Journal of Applied Physics

Detailed Balance Limit of Efficiency of *p-n* Junction Solar Cells

Cite as: Journal of Applied Physics **32**, 510 (1961); https://doi.org/10.1063/1.1736034
Submitted: 03 May 1960 . Published Online: 11 June 2004

William Shockley, and Hans J. Queisser

Detailed Balance Limit of Efficiency of *p-n* Junction Solar Cells*

WILLIAM SHOCKLEY AND HANS J. QUEISSER
Shockley Transistor, Unit of Clevite Transistor, Palo Alto, California

(Received May 3, 1960; in final form October 31, 1960)

In order to find an upper theoretical limit for the efficiency of *p-n* junction solar energy converters, a limiting efficiency, called the *detailed balance limit* of efficiency, has been calculated for an ideal case in which the only recombination mechanism of hole-electron pairs is radiative as required by the principle of detailed balance. The efficiency is also calculated for the case in which radiative recombination is only a fixed fraction f_c of the total recombination, the rest being nonradiative. Efficiencies at the matched loads have been calculated with band gap and f_c as parameters, the sun and cell being assumed to be blackbodies with temperatures of 6000°K and 300°K, respectively. The maximum efficiency is found to be 30% for an energy gap of 1.1 ev and $f_c=1$. Actual junctions do not obey the predicted current-voltage relationship, and reasons for the difference and its relevance to efficiency are discussed.

1. INTRODUCTION

MANY papers have been written about the efficiency of solar cells employing *p-n* junctions in semiconductors, the great potential of the silicon solar cell having been emphasized by Chapin, Fuller, and Pearson[1] in 1954. Also in 1954, Pfann and van Roosbroeck[2] gave a more detailed treatment including analytic expressions optimizing or matching the load. A further treatment was given by Prince[3] in 1955, in which the efficiency was calculated as a function of the energy gap. Loferski[4] has attempted to predict the dependence of efficiency upon energy gap in more detail. Review papers have recently appeared in two journals in this country.[5,6]

The treatments of efficiency presented in these papers are based on empirical values for the constants describing the characteristics of the solar cell.[7] They are in general in fairly good agreement with observed efficiencies, and predict certain limits. These predictions have become generally accepted as theoretical limits (see, for example, the review articles by Rappaport[5] and Wolf[6]).

It is the view of the present authors that the acceptance of this previously predicted limiting curve of efficiency vs energy gap is not theoretically justified since it is based on certain empirical values of lifetime, etc. We shall refer to it as the *semiempirical limit.*

There exists, however, a theoretically justifiable upper limit. This limit is a consequence of the nature of atomic processes required by the basic laws of physics, particularly the principle of detailed balance. In this paper this limit, called the *detailed balance limit,* is calculated and compared with the semiempirical limit in Fig. 1. Actually the two limits are not extremely different, the detailed balance limit being at most higher by about 50% in the range of energy gaps of chief interest. Thus, to some degree, this article is concerned with a matter of principle rather than practical values. The difference is much more significant, however, insofar as estimating potential for improvement is concerned. In fact, the detailed balance limit may lie more than twice as far above the achieved values as does the semiempirical limit, thus suggesting much greater possible improvement (see Fig. 1).

The situation at present may be understood by analogy with a steam power plant. If the second law of thermodynamics were unknown, there might still exist quite good calculations of the efficiency of any given configuration based on heats of combustion, etc. However, a serious gap would still exist since it would be impossible to say how much the efficiency might be improved by reduction of bearing friction, improving heat exchangers, etc. The second law of thermodynamics provides an upper limit in terms of more fundamental quantities such as the temperature of the exothermic reaction and the temperature of the heat sink. The merit of a given power plant can then be appraised in terms of the limit set by the second law.

A similar situation exists for the solar cell, the missing theoretical efficiency being, of course, in no way comparable in importance to the second law of thermodynamics. Factors such as series resistance and reflection losses correspond to friction in a power plant. There are even two temperatures, that of the sun T_s and that of the solar cell T_c. The efficiency of a solar converter can in principle be brought to the thermodynamic limit $(T_s-T_c)/T_c$ by using reflectors, etc.[8] However, a planar solar cell, without concentrators of radiation, cannot approach this limit. The limit it can approach depends on its energy gap and certain geometrical factors such as the angle subtended by the sun and the

* Research supported by Wright Air Development Center.

[1] D. M. Chapin, C. S. Fuller, and G. L. Pearson, J. Appl. Phys. 25, 676 (1954).
[2] W. G. Pfann and W. van Roosbroeck, J. Appl. Phys. 25, 1422 (1954).
[3] M. B. Prince, J. Appl. Phys. 26, 534 (1955).
[4] J. J. Loferski, J. Appl. Phys. 27, 777 (1956).
[5] P. Rappaport, RCA Rev. 20, 373 (1959).
[6] M. Wolf, Proc. I.R.E. 48, 1246 (1960).
[7] A treatment of photovoltage, but not solar-cell efficiency free of such limitations, has been carried out by A. L. Rose, J. Appl. Phys. 31, 1640 (1960).
[8] H. A. Müser, Z. Physik 148, 380 (1957), and A. L. Rose (see footnote 7) have used the second law of thermodynamics in their treatments of photovoltage.

angle of incidence of the radiation, and certain other less basic degrading factors, which in principle may approach unity, such as the absorption coefficient for solar energy striking the surface.

Among the factors which may approach unity, at least so far as the basic laws of physics are concerned, is the fraction of the recombination between holes and electrons which results in radiation. Radiative recombination sets an upper limit to minority carrier lifetime. The lifetimes due to this effect have been calculated using the principle of detailed balance.[9] It is this radiative recombination that determines the detailed balance limit for efficiency.[10] If radiative recombination is only a fraction f_c of all the recombination, then the efficiency is substantially reduced below the detailed balance limit.

How closely any existing material can approach the desirable limit of unity for f_c is not known. Existing silicon solar cells fail to fit the current-voltage characteristics predicted on the basis of any of the existing recombination models.[11] The extent of this discrepancy and one suggested explanation are discussed in Sec. 6.

