Band tails, path integrals, instantons, polarons, and all that

by M. H. Cohen
M. Y. Chou
E. N. Economou
S. John
C. M. Soukoulis

This paper reviews the explanations recently developed by the authors and their collaborators of how disorder leads to exponential band tails and to Urbach tails in optical absorption. It starts with the simplest single-potential-well models which, despite their simplicity, are remarkably successful in accounting for the experimental facts. It then identifies the weaknesses, hidden or explicit, in these models and shows, step by step, how they can be corrected by increasing the sophistication of the procedures used. Exact results are finally achieved through use of field-theoretic techniques, and appropriately formulated single-potential-well models are shown to reproduce these quite accurately. It is also shown that the probability distribution of the random potential must be close to Gaussian, with an autocorrelation function which cuts off fairly rapidly with distance for there to be a well-defined, broad energy range in which there are exponential band tails in the density of states and Urbach tails in the optical absorption.

1. Introduction

• *A historical perspective*

It is convenient to begin setting the problem of band and Urbach tails into historical perspective with Sommerfeld's free-electron theory of metals [1]. While that theory resolved through the introduction of quantum mechanics such paradoxes as the specific heat of metals, it revealed another difficulty. The mean free path of electrons in metals could not be understood as arising from scattering off each ion core. Bloch's solution to the mean-free-path problem lay in recognizing that the electrons diffracted rather than scattered in the perfectly periodic structure of the ideal crystal [2]. The Bloch–Floquet theorem he proved implied that electrons have quantum states in perfect crystals with energies which fall into allowed bands with sharp edges, separated by gaps. Wilson exploited this advance in developing his two-band model of semiconductors, a filled valence band separated by a band gap from the empty next higher band, the conduction band [3]. The main consequences of these ideas were worked out in the thirties, but substantial progress in semiconductor physics and technology did not take place until the immediate postwar period of the late forties and early fifties. As a consequence of extensive studies of doped and compensated semiconductors, it was realized during that period that the introduction of disorder caused tails to appear at band edges in the density of states. The earliest theories, semiclassical in nature, appeared in the sixties

[4, 5]. Finally, in a systematic and thorough study, Urbach documented the existence of exponential tails in the optical absorption of ionic crystals in 1953 [6].

Since Urbach, such exponential tails have been found in a wide range of materials, both in the optical absorption and in the density of states of individual bands. There have been many attempts to explain the existence of exponential tails. The best one can say of this literature is that it is confused, with a conflicting welter of contentions and interpretations. We show in this paper that the underlying physics is extremely simple, that the elements of an exact theory were already contained in the early work of Halperin and Lax [7], and that much of the confusion in the literature arose from a failure to appreciate the important effects of correlation in the random potential on the energy range over which exponential behavior occurs. This paper is in fact a review of the explanations developed by the authors and their collaborators in a recent series of papers [8–14] concerning the way in which disorder leads to exponential band tails and to Urbach tails; it puts that work in perspective.

• *Some facts about band tails*

Before discussing the experimental facts, it is necessary to recognize that two differing kinds of processes are studied experimentally. The first, or fast, processes involve induced transitions in which the transition time is much less than the relaxation times for both the electronic and atomic structures which remain, unrelaxed, in their initial configurations. Examples of such processes are optical absorption and transient photocurrents. The second, or slow, processes involve measurements taken on time scales such that the electronic and atomic structures are fully relaxed. These processes are much more difficult to understand, for reasons to be discussed later in this article, and we shall focus primarily on fast processes.

Ionic crystals, other insulators, crystalline semiconductors, and amorphous semiconductors all show exponential tails in the energy dependence of the optical absorption

$$\alpha = \alpha_0 \exp\left\{-\frac{h\omega - E_F}{E_0}\right\} \tag{1}$$

below a continuum of the expected type for transitions between extended states in the valence and conduction bands. In (1), $h\omega$ is the photon energy, E_0 is the width of the exponential tail, E_F is the energy of the Urbach focus, to be explained below, and α_0 is a prefactor weakly dependent on experimental conditions. The width E_0 appears to have additive contributions from thermal origin, E_{TH}, and structural origin, E_{ST},

$$E_0 = E_{TH} + E_{ST}. \tag{2}$$

The temperature dependence of E_{TH} is linear above the Debye temperature θ, then dropping rapidly at θ and passing over to power law variation at low temperatures, clearly implicating phonons. E_{ST} increases with structural disorder, implying that E_{TH} arises from the instantaneous thermal disorder introduced by phonons which appear frozen in fast processes. If that part of the optical absorption due to extended-state transitions is extrapolated with the aid of a suitable theory, it appears to vanish at a threshold energy E_T, as discussed by Tauc [15]. Cody [16] has shown that there is a remarkable correlation among E_T, E_0, and E_F for a-Si:H,

$$E_T = cE_0 + E_F. \tag{3}$$

It is clear that the Urbach tail originates from disorder, but how? Does the exponential energy dependence of α arise from the combined density of states, from the matrix elements, from the densities of states of the individual bands, or from some combination of these?

The universality of the Urbach tail obviates explanations specific to particular materials, as, e.g., that of Dow and Redfield [17], which is particular to ionic crystals. Other experimental studies which obtain information from fast processes involving a single band, in fact, suggest that the element of universality is the existence of exponential tails in the densities of states of the conduction and valence bands individually. For example, Monroe and Kastner have shown [18] via transient photocurrent measurements that the valence band of glassy As_2Se_3 has an exponential tail extending over five orders of magnitude in the density of states and 0.6 eV in energy, a significant fraction of the band gap. One can also argue plausibly that the one-electron Hamiltonians of typical materials contain many matrix elements, that the structural disorder, static and dynamic, has many components, and that these components affect different matrix elements differently. This leads to statistical independence of the states in the conduction and valence bands. The latter in turn implies that the dipole matrix element in the optical absorption is an energy-independent constant and that the Urbach tails are indeed due to exponential tails on the individual bands [10]. Indeed, there is much evidence that in amorphous semiconductors the valence and conduction bands individually have exponential tails in the density of states. In the following, therefore, we concentrate on understanding how exponential tails occur on individual bands.

