Journal of Statistical Physics, Vol. 38, Nos. 1/2, 1985

Lifschitz Tails for the Anderson Model[1]

Barry Simon[2]

Received May 7, 1984

We present, in an expository way, an elementary rigorous proof (patterned after an argument of Kirsch–Martinelli) that the Anderson model has Lifschitz tails in very great generality.

KEY WORDS: Random potentials; density of states; Anderson model; Lifschitz tails; Dirichlet Neumann bracketing.

1. INTRODUCTION

The Anderson model on Z^ν is defined as follows: Let $d\mu$ be a probability measure on R with compact support, say $a = \inf(\operatorname{supp} \mu)$, $b = \sup(\operatorname{supp} \mu)$. Let $V_\omega(n)$ $(n \in Z^\nu)$ be independent, identically distributed (iid) random variables with distribution $d\mu$. Let H_0 be the operator on Z^ν given by

$$(H_0 u)(n) = 2\nu u(n) - \sum_{j=1}^{\nu} [u(n+\delta_j) + u(n-\delta_j)] \tag{1.1}$$

where δ_j is the element of Z^ν with 1 in the jth coordinate and 0 otherwise. The Anderson model is the family of random Hamiltonians

$$H_\omega = H_0 + V_\omega \tag{1.2}$$

The integrated density of states, $k(E)$, has a number of equivalent definitions (see, e.g., Refs. 17 and 1). For example, let H_ω^L be the $(2L+1)^\nu \times (2L+1)^\nu$ matrix obtained by restricting H_ω to sites n with $|n_j| \leqslant L$. Let $\#(H_\omega^L \leqslant E)$

---

[1] Research partially supported by USNSF Grant No. MCS-81-20833.

[2] Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, California 91125

0022-4715/85/0100-0065$04.50/0

denote the number of eigenvalues of H_ω^L smaller than E and let $N_L \equiv (2L+1)^\nu$ be the number of sites $(=\#(H_\omega^L < \infty))$. Then

$$k(E) = \lim_{L\to\infty} N_L^{-1} \#(H_\omega^L \leqslant E) \tag{1.3}$$

where one can prove that for almost every ω, the limit in (1.3) exists for all E and is ω independent. We will make three assumptions on μ, the distribution of $V_\omega(0)$:

(1) $a \neq b$, i.e., $d\mu$ is not a δ function at a single point;

(2) $d\mu([a, a+\varepsilon)) \geqslant C\varepsilon^l$;

(3) $d\mu((b+\varepsilon, b]) \geqslant C\varepsilon^l$.

for some C, l. The basic result on Lifschitz tails for the Anderson model is:

Theorem 1. Let $k(E)$ be the integrated density of states for an Anderson model with $d\mu$ obeying hypotheses (1)–(3). Then

$$\lim_{E\downarrow a} \ln[-\ln k(E)]/\ln(E-a) = -\nu/2$$

$$\lim_{E\uparrow b+4\nu} \ln[-\ln(1-k(E))]/\ln(4\nu+b-E) = -\nu/2$$

Roughly speaking, this says that near $E=a$, $k(E)$ looks like $\exp[-c(E-a)^{-\nu/2}]$ and a similar result for $1-k$ near $E=4\nu+b-E$. It is known (see, e.g., Ref. 1) that $\operatorname{spec}(H_\omega) = \operatorname{supp}(dk)$ and that[14] $\operatorname{spec}(H_\omega) = [0, 4\nu] + \operatorname{supp}\mu$, so a and $b+4\nu$ are the two edges of the spectrum. Since $k\equiv 0$ if $E<a$ and $k\equiv 1$ if $E>b+4\nu$, these Lifschitz tails are consistent with $k(E)$ being C^∞ and suggest that this may be true, at least in many cases (there are indications[22] that when $d\mu(x) = \frac{1}{2}[\delta(x-a)+\delta(x-b)]$, $k(E)$ may not be C^∞).

