Band-Theoretic Model for the Photoelectric Effect in Silicon*

DAVID BRUST†
Department of Physics, Purdue University, Lafayette, Indiana
(Received 23 February 1965)

The pseudopotential method which was previously used to discuss the optical reflectivity of silicon has now been applied to study its photoelectric properties. A quadratic interpolation procedure has been used in conjunction with the pseudopotential method to get a very dense sampling of the Brillouin zone. This allows a detailed comparison of the band-theoretic results for the energy distribution of photoemitted electrons with experiment. Most of the structure in the kinetic-energy distributions in the optical and near-ultraviolet region of the electromagnetic spectrum can be understood in terms of direct interband transitions, although line shapes appear to be distorted by processes which are not treated in the present calculation. The photoelectric yield has been recomputed assuming total randomization of the crystal momentum for escaping electrons. The discussion of the optical properties has also been extended.

I. INTRODUCTION

A NUMBER of recent experiments have shed considerable light on the photoelectric effect in Si. The band-bending experiments of Gobeli and Allen[1] and of Van Laar and Scheer[2] emphatically demonstrate that photoemitted electrons originate in the bulk crystal states rather than from surface (Tamm) states. Further work indicates, that for energies above the appropriate interband thresholds, the electrons are produced by **k**-conserving excitations.[3–5] Experiments done on Si with a work function artificially reduced by deposit of less than one atomic layer of cesium on the surface[6,7] have provided rich spectral data. Such samples have shown very interesting structure in the kinetic-energy distributions for electrons emitted at various photon frequencies as well as in the spectral dependence of the yield efficiency.

Since it is believed that direct electronic transitions between bulk crystal states provide the electrons seen in photoemission, one expects that a study of the Si energy bands would be critical in interpreting the data. A semiempirical pseudopotential model was previously[8] used to construct the energy bands throughout the Brillouin zone. From the resulting joint density of states, the frequency dependence of the imaginary part of the dielectric constant was computed in the optical and near-ultraviolet regions. The good agreement between theory and experiment confirmed the pseudopotential band model. Thus, we have a realistic band structure to use in the study of the photoelectric problem. Below we shall apply it to gain an understanding of the Si photoelectric data, particularly the energy distributions. The problem is basically more complex than the case for the optical-absorption data. This is so, since electrons excited into the conduction bands can be scattered many times before ultimately being emitted from the crystal. However, as a first approximation we shall take a simple point of view in which dynamical electronic effects are neglected. We shall find that this gives a good qualitative understanding of the photoelectric data. This is essential to a further more quantitative analysis which would ultimately account for the dynamical processes as well.

II. DEFINITION OF PHOTOELECTRIC VARIABLES

The two principal quantities which have been measured in the photoelectric experiments for Si are the yield efficiency and the energy distributions. The first of these which we designate by $Y_B(\omega)$ = number of electrons emitted/quanta absorbed. Assuming that all electrons escape via the bulk conduction bands of the crystal,

$$Y_B(\omega)=\sum_n \int_{BZ} r_n(\mathbf{k},\omega) \times P(E,\mathbf{k})d^3\mathbf{k}/\sum \int_{BZ} r_n(\mathbf{k},\omega)d^3\mathbf{k}. \quad (1)$$

Here $r_n(\mathbf{k},\omega)$ is the rate at which electrons are excited from the valence bands to a state in the conduction band with index n and wave vector $\mathbf{k}$. $P(E,\mathbf{k})$ is the probability that an electron produced with energy E and wave vector $\mathbf{k}$ will escape, and the integral extends over the Brillouin zone (B.Z.). It is true, of course, that the probability factor should be dependent on the depth z below the surface at which the electron is produced. Neglecting dynamical processes, however, removes the z dependence from P, provided that we also assume a flat-band situation. All of our calculations will be done assuming no band bending, and will be

* Supported by the Advanced Research Projects Agency.
† Present address: Department of Physics, Northwestern University, Evanston, Illinois.
[1] G. W. Gobeli and F. G. Allen, *Proceedings of the International Conference on Semiconductor Physics, Exeter* (The Institute of Physics and the Physical Society, London, 1962), p. 818.
[2] J. Van Laar and J. J. Scheer, *Proceedings of the International Conference on Semiconductor Physics, Exeter* (The Institute of Physics and the Physical Society, London, 1962), p. 827.
[3] G. W. Gobeli and F. G. Allen, Phys. Rev. **127**, 141 (1962).
[4] E. O. Kane, Phys. Rev. **127**, 131 (1962).
[5] G. W. Gobeli, F. G. Allen, and E. O. Kane, Phys. Rev. Letters **12**, 94 (1964).
[6] W. E. Spicer and R. E. Simon, Phys. Rev. Letters **9**, 835 (1962).
[7] G. W. Gobeli and F. G. Allen (to be published).
[8] D. Brust, Phys. Rev. **134**, A1337 (1964).

compared to experiments in which band-bending effects are minor. Considering only vertical transitions (1) can be written as

$$Y_B(\omega)=\sum_{n,s}\int_{B,Z}\delta(\omega_{n,s}(\mathbf{k})-\omega)|M_{n,s}(\mathbf{k})|^2P \times d^3\mathbf{k}/\sum_{n,s}\int\delta(\omega_{n,s}(\mathbf{k})-\omega))|M_{n,s}(\mathbf{k})|^2d^3\mathbf{k}. \quad (2)$$

Here the delta function standing alone in the integrand would give the joint density of states for the conduction band (n) and the valence band (s); $\omega_{n,s}=(E_n-E_s)/\hbar$, and $|M|^2$ is a dipole matrix element connecting the bands.

The energy distribution function $N(E,\hbar\omega)$ is defined such that the number of electrons $\Delta\eta$ emitted with a range of kinetic energy ΔE is

$$\Delta\eta=N(E,\hbar\omega)\Delta E; \quad (3)$$

ω refers to the frequency of the incident light. N can be written as follows:

$$N(E,\hbar\omega)=C\sum_{n,s}\int_{BZ}\delta(\omega_{ns}(\mathbf{k})-\omega)|M_{n,s}(\mathbf{k})|^2P \times\delta(E-E_n(\mathbf{k}))d^3\mathbf{k}. \quad (4)$$

The first delta function picks out the set of states with a vertical energy gap $\hbar\omega$, and the second delta function selects from the first set those with a final conduction-band energy equal to E. Energies in this expression are measured relative to the vacuum level. C is a normalization constant. We should note that this expression assumes that an electron produced with a conduction-band energy $E_n(\mathbf{k})=E$ can only appear in vacuum with the energy E. If dynamical events such as loss to phonons and Auger pair production were to be considered, then electrons produced in state $E_n(\mathbf{k})$ would give rise to a distribution of emitted electron energies. We should then have to consider a function $g(E_n(\mathbf{k}),E)$ which would give the number of electrons seen with energy E as a result of producing one in state $E_n(\mathbf{k})$. For our simple picture $g(E_n(\mathbf{k}),E)\rightarrow P\delta(E-E_n(\mathbf{k}))$.

III. CONTRIBUTION OF THE ENERGY BANDS

As expressions (2) and (4) indicate we need to compute the integrals throughout the zone to get the energy-band contribution, the dipole matrix elements, and finally appropriate escape probability factors. In our work on the orbital absorption,[8] we found that the effect of including matrix elements could be taken as a minor correction to the contribution arising solely from density-of-states considerations. The same conclusion will also apply for the photoelectric properties. Furthermore, we shall see below that the results are not highly sensitive to the method of computing the escape probability (after we take account of energy conservation for the transmission of electrons through the crystal surface). In this section, therefore, attention is concentrated on the method of computing the energy-band contribution.

