varepsilon}{4V_T}\left[\left\langle\frac{E^3}{p}\right\rangle-2V_T^2\left\langle\frac{E''}{p}\right\rangle\right].$$

This equation can be further simplified by substitution of the expression (16) for E and subsequent integrations by parts, down to

$$\frac{\Delta V_1}{W}=\varepsilon V_T^2\frac{b(1-b)}{(1+b)^3}\left\langle\frac{p'^3}{p^4}\right\rangle. \tag{21}$$

It can be easily checked by simple trial distributions $p(x)$ having directionality that in general the average occurring in Eq. (21) is not zero. Note that since p' is of first order in the light intensity, $\frac{\Delta V_1}{W}=-<E_1>$ is of third order. Thus no contradiction exists between our results here and Theorem III.

Remark: The generalization to nonintrinsic samples of Eq. (16) is

$$\psi'=(1-b)\frac{p'}{b(n_0-p_0)+(1+b)p}, \tag{16'}$$

which still implies a functional relationship between ψ and ρ. Thus, in the zeroth order approximation (with respect to ε) there is no possibility of a periodic growth of ψ in a sample of uniform mobilities and carrier concentrations, irrespective of generation and recombination mechanisms. These facts may be regarded as well known.

In view of Theorem III the so-called Dember Effect (creation of the field ψ' by nonuniform irradiation) cannot be measured with ohmic probes to first order in the light intensity. In view of Eq. (16') and the boundary condition $p=p_0$ at ohmic contacts the zero order (in ε) Dember voltage vanishes between ohmic contacts, at least in linear geometrics. Equation (16') is of course also generalizable by replacing the derivative d/dx (represented by the primes) by the gradient, providing only that $j=0$ everywhere, so that this conclusion applies also to more general geometries.

VI. Solution[4] for $\mu_p/\mu_n << 1$

We can approach this case by expanding the solution in powers of μ_p. The zeroth order solution is clearly that applying to the case of strong intrinsic vanishing of ΔV. We have then to zeroth order

$$0=j_n=e\mu_n(-n\psi'+V_T n')$$

$$\psi'=V_T n'/n. \tag{22}$$

The generation and recombination are everywhere equal. Therefore

$$r(p,n,x)=g(x).$$

Presuming we can solve this equation for p, we have

$$p=s(n,x,g(x)). \tag{23}$$

Poisson's equation reads (assuming no traps)

$$\varepsilon\psi''=(n-\rho)+\bar{\rho}_0(x), \tag{24}$$

where $\bar{\rho}_0(x)$ is $\bar{\rho}_0(x)\equiv-\rho_0(x)-(n_0(x)-p_0(x))$,

where $e\rho_0(x)$ is the equilibrium charge distribution and $n_0(x)$ and $p_0(x)$ are the equilibrium carrier concentrations. Combining Eqs. (24) and (22) we find

$$\varepsilon V_T\left(\frac{n'}{n}\right)'+s(n,x,g(x))-n=0, \tag{25}$$

as the equation determining $n(x)$. This is a nonlinear second-order equation to be solved with periodic boundary conditions.

The *first-order* current $j_{n'}(x)$ is now easily found from the equation for hole current

$$j_{n'}(x)=-j_{p'}(x)=e\mu_p(p\psi'+V_T p'). \tag{26}$$

We have

$$\Delta V_1=-\int_0^W \frac{j_{n'}(x)}{\mu_n(x)n(x)}dx=\frac{\mu_p}{\mu_n}\int_0^W \frac{\rho\psi+V_T p'}{n}dx,$$

which, upon an integration by parts and substitution of (22), becomes

$$\Delta V_1=-2V_T\frac{\mu_p}{\mu_n}\int_0^W \frac{pn'}{n^2}dx.$$

Use of Eq. (25) reduces this to

$$\Delta V_1=\varepsilon V_T{}^2\frac{\mu_p}{\mu_n}\int_0^W \frac{1}{n}\left(\frac{n'}{n}\right)',$$

and an integration by parts to

$$\Delta V_1=\varepsilon V_T{}^2\frac{\mu_p}{\mu_n}\left\langle\frac{n'^3}{n^4}\right\rangle. \tag{27}$$

Substitution of the solution of Eq. (25) into (27) allows us to calculate ΔV. It should be noted that Poisson's equation plays an essential role in this calculation, since setting $\varepsilon=0$ in Eq. (27) gives $\Delta V_1=0$. Observe that despite the quite opposed points of departure Eq. (21) becomes identical to Eq. (27) under the substitution of n for p in Eq. (21) in the limit of small b.