• *Summary*

We begin in Section 2 with a demonstration that the band tail is very sensitive to the probability distribution of the random potential, and argue that the appropriate probability distribution is Gaussian [9, 10]. In Section 3 we review the results of our detailed study of six very simple models, all based on supposing that the band tails arise from localized states generated by potential fluctuations in the form of isolated potential wells of predetermined shape [13]. These calculations yield exponential band tails and optical

absorption in agreement with experiment and expose clearly the underlying physics, but the assumptions in the model are untested. In Section 4, the arbitrary choice of well size made in Section 3 is corrected with what will turn out to be remarkably accurate results [11]. In Section 5, the treatment of the tail as disjoint from the continuum is cured via an application of the coherent potential approximator (CPA), which leads to an understanding of the Urbach focus and of Equation (3) [9], but the Halperin–Lax regime [7] is lost. In Section 6, the Halperin–Lax regime is regained via a treatment based on path integrals which embraces all regimes in the tail as well as the continuum. In Section 7, the single-well assumption is shown to be asymptotically exact and the most probable well shape is found as an instanton in a classical field theory [14]. The exact results of the instanton theory were largely contained in the early work of Halperin and Lax [7], but failure of the earlier workers to appreciate the importance of correlation in the random potential was responsible for much of the confusion which appears in the literature. Some comments are made in Section 8 on the extent of our knowledge concerning the influence of lattice dynamics on the electron band tail. The non-adiabaticity of the electron-phonon interaction is treated using the Feynman path-integral method. In the adiabatic limit, the essential physics may again be recaptured by means of simple potential-well arguments, provided the electron oscillation frequency is large compared to the oscillation frequency of the well itself. Finally, we discuss some unresolved complications arising from electron–hole interactions important in optical processes. The main conclusions of our work are summarized in Section 9.

It should be emphasized once again that we are reviewing here primarily our own contributions to the subject, with the work of others discussed only to the extent necessary to set our own work into proper perspective. Moreover, our various contributions as well as those of others are presented in a logical order, in which the ordering element is increasing sophistication, completeness and accuracy, rather than in historical order. Thus, simple potential-well model calculations [13] are presented in Section 3, whereas a more sophisticated and accurate potential-well calculation [11] is presented in Section 4, even though it was done earlier, because the arbitrary choice of well size characteristic of the simplest model calculations [13, 19] is relaxed in Section 4. Similarly, the earliest calculation which clearly and unobjectionably displayed the exponential tail was the CPA calculation by Abe and Toyozawa [20], but the discussion of the CPA is deferred to Section 5, after the potential-well models are discussed.

2. The importance of the probability distribution

There are two particularly simple models of an electron moving in a random potential for which exact results are available. The first is the Lloyd model [21], a simple, single-band, tight-binding model in which the single-site energies are independent random variables having identical Lorentzian probability distributions of half-width Γ about vanishing average value. The density of states $N(E)$ is

$$N(E) = \int dE' N_0(E') \frac{1}{\pi} \frac{\Gamma}{(E' - E)^2 + \Gamma^2}, \tag{4}$$

where $N_0(E)$ is the density in the periodic case ($\Gamma = 0$). One sees that disorder introduces tails into the density of states at the band edges and that the tail width is Γ. However, the tails are clearly not exponential.

The second model is that of an electron moving in a continuous random potential $V(r)$, which has a correlation function

$$\langle V(r)V(r')\rangle = w^2 C(r - r'). \tag{5}$$

In (5), w^2 is the variance of $V(r)$ at a point ($\langle V(r)\rangle = 0$). If $C(R)$ is a smooth, monotonically decreasing function of R, one can define a characteristic length L through

$$\int d^3R C(R) = \frac{4\pi}{3} L^3. \tag{6}$$

If L is much greater than the deBroglie wavelength for the electron energies of interest, the motion of the electrons in $V(r)$ can be treated semiclassically [4, 5]. The density of states is then

$$N(E) = \int_{-\infty}^{E} dV N_{sc}(E - V)P(V), \tag{7}$$

where N_{sc} is the semiclassical, i.e., free-electron, density of states and $P(V)$ is the probability distribution of the random potential.

One learns from these two examples that the band tails introduced by disorder can be very sensitive to the details of the probability distribution of the associated random potential. It is therefore essential for what follows to have an accurate representation of $P(V)$.

Now in fast processes, the atoms appear to the electrons as frozen in their instantaneous positions, displaced by phonons from their equilibrium positions. These displacements, u, introduce into the Hamiltonian governing the electron motion a random potential V which, in general, is linear in u. This V has a multivariate Gaussian probability distribution because u does. That is, the random potential associated with thermal disorder has a Gaussian probability distribution.

In structurally disordered materials such as amorphous semiconductors, the structural disorder has many components. In covalently bonded materials, for example, there can be bond-length, bond-angle, and dihedral-angle variation; coordination defects; odd rings randomly dispersed in a matrix of even rings, etc. [22]. Each of these random elements influences in turn many matrix elements. Thus, loosely speaking, one can imagine something like a

central limit theorem holding so that $P(V)$ can be accurately represented once again by a multivariate Gaussian, this time because of the intrinsic complexity of the physical situation.

Finally, even when $P(V)$ is decidedly non-Gaussian, as for a binary alloy, there are regions in the probability space of substantial physical importance within which it is well approximated as a multivariate Gaussian. Thus, in the following, we shall confine ourselves to Gaussian random potentials. As we shall shortly see, *only* Gaussian random potentials produce exponential tails.

3. Simple potential-well models for band tails

It was understood correctly from the beginning of the subject that band tails were associated with localized states split off the top of the valence band by locally repulsive potential fluctuations or off the bottom of the conduction band by locally attractive potential fluctuations. The mystery has been why the tails are exponential (Urbach). We start our considerations by reviewing an examination of the behavior of localized states split off by specific simple, idealized potential fluctuations [13].

Six cases were studied in Reference [13]:

1. Free electron continuum with an imposed square well of depth ϵ and radius a (of atomic size).
2. The same, but with a well of Gaussian shape.
3. A simple tight-binding model with a semicircular density of states and single-site energies which are independent random variables.
4. The same, but with a simple cubic density of states.
5. The conduction band of a-Si:H via an accurate tight-binding model previously developed for the optical and transport properties [23, 24].
6. The same for the valence band [23, 24].

All cases show the same qualitative dependence of the binding energy $|E|$ of the first bound state on well depth ϵ, sketched in **Figure 1**. There are three distinct regions: a threshold region, an intermediate region, and an asymptotic region. If one translates the parameters of each model into quantities appropriate to the actual experimental situations, one sees that the threshold region is very narrow on the experimental scale, the intermediate region *is* on the experimental scale, and the asymptotic region is impossible to reach except for very narrow bands separated by very wide gaps.