A. Pseudopotential Calculation

In Ref. 8 we used a model Hamiltonian $H_p=$K.E. $+V_p$. Here K.E. is the kinetic-energy operator, and V_p is a semiempirical pseudopotential. V_p is a simple local potential. Since V_p depends only on the coordinate $\mathbf{r}$, it can be written as

$$V_p=\sum_J V_p(\mathbf{G}_J)\exp[(2\pi i/a)\mathbf{G}_J\cdot\mathbf{r}]. \quad (5)$$

In (5), $(2\pi/a)\mathbf{G}_J$ are the reciprocal lattice vectors of the diamond structure, and a is the lattice constant for Si. By using the high symmetry of the diamond lattice, it turned out that we could represent V_p adequately by using only the lowest three independent Fourier coefficients in the expansion. These three coefficients could then be treated as disposable parameters.[9,10] The three parameters were determined empirically from a set of model interband gap splittings at the symmetry points Γ, X, and L which were known through previous work.[11,12] The properties of the resulting eigenvalues of H_p are, first, that they give the proper interband gap splittings at the symmetry points, and second, that they represent a reliable interpolation for the energy bands throughout the rest of the zone. The reliability of the interpolation was well established by the good results H_p gave for the distribution of electronic states. The eigenfunctions of the Hamiltonian represent the smooth slowly varying part of the true crystal wave function outside of the atomic-core region.

It has been demonstrated[13,14] that V_p can be thought of as arising from the sum of two terms $V_{\text{XTAL}}+V_R$ where V_{XTAL} is the usual crystal potential, and V_R is a repulsive potential arising from the orthogonalization of the crystal wave function to the core states as in the orthogonalized-plane-wave (OPW) method. The effect of V_R is to cancel out the strong attractive part of the crystal potential near the atomic core. One can directly compute $V_{\text{XTAL}}+V_R$ and from it construct the OPW solutions. (For a discussion of the general theory of pseudopotentials see Ref. 15.) Errors, however, can then arise in some of the principal band gaps amounting to $\sim$3 eV.[16] Our V_p can be gotten from $V_{\text{XTAL}}+V_R$ of the OPW method by treating it as local, and by

[9] J. C. Phillips, Phys. Rev. **112**, 685 (1958).
[10] F. Bassani and V. Celli, J. Phys. Chem. Solids **20**, 64 (1961).
[11] H. Ehrenreich, H. R. Philipp, and J. C. Phillips, Phys. Rev. Letters **8**, 59 (1962).
[12] J. C. Phillips, Phys. Rev. **125**, 1931 (1962).
[13] J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287 (1959).
[14] M. H. Cohen and V. Heine, Phys. Rev. **122**, 1821 (1961).
[15] R. J. Austin, V. Heine, and L. J. Sham, Phys. Rev. **127**, 276 (1962).
[16] F. Bassani and M. Yoshimine, Phys. Rev. **130**, 20 (1963).

slightly changing the Fourier coefficients of the latter to make the interband splittings agree with the experimentally determined model. Since our potential can be derived from the theoretical one by a slight readjustment of the coefficients, we may refer to it as a semiempirical pseudopotential.

The method of solving for the eigenvalues of H_p was thoroughly discussed in Ref. 8. Briefly, we start with a plane-wave representation. All plane waves with a kinetic energy $\leq E_N$ are treated exactly in the expansion of the wave function, while plane waves with a kinetic energy $>E_N$ but $\leq E_\Gamma$ are treated by perturbation theory. E_N and E_Γ are chosen so as to insure reasonable convergence of the eigenvalues. Higher lying plane waves are ignored. This is rather like ordinary perturbation theory when a zero-order degeneracy is encountered; the degenerate states are treated exactly, and the effect of the other states is brought in through perturbation theory. Here the group of low-lying states are treated in a quasidegenerate fashion.

As one moves from a given point in the zone to a neighboring point the eigenvalues should vary continuously with $\mathbf{k}$. In general, this will be the case for our treatment of H_p; however, occasionally a discontinuous jump will occur when the kinetic energy of a plane wave crosses either the energy E_N or E_Γ. This situation is sketched in Fig. 1. When this happens, the manner in which the plane wave $\Phi_{\mathbf{k}+\mathbf{K}_i}=e^{i(\mathbf{k}+\mathbf{K}_i)\cdot r}$ is treated according to the present procedure changes, thus producing the discontinuity. We wished to reduce the discontinuous jumps from the value in our previous calculation. To do so we varied E_N and E_Γ and examined the effect on some representative eigenvalues. The results are plotted in Fig. 2. Our values for the choice of E_N and E_Γ were

$$E_N=9.0(2\pi^2\hbar^2/ma^2)\sim 45 \text{ eV},$$
$$E_\Gamma=24.0(2\pi^2\hbar^2/ma^2)\sim 120 \text{ eV}.$$

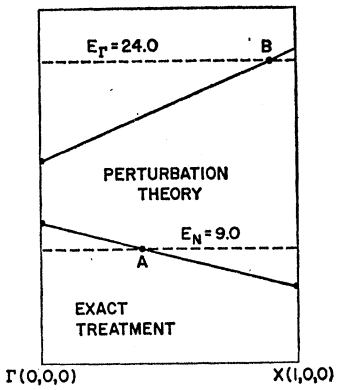


FIG. 1. The dependence of $|\mathbf{K}_i+\mathbf{k}|^2\cdot a^2/4\pi^2$ as $\mathbf{k}$ is varied across the Brillouin zone in the (1,0,0) direction. The lower curve is for $\mathbf{K}_i=(2\pi/a)\cdot(-3,1,1)$ and the upper curve for $\mathbf{K}_i=(2\pi/a)\cdot(4,0,0)$. The points A and B refer to locations in the Brillouin zone where there is a discontinuous change in the treatment of the plane wave.

Since it is desired to make the bands vary smoothly throughout the zone, it is of interest to estimate the discontinuities from Fig. 2. The two examples in 2(a) indicate that with $E_\Gamma=24.0$ moving ~ 6 plane waves from the set included by perturbation theory to the set being ignored changes the eigenvalues by no more than 0.01 eV (a change in E_Γ of 1.0 shifts ~ 6 plane waves across the boundary). This implies that the perturbation contribution from plane waves near 24.0 is <0.01 eV, and that we can therefore expect discontinuities associated with them to be small, also <0.01 eV. Looking at Fig. 2(b) we estimate the effect of moving plane waves from the set being treated by perturbation theory to the set entering into the exact calculation. Here the effect of shifting ~ 4 plane waves is ≤ 0.02 eV. Hence, the error made by treating plane waves with kinetic energy near 9.0 through perturbation theory rather than exactly is ~ 0.01 eV. Therefore, the discontinuities encountered when a plane wave moves across the boundary E_N in Fig. 1 is ~ 0.01 eV, and we conclude that with the parameters chosen our energy bands should be extremely smooth.

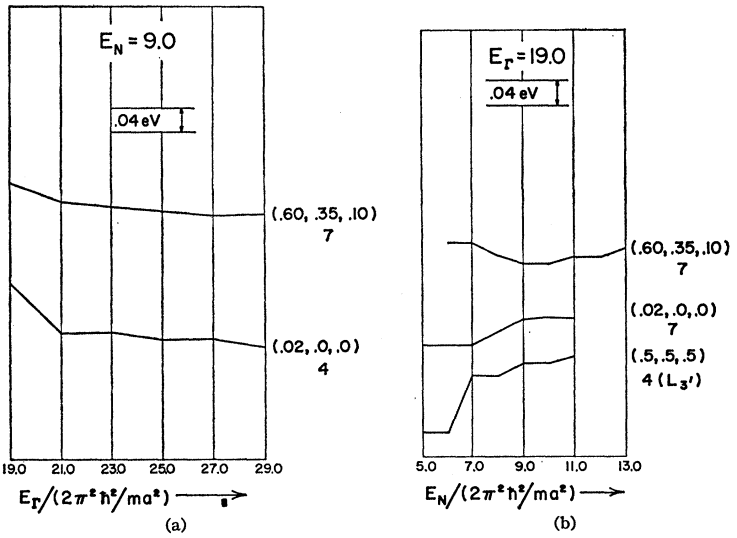


FIG. 2. Convergence tests for some representative eigenvalues. To the right of each curve we give the point $\mathbf{k}$ at which the eigenvalue was computed and the level number according to the ordering principle described in the text. (a) as a function of E_Γ, (b) as a function of E_N.