VII. A classical problem

In order to counteract the possible misleading impression given by the two preceding sections to the effect that photovoltages depend on $\varepsilon\neq 0$ it may be sufficient to point out that in drawing the relevant conclusions from Eq. (16′) (after Eq. (21)) in Section V it was necessary to assume that p_0 and n_0 were constants. However, to round out the discussion we recall to the reader's mind the "classical" discussion of junction photovoltages which dispenses entirely with Poisson's equation, and which we shall base almost entirely on Eq. (10).

Suppose that a p-n junction occurs at $x=0$, and that $-\frac{W}{2}<x<\frac{W}{2}$ is the fundamental interval. For simplicity we make the junction entirely symmetrical about the Fermi level, with equal mobilities for the two carriers. Let $r=0$ in the region $-d<x<d$, where $d<\frac{W}{2}$, (and all equivalent regions). Let the excess carrier lifetime be zero elsewhere. This is equivalent to placing ohmic contacts at $x=\pm d$. This implies that $F=0$ except in the interval $-d<x<d$. Let $g(x)$ be $A\delta(x)$, where $\delta(x)$ is the Dirac δ-function. Then, multiplying by $<\rho_p>=<\rho_n>$ in Eq. (10),

$$\Delta V=-\frac{A}{2}\int_0^d \rho_p dx.$$

Since the material from 0 to $W/2$ is assumed to be n-type, we have, providing the illumination is not too intense, $\rho_p>>\rho_n$, and thus

$$\Delta V=-\frac{A}{2}\int_0^d \rho_p dx.$$

To first order in the intensity of illumination, then

$$\Delta V=-\frac{A}{2}\int_0^d \rho_{0p} dx,$$

where ρ_{0p} is the equilibrium value of $\rho_p(x)$. It is clear, however, that $\int_0^d \rho_p dx$ decreases with illumination. Let us try to obtain an idea of the variation of ρ_p. In the vicinity of the junction the difference $\phi_p-\phi_n$ reaches its maximum and is approximately equal to ΔV. Recall that

$$n_p=n_0p_0e^{\frac{\phi_p-\phi_n}{V_T}}. \tag{28}$$

Next to the junction on the n-type side the fractional change in p is much larger than that in n, and we are justified in setting

$$\frac{\rho_p}{\rho_{0p}}=\frac{p_0}{p}=e^{\frac{\Delta V}{V_T}}.$$

The function $\rho_p(x)$ is as yet known only at $x=d$ and near the junction. Its exact nature is important in determining the value of the integral. If $p_0(x)=$ const., $0<x<d$, it is usually reasonable to suppose Eq. (1′) is linear (predominantly diffusion current), and we obtain

upon integration

$$\frac{\Delta V}{W}=\frac{Ad\rho_0}{2W}\frac{\Delta V/V_T}{\exp(\Delta V/V_T)-1}. \tag{29}$$

Thus

$$\Delta V=V_T\ln\left[1+\frac{Ap_0d}{2V_T}\right], \tag{30}$$

and the voltage is a slowly increasing function of the light intensity.

Note that nothing has been said about the short-circuit current—in general the increase in conductivity just compensates the saturation of $d\Delta V/dA$ (eventually ΔV itself saturates, but we shall not discuss this effect), so that the current remains proportional to A. The latter phenomenon is to be expected, since if the electrostatic potential configuration is more or less "stabilized" by short-circuiting the terminals, the ability of the existing junctions to separate positive and negative light-induced charges should remain virtually unimpaired over a large range of light intensity.