Let us therefore concentrate on the intermediate region, where

$$|E| \simeq A\epsilon^2 + B; \qquad E_1 < |E| < E_2. \tag{8}$$

Let us now suppose, according to the considerations of Section 2, that ϵ has a Gaussian distribution with variance w^2,

$$P(\epsilon) \propto e^{-\epsilon^2/2w^2}. \tag{9}$$

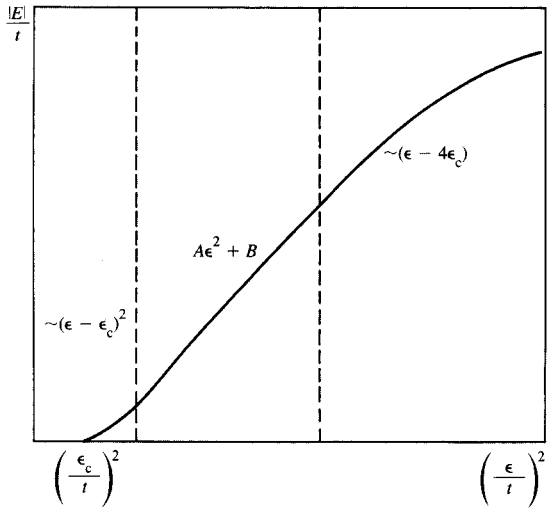

Figure 1

Schematic of dependence of binding energy $|E|$ on well depth ϵ for Cases 1–6. The asymptotic dependence is shown in the outer two regions. There is accurate linearity of $|E|$ in ϵ^2 in the intermediate region. t is a scaling energy which is different in each case. The horizontal energy scale is expanded in the left region and greatly compressed in the right.

Table 1 Parameters relating binding to well depth (after [13]).

Case	*m**	*A*	E_1	E_2	*Energy unit*
1	Yes	0.023	0.5	6	$\hbar^2/2m^*a^2$
2	Yes	0.010	0.3	6	$\hbar^2/2m^*a^2$
3	No	0.064	0.2	3	*V*
4	No	0.065	0.2	3	*V*
5	No	0.007	0.1	0.3	*eV*
6	No	0.003	0.1	0.3	*eV*

Insertion of (8) into (9) immediately yields exponential band tails,

$$N(E) \propto e^{-|E|/E_0}, \tag{10}$$

$$E_0 = 2Aw^2, \tag{11}$$

in a very simple and direct way. Moreover, the calculated values of A together with estimates of w yield values of E_0 in good quantitative agreement with experiment, as discussed below. The extent λ of the bound state is comparable to a in (E_1, E_2).

The detailed results of [13] are given in **Table 1**. The first column of the table lists the cases studied, numbered as above in the text. The second column states whether an effective mass approximation was used. The third column

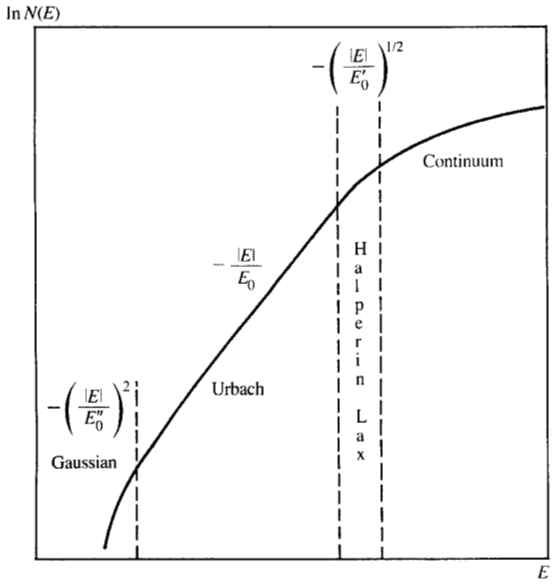

Figure 2

Schematic of energy (E) dependence of the density of states $N(E)$, showing the four different regions and the energy dependence of ln $N(E)$ up to an additive constant in each. The continuum and the Gaussian regions are contracted and the Halperin–Lax region expanded relative to the Urbach region.

lists A, which is dimensionless when ϵ is expressed in the units listed in the final column. A, E_1, and E_2 are defined in (8), and E_1 and E_2, given in the fourth and fifth columns, respectively, are also expressed in the units of the last column. In that column V is the nearest-neighbor electron transfer matrix element of the simple tight-binding models.

One sees that (1) the validity of the results is not dependent on making an effective mass approximation; (2) the form is universal, but the parameters are not; (3) exponential and Urbach tails follow from this universal form for the binding energy (8) *only* for Gaussian potentials; (4) $N(E)$ can drop by five orders of magnitude within (E_1, E_2); (5) there is quantitative agreement with experiment for specific materials [13]; (6) the range (E_1, E_2) is reduced in a-Si:H by the presence of a nearby interacting band.

Let us turn now to the other two regions. The threshold region begins at ϵ_c, the critical value of ϵ needed to bind one state. Inserting the dependence $|E| \propto (\epsilon - \epsilon_c)^2$ near ϵ_c into (9) gives

$$N(E) \propto e^{-(|E|/E_0')^{1/2}}, \tag{12}$$

with

$$E_0' \propto w^2/\epsilon_c. \tag{13}$$

Apart from numerical details, the result (12), (13) was first obtained by Halperin and Lax [7], and we call the region within which it holds the Halperin–Lax regime, and the corresponding part of the density of states the Halperin–Lax tail. Here, $\lambda \propto (\epsilon - \epsilon_c)^{-1}$ and is much larger than a.

In the deep tail, $|E| \sim \epsilon - 4\epsilon_c$ so that (9) yields

$$N(E) \propto e^{-[(|E|+4\epsilon_c)/E_0'']^2}, \tag{14}$$

$$E_0'' = \sqrt{2w}. \tag{15}$$

This asymptotic behavior is not reached until ϵ exceeds $10^2 t$, where t is the energy unit in Table 1. In Cases 5 and 6, it is never reached because of the presence of the other band. We call this regime the Gaussian regime because of (14), but in fact the central limit theorem may not hold in the deep tail, so that $N(E)$ may reflect the actual asymptotic form of $P(V)$ there. In the deep tail, $\lambda \ll a$ holds.

These results are summarized in **Figure 2**, where the three regions are clearly displayed. It should be pointed out that Halperin and Lax obtained all three regions, but did not recognize the intermediate region as an exponential tail because $(E_2 - E_1)/t$ happened to be particularly small in the case they studied. Plugging numbers into the detailed results of [13], one sees that the Halperin–Lax region is too narrow to see, the Gaussian region is too deep to see, but the Urbach region is observable.

Reference [13] contains comparisons with experiment for thermal disorder from acoustic and optical phonons, a comment on ionic crystals, and excellent values for E_0 for the conduction and valence bands of a-Si:H and of the valence band of As_2Se_3. One most significant result is that $E_0 \propto w^2$, Equation (11). Since for a Gaussian probability distribution the various contributions to the disorder contribute to w^2 by adding in the square, one immediately obtains Equation (2).