B. Application of Quadratic Interpolation

In order to evaluate expressions (2) and (4) we calculate the eigenvalues of H_p at a large number of points in the B.Z., and replace the integrals by discrete sums. In Ref. 8 we found eigenvalues to represent the bands at $\sim$1000 points in 1/48 of the B.Z. This gave $\sim$50 000 points for the whole zone. For the present case, however, the second delta function in (4) implies a second sampling such that values are now characterized by $E_n(\mathbf{k})$ as well as $E_{ns}(\mathbf{k})$. With this further subdivision histogram statistics based on a 1000-point sample would be inadequate. The sample size must be increased by an order of magnitude. It would lead to an excessively time-consuming calculation if the secular equation for H_p had to be diagonalized at so many points. Instead we chose to use the following scheme. First we found eigenvalues for H_p at a small number of discrete points in the shaded region of the B.Z. shown in Fig. 3. Then we used an algebraic interpolation to find the eigenvalues at neighboring points. The use of an algebraic interpolation is computationally very simple, and we are not severely restricted in terms of sample size.

For our current work we start by calculating the energy eigenvalues over a simple cubic mesh embedded in the 1/48 volume of the zone being used. This mesh has a lattice constant equal to 1/28 that of the reciprocal lattice of diamond. With this spacing we have $\sim$400 points at which to diagonalize H_p. We call this set of points S_p. Next we want to find eigenvalues (and matrix elements) for a much larger set of points S_A in order to evaluate (2) and (4). The details for constructing the set S_A are discussed below. Here the algebraic interpolation will be described. For simplifying the discussion we imagine k space to be two dimensional. In Fig. 4, A represents a point of the set S_A and the P's a neighboring group of points from S_p. In order to get the eigenvalues at A from the P's, we use a simple quadratic interpolation. The interpolation can be generated as follows. First do a quadratic interpolation using standard formulas for the one-dimensional problem along the solid lines to get a value at X_1, X_2, and X_3. Then use the values generated at X_1, X_2, and X_3 along the dotted line to get the value at A. In 3 dimensions the process involves 27 points neighboring A and one extra step.

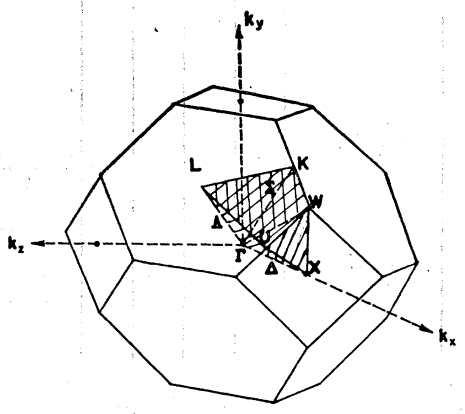


FIG. 3. Brillouin zone for Si.

The point A may lie near the boundary of the small region in which we are working. In this case some of the points P of Fig. 4 may not be included in the set S_p. There are three ways in which this can happen:

(i) The point P lies inside the first B.Z. but outside the shaded region of Fig. 3. In Ref. 8 this region was defined by the conditions

$$k_x \geqq k_y \geqq k_z \geqq 0. \tag{6}$$

Hence, if a point lies outside the region we simply reorder the values of k_x, k_y, k_z so as to satisfy condition (6). This is equivalent to an application of the group elements of the diamond lattice.[17]

(ii) The point P may lie outside the UXW boundary of the B.Z. In this case all that is necessary is to use the reflection symmetry of the plane to get an equivalent point in the shaded wedge.

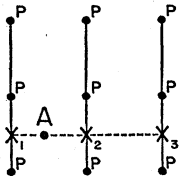


FIG. 4. Schematic diagram for the algebraic interpolation procedure.

(iii) The point can lie just outside of the boundary plane $LUWK$. Since this plane is defined by the equation

$$(k_x+k_y+k_z)=(\tfrac{3}{2})(2\pi/a),$$

a point which lies just outside of the plane will be given by

$$(k_x+k_y+k_z)a/2\pi=\tfrac{3}{2}+\Delta.$$

To get a corresponding point inside the wedge, construct the quantity

$$\left(\frac{2\pi}{a}-k_z, \frac{2\pi}{a}-k_y, \frac{2\pi}{a}-k_x\right).$$

This corresponds to first reordering to get k_z, k_y, k_x, then inverting to get $-k_z$, $-k_y$, and $-k_x$, and finally adding the reciprocal lattice vector $(2\pi/a, 2\pi/a, 2\pi/a)$. All of the operations are group operations; therefore, the resulting point will be equivalent to the initial one, and will be inside of the wedge since

$$(k_x'+k_y'+k_z')a/2\pi=\tfrac{3}{2}-\Delta.$$

Of course, we may have to do more than one of the

[17] Reflections in the (110) planes will interchange a pair of values for the components of $\mathbf{k}$.

operations indicated by (i), (ii), and (iii) in some instances.

C. Monte Carlo-Generated Sample

The next question is the generation of the sample S_A which will be used for constructing the histograms. In the previous work the points S_A lay on a lattice having the bcc structure, so that the sampling points were just contained within the elementary region of the B.Z. To get some idea of the statistical fluctuations expected with this spatially uniform sample we examine Fig. 5(a). Here we have a two-dimensional problem in which we wish to estimate the area of the square by the number of sample points within it (this corresponds to finding the size of a region for which the energy lies in some range ΔE). For this case the error is approximated by the total number of points on the boundary (N_b) divided by the total number of points in the square (N); i.e., error $\sim N_b/N$. For a three-dimensional problem $N\sim 1/l^3$ and $N_b\sim 1/l^2$, where l is the dimension size of the sampling lattice. We then have

$$\text{error}\sim l^3/l^2\sim l\sim 1/N^{1/3}. \tag{7}$$

We can imagine another situation. Suppose the boundary of the region whose area we want follows a complex path through k space as in Fig. 5(b). This we can suppose to be essentially a random path with respect to a rectangular sample. There will then be no correlation between the position of the boundary at one surface sample point and a neighboring one. In this case,

$$\text{error}\sim N_b{}^{1/2}/N\sim l^3/l\sim l^2\sim 1/N^{2/3}. \tag{8}$$

If, on the other hand, we use a randomly generated sample, then irrespective of the shape of the region being sampled the fluctuations will go like

$$\text{error}\sim 1/N^{1/2}. \tag{9}$$

We see that (9) in its rate of convergence as a function of N is superior to (7) but inferior to (8). Previously we found a good deal of boundary correlation error when a rectangular sampling grid was used. Hence, in our present work, we generate S_A by a Monte Carlo technique.

IV. ESCAPE PROBABILITY

In expressions (2) and (4) there appears the escape probability factor $P(E,\mathbf{k})$. To treat this exactly would necessitate having rather detailed information about the surface potential. For the purposes of the present study we neglect partial reflection of electrons from the surface barrier which is probably small throughout most of the Brillouin zone.[18] With this model an electron impinging on the surface is either totally reflected or transmitted. We then compute the escape probability according to two different assumptions.