In determining the detailed balance limit of efficiency, the efficiency η calculated below is defined in the usual way as the ratio of power delivered to a matched load to the incident solar power impinging on the cell. The following sections present a step-by-step calculation of this efficiency as a function of the essential variables, including several which may reduce the efficiency below the detailed balance limit. Three of these variables have the dimensions of energy and can be expressed as temperatures, voltages or frequencies. These variables are: the temperature of the sun T_s,

$$kT_s=qV_s; \tag{1.1}$$

the temperature of the solar cell T_c,

$$kT_c=qV_c; \tag{1.2}$$

and the energy gap E_g,

$$E_g=h\nu_g=qV_g, \tag{1.3}$$

where k is Boltzmann's constant, $q=|q|$ is the electronic charge, and h is Planck's constant. The efficiency is found to involve only the two ratios

$$x_g=E_g/kT_s \tag{1.4}$$

$$x_c=T_c/T_s. \tag{1.5}$$

The efficiency also depends strongly upon t_s, which is

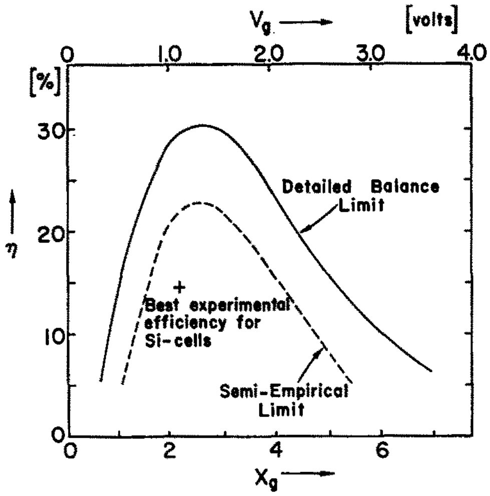

FIG. 1. Comparison of the "semiempirical limit" of efficiency of solar cells with the "detailed balance limit," derived in this paper. + represents the "best experiment efficiency to date" for silicon cells. (See footnote 6.)

defined as

t_s = the probability that a photon with $h\nu>E_g$ incident on the surface will produce a hole-electron pair. (1.6)

For the detailed balance efficiency limit to be reached, t_s must be unity.

Other parameters involving transmission of radiative recombination out of the cell and the solid angle subtended by the sun enter as factors in a quantity f discussed in Eq. (3.20). The value of f for the highest efficiency, corresponding to the detailed balance limit, is determined by the solid angle subtended by the sun, the other factors related to material properties being given their maximum values, which are unity.

To a very good approximation the efficiency is a function $\eta(x_g, x_c, t_s, f)$ of four variables just discussed. It can be expressed in terms of analytic functions based on the Planck distribution and other known functions. The development of this relationship is carried out in Secs. 2–5. Section 6 compares calculations of the detailed balance limit with the semiempirical limit.

2. ULTIMATE EFFICIENCY: $u(x_g)$

There is an ultimate efficiency for any device employing a photoelectric process which has a single cutoff frequency ν_g.

We shall consider a cell in which photons with energy greater than $h\nu_g$ produce precisely the same effect as photons of energy $h\nu_g$, while photons of lower energy will produce no effect. We shall calculate the maximum efficiency which can be obtained from such a cell subjected to blackbody radiation.

Figure 2 (a) illustrates an idealized solar cell model which we shall consider in this connection. It represents a p-n junction at temperature $T_c=0$, surrounded by a blackbody at temperature T_s. In a later discussion

[9] W. van Roosbroeck and W. Shockley, Phys. Rev. 94, 1558 (1954).

[10] A preliminary report of the analysis of this paper was presented at the Detroit meeting of the American Physical Society: H. J. Queisser and W. Shockley, Bull. Am. Phys. Soc. Ser. II, 5, 160 (1960).

[11] This discrepancy appears to have been first emphasized by Pfann and van Roosbroeck (see footnote 2), who point out that the forward current varies as $\exp(qV/AkT)$ with values of A as large as three.

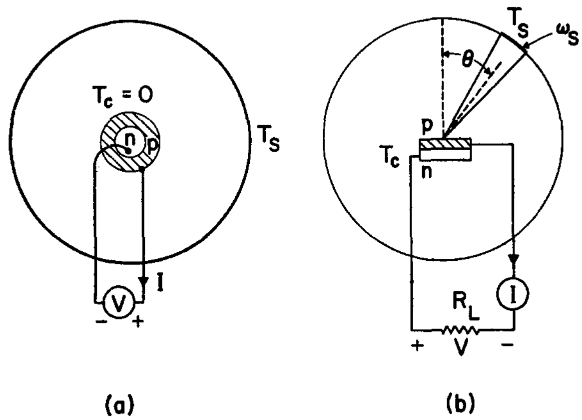

FIG. 2. Schematic representation of the solar battery considered. (a) A spherical solar battery surrounded by a blackbody of temperature T_s; the solar battery is at temperature $T_c=0$. (b) A planar cell irradiated by a spherical sun subtending a solid angle ω_s at angle of incidence θ.

we shall allow a finite T_c and replace the surrounding body at temperature T_s by radiation coming from the sun at a small solid angle ω_s as represented in Fig. 2 (b). We shall assume that some means not indicated in the figure are present for maintaining the solar cell at temperature $T_c=0$ so that only steady state conditions need be considered. According to the ultimate efficiency hypothesis[12]:

> Each photon with energy greater than $h\nu_g$ produces one electronic charge q at a voltage of $V_g=h\nu_g/q$. (2.1)

The number of photons incident from the solar radiation in Fig. 2 is readily calculated in accordance with the formulas of the Planck distribution. We denote by Q_s the number of quanta of frequency greater than ν_g incident per unit area per unit time for blackbody radiation of temperature T_s. For later purposes we shall also introduce the symbol $Q(\nu_g,T_s)$ in order to be able to represent situations for different values of the limiting frequency. In accordance with this notation and well-known formulas, we have

$$Q_s \equiv Q(\nu_g,T_s)=(2\pi/c^2)\int_{\nu_g}^{\infty}[\exp(h\nu/kT_s)-1]^{-1}\nu^2 d\nu$$
$$=[2\pi(kT_s)^3/h^3c^2]\int_{x_g}^{\infty}x^2dx/(e^x-1), \quad (2.2)$$

in which the symbol x_g is that of Eq. (1.4),

$$x_g kT_s = h\nu_g = qV_g. \quad (2.3)$$

Q_s is seen to be a function of the form T_s^3 times a function of x_g.

If the surface subject to the radiation in Fig. 2 has an area A, then in accordance with the ultimate efficiency hypothesis, the output power will be given by:

$$\text{output power} = h\nu_g A Q_s. \quad (2.4)$$

The incident power, due to the radiation at T_s falling upon the device of Fig. 2, will evidently be:

$$\text{incident power} = AP_s. \quad (2.5)$$

P_s is the total energy density falling upon unit area in unit time for blackbody radiation at temperature T_s. In accordance with well-known formulas for the Planck distribution, P_s is given by

$$P_s = 2\pi h/c^2\int_0^{\infty}\nu^3 d\nu/[\exp(h\nu/kT_s)-1]$$
$$=2\pi(kT_s)^4/h^3c^2\int_0^{\infty}x^3dx/(e^x-1)$$
$$=2\pi^5(kT_s)^4/15h^3c^2. \quad (2.6)$$

It is instructive to compare P_s with the total number of incident photons per unit time $Q(0,T_s)$ so as to obtain the average energy per photon:

$$P_s=\left\{\left[\int_0^{\infty}x^3dx/(e^x-1)\right]\Big/\int_0^{\infty}x^2dx/(e^x-1)\right\}$$
$$\times[kT_sQ(0,T_s)]$$
$$=[3!\zeta(4)/2!\zeta(3)][kT_sQ(0,T_s)]$$
$$=[(3\pi^4/90)/(\pi^3/25.794\cdots)][kT_sQ(0,T_s)]$$
$$=2.701\cdots kT_sQ(0,T_s). \quad (2.7)$$