4. Correcting the arbitrary choice of *a*

The simple physical models discussed in the preceding section give clear physical insight into and a good quantitative account of the Urbach tail. Nevertheless, they contain five questionable features:

1. Are the wells really isolated?
2. a is chosen arbitrarily.
3. The well shape is chosen arbitrarily.
4. The tail is treated as disjoint from the continuum.
5. Correlation in the random potential was not considered explicitly.

In this section we summarize the results of an equally simple earlier investigation which eliminates questionable point 2 above [11].

Model 2 of the last section was studied, but with the generalization that the width a of the well was allowed to be a free variable. The best value of a was fixed by a variational argument separately for each energy. Point 5 was dealt with

by explicitly using a correlated Gaussian random potential. The results were identical to those of [13] qualitatively, but there were significant improvements. In particular, the results were exact in the Gaussian and Urbach regions but were 3% off the exact value of E_0' in the Halperin–Lax region. How we know what the exact answers are is discussed in Section 7. To be off by 3% in an essentially unobservable regime and to be exact elsewhere is a considerable achievement for a simple potential-well calculation.

In analogy with the discussion of Section 3, we consider a probability distribution for the Fourier components $V(k)$ of the random potential which takes the form

$$P\{V(k)\} \propto \exp[-S\{V\}], \tag{16a}$$

where

$$S = 1/2 \int \frac{d^d k}{(2\pi)^d} V(k) B^{-1}(k) V(-k). \tag{16b}$$

The choice of autocorrelation function $B(k) = \langle V(k)V(-k)\rangle$ is in fact crucial for the occurrence of a broad range of linear exponential behavior in the band tail. For convenience, we take

$$B(k) = V_{\text{rms}}^2 (\pi L^2)^{d/2} \exp(-k^2 L^2/4), \tag{17}$$

where we have introduced a correlation length L describing the spatial extent of short-range order in the solid, typically of the order of the interatomic spacing. By working in general dimension d, we can illustrate the role of dimensionality in pinching the Halperin–Lax region into an unobservably small energy window near the continuum edge. The probability of occurrence of a potential fluctuation of the form

$$V(x) = -V_0 \exp(-x^2/a^2) \tag{18}$$

of variable depth V_0 and range a is then given by (16), with

$$S = \frac{V_0^2}{2V_{\text{rms}}^2} \left(\frac{a}{L}\right)^d \left[2 - \left(\frac{L}{a}\right)^2\right]^{-d/2}. \tag{19}$$

The dominant contribution to the density of states (DOS) in the tail at an energy $-|E|$ arises from the most probable of the fluctuations of the form (18), with the additional constraint that the potential well (18) possess a ground state at the energy $-|E|$. Accordingly, we minimize (19) with respect to V_0 and subject to this constraint. Two limiting cases are worth noting. In the very shallow tail the scale of the localized state and the range a are large compared to L. It follows that the quantity in square brackets in (19) may be approximated by unity and $S \propto V_0^2 a^d$. The constraint on V_0 and a to produce a bound state at energy $-|E|$ takes the form $|E| + 1/a^2 = V_0$. For $d < 4$, the approximate S has a local minimum when the ratio of the range of the potential fluctuation to the electron deBroglie wavelength $\lambda = h/\sqrt{2m|E|}$ is given by $(a/\lambda)^2 = (4-d)/d$. Since this ratio vanishes as $d \to 4$ for any fixed E, it is evident that the influence of the correlation length L is felt even for relatively shallow band tail states in high dimensions. In particular, the Halperin–Lax region for $d = 3$ is pinched into a narrow energy window near the continuum edge. In the very deep tail the binding energy $-|E| \simeq -V_0$. Substituting into (19) and minimizing with respect to a gives $a = L$ and $S_{\text{min}} = |E|^2/2V_{\text{rms}}^2$, the well-known semiclassical result. If no approximation for S is made and the exact constraint between V_0 and a for $d = 3$ is used, the result illustrated in **Figure 3** is obtained. The energy scale $\epsilon_L = \hbar^2/(2mL^2)$ is typically ~0.5 eV, so that essentially the entire experimentally observable band tail lies in the crossover regime between the Gaussian tail, which is unobservable due to the presence of the valence band, and the Halperin–Lax region, which occurs over a range

$$0 < |E| < \left(1 - \frac{d}{4}\right)^2 \epsilon_L.$$

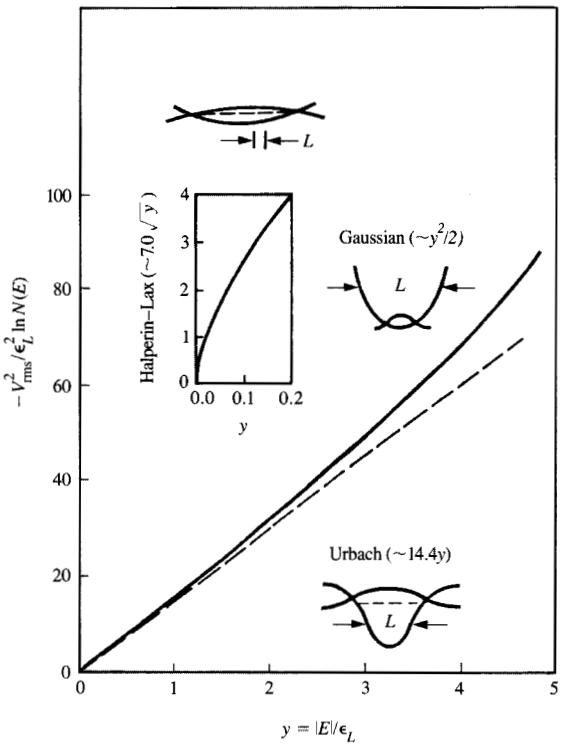

Figure 3

Exponential part of density of states as a function of $y = |E|/\epsilon_L$ exhibiting the Gaussian tail ($y >> 2$), the Urbach tail ($0.1 < y < 2$) and the Halperin–Lax tail ($0 < y < 0.1$; see inset). The nature of the most probable potential fluctuation and wavefunction in each regime is illustrated schematically.

The accurately linear exponential tail over the energy interval $0.1 < |E|/\epsilon_L < 2$ easily spans five decades in the DOS, and we associate this with the experimentally observed Urbach tail. As we show in Section 7, the restriction of potential fluctuations (18) to Gaussian form does not substantially affect the numerical accuracy of these results. In fact, it allows us to recapture to a high degree of accuracy the essential physics of the more formal path-integral and replica methods discussed in Sections 6 and 7.