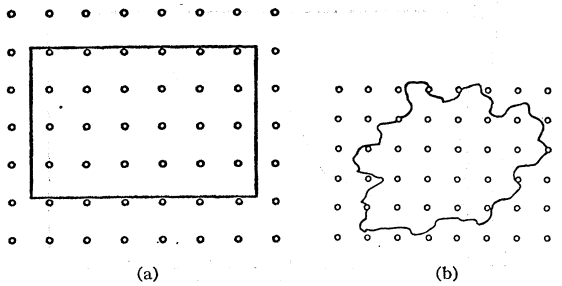


FIG. 5. Evaluation of the area of a given region by counting the number of sampling points it contains (a) for a region with a simple rectangular boundary, showing how a large number of boundary points can be moved as a group across the boundary if the sampling mesh dimension is changed slightly, (b) for a region with a highly irregular boundary.

(i) We assume that the electron after being excited into the conduction band propagates with *no scattering*. With this assumption all electrons which initially travel with a group velocity away from the surface continue to do so and cannot possibly escape. Electrons whose group velocity is directed toward the surface may, under the correct conditions, be emitted into the vacuum. For these electrons we write

$$\begin{aligned} P(E,\mathbf{k}) &= \text{Constant, if } E_n(\mathbf{k})>\hbar^2k_T{}^2/2m \\ &= 0, \text{ otherwise} \end{aligned} \tag{10}$$

with all energies measured relative to the vacuum level. In Eq. (10), k_T is the component of the received wave vector parallel to the crystal surface. The factor $\hbar^2k_T{}^2/2m$ comes from the wave-mechanical boundary condition. It gives the minimum kinetic energy with which the electron is able to appear in vacuum.

(ii) We now assume that *elastic*-scattering processes randomize the crystal momentum before the electron reaches the surfaces.[19] Doing this removes the k dependence from P. The resulting $P(E)$ as found from the energy-band model of Si is shown in Fig. 6, where the vacuum level W was taken as 2.5 eV above the top of the valence band.

V. RESULTS

A. Yield Efficiency

We previously analyzed the structure of $Y_B(\omega)$ on the 0% scattering model.[8,20] Here we have recomputed it on both the 0 and 100% scattering model. In Figs. 7(a) and 7(b) we compare the results for the two limiting cases with the experiment of Allen and Gobeli.[21]

[18] C. Herring and M. H. Nichols, Rev. Mod. Phys. **21**, 185 (1949).

[19] This does not contradict the earlier statement concerning the neglect of dynamical events. The 100% scattering as treated here assumes no energy transfer, and is equivalent to taking the electron's crystal momentum as being random immediately after excitation into the conduction band.

[20] D. Brust, M. L. Cohen, and J. C. Phillips, Phys. Rev. Letters **9**, 389 (1962).

[21] F. G. Allen, Bull. Am. Phys. Soc. **8**, 422 (1963).

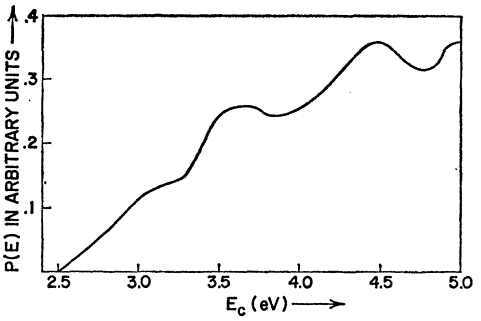


FIG. 6. Sketch of the escape probability assuming total randomization of the electrons before they reach the surface. E_c is the conduction-band energy measured from the top of the valence band. The vacuum level is taken as 2.5 eV above the valence band. As E becomes large $P(E) \rightarrow 0.5$, which means that all electrons with a group velocity directed toward the surface will escape. It has also been assumed that the crystal is cut on a (111) plane.

$Y_B(\omega)$ has been constructed using interband transitions $4 \rightarrow 5$ and $4 \rightarrow 6$ only. The dipole matrix element has been taken as a constant.

Examining the figures indicates that treating the problem in the 100% scattering limit gives somewhat better results. This is what one might expect, as the experimental curves were taken with nearly one monolayer of C_s on the surface. Experiment indicates that deposition of more than about 0.3 of a monolayer of C_s on a silicon surface will result in total randomization of escaping electrons.[5] We note, however, that the principal features in the structure are similar to that obtained before, and we have made no changes in the previous identifications.

B. Kinetic-Energy Distribution

We have computed the kinetic energy distributions for escaping electrons with photon energies between 3.2 and 6.2 eV. The kinetic-energy range is 0.0–2.8 eV. In Fig. 8 we present the results of the calculation using only the 100% scattering model. The results are quite similar if the 0% scattering situation is assumed except that the peaks are sharper (particularly for $E=2.0$ eV and $\hbar\omega=5.0$ eV). In computing the dipole matrix element appearing in expression (4) for one of the randomly generated points in our sample we substituted the matrix element at a neighboring point of the simple cubic interpolating mesh. This simplification will have a very minor effect on the results presented here. In order to evaluate the matrix elements the eigenfunctions of the pseudopotential Hamiltonian were used. This corresponds to using only the smooth part of the wave function, and as was previously discussed[8] generally should lead to errors no more than $\sim 20\%$.

In arriving at the graphs of Fig. 8 we used only transitions involving bands 3, 4, 5, 6, and 7, i.e., $(3,4) \rightarrow (5,6,7)$. For the photon and kinetic-energy range under discussion these transitions will dominate the bulk photoelectric response. This is apparent from examination of Fig. 9 where the symmetry directions Δ, λ, and Σ are drawn. It should be noted that we have shown all our results with a vacuum level 2.5 eV above the top of the valence band. We have done the calculations for other values, but the results can essentially be generated by shifting the curves to higher or lower energies as the vacuum level is moved down or up.

VI. DISCUSSION OF THE KINETIC-ENERGY DISTRIBUTIONS

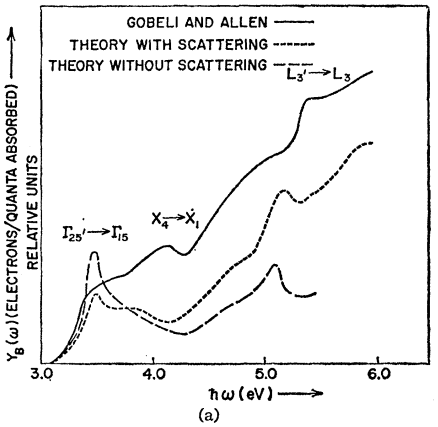
A. Role of the Conduction-Band Structure

Before comparing theory with experiment, we find it valuable to develop some basis for understanding the results of the calculations. The discussion of the results, however, should be preceded by a brief digression on the question of normalization. From Eq. (4) we have

$$\int N(E,\hbar\omega)dE = C\sum_{n,s}\int_{BZ}\delta(\omega_{n,s}(\mathbf{k})-\omega) \times |M_{n,s}(\mathbf{k})|^2 P. \quad (11)$$

We recall that

$$\epsilon_2(\omega) \sim \frac{1}{\omega^2}\sum_{n,s}\delta[\omega_{n,s}(\mathbf{k})-\omega]|M_{n,s}(\mathbf{k})|^2 d^3\mathbf{k}, \quad (12)$$



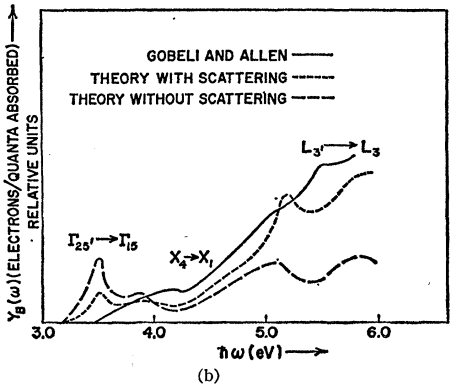


FIG. 7. Spectral dependence of the yield efficiency (a) experimental vacuum level =2.7 eV, theoretical vacuum level=2.5 eV, (b) experimental vacuum level=3.2 eV, theoretical level=3.0 eV.