The integrals in Eqs. (2.6) and (2.7), one of which is the limiting form for $x=0$ in Eq. (2.2), may be expressed by products of the gamma function and the Riemann zeta function. The mathematical relations involved and numerical values are found in standard references.[13]

In accordance with the above definitions, the ultimate efficiency is a function only of x_g and is

$$u(x_g)=h\nu_gQ_s/P_s$$
$$=\left[x_g\int_{x_g}^{\infty}x^2dx/(e^x-1)\right]\Big/\int_0^{\infty}x^3dx/(e^x-1). \quad (2.8)$$

[12] Once a photon exceeds about three times the energy gap E_g, the probability of producing two or more hole-electron pairs becomes appreciable: V. S. Vavilov, J. Phys. Chem. Solids **8**, 223 (1959), and J. Tauc, J. Phys. Chem. Solids **8**, 219 (1959). These authors interpret this result in terms of a threshold of about $2E_g$ for an electron to produce a pair. However, the data can be well fitted up to quantum yields greater than two by assuming a threshold of only slightly more than E_g and assuming the energy divides equally between the photohole and the photoelectron. This effect would slightly increase the possible quantum efficiency; however, we shall not consider it further in this article. See also W. Shockley, Solid State Electronics **2**, 35 (1961).

[13] For example: I. M. Ryshik and I. S. Gradstein, *Tables of Series, Products and Integrals* (Deutscher Verlag d. Wissensch., Berlin, 1957), pp. 149, 413; E. Jahnke and F. Emde, *Tables of Functions*, (Dover Publications, New York, 1945) 4th ed., pp. 269, 273.

It is evident that $u(x_g)$ has a maximum value, since the numerator in Eq. (2.8) is finite and vanishes both as x_g approaches zero and as it approaches infinity.[14] Figure 3 shows the dependence of $u(x_g)$ as a function of x_g.[15] It is seen that the maximum efficiency is approximately 44% and comes for an x_g value of 2.2 in terms of a temperature of 6000°K for the sun. This corresponds to an energy gap value given by Eq. (2.3) of 1.1 ev.

3. CURRENT-VOLTAGE RELATIONSHIP FOR A SOLAR CELL

In this section we shall consider a solar cell subjected to radiation from the sun, which is considered to subtend a small solid angle as represented in Fig. 2 (b). The treatment will be based upon determining the steady state current-voltage condition which prevails on the basis of requiring that hole-electron pairs are eliminated as rapidly as they are produced. In order to carry out the calculation, five processes must be considered: (1) generation of hole-electron pairs by the incident solar radiation, the rate for the entire device being F_s; (2) the radiative recombination of hole-electron pairs with resultant emission of photons, the rate being F_c; (3) other nonradiative processes which result in generation and (4) recombination of hole-electron pairs; and (5) removal of holes from the p-type region and electrons from the n-type region in the form of a current I which withdraws hole-electron pairs at a rate I/q. The steady state current-voltage relationship is obtained by setting the sum of these five processes equal to zero.

Consider first the net rate of generation of hole-electron pairs for the solar battery of Fig. 2 (a) under the condition in which it is surrounded by a blackbody at its own temperature, $T_c \neq 0$. Under these conditions photons with frequencies higher than ν_g will be incident per unit area per unit time on the surface at a rate Q_c, where $Q_c = Q(\nu_g, T_c)$ as given by Eq. (2.2). Evidently Q_c is a function of the form T_c^3 times a function of x_g/x_c. The number of these photons which enter the cell and produce hole-electron pairs is represented by F_{c0}, where

$$F_{c0} = A t_c Q_c \equiv A t_c Q(\nu_g, T_c). \tag{3.1}$$

In this expression, t_c represents the probability that an incident photon of energy greater than E_g will enter the body and produce a hole-electron pair. A is the area of the body.

The total rate of generation of hole-electron pairs due to the solar radiation falling upon the body is given by

$$F_s = A f_\omega t_s Q_s, \tag{3.2}$$

in which the factor f_ω is a geometrical factor, taking into account the limited angle from which the solar energy falls upon the body. t_s is the probability that incident photons will produce a hole-electron pair and may differ from t_c because of the difference in the spectral distribution of the blackbody radiation at temperature T_c and T_s, and the dispersion of the reflection coefficient or transmission coefficients for the surface of the battery.

The geometrical factor f_ω is dependent upon the solid angle subtended by the sun and the angle of incidence upon the solar battery. The solid angle subtended by the sun is denoted by ω_s, where

$$\begin{aligned}\omega_s &= \pi(D/L)^2/4 = \pi(1.39/149)^2/4 \\ &= 6.85\times 10^{-5}\ \text{sr},\end{aligned} \tag{3.3}$$

and D, L are, respectively, the diameter and distance of the sun, taken as 1.39 and 149 million km.

If the solar cell is isotropic (i.e., is itself a sphere) then it is evident that f_ω should be simply the fraction of the solid angle about the sphere subtended by the sun, so that

$$f_\omega = \omega_s/4\pi. \tag{3.4}$$

If the cell is a flat plate with projected area A_p, then it is more natural to deal with incident energy on the basis of the projected area A_p rather than the total

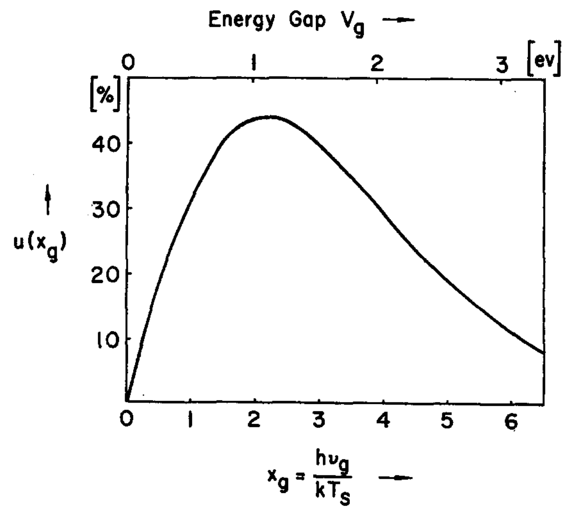

FIG. 3. Dependence of the ultimate efficiency $u(x_g)$ upon the energy gap V_g of the semiconductor.

[14] For the calculations, numerical tables of the integrals involved were used as given by K. H. Böhm and B. Schlender, Z. Astrophysik 43, 95 (1957). We are indebted to A. Unsöld who directed our attention to this publication. A convenient aid to such calculations is a slide rule, manufactured by A. G. Thornton, Ltd., Manchester, England. It is described by W. Makowski, Rev. Sci. Instr. 20, 884 (1949).