5. Curing the disjuncture from the continuum: The CPA

Here we consider the results of the coherent potential approximation (CPA). The DOS can be expressed in terms of the single-particle Green's function

$$N(E) = -\frac{1}{\pi}\text{tr Im}\,G(E^-), \tag{20}$$

$$G(E^-) = \lim_{\delta\to 0^+}\frac{1}{E - i\delta - H}. \tag{21}$$

The Hamiltonian is supposed to have a regular, i.e., periodic, part H_0 and a random part V. Taking the trace in (20) is tantamount to taking an average over the random potential V, or

$$N(E) = -\frac{1}{\pi}\text{tr Im}\langle G(E^-)\rangle_V, \tag{22}$$

provided $P(V)$ satisfies certain nonrestrictive conditions. The effect of averaging over V can be represented as replacing V in H in (21) with a complex energy-independent self-energy,

$$\langle G(E^-)\rangle_V = \frac{1}{E - i\delta - H_0 - \Sigma(E^-)}. \tag{23}$$

In the CPA $\Sigma(E^-)$ is calculated by treating scattering of local potential fluctuations exactly but independently of all other potential fluctuations, which are treated in a mean field approximation through their replacement by Σ itself.

The CPA is easiest to implement for simple tight-binding models in which the single-site energies are treated as independent random variables. Thus, the only improvement afforded by the CPA over the simple potential-well calculations of Sections 3 and 4 is the capability of treating the tail and the continuum on the same footing by using a single method, and, as we shall see, some things are lost.

The CPA was first applied to this problem by Abe and Toyozawa [20], who showed that it yields Urbach tails for Gaussian random single-site energies. Their work was amplified considerably in [9]. We can summarize the results of the CPA as follows:

1. The Halperin–Lax region is lost because of the single-site feature of the CPA.
2. In the Urbach and Gaussian regions, the relationship between ϵ and E becomes

$$\frac{1}{\epsilon} = -G(E). \tag{24}$$

Thus, the ϵ dependence of $|E|$ in the simple models is seen to derive from the energy dependence of G, that is, from fundamental band properties.

3. Urbach tails occur only for Gaussian probability distributions $P(V)$.
4. The quantitative results for the Urbach and Gaussian regions are the same as before.
5. The Urbach focus corresponds to the bottom of the unperturbed band shifted to $\langle V\rangle_V$, the virtual crystal approximation. This works well for crystalline materials, and for a-Si:H the theory gives 2.0 eV for E_F [23, 24], whereas experiment gives 2.1 eV [16].
6. Finally, Equation (3) is obtained, with a value of c in agreement with that found by Cody [23, 24].

6. Finding the Halperin–Lax regime again; path integrals

The CPA catches the continuum, the Urbach, and the Gaussian regions, but it misses the Halperin–Lax regime, the well shape is not general, a is restricted, and the Gaussian potential is uncorrelated. The question of isolated potential wells is not addressed. Feynman's-path integral formalism does catch all of the regimes. It keeps all of the advances made by the CPA, and, in addition, a is not restricted. However, the well shape is not general [14], and isolation is again not addressed.

Sa-yakanit and collaborators, especially Glyde, have developed the application of Feynman's methods to the problem of an electron moving in a random potential in considerable detail [25–28]. Their procedure can be summarized as follows. The Green's function for propagation from a point x_1 at time t_1 to a point x_2 at time t_2 is represented by a path integral:

$$G(x_1, t_1; x_2, t_2) = \int D[X(t)]e^{\frac{i}{\hbar}A[x(t)]}, \tag{25}$$

$$A = \int_{t_1}^{t_2} dtL, \tag{26}$$

$$L = 1/2\ m^*v^2 - V(x). \tag{27}$$

The integration in (25) is over all paths $x(t)$ for which

$$x(t_1) = x_1,\ x(t_2) = x_2. \tag{28}$$

The Action A and the Lagrangian L are those of a particle of effective mass m^* moving in a continuum through a random potential V.

G is thus represented by a functional integral. However, the only functional integrals which can be evaluated are Gaussian; i.e., the action must be quadratic in the position x and the velocity $v = \dot{x}$. Feynman showed in his polaron theory how to replace a more complex A with a Gaussian A via a variational principle [27, 29]. In the present problem,

however, one can prove that a corresponding variational principle holds only in the deep tail. In the other energy regions, one blindly uses the Lloyd and Best variational principle [30] and hopes for the best.

Carrying out this program as was done by Sa-yakanit and Glyde, one gets results equivalent to the jointure of the CPA and the potential-well theories. The results are exact in the Gaussian and Urbach regions, the continuum is excellent, and the Halperin–Lax regime is present but with a value of E_0', which is about 10% off the exact value. Thus, the simple potential-well theory of Section 4 is somewhat more accurate than the path-integral theory in the tail, but it lacks the smooth transition to the continuum of the latter. One very important feature of the path-integral theory for the tail is that the results come out in analytic form. As this form is much easier to derive from the field theory of the next section, we postpone discussion of it until then.

7. Introducing asymptotically exact arguments, field theory, and instantons

An alternative functional-integral representation of the averaged Green's function is given by classical field theory:

$$\langle G(x, x'; E^{\pm})\rangle = \lim_{n\to 0} \int \tilde{D}[\phi]\phi^{(1)}(x)\phi^{(1)}(x')e^{-S[\vec{\phi}]}, \tag{29}$$

$$S[\vec{\phi}] = 1/2(\vec{\phi} \cdot (E^{\pm} - H_0)\vec{\phi}) - 1/8W^2 \int\int dxdx'\phi^2(x)C(x - x')\phi^2(x'). \tag{30}$$

The right-hand side of (29) is a functional integral over the n-component vector field $\vec{\phi}(x)$ in the limit $n \to 0$. $\phi^{(1)}(x)$ is a single component of the full vector $\vec{\phi}(x)$. The action $S[\vec{\phi}]$ contains a scalar and not a Hermitian product of $\vec{\phi}$ with $(E^{\pm} - H_0)\vec{\phi}$ as indicated by the dot. The random potential has been assumed to have a Gaussian multivariate distribution with correlation function given by (5).

Just as in the path-integral case, one can do only Gaussian functional integrals. Therefore, one evaluates (29) approximately by a saddle-point integration about a field $\vec{\phi}_0$ for which

$$\left.\frac{\delta S}{\phi\vec{\phi}}\right|_{\vec{\phi}_0} = 0. \tag{31}$$

For energies in the continuum, the only solution of (31) is at $\vec{\phi} = 0$. However, in the tail, as pointed out and worked out by Cardy [31], there are additional solutions of (31) for nonzero $\vec{\phi}_0$, the instantons of field theory.