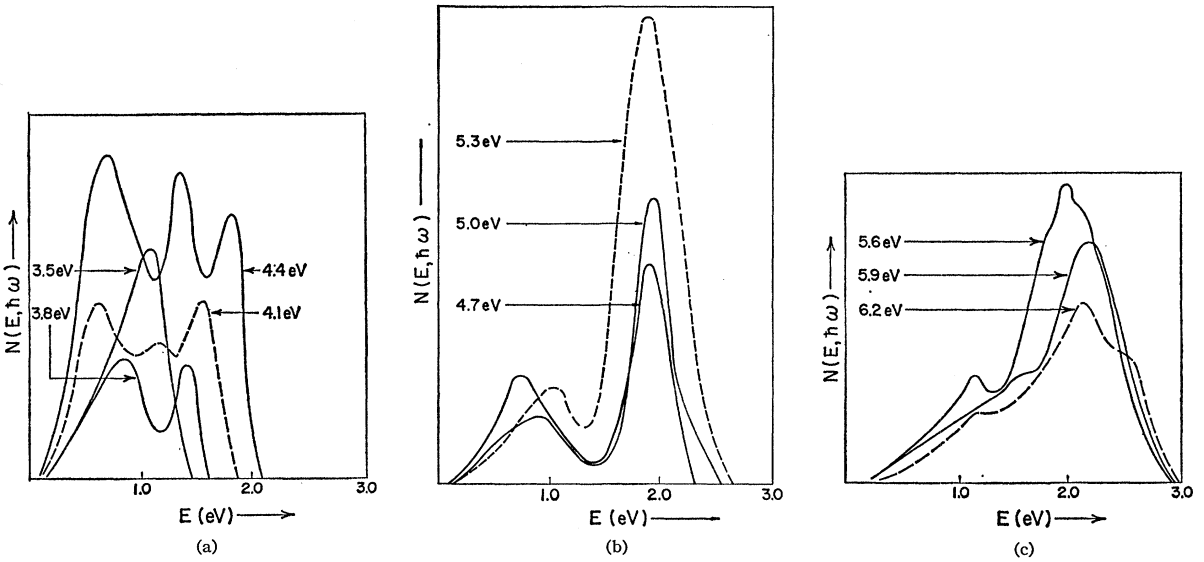


FIG. 8. Kinetic-energy distributions with the vacuum level taken 2.5 eV above the top of the valence band. E refers to the kinetic energy in vacuum. The curves are plotted for several values of $\hbar\omega$ as a computational parameter. For a discussion of normalization refer to the text.

from which follows

$$\int N(E,\hbar\omega)dE = DY_B(\omega)\epsilon_2(\omega)\omega^2. \quad (13)$$

D is an unspecified constant whose value may be set by finding the area under any one of the curves of Fig. 8, and then satisfying condition (13). The constant D will be the same for all other curves. This has essentially been done in Fig. 8 except that we have decreased D by a factor of 2 upon going from 8(a) to 8(b) which corresponds to shrinking the vertical scale by the same factor.

Examination of Fig. 8 indicates that the most prominent and persistent features of the computed kinetic energy spectra are the peaks near 1.0 and 2.0 eV (we shall refer to these as series A and B, respectively). The origin of these two series can be qualitatively understood in terms of the conduction band density of states. Figure 10 shows a sharp maximum near 4.5 eV (which corresponds to a vacuum kinetic energy of 2.0 eV when the vacuum is 2.5 eV above the top of the valence band).[22] Broadly speaking, series B results from transitions between the valence bands and states near the maximum. Series A results from the maximum near 3.0 eV in the conduction-band density of states. However, since this maximum occurs just above the 2.5-eV vacuum level the peaks associated with it are somewhat more complex. At this energy $P(E)$ is still rapidly

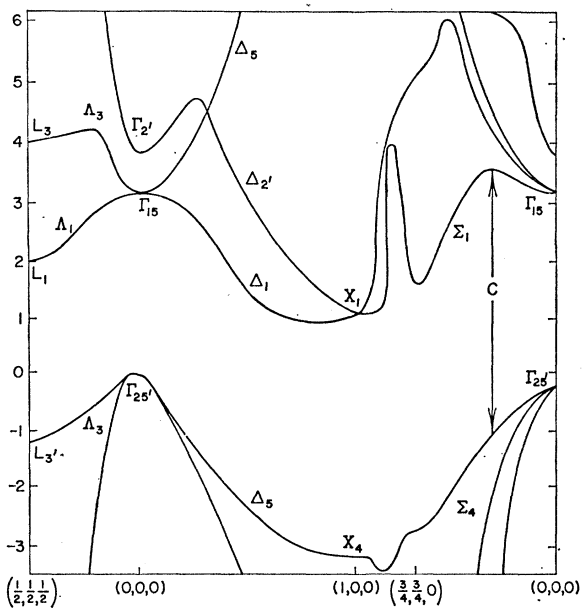


FIG. 9. The energy bands of Si from the pseudopotential model. The bands between X and K are drawn along the straight line connecting them in the plane defined by $k_Z=0$.

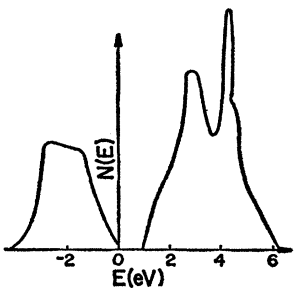


FIG. 10. Electronic density of states for Si. In this case E refers to the energy measured relative to the top of the valence band. This includes only the highest valence band (4) and the lowest conduction bands (5,6).

[22] The maximum at 4.5 eV appears somewhat sharper than it should due to our omission of band 7 in computing the density of states.

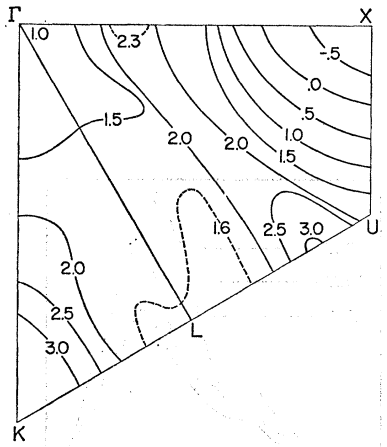


FIG. 11. Energy contours of band 6 plotted in a (110) reflection plane. The energies are measured relative to the vacuum level of 2.5 eV.

increasing as E is increased. Hence, the peaks in the kinetic-energy distribution occur at slightly higher energies than the maximum in the density of states would suggest.

The origin of the 4.5-eV maximum in the density of states can itself be understood after examination of Fig. 11 in which the energy contours of band 6 are drawn in a (110) reflection plane. The band is extremely flat near the Λ symmetry line giving rise to the sharp maximum. Therefore, we can associate our corresponding peaks in series B with Λ electrons. The lower maximum appears to come from states distributed throughout a large part of the zone, and there does not seem to be a simple way of identifying it with a definite location.

Further analysis of the structure requires a much more rigorous examination of the energy-band contours. In the next section some further aspects of the energy distributions are examined. We will develop a somewhat more quantitative understanding of the results; however, the basic picture described in this section will remain unaltered.

B. Analysis in Terms of Critical Lines

Before further analyzing the results of the computed kinetic-energy distributions presented in Fig. 8, we briefly discuss the analytical tools which are helpful in giving meaning to the structure. In the discussion of the dielectric constants of Si and Ge,[8] the concept of a critical point was extremely useful. Most of the computed structure in that case was attributed to critical points (c.p.) in the joint energy bands, i.e., points where $|\nabla_{\mathbf{k}} E_{ns}(\mathbf{k})|=0$. The c.p. gave rise to Van Hove singularities in the integral for the density of transitions

$$J_{ns}(\omega)\sim\int_{E_{ns}(\mathbf{k})=\hbar\omega} ds/|\nabla_{\mathbf{k}} E_{ns}(\mathbf{k})|\,.$$

The theoretical and corresponding experimental line shapes were then well understood in terms of the behavior of the joint density of states in the vicinity of the c.p.