[15] Similar conclusions have been reached by H. A. Müser, Z. Physik, 148, 385 (1957), who estimates approximately 47% for the maximum of $u(x_g)$, but does not show a curve. Results similar to those described above have also been derived by W. Teutsch, in an internal report of General Atomic Division of General Dynamics, and by H. Ehrenreich and E. O. Kane, in an internal report of the General Electric Research Laboratories. A curve which is quantitatively nearly the same has also been published, since the submission of this article, by M. Wolf (see footnote 6) who defines the ordinate as "portion of sun's energy which is utilized in pair production," a definition having the same quantitative significance but a different interpretation from our quantity $u(x_g)$.

area of both sides, which is $2A_p$. In terms of A_p the total power falling on the cell is:

$$\text{incident power}=A_pP_s\omega_s\cos\theta/\pi, \tag{3.5}$$

where θ is the angle between the normal to the cell and the direction of the sun. This expression integrates as it should to A_pP_s when ω_s is integrated over a hemisphere (2π steradians) since $\cos\theta$ has an average value of $\frac{1}{2}$. For normal incidence, the incident power is thus:

$$\text{incident power}=A_pP_sf_\omega, \tag{3.6}$$

where

$$f_\omega=\omega_s/\pi=2.18\times10^{-5}. \tag{3.7}$$

It is evident that the rate of generation of hole-electron pairs by solar photons involves the same factor so that this value of f_ω should be used in Eq. (3.2). The blackbody radiation from the cell comes from an area of $2A_p$ so that

$$F_{c0}=2A_pt_cQ_c. \tag{3.8}$$

The rate of recombination, with resultant radiation, of hole-electron pairs depends upon the disturbance from equilibrium. For the case in which the battery is in equilibrium, and is surrounded by a blackbody at temperature T_c, the rate of emission of photons due to recombination must be exactly equal to the rate of absorption of photons which produce recombination. As discussed above, this is given by F_{c0} in Eq. (3.1). To begin with, we shall consider that the only radiative recombination of importance is direct recombination between free holes and electrons and is accordingly proportional to the product of the hole and electron density, i.e., to the product np. When this product is equal to the thermal equilibrium value n_i^2, the rate of recombination will be F_{c0}. Accordingly we may write for F_c, the rate of radiative recombination throughout the cell,

$$F_c(V)=F_{c0}np/n_i^2=F_{c0}\exp(V/V_c), \tag{3.9}$$

in which V represents the difference in imrefs or quasi-Fermi levels for holes and electrons, and the product np is proportional to the Boltzmann factor for this difference expressed as a voltage.[16] V is evidently the voltage between the terminals connected to the p- and n-regions of the solar cell[17]; V_c stands for kT_c/q.

The net rate of increase of hole-electron pairs involves, in addition to generation, corresponding to F_s, and recombination, corresponding to F_c, nonradiative processes and removal of hole-electron pairs by current to the external circuit. The nonradiative recombination and generation processes are represented by $R(V)$ and $R(0)$ respectively. They will be equal for $V=0$, the thermal equilibrium condition. The algebraic sum of the rates of increase of hole-electron pairs must vanish for the steady state condition. This leads to

$$\begin{aligned}0&=F_s-F_c(V)+R(0)-R(V)-I/q\\&=F_s-F_{c0}+[F_{c0}-F_c(V)+R(0)-R(V)]-I/q.\end{aligned} \tag{3.10}$$

In Eq. (3.10) the quantity in square brackets represents the net rate of generation of hole-electron pairs when the cell is surrounded by a blackbody at temperature T_c. If the cell is so surrounded, it is evident that the term F_s-F_{c0} vanishes. The steady state condition, under these circumstances, gives the current-voltage characteristic of the cell in the absence of a disturbance in the radiation field. On the other hand, if the cell is surrounded by cold space, it will generate a small open-circuit reverse voltage due to the $-F_{c0}$ term.

In order to describe the current-voltage characteristics of the cell we introduce the quantity f_c, which represents the fraction of the recombination-generation current which is radiative. This leads to

$$F_{c0}-F_c(V)=f_c[F_{c0}-F_c(V)+R(0)-R(V)]. \tag{3.11}$$

For the particularly simple case, which occurs in germanium p-n junctions, that the nonradiative recombination fits the ideal rectifier equation, we can write

$$R(V)=R(0)\exp(V/V_c). \tag{3.12}$$

For this condition f_c is a constant independent of voltage, and is given by

$$f_c=F_{c0}/[F_{c0}+R(0)]. \tag{3.13}$$

Under these conditions the current-voltage characteristic for the cell in the absence of radiative disturbance is given by

$$I=I_0[1-\exp(V/V_c)], \tag{3.14}$$

where

$$I_0\equiv q[F_{c0}+R(0)] \tag{3.15}$$

is the reverse saturation current.

It is noted that this equation differs in sign from the usual rectifier equation, the convention chosen in this paper being that current flowing into the cell in what is normally the reverse direction is regarded as positive, and voltage across the cell in the normally forward direction is also regarded as positive. These are the polarities existing when the illuminated cell is furnishing power to an external load, so that positive values of I and V correspond to the cell acting as a power source.

For an energy gap of 1.09 ev and a temperature of 300°K, Q_c is 1.7×10^3 cm^{-2} sec^{-1}. Thus per cm^2 of surface the recombination current is 2.7×10^{-16} amp. For a planar cell with $t_c=1$ radiating from both sides this leads to a contribution to I_0 of 5.4×10^{-16} amp/cm^2 of junction area, according to Eq. (3.8). As discussed in Sec. 6, actual cells have currents larger by about 10 orders of magnitude, so that $f_c\approx10^{-10}$.

In the event that $R(V)$ does not obey Eq. (3.12), then the quantity f_c must be regarded as a function of

---

[16] For example: W. Shockley, *Electrons and Holes in Semiconductors* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1950) p. 308; the product of Eqs. (18) and (19).

[17] See footnote 16, p. 305; also W. Shockley, Bell System Tech. J. **28**, 435 (1949), Sec. 5.

the voltage, so that I_0 must be regarded as voltage dependent.

The current-voltage relationship for the cell when subjected to radiant energy may be obtained by solving Eq. (3.10) for I. This leads to[18]:

$$\begin{aligned} I &= q(F_s - F_{c0}) + (qF_{c0}/f_c)[1-\exp(V/V_c)] \\ &= I_{sh} + I_0[1-\exp(V/V_c)], \end{aligned} \tag{3.16}$$

in which the symbol I_{sh} represents the short circuit current corresponding to $V=0$, and for the case of a planar cell of projected area A_p is given by

$$\begin{aligned} I_{sh} &\equiv q(F_s - F_{c0}) = qA_p(f_\omega t_s Q_s - 2t_c Q_c) \\ &\approx qF_s \equiv qA_p(f_\omega t_s Q_s). \end{aligned} \tag{3.17}$$

The last form in Eq. (3.17) corresponds to the approximation that in most conditions of interest the solar energy falling upon the body produces hole-electron pairs at a rate that is so much larger than would blackbody radiation at the cell's temperature that the latter term can be neglected in comparison with the former.

The open-circuit voltage V_{op} which the cell would exhibit is obtained by solving Eq. (3.16) for the case of $I=0$. This leads to

$$\begin{aligned} V_{op} &= V_c \ln[(I_{sh}/I_0)+1] \\ &= V_c \ln[(f_c F_s/F_{c0}) - f_c + 1]. \end{aligned} \tag{3.18}$$

This particular solution is valid for the case in which R depends upon voltage as given in Eq. (3.12). Otherwise, Eq. (3.18) will contain the open-circuit voltage in the term I_0 on the right side of the equation.