The explicit form of (31) is

$$(E^{\pm} - H_0)\vec{\phi}_0(x) - 1/2w^2 \int dx' C(x - x')\vec{\phi}_0^2(x')\vec{\phi}_0(x) = 0. \tag{32}$$

Cardy confined himself to white-noise potentials for which $C(x - x')$ is a delta function. As a consequence, he was able to find only the Halperin–Lax region. John and Stephen found the deep tail by introducing a finite correlation length into $C(x - x')$ [32]. However, it is possible to solve (32) for all energies in the tail [14]. The essential point is that (32) has the same structure as the Hartree equation, so that all the advanced numerical techniques of electronic structure theory are available for its solution.

After the saddle-point integration is carried out, one obtains

$$N(E) = N_0(E)e^{-S[\vec{\phi}_0]}, \tag{33}$$

with the dominant energy dependence coming from the second factor. If, instead of inserting the true solution of (32) for $\vec{\phi}_0$ into S, one inserts a $\vec{\phi}_0$ of Gaussian form and minimizes S with respect to the amplitude and width of the Gaussian, one obtains the path-integral results of Sa-yakanit and Glyde [25–28] identically [14]. The resulting analytic form for the density of states is given explicitly in Section 3 of Reference [14]. There is a close analogy between the true instanton and the most probable potential. Thus, by solving the instanton equation of motion (32) exactly, one allows for the most general well shape possible.

Equation (32) appears over and over again in the theoretical literature on band tails. One sees it, e.g., in the work of Halperin and Lax [7], of Zittartz and Langer [33], and in various papers by Lifshitz and collaborators [34]. The use of field theory greatly simplifies its derivation, allows for the effect of spatial correlation in the random potential, simplifies to some extent the derivation and calculation of the prefactor $N_0(E)$ of Equation (33), and, perhaps most important, provides a basis for showing that the single-potential well approximation becomes asymptotically exact when $S[\vec{\phi}_0) \gg 1$. This last proceeds by carrying out formally a cumulant expansion for $N(E)$ around $\vec{\phi}_0$ [14], and $S[\vec{\phi}_0]$ becomes $\gg 1$ already in the Halperin–Lax region. Thus, the field-theoretic results for the tail are exact starting with the Halperin–Lax region in the tail and going deeper. Exact values E_0, E_0', and E_0'' are to be found in Reference [14] for various forms of the correlation function $C(x - x')$. On the basis of these last results we were able to assign errors to the results of the theories discussed in the previous sections.

By studying the influence of the shape of $C(x - x')$ on the energy dependence of $S(\vec{\phi}_0)$, one can understand the sources of much of the confusion in the literature about the origin of Urbach and exponential tails. It has often been stated that the exponential region in the density of states is simply a transition region between the Halperin–Lax and the Gaussian regions in which $\ln N(E)$ has a point of inflection and thus looks linear. This point of view, while correct in essence, is seriously misleading. It does not recognize that the Halperin–Lax regime is too narrow to be seen experimentally and that Gaussian behavior is established only at unobservable deep energies, if at all. Moreover, it carries the implication that strict exponential behavior can

occur only in a narrow energy range about the point of inflection. In fact, the range of energies over which $S[\vec{\phi}_0]$ or the single-well binding energy $|E|$ varies in such a way as to give an exponential $N(E)$ depends strongly on the shape of $C(R)$ [remember that $C(0) = 1$]. The broadest range occurs when $C(R)$ is nearly constant out to $R = a$ and then drops rapidly to zero [14]. A long tail in $C(R)$ narrows the energy range. In the model considered by Dow and Redfield [17], $C(R) \sim 1/R$ and $N(E)$ lacks both the Halperin–Lax and the Urbach regions. On the basis of the Halperin–Lax result for random-screened Coulomb potentials, corresponding to $C(R) \sim e^{-R/a}$ and a rather narrow Urbach range, the essential identity of the related family of theories (Halperin and Lax, Zittartz, and Langer, path integral, field theory) and the simple potential-well models has been denied [19]. The Cardy result of a Halperin–Lax region only did not help in this regard.

One can conclude that short-range order in the random potential favors Urbach tails, that one can get all of the physics and accurate quantitative results out of simple potential-well models, but that one needs the field theory to establish the accuracy of the simple potential-well models and give approximate analytic results. It is also important that the field-theoretic methods make computations feasible in the tight-binding representation so that models which represent real materials much more accurately than the simple ones examined thus far are feasible to tackle.

8. Additional complications: Polarons and excitons

The electron-phonon interaction has been taken into account only in the static limit in all the calculations described thus far. However, in slow processes, the dynamic effects of the electron-phonon interaction are important. These include inelastic scattering, phonon-induced hopping, and polaron formation. It is already known that disorder has a strong effect on polaron formation [35], and that the effects of disorder are strongly affected by the dynamic aspects of the electron-phonon interaction [35–37]. The problem is a very difficult one, as witnessed by the fact that obtaining the full density of states of, say, the acoustic polaron in a crystal has only just been addressed [38, 39].

Some results on the influence of lattice dynamics at zero temperature on band tailing have recently been obtained [38, 39]. In the absence of static disorder [38] and in the presence of strong electron-acoustic phonon interaction, the one-electron DOS projected onto the phonon vacuum exhibits a linear exponential band tail which terminates at the small polaron ground state. This tail arises from quantum fluctuations of the lattice ground state, which by themselves provide potential wells in which the electron may localize provided the oscillation frequency of the electron in the well is fast compared to the oscillation frequency of the well itself. In particular, there are three physically distinct energy regions: (1) At shallow energies there is a shift of the continuum edge associated with the emission and reabsorption of virtual phonons. (2) This is followed by a linear exponential tail in which the electron finds and stabilizes potential wells described above arising from quantum fluctuations of the lattice. (3) Finally, at deep energies, the band tail terminates at the polaron ground state. States in this last energy regime are truly self-trapped states in that the electron must create its own potential well rather than simply finding or stabilizing a pre-existing quantum fluctuation.

In most real materials the electron-phonon coupling is not sufficiently strong to give rise to appreciable band tailing from polaron formation alone. In many materials, however, the electron-phonon coupling is near threshold for small-polaron formation. With the addition of a small amount of static disorder we have shown numerically [39] that there is a substantial synergistic interplay between static localization and polaronic effects. Static disorder provides nucleation centers for small-polaron formation, and the resulting DOS in the tail can be considerably larger than that arising from either effect individually. In all cases there is a broad range of linear exponential behavior, although the actual slope of the exponential tail may have a large component arising from polaron formation. This is particularly evident when the electron-phonon coupling is just below the small-polaron threshold in the corresponding crystal.

