Exponential Band Tails in Random Systems

C. M. Soukoulis,[a] M. H. Cohen, and E. N. Economou[b]
Corporate Research-Science Laboratory, Exxon Research and Engineering Company, Annandale, New Jersey 08801
(Received 21 March 1984)

We present a simple derivation of the exponential band tails universally observed in three-dimensional disordered materials. The physical picture employed is that of Halperin and Lax in which states are localized by long-wavelength potential fluctuations. When the effect of the small-scale fluctuations of the potential are included as well, via the scaling arguments of Thouless, there results an exponential dependence of the density of states on energy below an energy E_1. The magnitudes and dependences of E_1 and the width of the tail on disorder agree with the experiment.

PACS numbers: 72.80.Ng, 71.50.+t, 71.55.Jv

An exponential relationship between the absorption coefficients and the photon energy, known as the Urbach rule, was first enunciated in 1953 for the observed optical-absorption edge in AgBr.[1] In many insulators and crystalline semiconductors[2] similar tails have subsequently been observed in the optical absorption. More recent experiments on optical-absorption coefficients[3] have yielded indirect experimental evidence for exponential tails in the densities of states (DOS) of amorphous semiconductors through a demonstration of the internal consistency of the assumption of constant dipole matrix elements.[3] Stronger and more direct experimental evidence for exponential tails in individual bands in those materials derives from a range of studies of trap densities.[4] Usually attributed to disorder, i.e., defects, impurities, and thermal disorder, the physical origin of this essentially universal behavior is still not well understood despite extensive experimental[2,3] and theoretical studies[5-11] in recent years.

Various attempts[11-21] have been made to calculate the low-energy tail of the density of states of a disordered system within the physical picture of independent electrons without mutual interactions, the electron-phonon interaction, or spin-flip scattering. In all these attempts it has been recognized that the density of localized states is several orders of magnitude smaller than the density of extended states, implying that the former arise from special atomic configurations. Consequently, one is forced to abandon mean-field-like or coherent-potential-approximation-like theories and even numerical simulations to obtain the density of localized states. They arise from potential fluctuations on wavelength scales of at least several atomic distances due to a variation in the physical parameters of the disordered system on the same scale. When the potential varies slowly enough the fluctuations in the energies of states mirror the fluctuations in the potential energy. This approach has been discussed by Lifshitz[12] and by Bonch-Bruevich.[13] Kane[14] and Eggarter and Cohen[15] have combined the fluctuations with the semiclassical Thomas-Fermi method to calculate the DOS. Since the potential-energy fluctuations are Gaussian, the tail found by Kane and by Eggarter and Cohen is Gaussian. By including the kinetic energy of localization omitted in the Fermi-Thomas procedure, Halperin and Lax[16] obtained tails falling off less rapidly than Gaussian. Feynman path integrals,[17] field-theoretic treatments,[18] and integration in other function spaces[19] as well as variational calculations[20] yield essentially the same results as those of Halperin and Lax. It is clearly established by any of the above treatments[21] that wells of size λ in which the minimum kinetic energy of localization is proportional to χ_d/λ^2 gives a DOS $N(E)$ which behaves as

$$\ln N(E) = \ln(N_0) - \left(\frac{4}{4-d}\right)^2\left(\frac{4-d}{d}\right)^{d/2} \times \frac{\chi_d^{d/2}}{2W^2L^d}|E|^{(2-d/2)}, \quad (1)$$

where N_0 is the preexponential, $\chi_d = d\pi^2\hbar^2/2m$, L is the correlation length of the potential fluctuations and is of atomic size, d is the space dimensionality, and W^2 is the variance of the random potential. $|E|$ is measured from the bottom of the conduction band. The energy dependence in Eq. (1) can be understood as follows: The factor $|E|^2$ comes from the amplitude of the potential fluctuation which becomes Gaussian on the length scale $\lambda >> L$, and the factor $|E|^{-d/2}$ comes from the spatial extent of the potential fluctuations λ_d, which scales as $|E|^{-d/2}$ when one assumes that the kinetic energy $|E|$ of confinement of the wave function for any dimension goes like $1/\lambda^2$.

The above treatments, which give the exact one-dimenisonal (1D) asymptotic behavior for the DOS,[16] give an $|E|^{1/2}$ dependence in the logarithm of the DOS for 3D in disagreement with the linear dependence seen experimentally. Another puzzling

feature is the hierarchy of magnitudes of the characteristic energies: For example, from the broadening of the main peak in the absorption coefficient in passing from the crystalline[3] to the amorphous[3] (a) form of silicon, one estimates W to be of order electronvolts. The quantity $E_c - E_1$ can be estimated[4] to be of order tenths of electronvolts in a-Si, E_c being the mobility edge and E_1 being the energy below which the exponential behavior occurs. E_0, the slope of the exponential tail,[3,4] is of order hundredths of electronvolts in a-Si.