As discussed in connection with Eq. (3.17), for most cases of interest the solar energy falling on the cell will be very large compared to blackbody radiation at the temperature of the cell, and accordingly the terms which do not involve F_s in Eq. (3.18) can be neglected in comparison, as long as f_c is not too small. This leads to the approximate result

$$V_{op} \cong V_c \ln(f_c f_\omega t_s Q_s/2t_c Q_c) = V_c \ln(fQ_s/Q_c), \tag{3.19}$$

in which it is seen that the geometrical and transmission factors together with the effect of excess recombination over radiative recombination may be lumped together in the single expression f, where

$$f \equiv f_c f_\omega t_s/2t_c. \tag{3.20}$$

The factor 2 comes from the fact that sunlight falls on only one of the two sides of the planar cell. [See Eq. (3.8).]

It is thus evident that as far as open-circuit voltage is concerned, similar results are produced by any of the four following variations: (1) reducing the efficiency of transmission of solar photons into the cell; (2) reducing the solid angle subtended by the sun, or (3) its angle of incidence upon the solar cell; or (4) introducing additional nonradiative recombination processes, thus making smaller the fraction of the recombination which is radiative.

The maximum open-circuit voltage which may be obtained from the cell, in accordance with the theory presented, is the energy gap V_g. This occurs as the temperature of the cell is reduced towards zero. Under these circumstances the quantity Q_c tends towards zero and the logarithm in Eq. (3.19) to large values. The limiting behavior can be understood by noting that in accordance with Eq. (2.2) we have

$$\begin{aligned} -\ln Q_c &= h\nu_g/kT_c + \text{order of } \ln T_c \\ &= V_g/V_c + \text{order of } \ln T_c. \end{aligned} \tag{3.21}$$

The terms which are of the order $V_c \ln T_c$ vanish as T_c and V_c approach zero in Eq. (3.19), so that

$$\lim(V_c \to 0) V_{op} = V_g. \tag{3.22}$$

At higher temperatures the voltage is only a fraction of V_g. This fraction[19] denoted by v may be expressed as a function of three of the four variables discussed in the introduction for the important case in which the last two terms in the ln term of Eq. (3.18) can be neglected. The necessary manipulations to establish this relationship are as follows:

$$\begin{aligned} v(x_g,x_c,f) &\equiv V_{op}/V_g = (V_c/V_g)\ln(fQ_s/Q_c) \\ &= (x_c/x_g) \text{ function of } (x_g, x_c \text{ and } f) \end{aligned} \tag{3.23}$$

$$Q_s/Q_c = x_c^{-3}\int_{x_g}^{\infty}\cdots \Big/ \int_{x_g/x_c}^{\infty}\cdots, \tag{3.24}$$

where the integrands are each that of Eq. (2.2). [For cases of very low illumination in which the approximation of Eq. (3.19) would not hold, v also depends explicitly on f_c.]

In the following two sections we shall consider expressions for the output power in terms of the open-circuit voltage and short circuit currents just discussed.

4. NOMINAL EFFICIENCY

For the geometrical configuration represented in Fig. 2 (b), the incident power falling from the sun upon the solar cell may evidently be written in the form

$$P_{inc} = f_\omega A P_s = A f_\omega h\nu_g Q_s/u(x_g), \tag{4.1}$$

in which Eq. (2.8) has been used to introduce the ultimate efficiency function $u(x_g)$ for purposes of simplifying subsequent manipulations.

A nominal efficiency can be defined in terms of the

[18] Equations like (3.16) occur in published treatments of solar-cell efficiency. The difference is that the term in I_{sh} due to F_{c0}, which is small but required by the principle of detailed balance, is included, and the coefficient of I_0 is related to the fundamental minimum reverse saturation current rather than to a semiempirical value.

[19] As for Eq. (3.16), factors like v have been introduced by various authors, most recently by M. Wolf (see footnote 6). However, the forms are dependent upon additional semiempirical quantities so that they cannot be used for the purposes given in the introduction.

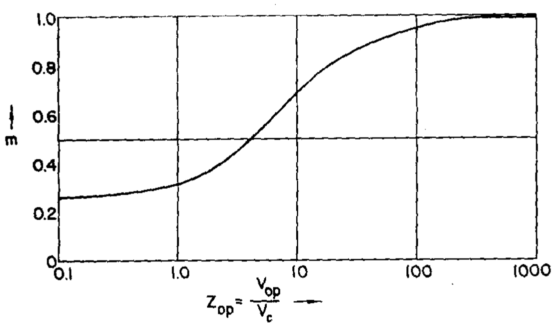

FIG. 4. Relationship between the impedance matching factor m and the open circuit voltage of a solar cell.

incident power and the product of the open-circuit voltage V_{op} and the short circuit current I_{sh}. The actual efficiency will be somewhat lower since the current-voltage characteristic is not perfectly rectangular. We shall consider the problem of matching the impedance in the following section.

The nominal efficiency in terms of open-circuit voltage and short circuit current is evidently given by

$$\begin{aligned} V_{op}I_{sh}/P_{inc} &= V_{op}Aqf_\omega t_sQ_s/[Af_\omega h\nu_gQ_s/u(x_g)] \\ &= (V_{op}/V_g)u(x_g)t_s \\ &= v(x_g,x_c,f)u(x_g)t_s, \end{aligned} \tag{4.2}$$

in which the symbol v is the ratio of Eq. (3.23) of the open-circuit voltage V_{op} to the ultimate voltage V_g that could be obtained if the battery were at zero temperature.

5. DETAILED BALANCE LIMIT OF EFFICIENCY AND $\eta(x_g,x_c,t_s,f)$

The maximum power output from the solar battery is obtained by choosing the voltage V so that the product IV is a maximum. In accordance with the current-voltage relationship, Eq. (3.16), and Eq. (3.18) for the open-circuit voltage, the current-voltage relationship may be rewritten in the form

$$\begin{aligned} I &= I_{sh}+I_0-I_0\exp(V/V_c) \\ &= I_0[\exp(V_{op}/V_c)-\exp(V/V_c)]. \end{aligned} \tag{5.1}$$

The maximum power occurs when[20]:

$$\begin{aligned} &d(IV)/dV=0 \\ &I_0\{\exp(V_{op}/V_c)-[(V+V_c)/V_c]\exp(V/V_c)\}=0. \end{aligned} \tag{5.2}$$

This equation may be conveniently rewritten by introducing the symbols

$$z_{op}=V_{op}/V_c=vx_g/x_c,\quad z_m=V(\max)/V_c, \tag{5.3}$$

in which $V(\max)$ is the voltage which satisfies Eq. (5.2). Substituting the symbols introduced in Eq. (5.3) into Eq. (5.2) leads readily to the relationship

$$z_{op}=z_m+\ln(1+z_m). \tag{5.4}$$

This gives the functional relation between the open-circuit voltage and the voltage at which maximum power is obtained. In effect it establishes a functional relationship between z_m and z_{op}, and thus between z_m and the variables f, x_c and x_g.