The idea that $k(E)\sim\exp[-c(E-a)^{-\nu/2}]$ near the bottom of the spectrum is due to E. M. Lifschitz[8] with a cogent intuition which we will describe in a moment. There have been numerous rigorous proofs for continuum and discrete models using the method of large deviations of Donsker–Varadhan or some other large deviations method (see Pastur,[18] Benderskii and Pastur,[2] Fukushima,[5] Nagai,[15] Fukushima *et al.*,[6] Pastur,[19] Nakao,[16] Luttinger,[4,9] and Romerio and Wrezinski.[21]) These proofs are sometimes able to evaluate the constant c in the asymptotics by making rather specific assumptions on the process $V_\omega(x)$ or $V_\omega(n)$ but have the disadvantage of being rather involved and applying only to rather specialized situations. Recently, Kirsch and Martinelli[12] gave a proof for continuum models which is quite elementary, and has the advantage of being very close to Lifschitz's intuition. Our goal here is to describe their proof in the discrete (Z^ν) case. In part, we wish to describe some technical aspects special to the discrete case where Dirichlet–Neumann bracketing is not so commonplace (in fact, we will see it is extremely elementary—in many ways simpler than the continuum analog; see Ref. 23 for a discussion of Dirichlet–Neumann bracketing). But our main goal is to advertise the Kirsch–Martinelli proof while honoring Lifschitz's memory, and to present it in an expository light which may be more accessible to theoretical physicists. We will also exploit Temple's inequality in a place that they use an inequality of Thirring. Kirsch and Simon[13] have extended their proof to some other situations. We note that[6,15,21] in particular have previously discussed the discrete case.

As a preliminary, we note that we can suppose that $a = 0$ (which we do henceforth) since $H_\omega - a1$ is a random Hamiltonian with a shifted to zero. Moreover, we need only consider the case $E \sim a$ for the case $E \sim b + 4\nu$ follows from this. This is because the replacement V_ω by $-V_\omega$ followed by the unitary map $u(n) \to (-1)^{|n|}\, u(n)$ (where $|n| = \sum_{j=1}^{\nu} |n_j|$) takes H_ω into $4\nu - H_\omega$ and the Lifschitz tail estimate for $4\nu - H_\omega$ at the bottom of its spectrum is equivalent to the Lifschitz tail estimate for H_ω at the top of its spectrum.

Thus, we henceforth suppose that $a = 0$ and concentrate on $k(E)$ near $E = 0$. Here is Lifschitz's intuition: How can H_ω have an eigenvalue below E_0 a very small number? Let u be the corresponding eigenvector. Since $V \geqslant 0$ (since $a = 0$) both $(u, H_0 u)$ and (u, Vu) must be smaller than E_0. For $(u, H_0 u)$ to be smaller than E_0, u must be spread out at least over a region of size R where $E_0 = R^{-2}$ which has $R^\nu = E_0^{-\nu/2}$ sites. On most of these sites, V must have a very small value so R^ν independent events of probability e^{-c} must occur, i.e., the probability is $O(\exp(-cE^{-\nu/2}))$.

The key to proving Theorem 1 will be the use of Dirichlet–Neumann bracketing. We discuss this in Section 2. Lower bounds on $k(E)$ (equivalent to upper bounds on eigenvalues) are obtained using Dirichlet comparisons in Section 3, and upper bounds on $k(E)$ using Neumann comparisons and Temple's inequality in Section 4.