Thus, a major problem exists. A well established theory gives results in contradiction to experiment in three dimensions but exact in one dimension. One can attribute this difficulty either to the omission of interactions or to an inadequacy in the Halperin and Lax analysis in the absence of interactions. In this Letter, we take the latter approach and identify a crucial omission in the previous treatments of the DOS calculation and show that, when corrected, the Halperin and Lax class of theories leads to an $\exp(-|E|/E_0)$ dependence in the DOS for 3D, while still giving the exact result[16] in 1D. The correct hierarchy of magnitudes for W, $E_c - E_1$, and E_0 is also obtained. The effects of interactions remain to be assessed.

The discrepancy between theory and experiment for the DOS in the 3D cases derives from the assumption that the energy needed to localize a wave function to a potential well of size λ goes as $1/\lambda^2$ in a disordered system. This assumption ignores the short-wavelength fluctuations of the random potential, i.e., it ignores the effects of disorder on the energy of localization. To obtain the energy change ΔE due to constraining an extended state of energy E' into a volume of linear dimensions λ in the presence of disorder, we observe that ΔE equals (within a numerical factor) the shift of the energy due to changing the boundary conditions from periodic to antiperiodic. Thouless[22,23] has shown that this shift can be expressed in terms of the density of states per unit volume $N(E')$ and the dimensionless conductance $g(\lambda,E')$. Thus we obtain for the energy of localization ΔE in a three-dimensional material

$$\Delta E = Bg(\lambda,E')/\lambda^d N(E'), \tag{2}$$

with $d=3$, where B is a constant. For lengths much higher than interatomic distances we can use the Vollhardt and Wolfle[24] expression

$$g(\lambda,E') = \frac{1}{\pi^3}\left[1+\frac{\pi}{2}\frac{\lambda}{\xi(E')}\right]$$

in three dimensions, where ξ is the correlation length which characterizes the spatial extent of the amplitude fluctuations of extended states of energy E' above the mobility edge. As E increases above E_c, these amplitude fluctuations become less severe and finally disappear near an energy E_u at which ξ has reached ξ_0.[25] At E_u, $\Delta E = \chi_3/\lambda^2$ for $\lambda >> \xi_0$ because of the uniformity of the wave function. Comparing this with (2) gives $2\pi^2\chi_3\xi_0 N(E_u)$ for B. For $E_c \leq E' \leq E_u$, it is possible that $\xi_0 << \xi$, when $g(\lambda) \rightarrow 1/\pi^3$ and $\Delta E \cong 2\chi_3\xi_0/\pi\lambda^3$ [$N(E_u) \cong N(E')$]. This value of ΔE is much smaller than χ_3/λ^2 because it costs much less energy to compress a highly fluctuating eigenfunction than a uniform one.

It is easy to show from the Lloyd and Best variational[20] principle that the optimal choice of E' is $E' \simeq E_c$ so that $\lambda << \xi$ for any λ, which enables us to use $\Delta E \simeq 2\chi_3\xi_0/\pi\lambda^3$. By repeating any of the previous calculations of the Lax and Halperin type for finding the DOS in the tail with this new expression for ΔE we get

$$N(E) \sim \exp\left(-\frac{16\chi_3\xi_0}{3\pi W^2 L^3}|E|\right), \tag{3}$$

in 3D, where $|E|$ is measured from the mobility edge E_c. The inverse slope E_0 of the exponential is given by[26]

$$E_0 = 3\pi W^2 L^3/16\chi_3\xi_0. \tag{4}$$

Clearly E_0 is only a small fraction of W. The variational principle of Lloyd and Best[20] gives

$$\lambda^3 = \tfrac{1}{3}\times 2\chi_3\xi_0/\pi|E|, \tag{5}$$

for the optimal value of λ.

For the argument leading to Eq. (3) to hold, the potential well, of extent λ and of depth $|E|+\Delta E$, must be isolated. Isolation is guaranteed in 3D if the probability that such a potential fluctuation will occur is less than 10^{-2}. Thus the smallest value of $|E|$ for which this condition can be satisfied, $E_c - E_1$, satisfies

$$\int_{-\infty}^{-(E_c - E_1 + \Delta E)} P(V)dV \cong 10^{-2}, \tag{6}$$

where $P(V)$ is the probability distribution of the average potential in a region of volume λ^3. It has zero mean and a variance $W^2(L/\lambda)^3$. From Eq. (6) we have $E_c - E_1 = (0.6 \text{ to } 0.8)W(L/\lambda)^{3/2}$ so that from Eq. (5) we obtain

$$(E_c - E_1) = (1.2 \text{ to } 2.3)\frac{\pi W^2 L^3}{2\chi_3\xi_0}. \tag{7}$$

Note that $E_c - E_1$ is of order $10^{-1}W$, while E_0 from Eq. (4) is of order 10^{-2} of W. The ratio $|E_1|/E_0 \simeq 3.2$–6.2 is in agreement with experiment.[3]