It is seen that the open-circuit voltage is always larger than the voltage for maximum output, and when both voltages are small compared to thermal voltage V_c, then Eq. (5.4) leads to a maximum power voltage equal to one-half the open-circuit voltage, the situation corresponding to a battery with an ohmic internal resistance. On the other hand, when either z_{op} or z_m is large compared to unity, then the ratio between the two approaches unity.[20]

The maximum power is smaller than the nominal power $I_{sh}V_{op}$ by the impedance matching factor m, where m is given by

$$\begin{aligned} m &= I[V(\max)]V(\max)/I_{sh}V_{op} \\ &= z_m^2/(1+z_m-e^{-z_m})[z_m+ln(1+z_m)] \\ &= m(vx_g/x_c)=m(x_g,x_c,f). \end{aligned} \tag{5.5}$$

Figure 4 shows the dependence of m upon z_{op} obtained by computing pairs of values of $m(z_m)$ and $z_{op}(z_m)$ for various values of z_m. The limits of m are 0.25 and 1.0 for small and large values of z_{op}.

In terms of m the efficiency η can now be expressed as a function of the four variables x_g, x_c, t_s, and f introduced in Sec. 1. The detailed balance limit corresponds to setting $t_s=1$ and $f=f_\omega/2$. The efficiency η may be written as

$$\begin{aligned} \eta(x_g,x_c,t_s,f) &= I[V(\max)]V(\max)/P_{inc} \\ &= t_su(x_g)v(f,x_c,x_g)m(vx_g/x_c). \end{aligned} \tag{5.6}$$

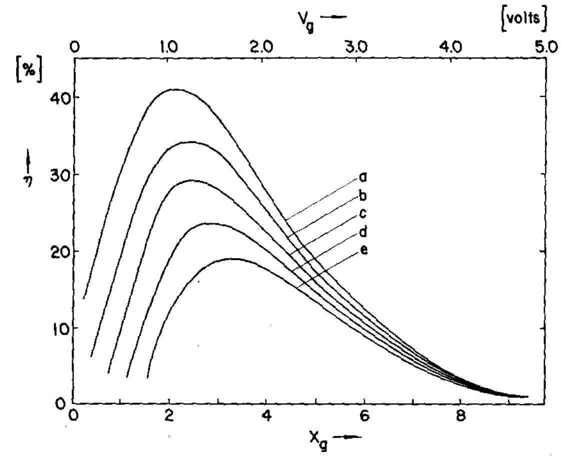

FIG. 5. Efficiency η for a blackbody solar cell at $T_c=300°$K, with sun at $T_s=6000°$C, as a function of energy gap for different values of the parameter f: curve (a) $f=1$; (b) $f=10^{-3}$; (c) $f=10^{-6}$; (d) $f=10^{-9}$; (e) $f=10^{-12}$.

[20] Similar maximization of the output power has been carried out in terms of the maximum area rectangle on the I-V plot by various authors, in particular W. G. Pfann and W. van Roosbroeck (See footnote 2). The results do not, however, appear to have been published in analytic form in which the matching factor m is shown to be a function solely of the variable $z_{op} = V_{op}/V_c=v(x_g,x_c,f)(x_g/x_c)$.

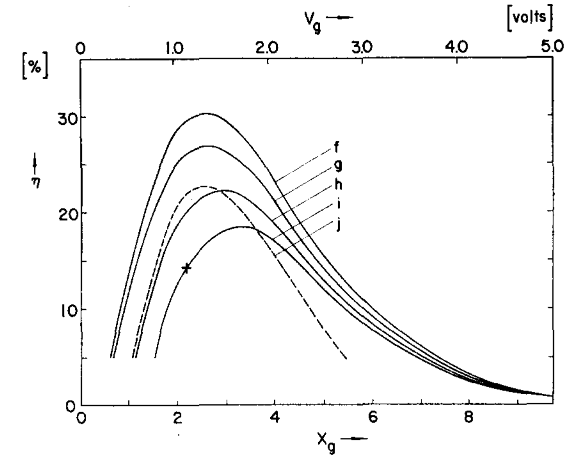

FIG. 6. Efficiency η for a solar cell at temperature $T_c=300°K$ exposed to a blackbody sun at temperature $T_s=6000°K$. Curve (f) is the detailed balance limit of efficiency, assuming the cell is a blackbody (i.e., $t_s=t_c=1$). Curve (j) is the semiempirical limit, or limit conversion efficiency of Prince (see footnote 3). + represents the "best experimental efficiency obtained to date" for Si (see footnote 6). Curves (g), (h), and (i) are modified to correspond to 90% absorption of radiation (i.e., $t_s=t_c=0.9$) and 100-mw incident solar energy. The values for the f quantities discussed in Sec. 6 are: (f) $f=1.09\times10^{-5}$ ($f_\omega=2.18\times10^{-5}$, $f_c=1$) $t_s=t_c=1$; (g) $f=0.68\times10^{-5}$ ($f_\omega=1.36\times10^{-5}$, $f_c=1$) $t_s=t_c=0.9$; (h) $f=0.68\times10^{-8}$ ($f_\omega=1.36\times10^{-5}$, $f_c=10^{-3}$) $t_s=t_c=0.9$; (i) $f=0.68\times10^{-11}$ ($f_\omega=1.36\times10^{-5}$, $f_c=10^{-6}$) $t_s=t_c=0.9$.

To summarize, the efficiency is defined as the electrical power out of the cell into a matched load, divided by the incident solar energy falling on the cell. The factors in Eq. (5.6) are as follows: t_s is the probability, averaged over incident solar photons having sufficient energy to produce electron-hole pairs, that a photon will produce an electron-hole pair; $u(x_g)$ is the ultimate efficiency in accordance with Eq. (2.1); $v(x_g,x_c,f)$ is the ratio of the open-circuit voltage to the energy gap of the cell; and $m(vx_g/x_c)$ is the impedance matching factor, which is a function of the ratio of the open-circuit voltage to thermal voltage for the cell.

In the following section results of calculation of η are presented and compared with the semiempirical limit of efficiency.

6. CONCLUSIONS

Efficiencies computed on the basis of Eq. (5.6) are shown in Figs. 5 and 6 as a function of V_g, based upon the following values for the parameters:

$$x_c=T_c/T_s=0.05 \quad (T_s=6000°K,\ T_c=300°K)$$
$$x_g=qV_g/kT_s=1.94V_g. \qquad (6.1)$$

Curves are given for different values of the parameter f. Figure 5 shows the decrease of the efficiency with f lowered from its maximum value 1 by factors of 10^{-3}. For Fig. 6, f was calculated according to its definition (3.20). Curve (f) corresponds to a perfectly absorbing cell ($t_c=t_s\equiv1$) with normal incidence ($f_\omega=\omega_s/\pi$) and no nonradiative recombination ($f_c=1$). Curves (g), (h) and (i) are calculated on the assumption of 90% absorption with a value of 1.36×10^{-5} for f_ω, which gives 100 mw/cm² incident power, and different values of f_c.