It is not to be expected that we can achieve as good an account of the experimental line shapes for the energy distributions as for the optical constants, since we are neglecting dynamical processes and are using an approximation to the surface reflection problem. Nevertheless, further analysis of the theoretical structure can be useful. Below we shall use an approach due to Kane[23] to examine the analytic behavior of the bands in a way similar to the c.p. analysis.

In order to understand the contribution of a given pair of bands to the kinetic-energy distribution $N(E,\hbar\omega)$, consider the frequency fixed. This then defines a surface $S_{\hbar\omega}$ in k space given by the condition $E_{ns}(\mathbf{k})=\hbar\omega$. Next attention is fixed on the energy contours of the conduction band $E_n(\mathbf{k})$ imbedded in $S_{\hbar\omega}$. To compute the density of states satisfying the conditions $E_{ns}(\mathbf{k})=\hbar\omega$ and $E_n(\mathbf{k})=E$, one must integrate along a line l in $S_{\hbar\omega}$ defined by $E_n(\mathbf{k})=E$. The density of states will then be

$$N'(E,\hbar\omega)\sim\sum_{n,s}\int_{[E_n(\mathbf{k})=E_0,E_{ns}(\mathbf{k})=\hbar\omega]} dl/|\nabla_{\mathbf{k}}(S_{\hbar\omega})E_n(\mathbf{k})|\,|\nabla_{\mathbf{k}} E_{ns}(\mathbf{k})|\,. \quad (14)$$

Here $\nabla_{\mathbf{k}}(S_{\hbar\omega})E_n(\mathbf{k})$ gives the components of the $\nabla_k E_n(\mathbf{k})$ lying within the surface $S_{\hbar\omega}$. This two-dimensional gradient is proportional to the distance Δk_c between two contour lines differing in energy by ΔE_c. This is shown in Fig. 12. The $|\nabla_{\mathbf{k}} E_{ns}(\mathbf{k})|$ is a measure of the distance between two surfaces $S_{\hbar\omega}$ and $S_{\hbar(\omega+\Delta\omega)}$. Expression (14) is equivalent to (4), the kinetic-energy distribution, except for the omission of the matrix element $|M_{ns}(\mathbf{k})|^2$ and the probability-of-escape function $P(E)$. Points on $S_{\hbar\omega}$ where $|\nabla_{\mathbf{k}}(S_{\hbar\omega})E_n(\mathbf{k})|=0$ give rise to two-dimensional critical points. The functional dependence of $E_n(\mathbf{k})$ within $S_{\hbar\omega}$ can be of three types near a two-dimensional c.p., i.e., a minimum, maximum, or saddle point. Kane shows that near a minimum or maximum we expect a discontinuity in N' of expression (14) as a function of E; whereas, near a saddle point a logarithmic singularity is expected. This behavior is shown in Fig. 13.

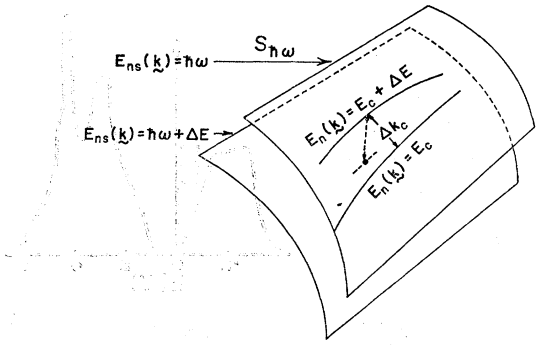


FIG. 12. Intersection of the $E_{ns}(\mathbf{k})$ surfaces with $E_n(\mathbf{k})$ surfaces.

[23] E. O. Kane (to be published).

If the photon energy is changed slightly we define a new surface $S_{\hbar(\omega+\Delta\omega)}$. According to Kane the two-dimensional c.p.'s on the first surface will be in one to one correspondence with a neighboring set on the second surface. The trajectory which is a two-dimensional c.p. follows in k space as $\hbar\omega$ is varied is called a critical line (c.l.) (we use Kane's terminology). That is, if $\mathbf{k}_0(\hbar\omega)$, defines a c.l., the energy $E_n(k_0(\hbar\omega))$ plotted as a function of $\hbar\omega$ is called an $E-\hbar\omega$ image of the c.l. Kane also shows that every three-dimensional c.p. either in the joint energy bands or in the conduction bands (OCP or ECP) has at least one critical line passing through it.

Below we shall draw $E-\hbar\omega$ images for some of the principal peaks in our energy distributions. These can be compared with the analysis recently completed by Kane for the 110 symmetry plane. In particular, we shall find that the structure can be partially related to c.l.'s in the 110 plane. It should be emphasized that our bands are all ordered by increasing energy. Kane, however, takes advantage of the reflection symmetry of the 110 plane to classify his states according to the reflection parity of the wave functions. The even and odd states after having been separated are then ordered by increasing energy. Hence, when an even band crosses an odd band in the 110 plane there will be a discontinuity between our definition and Kane's.

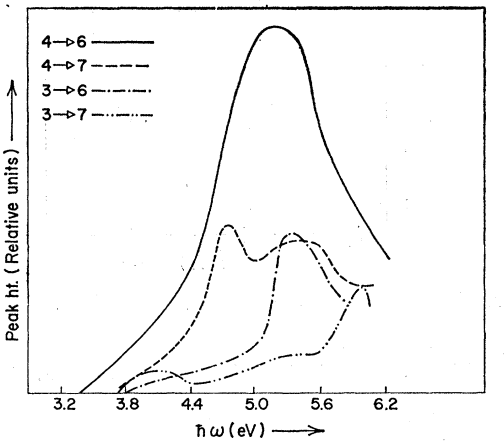


FIG. 14. Strength plots for the separate histograms contributing to $N(E,\hbar\omega)$. The peak contribution for each histogram (near $E=2.0$ eV) has been selected and plotted as a function of $\hbar\omega$.

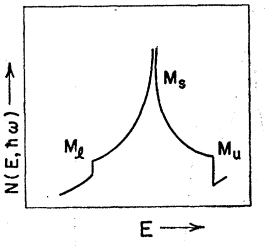


FIG. 13. Behavior of the density of states near a two-dimensional critical point according to Kane. M_l comes from an absolute minimum, M_s from a saddle point, and M_u an absolute maximum. E refers to the conduction-band energy and $\hbar\omega$ is the interband-gap energy (taken as fixed).

We are now in a position to re-examine our computed kinetic-energy spectra in terms of their analytical properties. The first thing to be discussed are the peaks of series B. Examination of the individual histograms shows that the largest contribution to $N(E,\hbar\omega)$ near $E=2.0$ eV comes from the $4\rightarrow 6$ transitions (see Fig. 14), i.e., transitions from the topmost valence band to the second lowest conduction band. Attention is therefore focused on this pair of bands. In Fig. 15, $N(E,\hbar\omega)$ is drawn including only the contribution from the $4\rightarrow 6$ transitions for a number of different frequencies. We have attempted to resolve the structure in terms of the singularities expected from Kane's analysis of two-dimensional critical points; although, this is somewhat difficult to do in a categorical way. This is particularly difficult when several c.p. fall near the same energy as is apparently the case for the higher frequencies of Fig. 15. The three-point averaging which is used to reduce statistical scatter tends to blur the c.p.'s into one another. Assuming that the analytical assignments of Fig. 15 are basically correct, we can draw $E-\hbar\omega$ images. This has been done, and the result