It must be noted, however, that Eq. (3) does not hold arbitrarily deep into the tail. Equation (5) predicts that λ decreases continuously into the tail. Ultimately λ would become smaller than the correlation length L of the potential or the length η below which λ^{-2} scaling of ΔE sets in.[23] If L and η were both larger than interatomic distances with $\eta > L$, there would be a transition to Halperin-Lax behavior followed by a transition to a Gaussian. With $L > \eta$, there is only a transition to a Gaussian. If L and η are of atomic size, as is to be expected, the behavior in the deep tail depends on the particular system.

Our arguments for the exponential behavior of the DOS are very general and can be applied to any type of disordered system whether the disorder is of structural, compositional, or thermal origin.[27] The actual values of E_0 and E_1 we obtain for a-Si are in agreement with experiment,[3,4] but the true significance of the above results is their essentially universal applicability. However, as the previous theories gave the exact result in 1D and failed in 3D, we must confirm that our method of calculation does not fail in 1D while yielding the correct result in 3D. In 1D the conductance is given by $(e^{2\lambda/l(E')}-1)^{-1}$, where $l(E')$ is the localization length. Insertion into (2) gives $\Delta E \sim \lambda^{-2}$ only if $\lambda \ll l(E')$. Use of the variation principle shows that the optimal choice of E' is at the position of the band edge in the absence of disorder, where $N(E')$ peaks as long as W is sufficiently smaller than the bandwidth. Further use of the variational principle with $\Delta E \sim \lambda^{-2}$ gives $\lambda \sim |E|^{-1/2}$. Thus, the condition $\lambda \ll l(|E'|)$ becomes a condition on $|E_1|$ which is more stringent than the condition that the potential fluctuate be isolated. If the former condition is satisfied, the correct asymptotic behavior, $N(E) \sim \exp(-|E/E_0|^{3/2})$, $|E| < (E_c - E_1)$ is obtained by other methods in 1D, with E_0 given by Eq. (1) and

$$E_c - E_1 \simeq 3\pi^2 W \tag{8}$$

on the assumption that $\lambda/l(E_1) \simeq 0.1$.

The above results for the power y of magnitude $|E|$ in the exponential can be generalized to continuous dimension:

$y = 1$ for $d \geq 2$ and $y = 2 - d/2$ for $d < 2$.

In sum, we find that one feature of the physical picture of band-tail states embodied in the Halperin-Lax theory remains essentially correct. That is, below an energy E_1, localized states have a physical extent governed by the size of the potential fluctuation localizing them. The fact that Halperin and Lax and their successors obtained the wrong energy dependence of the density of states in $2+\epsilon$ dimensions derived from their use of the incorrect scaling between energy of localization and size of the localized state. The Halperin and Lax arguments do not apply in any way to localized states with energies $E_c > E > E_1$. In that energy range, the long-wavelength potential fluctuations merge, and the physical extent of the localized states is determined by the localization length, which diverges at E_c. Arguments analogous to those above can also be used to estimate the energy E_u at which ξ decreases to ξ_0 and fluctuations in the extended states lose their importance. Just as the attractive potential fluctuations can bind a state of sufficiently low energy, they can also lead to resonances in the extended states. Moreover, repulsive-potential fluctuations lead to an exclusion of the extended state. The amplitude fluctuations thus disappear when the energy is high enough for there to be only isolated positive potential fluctuations. Thus $E_u - E_c$ is given by an expression like the right-hand side of (7) and, for a-Si, for example, is of order tenths of electronvolts.

We would like to thank G. D. Cody for useful discussions about the experimental situation and Professor B. I. Halperin for useful correspondence.

[a]Present address: Department of Physics, Iowa State University, Iowa 50011.

[b]Permanent address: Department of Physics and Research Center of Crete, University of Crete, Heraklio, Crete, Greece.

[1]F. Urbach, Phys. Rev. **92**, 1324 (1953).

[2]D. J. Dunstan, J. Phys. C **30**, L419 (1982).

[3]G. D. Cody, T. Tiedje, B. Abeles, B. Brooks, and Y. Goldstein, Phys. Rev. Lett. **47**, 1480 (1981); G. D. Cody, *The Optical Absorption Edge of a-Si:H_x in Amorphous Silicon Hydride,* Semiconductors and Semimetals Vol. 21B, edited by J. Pankove (Academic, New York, 1984), p. 11, and references therein.