Also on Fig. 6 is shown the generally accepted curve for the "limit conversion efficiency." This curve is in agreement with Prince[3] and Loferski[4] as reported by Wolf.[6] Also shown is the value of 14% for silicon solar cells, which Wolf[6] reports as the best achieved to date.

On the basis of the semiempirical limit it would appear that silicon solar cells might be improved by from 14 to 21.7%, an improvement of a factor of 1.55. On the basis of the detailed balance limit, the improvement might be 14 to 26% [curve (g)], or a factor of 1.9. The true physical limit for silicon must lie somewhere between these two limits.

Figure 7 shows the current voltage characteristics for several silicon solar cells. These are discussed further in the Appendix. The figure also shows the minimum forward current characteristic for a planar cell with $t_c=1$ as discussed in Sec. 3.

Somewhere between the empirical curves and the limit set by detailed balance is a true limit determined basically by the fact that silicon is element 14 in the Periodic Table and has a certain rate of "unavoidable"[21] nonradiative transitions.

On the basis of the preceding paragraphs two questions are obvious: (1) Where is the true physical limit and what processes determine it? (2) What determines the location of the present experimental curves?

It is evident that question (1) will be difficult if not impossible to answer before question (2) is answered. We shall first discuss question (2).

It has been noted by many writers[2,4,22] on solar-cell

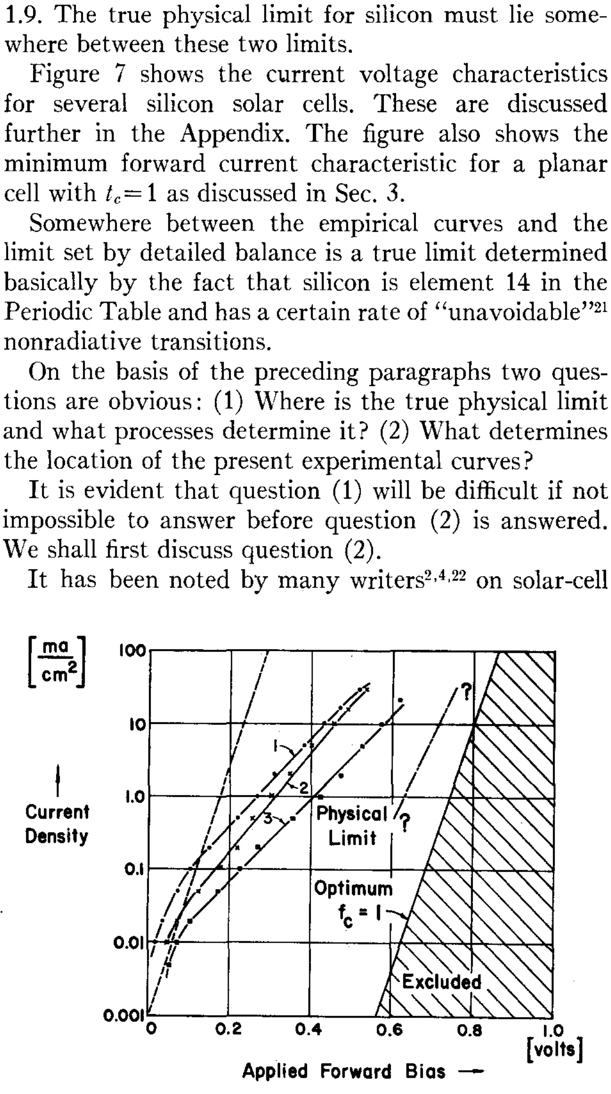

FIG. 7. Current-voltage relationships at room temperature for silicon *p-n* junctions used as solar energy converters. Curves 1–3 are empirical; the dashed line on the left represents the slope for an exponent kT/q; the heavy line on the right gives the optimum for radiative recombination only; and the question-marked line simulates a hypothetical maximum efficient junction, limited by the inherent properties of silicon.

[21] P. T. Landsberg, Proc. Inst. Elec. Engrs. (London) **106**, **Pt. II**, Suppl. No. 17, 908 (1959).

[22] V. M. Tuchkevich and V. E. Chelnokov, J. Tech. Phys. (U.S.S.R.) **28**, 2115 (1958).

efficiency that the empirical data are not in agreement with the diffusion theory of p-n junctions.[17] This theory predicts that forward currents vary as exp (qV/AkT) with $A=1$. The experimental data conform typically to A values of about two to three at room temperature (see Appendix). There is a theory which predicts an A value of two over a range of voltages[23] and this theory has recently been confirmed for heavily gold-doped silicon.[24] Nevertheless, the calculations of efficiency which give the semiempirical curve of Fig. 6 are based upon the diffusion theory.

The origin of A values as high as three is at present largely a mystery. As discussed in the Appendix, two explanations involving series resistance and surface currents can be rejected.

Wolf[25] proposes that the explanation may be internal field emission of the form reported by Chynoweth and McKay.[26] Such a current will clearly be of the nonradiative type and will reduce f_c. Thus, if it occurs, it is clear that solar cells can be improved by widening the transition region slightly since tunneling decreases exponentially with the width. If Wolf's proposed explanation were correct, therefore, it seems improbable that it would not have been eliminated in the large amount of development directed towards improving existing solar cells.

It is the conjecture of the present authors that the large values of A and large reverse currents both arise from recombination centers.

One recent observation which appears to support this view is the finding by Wolf and Prince[27] that "optimization of power conversion from light into electrical power resulted in devices with extremely soft reverse characteristics." Since soft reverse characteristics are clearly evidence of unnecessary current paths in p-n junctions, it is clear that they cannot of themselves contribute to the efficiency. An attractive explanation of this seemingly contradictory observation is that the softness results from the precipitation of impurities. Evidence that such precipitates cause soft reverse characteristics with current proportional to V^n, with n ranging from about four to seven, have been reported by Goetzberger and Shockley,[28] as well as means of removing certain metal impurities by "gettering."[29] This explanation cannot be checked from the publication of Wolf and Prince, who give no data on reverse characteristics and who state, "Since the solar cell is operated exclusively in the forward direction, no attention need be paid to the reverse characteristic of the device." This is not in agreement with the views of the present authors that an attempt should be made to understand as fully as possible the physics of the p-n junctions involved.

A theory based on recombination centers by which A values as high as three may be explained might possibly be developed along the following lines: In a forward biased junction the recombination occurs predominantly in a very narrow region in which the electrostatic potential varies by about $2kT/q$.[23] As the potential is varied, this region should move in position. If it moves into a region of lower recombination center density, values of A greater than two will arise. If the recombination centers are highly charged, they may be distributed in a very nonuniform manner through the junction.[30,31] Further investigations, which are currently being undertaken, are required to appraise this theory. If the theory does prove to be correct, important improvements in solar cells can probably be made by reducing contamination by chemical impurities.