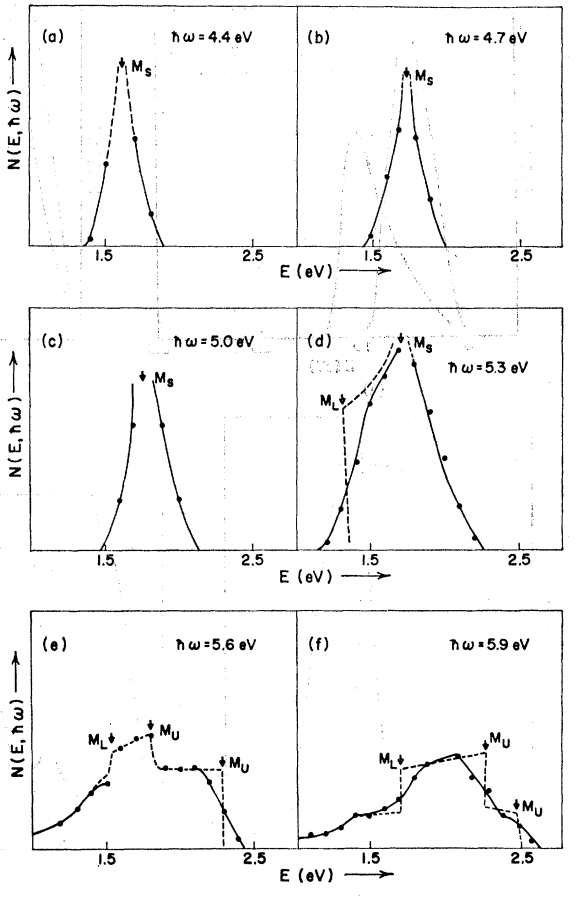


FIG. 15. The contribution of the $4\rightarrow 6$ histograms to $N(E,\hbar\omega)$ for several choices of $\hbar\omega$. The dotted lines represent analytical behavior associated with two dimensional critical points (c.p.).

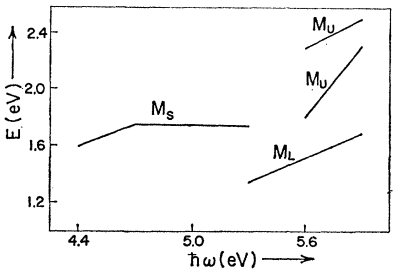


FIG. 16. $E-\hbar\omega$ images for the two-dimensional c.p. of Fig. 15.

is shown in Fig. 16. We can compare our results with the $E-\hbar\omega$ images of the strongest c.l.'s which Kane finds for the $4\rightarrow 6$ transitions in the (110) plane. The direct comparison of our $E-\hbar\omega$ images with Kane's does not appear entirely satisfactory. Our strong saddle point does seem to have approximately the same features as the $E-\hbar\omega$ image of Kane's strong saddle point for $\hbar\omega<5.3$ eV. It should be mentioned that the path of Kane's strong saddle type of critical line in k space is near to Λ. This agrees with our earlier identification for the series B peak. Furthermore, we notice in Fig. 14 that the strength curve for the $4\rightarrow 6$ transition peak has a maximum near $\hbar\omega=5.3$ eV. This corresponds to the $L_3'\rightarrow L_3$ direct energy gap. The kinetic energy of the peak is also what one would expect for L_3 electrons. This is satisfactory, since Kane's strong c.l. passes through L, and one expects a maximum in the strength curve as a c.l. approaches a symmetry point (actually Kane shows that the strength curve tends to infinity as the symmetry point is approached).

We can also analyze the peak occurring in the distributions for $\hbar\omega=4.1$ and 4.4 eV with kinetic energy $\sim$1.1 eV. This can be traced to a c.l. which apparently includes the transition marked "C" in Fig. 9. Checking Kane's results we corroborate our interpretation. He shows a transition between his V_- and C_{1+} bands corresponding to a $4\rightarrow 5$ transition in our case. The characteristic of this peak is that it has a very rapid rise in intensity near $\hbar\omega=4.4$ eV and then disappears.

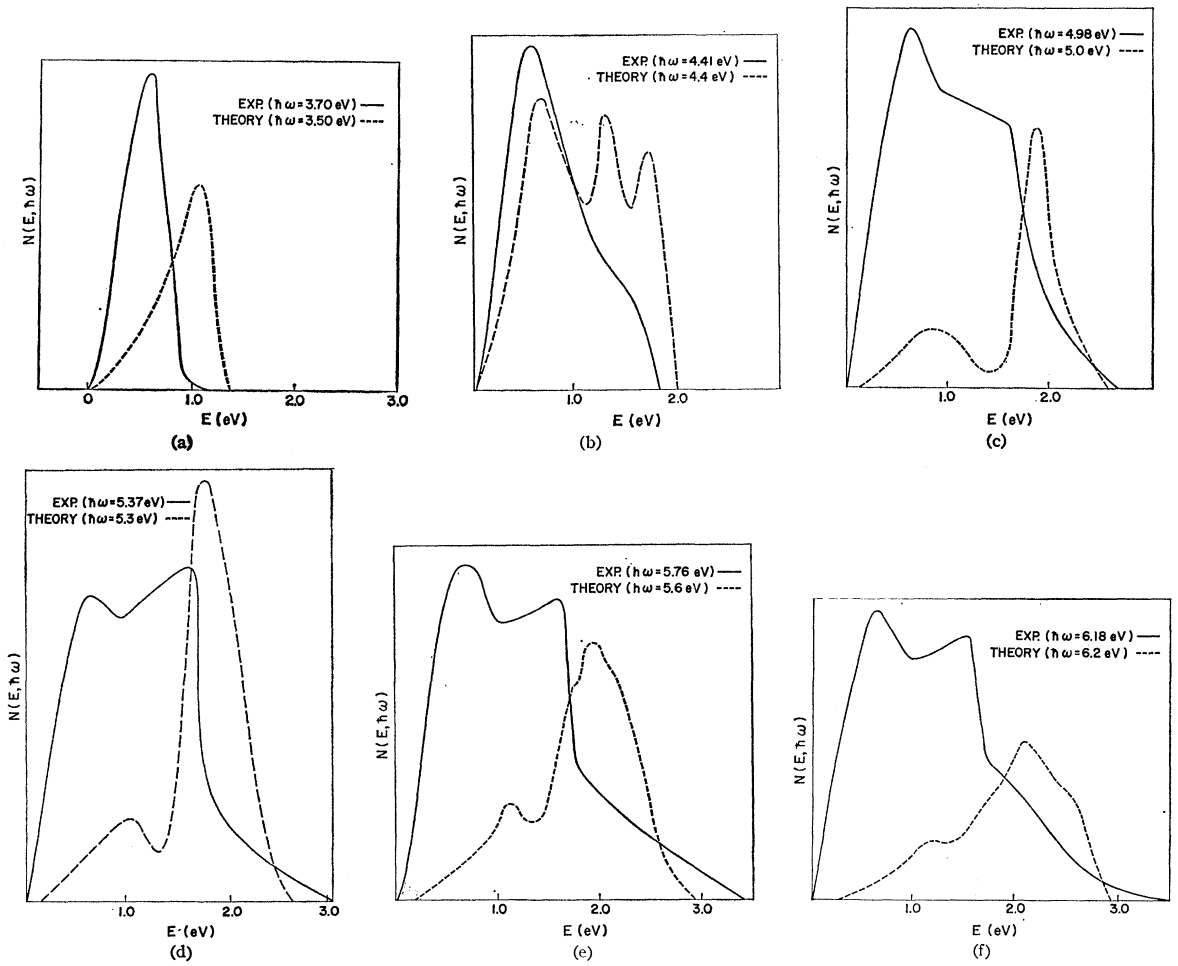


FIG. 17. Comparison of the preliminary energy distributions of Gobeli and Allen with the results of computation. The experimental vacuum level is $\sim$2.7 eV and the theoretical one is 2.5 eV.