We note that while we have supposed that the process $V_\omega(n)$ is iid, the proof shows that much less is needed. With the normalization $a = 0$, it suffices that (i) $V_\omega(n)$ have exponential mixing in the sense that if $A \subset Z^\nu$ and P_A is the conditional expectation induced by $\{V_\omega(n) \mid n \in A\}$, then for convex sets A, B, $\|P_A P_B\| \leqslant Ce^{-\alpha\, \mathrm{dist}(A,B)}$ for some $\alpha > 0$; (ii) $V_\omega(n) \geqslant 0$; (iii) for a box, B, of side l, small ε, large l and any sufficiently small f including $f = 0$, $\mathrm{Prob}(V_\omega(n) > \varepsilon$ at a fraction of sites f or less in $B) \sim C_{f,\varepsilon} \exp(-l^\nu D(\varepsilon, f))$ in the sense of upper and lower bounds of this form.

2. DIRICHLET–NEUMANN BRACKETING

A bond in Z^ν will denote a pair of neighboring sites. Given a Hamiltonian of the form (1.2) and a set of bonds, B, we will want to define two Hamiltonians $H_\omega^{B;N}$ and $H_\omega^{B;D}$ so that

$$H_\omega^{B;N} \leqslant H_\omega \leqslant H_\omega^{B;D} \tag{2.1}$$

and so that if B disconnects Z^ν into subsets $\{S_\alpha\}$, then under the direct sum decomposition $l^2(Z^\nu) = \oplus_\alpha l^2(S_\alpha)$, $H_\omega^{B;\#}$ is a direct sum ($\#$ will denote either D or N), i.e., so that if i, j are in distinct S_α, then $(H_\omega^{B;\#})_{ij} = 0$.

Because of the fact that the analog of (2.1) in the continuum case are Dirichlet and Neumann boundary conditions (see the reviews in Refs. 23 and 20), we use those names and the letters D, N. However, as we should now like to explain, H^D is *not* what one usualy chooses for a discrete Dirichlet Hamiltonian as discussed, for example, in Ref. 7. To explain what we mean, we set $V_\omega = 0$ and try $\nu = 1$ with $B = [0, 1]$, $[l, l+1]$. The usual Dirichlet Laplacian on $\{1,..., l\}$ is just the restriction of H_0 to those sites, i.e.,

$$\begin{aligned}(\tilde{H}_0^D)_{ij} &= 2, && 1 \leqslant i = j \leqslant l\\ &= -1, && |i-j| = 1, \quad i, j \in \{1,..., l\}\\ &= 0, && |i-j| \geqslant 2\end{aligned}$$

Eigenfunction of $\tilde{H}_0^D$ obey

$$2u(n) - u(n-1) - u(n+1) = eu(n) \tag{2.2}$$

at $n = 1,..., l$ if one takes the boundary conditions

$$u(0) = u(l+1) = 0, \qquad \tilde{D} \text{ b.c.} \tag{2.3}$$

For example, the operator H_ω^L used in Section 1 to define $k(E)$ has $\tilde{D}$ b.c.

The problem with $\tilde{D}$ b.c. is that if we try to decompose Z into $\{1,..., l\} = S_0,..., S_\alpha = \{l\alpha + 1,..., l\alpha + l\},...$ $(\alpha \in Z)$, then $\oplus_\alpha \tilde{H}_0^D$ is just H_0 with certain off-diagonal elements set equal to zero and $H_0 - \oplus_\alpha \tilde{H}_0^D$ is neither positive nor negative definite and bracketing fails.

We will, instead, make a choice for a Dirichlet Laplacian on $\{1,..., l\}$ with

$$\begin{aligned}(H_0^D)_{ij} &= 3, && i = j = 1 \text{ or } i = j = l\\ &= 2, && 2 \leqslant i = j \leqslant l-1\\ &= -1, && |i-j| = 1, \quad i, j \in \{1,..., l\}\\ &= 0, && |i-j| \geqslant 2\end{aligned}$$

With this choice, $\oplus_\alpha H_0^D \geqslant H_0$ as we shall see. We will also see that eigenfunctions of H_0^D obey (2.2) at $n = 1,..., l$ with the boundary conditions

$$-u(1) = u(0), \qquad -u(l+1) = u(l), \qquad D \text{ b.c.} \tag{2.4}$$

so that D b.c. corresponds to choosing the linear interpolation[3] to vanish at $\frac{1}{2}$ and $l + \frac{1}{2}$. Some thought will even suggest that if one wishes to consider touching regions, this is a more reasonable choice of D b.c. than the usual one. Our choice of D b.c. also has the advantage of turning into N b.c. under the unitary transformation $u(n) \to (-1)^n u(n)$.