VIII. The short-circuit current: upper and lower limits on the period

The first goal of this section is to deduce an equation analogous to Eq. (10) for the short-circuited case. In a sense this case is aesthetically superior to the open-circuit situation in that *all* quantities are periodic, including ϕ_p, ϕ_n, and ψ. We start by defining new quantities

$$j_+=\frac{1}{2}(j_p+j_n)=\frac{j}{2}, \tag{31}$$

$$j_-=\frac{j}{2}(j_p-j_n), \tag{32}$$

in terms of which

$$j_p=j_++j_-,$$

$$j_n=j_+-j_-.$$

As equations replacing Eqs. (3) and (4) we have

$$\nabla\cdot j_-=G, \tag{33}$$

$$j_+=\text{const.} \tag{34}$$

We shall assume a total length L, composed of N periodic structures (or cells) of length W.

In view of the periodicity of ϕ_p and ϕ_n we have from Eqs. (1) and (2)

$$\int_0^W\rho_pj_pdx=\int_0^W\rho_nj_ndx=0. \tag{35}$$

Thus

$$0=\int_0^W[a\rho_pj_p+b\rho_nj_n]dx$$

$$=\int_0^W(a\rho_p+b\rho_n)j_+dx+\int_0^W(a\rho_p-b\rho_n)j_-dx. \tag{36}$$

Choose $a=<\rho_p>^{-1}$, $b=<\rho_n>^{-1}$. Then

$$j_+=\left\langle F\left(\frac{\rho_p}{<\rho_p>}-\frac{\rho_n}{<\rho_n>}\right)\right\rangle, \tag{37}$$

or

$$j=-2<FP>. \tag{38}$$

Eq. (37) is the short-circuit analogue of Eq. (10). (Naturally $<FP>$ need not have the same value for open-circuited as for closed-circuited conditions.)

The measured short-circuit current permits us to place a minimum on the possible length of a cell. Let N_e be the total short-circuit current in a sample divided by the electronic charge (magnitude), and let N_p be the total number of photons (in the range of energies exciting hole-electron pairs) absorbed by the entire sample. Let N_{pc} be the total number of photons absorbed by a cell (fundamental period). Let

$$Q\equiv N_e/N_p, \tag{39}$$

and

$$Q_c\equiv N_e/N_{pc}, \tag{40}$$

be the effective quantum efficiencies per sample and per cell, respectively. Let N be the number of cells in a sample. Clearly

$$N=Q_c/Q=Q_cN_p/N_e. \tag{41}$$

Q_c denotes the number of electron charges passing through a fundamental period per photon incident on the period. Most reasonable physical models limit Q_c to a maximum value 1, obtained by letting the electron from one period recombine with the hole from the next period. Invoking this criterion yields

$$N\leq N_p/N_e. \tag{42}$$

If L is the length of a sample and W the length of a cell, it follows that

$$W\geq N_eL/N_p. \tag{43}$$

In terms of experimental values, let J be the current in amperes, L the length in centimeters, I the total power contained in a uniform monochromatic light beam in watts, f the fraction intercepted by the sample, E_q the energy of one quantum in electron volts; then

$$W\geq\frac{|J|E_q}{If}\text{L}. \tag{44}$$

It would be desirable also to have an experimental upper limit on W. Such a limit is set by the equation

$$W<V_G\frac{L}{V}, \tag{45}$$

where V_G is the forbidden gap width in electron volts, and L/V is the reciprocal of the average electric field. Eq. (45) follows from the inequality

$$\Delta V<V_G. \tag{46}$$

It is often possible to strengthen inequality (46) by replacing V_G by V_b, where V_b is the height of the barrier

effective in separating holes and electrons, if V_b is known or can be estimated. Equations (44) and (45) place upper and lower limits on the number of fundamental periods which can exist in a given crystal.