C. Comparison of Theory and Experiment

In Fig. 17 we directly compare our current theoretical-energy distributions with the preliminary results of experiments by Gobeli and Allen[24] which are similar, apart from peak heights, to those of Spicer.[25] We find that all of our structure is in good one-to-one relationship with experiment for the values of $\hbar\omega$ under study. The only disappointment in this respect is the failure of the experiment to show any evidence for the $\Sigma_4 \rightarrow \Sigma_1$ peak we see in our 4.4-eV distribution. There is a possibility that the experiment missed the narrow-frequency range required to observe it. From the experiment it appears that the series-B peaks are rigidly fixed with respect to kinetic energy for $\hbar\omega > 5.0$ eV, whereas our results show a slight movement of the peak. The experimental resolution, however, is such as to not rule out a slight variation in kinetic energy as $\hbar\omega$ is changed.[26] The point of greatest disagreement between theory and experiment is in the difference between our peak intensities and those seen experimentally. The main discrepancy is the relative heights of peaks of series A and B. We might assume that a large number of Λ electrons are degraded in energy, as, for example, by Auger processes. This would lower the peak heights of series B with respect to those of series A. For the sake of discussion, we may assume that the heights of the peaks of series B are reduced by a fixed fraction. In Fig. 18 we compare the experimental strength of series B with theory (the latter reduced by a fixed factor), where both have been normalized to the peak heights of series A. This procedure allows us to eliminate errors which enter the problem due to an inapprorpiate treatment of $P(E)$. It is clear that $P(E)$ may be substantially reduced at energies corresponding to that of series B by energy-dependent scattering effects. This method of comparison eliminates $P(E)$ from the problem by comparing intensities at nearly fixed energies (i.e., series A and series B). We therefore get a good test of the density-of-states arguments upon which the analysis of the photoelectric emission depends.

It would appear then that our model of the photoelectric process in Si including only direct transitions between bulk valence and conduction bands is able to explain the basic features of the data. A more complete calculation to give better results for the line shapes should include dynamical effects that we have neglected. From the present calculation it seems that indirect transitions are considerably weaker than for cadmium sulfide where Spicer[27] has found that they must be included to explain his data.

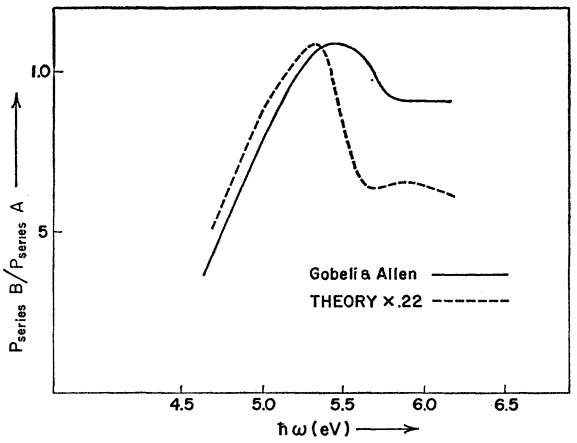


FIG. 18. Ratio of peak heights in series B to those of series A. Series B has been multiplied by an arbitrary factor to approximately normalize the two curves.

VII. FURTHER DISCUSSION OF THE DIELECTRIC CONSTANT

In Ref. 8 we considered $\epsilon_2(\omega)$, the imaginary part of the dielectric constant. We found good over-all agreement between theory and experiment for Si, however, there did not seem to be any evidence for the sharp peak at 3.4 eV seen in the measurements.[28] Our present study including a sample 20 times larger than previously employed does clearly show the peak (see Fig. 19). The peak is associated with a critical point at $\Gamma(\Gamma_{25} \rightarrow \Gamma_{15})$, and perhaps another one on Δ as seen in Fig. 20; although, an increase in our interband energy ~ 0.02 eV at the critical point would serve to flatten out the kink in the bands

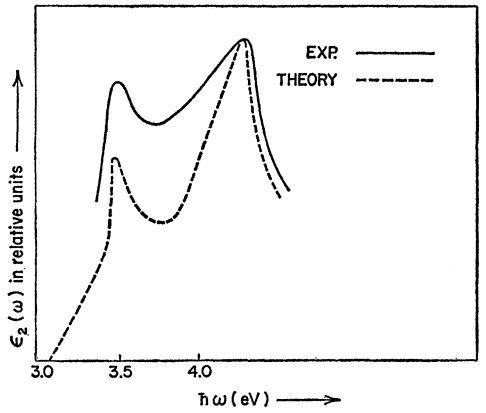


FIG. 19. Comparison of $\epsilon_2(\omega)$ as presently computed with experiment (Ref. 28). We have taken a constant dipole matrix element in constructing $\epsilon_2(\omega)$, and have normalized the two curves at 4.2 eV. Only the $4 \rightarrow 5$ contribution was included since for this energy range contributions from all other pairs of bands are negligible (Ref. 8).

[24] G. W. Gobeli and F. G. Allen (to be published).
[25] W. E. Spicer and R. E. Simon, Phys. Rev. Letters **9**, 385 (1962).
[26] F. G. Allen (private communication).
[27] N. B. Kindig and W. E. Spicer (to be published).
[28] H. R. Philipp and E. A. Taft, Phys. Rev. **120**, 37 (1960).

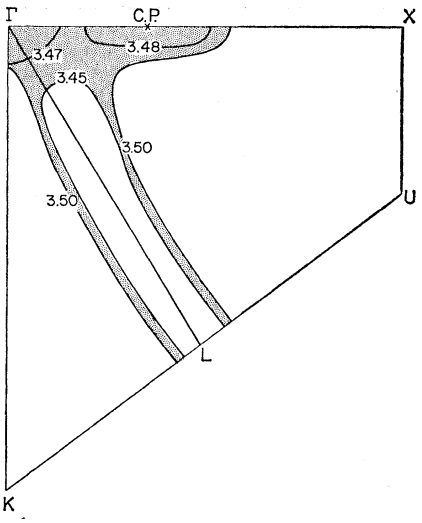


FIG. 20. The 4 → 5 energy bands in the 110 plane for interband gap energies near ~3.5 eV. The shaded portion represents the region of the 110 plane which contributes to the 3.5 peak. The two narrow strips along Λ are drawn considerably wider than they actually appear.

along Δ enough to eliminate this c.p.[29] The important thing to notice in Fig. 20 is that most of the states contributing to the 3.4-eV peak come from points near Δ; although, we may speak of the c.p. from the topological viewpoint as being located at Γ.

This interpretation is supported by reflectance measurements on stressed Si samples by Gerhardt.[30] Gerhardt finds, when he applies a stress along the (111) direction, that the effect is to shift the optical peak in energy but to leave its shape the same. This is what we expect with our model, since all the Δ axes should be equally affected by a (111) stress. When Gerhardt applies a (100) stress, however, the peak broadens. Again this is what our model should predict since a (100) stress will move states on different Δ axes by different amounts depending on whether the Δ axis is perpendicular or parallel to the stress direction.

[29] Actually a thorough survey of k space near Γ might very well show a cluster of weak critical points. Since the positive identification of these is probably beyond the limits of our present procedure, the search for them does not appear to be warranted.

ACKNOWLEDGMENTS

The author is grateful to F. G. Allen and G. W. Gobeli of Bell Telephone Laboratory for extensive discussion of their data prior to publication. He would particularly like to thank E. O. Kane for many stimulating discussions and for allowing the author to read his manuscript before publication. Thanks are due to H. Y. Fan and S. Rodriguez for useful conversations, to H. M. James for providing the author with ideal working conditions during his stay at Purdue University, and to the staff at Northwestern University for final preparation of the manuscript. He would also like to acknowledge the 7090 staff at Purdue for technical assistance. Finally, the acknowledgments could not possibly be complete without the author expressing his gratefulness to J. C. Phillips for having introduced him to the study of energy bands and having encouraged him to undertake this problem.

[30] U. Gerhardt, Phys. Letters 9, 117 (1964).