An example of a new area in which a detailed balance treatment is needed is the proposal that a solar cell may be improved by adding traps to it to absorb the longer wavelength radiation.[32] It appears to the present authors that this may well be equivalent to shunting one cell with a threshold $\nu_g=E_g/h$ with another cell with a threshold much lower. Such a combination of cells would appear more likely to lower than to raise efficiency. A detailed balance argument, involving only radiative transitions for the traps, would set an upper limit for such a model like that of the curve for $f_c=1$ on Fig. 5. The present authors anticipate that traps will probably lower this limit; traps in general contribute strongly to recombination because they facilitate delivering energy to phonons. This implies that traps inherently have low f_c values, so that it is improbable that they would improve efficiency.

Returning briefly to question (1), we may note that one inherent process which may reduce f_c is the Auger effect, in which the energy of recombination is carried off by a hole or an electron.[33] Another mechanism that may have to be considered is the formation of donor and acceptor complexes at the high doping levels employed. These may also act as recombination centers. In any junction formed at higher temperatures, there will be certain densities of vacancies and even dislocation loops. These imperfections do not appear to be

[23] C. T. Sah, R. Noyce and W. Shockley, Proc. I.R.E. **45**, 1228 (1957).

[24] A. E. Bakanowski and J. H. Forster, Bell System Tech. J. **39**, 87 (1960).

[25] M. Wolf (footnote 6, p. 1252) reports agreement with this model, but his data is apparently not available in the literature.

[26] A. G. Chynoweth and K. G. McKay, Phys. Rev. **106**, 418 (1957).

[27] M. Wolf and M. B. Prince, Brussels Conference 1958, in *Solid State Physics* (Academic Press, Inc., New York, 1960) Vol. 2, Part 2, p. 1180.

[28] A. Goetzberger and W. Shockley, *Structure and Properties of Thin Films*, edited by C. A. Neugebauer, J. B. Newkirk, and D. A. Vermilyea (John Wiley & Sons, Inc., New York, 1959), p. 298; Bull. Am. Phys. Soc. Ser. II, 4, 409 (1959).

[29] A. Goetzberger, Bull. Am. Phys. Soc. Ser. II, **5**, 160 (1960); A. Goetzberger and W. Shockley, J. Appl. Phys. **31**, 1821 (1960).

[30] H. Reiss, C. S. Fuller, and F. J. Morin, Bell System Tech. J. **35**, 535 (1956).

[31] W. Shockley and J. L. Moll, Phys. Rev. **119**, 1480 (1960).

[32] M. Wolf, Proc. I.R.E. **48**, 1259 (1960).

[33] L. Pincherle, Proc. Phys. Soc. (London) **B68**, 319 (1955); L. Bess, Phys. Rev. **105**, 1469 (1957); A. R. Beattie and P. T. Landsberg, Proc. Roy. Soc. (London) **A249**, 16 (1958).

inherently necessary, and their densities may be governed by rate effects involved in fabrication.

Graded energy gap structures for solar cells to aid in collection efficiency are discussed by Wolf[6] following previous proposals for increasing emitter efficiency.[34] Such structures will in general also involve a gradation of lattice constant, and for minimum free energy subject to a fixed concentration gradient, there will be an equilibrium density of dislocations which reduce the stored elastic energy. These dislocations will, of course, have a deleterious effect on lifetime, diffusion length, and collection efficiency. The probable necessity of their presence does not appear to have been noted in the cited publications.

Note added in proof. The presence of slip bands due to stress in thin, heavily doped, diffused layers in silicon has recently been observed.[35] The associated dislocations may be important in reducing solar-cell efficiency.

APPENDIX

Empirical Current Voltage Relationships

The data shown in Fig. 7 were obtained for three commercial solar cells (curves 1–3). The same curves were obtained by plotting V_{op} vs I_{sh} for varying levels of illumination up to about 10 ma/cm² as by plotting V vs I; this agreement shows that series resistance plays no important role in determining the shape of the V-I lines up to 10 ma.

The distribution of forward current over the area of the cells was also studied by using a potential probing technique.[28] This study showed that the current was uniformly distributed over the area and not concentrated at the edges where the junctions were exposed.

TABLE I. Calculation of solar-cell efficiency from junction forward characteristics.

Cell no. (see Fig. 7)	I^*/A μa/cm²	V_c^* mv	$10^{-11}f_c^*$	V_c^*/V_c	v	m	η_{calc}^* %	η_{meas}[a] %
1	20	68	1.4	2.7	0.18	0.64	13.3	8
2	6	60	4.4	2.4	0.21	0.68	14.2	6
3	4	72	6.8	2.8	0.22	0.69	17.8	8

[a] This represents the average value, as given by the manufacturer.

[34] H. K. Kroemer, Proc. I.R.E. **45**, 1535 (1957); W. Shockley, U. S. Patent 2,569,347, issued September 25, 1951.

[35] H. J. Queisser, Bull. Am. Phys. Soc. **6**, 106 (1961).

The reciprocal slopes $(dV/d\ln I)$ of the straight line portions of the V-I lines are tabulated in Table I as V_c^*. [The ratio V_c^*/V_c corresponds to the quantity A in $\exp(qV/AkT)$ used by other authors. See footnote 11.]

If the forward currents for such cells are represented by

$$-I=I^*\exp(V/V_c^*),$$

then the calculations of Sec. 5 may be carried forward in a straightforward way to obtain an expression for the efficiency in terms of the same functions as in Sec. 5. The steps involve introducing the quantities

$$f_c^*=qQ_cA/I^*$$
$$f^*=f_c^*f_\omega t_s/2t_c$$

It then follows that

$$V_{op}/V_g=(V_c^*/V_c)v(x_g,x_c,f^*),$$

where v is the function of three variables defined in Sec. 4. The final expression for efficiency is a function of five variables:

$$\eta^*(x_g,x_c,t_s,f^*,V_c^*/V_c) = (t_xV_c^*/_c)Vu(x_g)v(x_g,x_c,f^*)m(vx_g/x_c).$$

Values of η^* calculated for the three cells of Fig. 7 are shown in Table I. The values for the five variables needed for the calculation are as follows: $t_s=0.97=t_c$[36]; V_c^* from Fig. 6, $V_c=25.9$ mv corresponding to 300°K; $f^*=f_\omega f_c^*t_s/2t_c$ with $f_\omega=1.36\times10^{-5}$ in order to give 100 mw/cm² on the exposed surface; $f_c^*=qQ_cA_p/I^*$ where $qQ_c=1.6\times10^{-19}\times1.7\times10^3=2.7\times10^{-16}$ amp/cm²; and I^*/A_p is taken from Fig. 7.

The fact that the calculated values of η^* are about twice as large as the measured values is evidence, as has been pointed out in the references cited, of losses due to collection efficiency, series resistance, etc. If these losses could be avoided, the cells would have an efficiency of about 14%, which is the highest value obtained to date.[6]

This value is about half the limiting efficiency of Fig. 6. Thus a significant potential for improvement of cells appears to exist by improving junction characteristics. The first step toward this improvement appears to be an understanding of the physics behind the high V_c^* values.

[36] M. Wolf and M. B. Prince (see footnote 27), p. 1186.