IX. The stacking-fault model: dipole junctions

Our discussion would not be complete without some mention of how the preceding work applies (or does not apply) to dipole junctions. Nevertheless we specifically wish to avoid involvement with the details of various mechanisms that have been informally proposed at this laboratory in connection with dipole junctions to explain the observations by Cheroff and Keller on sign reversals of ΔV, et cetera.

By a dipole junction we mean an abrupt junction caused by some dipole layer associated with an interface in the crystal, such as a stacking fault. The principal characteristic of a dipole junction which differentiates it from other junctions is that it cannot be removed by rearrangement of free carriers or by changes in the occupation of impurity states. The appearance of the band edges in an insulator possessing such stacking faults might be something like configuration *A* before illumination. Illumination by producing free carriers shortens the Debye length, and the appearance changes to configuration *B*. (See Fig. 2.)

There is an important fact to be noted about the configurations *A* and *B:* the tendency of the electric field is always to sweep the electrons to the right and holes to the left, except at isolated points. This means that it is *a priori* more difficult to produce photocurrents flowing to the right. Diffusion currents (of the type needed) in Case *B* would be weak except very near the junction. However another mechanism is available to make current flow to the right. Consider electron-hole pairs created with excess energy within a mean free path of the dipole junction; electrons created on the right of a junction may possibly pass to the left, holes created on the left may possibly pass to the right—but the partner particle is stopped. This is an example of a mechanism depending essentially on the departure from thermal equilibrium in each band.

Another possible mechanism depending on departure from thermal equilibrium, but producing electron current in the opposite direction, is the direct excitation of electrons over the dipole barrier from the adjacent "trough." This mechanism could create a current even though the hole mobility were zero.

This brief discussion serves to illustrate the limitations of the foregoing treatment. However one point emphasized in that treatment remains valid; namely, the great importance of the recombination mechanism. Consider,

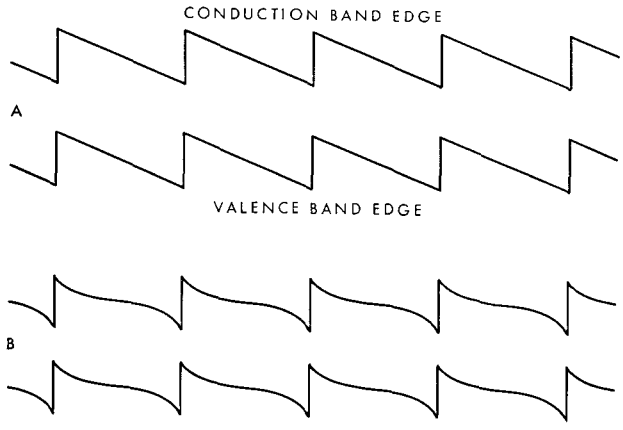

Figure 2

for example, the polarity of the normal short-circuit current in two cases (*a*) and (*b*).

Case (*a*): Recombination only at stacking faults.

Case (*b*): No recombination at stacking faults.

The recombination referred to is that of electrons in the trough next to the stacking fault with holes in the corresponding inverted trough to the right of the dipole layer. By normal current we mean that produced by generation across the gap in the bulk of the material. Then in Case (*a*) the positive current flows to the left, in Case (*b*) to the right (if at all).

Acknowledgment

The writer wishes to acknowledge the invaluable contribution of G. Cheroff and S. Keller, the discoverers of the effect; the introduction to stacking faults given him by R. Landauer; and illuminating conversations with R. C. Casella, D. C. Mattis, and R. Landauer.

Editor's note

While not specifically cited in any of the arguments, the reader's attention is directed to the author's closely related earlier work "Clarification of First-Order Semiconduction Effects Through Use of Electrochemical Potentials," *IBM Journal* **1,** 39 (1957).

References and footnotes

1. G. Cheroff and S. P. Keller, *Phys. Rev.* **111,** 98 (1958).
2. W. J. Merz, *Helv. Phys. Acta* **31,** 625 (1958).
3. A structure or function of position has "directionality" if there exists no reflection symmetry about any point on the x-axis.
4. Similar considerations have been undertaken by D. C. Mattis and R. C. Casella.

Received February 25, 1960