[4]T. Tiedje, J. M. Cebulka, D. L. Morel, and B. Abeles, Phys. Rev. Lett. **46**, 1425 (1981).

[5]M. V. Kurik, Phys. Status Solidi (a) **8**, 9 (1971), and references therein.

[6]D. Redfield, Phys. Rev. **130**, 916 (1963); J. D. Dow and D. Redfield, Phys. Rev. B **5**, 594 (1972).

[7]J. J. Hopfield, Comments Solid State Phys. **1**, 16 (1968).

[8]H. Sumi and Y. Toyozawa, J. Phys. Soc. Jpn. **31**, 342 (1971).

[9]J. Skettrup, Phys. Rev. B **18**, 2622 (1978).

[10]M. Schreiber and Y. Toyozawa, J. Phys. Soc. Jpn. **51**, 1528, 1537, and 1544 (1982).

[11]S. Abe and Y. Toyozawa, J. Phys. Soc. Jpn. **50**, 2185

(1981).

[12]I. M. Lifshitz, Adv. Phys. **13**, 483 (1964).

[13]V. L. Bonch-Bruevich and A. G. Mironov, Fiz. Tverd. Tela. **3**, 3009 (1962) [Sov. Phys. Solid State **3**, 2194 (1962)].

[14]E. O. Kane, Phys. Rev. **131**, 79 (1963).

[15]T. P. Eggarter and M. H. Cohen, Phys. Rev. Lett. **25**, 807 (1970).

[16]B. I. Halperin and M. Lax, Phys. Rev. **148**, 722 (1966), and **153**, 802 (1967).

[17]S. F. Edwards, J. Phys. C **3**, L30 (1970), and J. Non-Cryst. Solids **4**, 417 (1970).

[18]J. L. Cardy, J. Phys. C **11**, L321 (1978).

[19]J. Zittartz and J. S. Langer, Phys. Rev. **148**, 741 (1966).

[20]P. Lloyd and P. R. Best, J. Phys. C **8**, 3752 (1975).

[21]Z. Jiman, *Models of Disorder* (Cambridge Univ. Press, Cambridge, England, 1979).

[22]D. J. Thouless, Phys. Rev. Lett. **39**, 1167 (1977).

[23]M. H. Cohen, E. N. Economou, and C. M. Soukoulis, Phys. Rev. Lett. **51**, 1202 (1983).

[24]D. Vollhardt and P. Wolfle, Phys. Rev. Lett. **48**, 699 (1982).

[25]When all energies of interest are small compared to the bandwidth and all lengths large compared to interatomic separations, the Hamiltonian becomes dimensionless and universal when expressed in units of $E=\gamma^2(2m/\hbar^2)^3$ for energy and $l=(\hbar^2/2m)^2/\gamma$ for length, where $\gamma=W^2L^3$, in three dimensions. Any characteristic length such as ξ_0 is a universal number times l under those circumstances, i.e, $\xi_0\sim l\sim W^{-2}$. When W increases to the point that l becomes of order interatomic separations a, then ξ_0 becomes of order a and remains there as disorder increases further.

[26]In the event that $L\approx\xi_0$, as for strong disorder and short correlation length, Eq. (4) simplifies to $E_0=(1/16\pi)W^2(\hbar^2/2m\xi_0^2)^{-1}$ for which the smallness of E_0/W is evident.

[27]When many independent sources of disorder coexist, each one producing a variance W_i^2, the variance $W^2=\sum_i W_i^2$ [see Ref. 3 and B. A. Vaid, K. C. Sarma, and D. R. Sharma, Phys. Lett. **102A**, 373 (1984)]. In particular, the thermal lattice vibrations produce through electron-phonon interactions a $W_i^2\sim T$ for high temperatures. For disorder large enough that ξ_0 is of atomic size (Ref. 25), Eq. (4) implies that $E_0\sim W^2$ and therefore linear in T. For smaller disorder $\xi_0\sim W^{-2}$ and $E_0\sim W^4$ holds (Ref. 25), where linearity in T follows only if the thermal disorder is a relatively small contribution to W^2. In the case of large disorder dominated by phonons as in crystals at high temperature, proportionality to T holds, a result obtained by previous authors from a totally different approach, optical absorption of excitons interacting with phonons (Ref. 8) or simple electron-phonon interactions without consideration of localization (Ref. 9). It is now known that the electron-phonon interaction can cause polaron formation above the mobility edge E_c and that localized states below E_c are localized further by the electron-phonon interaction (Ref. 23). While no one has yet done a full theory of the electron-phonon interactions in disordered systems, we have preliminary results which show that the polaron effects add an adiabatic term to E_0, which decreases with increasing T, whereas growing thermal disorder contributes a countervailing increase, as just discussed.