A similar situation exists for excitons. Just as the electron-phonon interaction is always present and always important in the problem of band tails for slow processes, so the Coulomb interaction between a hole in the valence band and an electron in the conduction band is always present and always important in interband optical absorption. There are some known results for special cases for the exciton problem. In the case of the Frenkel exciton, where the electron and hole propagate as a tightly bound pair with the entire exciton band well separated from valence or conduction bands, the problem of the optical absorption can be reduced to a single-particle problem, and the results reviewed here for individual bands can be taken over directly. In the case of a generalized Lloyd model, the combined density of states for the electron–hole pair can be worked out exactly [40] and the effect of disorder on the interplay between exciton absorption and interband absorption understood. In general, however, the effects of the Coulomb interaction are not understood.

9. Conclusions

We now have an accurate, simple physical picture of the nature and origin of a large and common class of band tails. This physical picture is backed up by a powerful and varied set of formal arguments which pull together the entire 25-year development of the theory. Any system amenable to a tight-binding representation of its electronic structure and

possessing a reference system with normal band edges [41] can be studied quantitatively via the instanton theory. Nevertheless, much remains to be done. The pre-exponential factor $N_0(E)$ needs to be calculated via the instanton theory for correlated random potentials. Explicit instanton calculations need to be done for tight-binding models. The field theory as it is currently developed requires different computational techniques for the tail and the continuum. A field-theoretic variational principle valid in the presence of complex fields and suitable for both continuum and tail is needed and has long been lacking. The interplay between the dynamic aspects of the electron-phonon interaction and disorder must be addressed for slow processes, as discussed in Section 8. The interplay between the electron–hole interaction and disorder must be addressed for fast interband processes, as is also discussed in Section 8.

Acknowledgment

C. M. Soukoulis acknowledges support from the U.S. Department of Energy under Grant No. W-7405-Eng.-82.

References

1. A. Sommerfeld, "Zur Elektronentheorie der Metalle auf Grund der Fermischen Statistik," *Z. Phys.* **47,** 43 (1928).
2. F. Bloch, "Uber die Quantenmechanik der Elektronen in Kristallgittern," *Z. Phys.* **52,** 555 (1928).
3. A. H. Wilson, "The Theory of Electronic Semiconductors. I," *Proc. Roy. Soc. A* **133,** 458 (1931); "The Theory of Electronic Semiconductors. II," *Proc. Roy. Soc. A* **134,** 277 (1931).
4. E. O. Kane, "Thomas–Fermi Approach to Impure Semiconductor Band Structure," *Phys. Rev.* **131,** 79 (1963).
5. V. L. Bonch-Bruevich and A. G. Mironov, "Influence of an Impurity on the Energy Spectrum of Current Carriers," *Fiz. Tverd. Tela* (Leningrad) **3,** 2194 (1962) [*Sov. Phys. Solid State* **3,** 2194 (1962)].
6. F. Urbach, "The Long-Wavelength Edge of Photographic Sensitivity and of the Electronic Absorption of Solids," *Phys. Rev.* **92,** 1324 (1953).
7. B. I. Halperin and M. Lax, "Impurity-Band Tails in the High-Density Limit. I. Minimum Counting Methods," *Phys. Rev.* **148,** 722 (1966).
8. M. H. Cohen, E. N. Economou, and C. M. Soukoulis, "Band-Edge Features in Disordered Systems," *Phys. Rev. B* **32,** 8268 (1985).
9. E. N. Economou, C. M. Soukoulis, M. H. Cohen, and A. D. Zdetsis, "Quantitative Results Near the Band Edges of Disordered Systems," *Phys. Rev. B* **31,** 6172 (1985).
10. M. H. Cohen, C. M. Soukoulis, and E. N. Economou, "Optical Effects in Amorphous Semiconductors," P. C. Taylor and S. Bishop, Eds., *AIP Conf. Proc.* **120,** 371 (1984).
11. S. John, C. Soukoulis, M. H. Cohen, and E. N. Economou, "Theory of Electron Band Tails and the Urbach Optical-Absorption Edge," *Phys. Rev. Lett.* **57,** 1777 (1986).
12. S. John, C. Soukoulis, M. H. Cohen, and E. N. Economou, in *Proceedings of the 18th International Conference on the Physics of Semiconductors,* World Scientific Press, Singapore, 1987, p. 1025.
13. N. Bacalis, E. N. Economou, and M. H. Cohen, "A Simple Derivation of Exponential Tails in the Density of States," *Phys. Rev. B* (in press).
14. S. John, M. Y. Chou, M. H. Cohen, and C. Soukoulis, "Density of States for an Electron in a Correlated Gaussian Random Potential: Theory of the Urbach Tail," *Phys. Rev.* (submitted).
15. J. Tauc, in *Optical Properties of Solids,* B. Abeles, Ed., North-Holland Publishing Co., Amsterdam, 1972, p. 279; *Amorphous and Liquid Semiconductors,* J. Tauc, Ed., Plenum Press, London, 1974, Ch. 4.
16. G. D. Cody, "Hydrogenated Amorphous Silicon," *Semiconductors and Semimetals,* Vol. 21B, J. Pankove, Ed., Academic Press, Inc., New York, 1984, p. 11.
17. J. D. Dow and D. Redfield, "Toward a Unified Theory of Urbach's Rule and Exponential Absorption Edges," *Phys. Rev. B* **5,** 594 (1971).
18. D. Monroe and M. A. Kastner, "Exactly Exponential Band Tail in a Glassy Semiconductor," *Phys. Rev. B* **33,** 8881 (1986).
19. C. T. Chan, S. G. Louie, and J. C. Phillips, "Potential Fluctuations and Density of Gap States in Amorphous Semiconductors," *Phys. Rev. B* **35,** 2744 (1987).
20. S. Abe and Y. Toyozawa, "Interband Absorption Spectra of Disordered Semiconductors in the Coherent Potential Approximation," *J. Phys. Soc. Jpn.* **50,** 2185 (1981).
21. P. Lloyd, "Exactly Solvable Model of Electronic States in a Three-Dimensional Disordered Hamiltonian: Non-existence of Localized States," *J. Phys. C* **2,** 1717 (1969).
22. M. H. Cohen, H. Fritzsche, J. Singh, and F. Yonezawa, "Kinds of Disorder and the Electronic Structures of Tetrahedral Amorphous Semiconductors," *J. Phys. Soc. Jpn.* **49,** Suppl. A, 1175 (1980).
23. D. A. Papaconstantopoulos and E. N. Economou, "Calculations of the Electronic Properties of Hydrogenated Silicon," *Phys. Rev. B* **24,** 7233 (1981).
24. A. D. Zdetsis, E. N. Economou, D. A. Papaconstantopoulos, and N. Flytzanis, "Electronic and Transport Properties of Hydrogenated Amorphous Silicon," *Phys. Rev. B* **31,** 2410 (1985).
25. V. Sa-yakanit, "Electron Density of States in a Gaussian Random Potential: Path-Integral Approach," *Phys. Rev. B* **19,** 2266 (1979).
26. V. Sa-yakanit and H. R. Glyde, "Impurity-Band Density of States in Heavily Doped Semiconductors: A Variational Approach," *Phys. Rev. B* **22,** 6222 (1980).
27. W. Sritrakool, V. Sa-yakanit, and H. R. Glyde, "Band Tails in Disordered Systems," *Phys. Rev. B* **33,** 1199 (1986); "Absorption Near Band Edges in Heavily Doped GaAs," *ibid.* **32,** 1090 (1985).
28. V. Sa-yakanit and H. R. Glyde, "Urbach Tails and Disorder," *Comm. in Cond. Matter Phys.* **13,** 35 (1987).
29. R. P. Feynman, "Slow Electrons in a Polar Crystal," *Phys. Rev.* **97,** 660 (1955).
30. P. Lloyd and P. R. Best, "A Variational Approach to Disordered Systems," *J. Phys. C* **8,** 3752 (1975).
31. J. L. Cardy, "Electron Localization in Disordered Systems and Classical Solutions in Ginzburg–Landau Field Theory," *J. Phys. C* **11,** L321 (1978).
32. S. John and M. J. Stephen, "Electron Density of States in a Long-ranged Correlated Potential," *J. Phys. C* **17,** L559 (1984).
33. J. Zittartz and J. S. Langer, "Theory of Bound States in a Random Potential," *Phys. Rev.* **148,** 741 (1966).
34. I. M. Lifshitz, "The Energy Spectrum of Disordered Systems," *Adv. Phys.* **13,** 483 (1964).
35. M. H. Cohen, E. N. Economou, and C. M. Soukoulis, "Polaron Formation Near a Mobility Edge," *Phys. Rev. Lett.* **51,** 1202 (1983).
36. D. J. Thouless, "Maximum Metallic Resistance in Thin Wires," *Phys. Rev. Lett.* **39,** 1167 (1977).
37. N. F. Mott, "Electrons in Disordered Structures," *Adv. Phys.* **16,** 49 (1967); N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials,* Clarendon Press, Oxford, 1971.
38. S. John, "Localization and the Density of States for an Electron in a Quantized Elastic Continuum," *Phys. Rev. B* **35,** 9291 (1987).
39. C. H. Grein and S. John, "Polaronic Band Tails in Disordered Solids: Combined Effects of Electron-Phonon Interactions and Static Disorder," *Phys. Rev. B,* in press.
40. M. H. Cohen, in *Melting, Localization, and Chaos,* R. K. Kalia and P. Vashishta, Eds., Elsevier, New York, 1982, p. 125.

41. M. H. Cohen, C. M. Soukoulis, and E. N. Economou, in *Physics of Disordered Materials,* D. Adler, H. Fritzsche, and S. R. Ovshinsky, Eds., Plenum Press, New York, 1985.

Received August 24, 1987; accepted for publication August 27, 1987

Morrel H. Cohen *Exxon Research and Engineering Company, Annandale, New Jersey 08801.* Dr. Cohen is currently a Senior Scientific Advisor at Exxon Research and Engineering Company. He received a B.S. in 1947 and a D.Sc. in 1973 from Worcester Polytechnic Institute, an M.A. from Dartmouth College in 1948, and a Ph.D. in physics from the University of California in 1952. Before joining Exxon, Dr. Cohen was a faculty member of the Physics Department and the James Franck Institute at the University of Chicago from 1952 to 1981. He was also the Louis Block professor of physics and theoretical biology from 1972 to 1981, the director of the James Franck Institute from 1968 to 1971, and the director of the NSF Materials Research Laboratory at the University of Chicago from 1977 to 1981. Dr. Cohen has received many scientific honors and awards. He is a member of the National Academy of Sciences and a Fellow of the American Physical Society. His research fields include theoretical physics of condensed matter, developmental biology, quantum theory of solids, and general physics of solids.

Mei-Yin Chou *Exxon Research and Engineering Company, Annandale, New Jersey 08801.* Dr. Chou has been a postdoctoral research physicist at Exxon since 1986. She received a B.S. from the National Taiwan University, Taipei, Taiwan, in 1980, and an M.S. and a Ph.D. in physics from the University of California at Berkeley in 1983 and 1986, respectively. Dr. Chou's main research interests are the studies of the structural and electronic properties of clusters, surfaces, and semiconductors.

Eleftherios N. Economou *Research Center of Crete and Department of Physics, University of Crete, Heraklio, Crete, Greece.* Professor Economou received a B.S. from Polytech Institute of Athens, Greece, in 1963, and an M.S. and a Ph.D. in physics from the University of Chicago in 1967 and 1969, respectively. Before joining the faculty of the University of Crete, he was a faculty member at the University of Virginia. Professor Economou was also a consultant of the Naval Research Laboratory and a visiting associate professor at the University of Chicago (1975–76). He is currently serving as Secretary General of the General Secretariat of Research and Technology of the Greek Government. His research interests include properties of disordered systems, surface plasmons, electron–electron correlations, and magnetic properties in amorphous and crystalline semiconductors.

Sajeev John *Department of Physics, Princeton University, Princeton, New Jersey 08544.* In 1986 Dr. John joined the physics faculty at Princeton, where he is currently an assistant professor. During the past several years, he has also served as a consultant to the Exxon Research Laboratory in Annandale, New Jersey. He received his S.B. in 1979 from M.I.T. and his Ph.D. in physics from Harvard in 1984. He was a Natural Sciences and Engineering Research Council of Canada postdoctoral Fellow at the University of Pennsylvania from 1984–86. Dr. John's primary contributions have been in the development of a theory of optical and acoustical wave localization in disordered media, the prediction of a spin-glass phase of granular superconductors, and the development of a theory of the Urbach optical absorption edge in amorphous semiconductors.

Costas M. Soukoulis *Department of Physics, Iowa State University, Ames, Iowa 50011.* Professor Soukoulis has been a faculty member of the Physics Department at the Iowa State University since 1984. He received a B.S. from the University of Athens, Greece, in 1974, and an M.S. and a Ph.D. in physics from the University of Chicago in 1975 and 1978, respectively. He was a visiting assistant professor at the University of Virginia from 1978 to 1981 and a research physicist at Exxon Research and Engineering Company from 1981 to 1984. Dr. Soukoulis' current research interests include theory of disordered systems, spin glasses, amorphous magnets, amorphous semiconductors, optical localization, and nonlinearities in periodic and nonperiodic media.