We note that there is a sense in which $\tilde{D}$ b.c. dominates H_0; namely, if we define $\tilde{H}_0^D$ on Z to be $\tilde{H}_0^D$ on $S_\alpha = \{(l+1)\alpha + 1,..., (l+1)(\alpha+1) - 1\}$ and ∞ on $\{(l+1)\alpha\}$ then $\tilde{H}_0^D \geqslant H_0$. Of course, one has a density of eigenvalues $1/l + 1$ at $+\infty$, but since we take $l \to \infty$ in our proof in Section 3, one could use $\tilde{D}$ b.c. there, but we find D b.c. much more natural.

With this purple prose out of the way, we turn to the precise definitions. Given a pair of indices $m, n \in Z^\nu$, we define two operators on $l^2(Z^\nu)$:

$$\begin{aligned} (L^{mn}u)(i) &= 0, & (S^{mn}u)(i) &= 0, & i &\neq m, n \\ &= u(m) - u(n), & &= u(m) + u(n), & i &= m \\ &= u(n) - u(m), & &= u(n) + u(m), & i &= n \end{aligned}$$

so $(u, L^{mn}u) = (u(m) - u(n))^2$ and $(u, S^{mn}u) = (u(m) + u(n))^2$ and we have

$$L^{mn} \geqslant 0, \qquad S^{mn} \geqslant 0 \tag{2.5}$$

corresponding to the positivity of the matrices

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

If $\sum_{\langle mn \rangle, \cdots}$ denotes the sum over all pairs obeying $\cdots$ with each pair counted once, then clearly

$$H_0 = \sum_{\langle mn \rangle; |m-n| = 1} L^{mn}$$

and we define, given a set of bonds, B:

$$H_0^{B,N} = H_0 - \sum_{\langle mn \rangle; (mn) \in B} L^{mn} = \sum_{\substack{\langle mn \rangle; |m-n| = 1 \\ (m,n) \notin B}} L^{mn} \tag{2.6}$$

and

$$H_0^{B,D} = H_0 + \sum_{\langle mn \rangle; (mn) \in B} S^{mn} \tag{2.7}$$

By (2.5) we trivially have the following:

Proposition 2.1. For any set of bonds, B, we have

$$H_0^{B,N} \leqslant H_0 \leqslant H_0^{B,D}$$

Moreover, we have the following:

Proposition 2.2. Suppose that removing the bonds, B, disconnects Z^ν into disjoint sets S_α. Then under the decomposition $l^2(Z^\nu) = \oplus\, l^2(S_\alpha)$ both $H_0^{B,D}$ and $H_0^{B,N}$ are direct sums.

Proof. Clearly $H_0^{B,N}$ has no matrix elements involving sites in distinct S_α since such matrix elements in H_0 come only from L^{mn} with $(mn) \in B$. For $H_0^{B,D}$, the same argument is applicable if one notes that $L^{mn} + S^{mn}$ is a diagonal matrix. ∎

Now fix $L = 1, 2, \ldots$ and for $\alpha \in Z^\nu$, let $S_\alpha^{(L)}$ be the set of $n \in Z^\nu$ with $\alpha_i L + 1 \leqslant n_i \leqslant \alpha_i(L+1)$ so $\{S_\alpha^{(L)}\}$ is a partition of Z^ν into disjoint boxes with L^ν sites. When no confusion can result, we will drop the superscript (L). Let $B(L)$ denote the bonds coupling distinct S_α. Then, by the above $H_0^{B(L),\#}$ is a direct sum of operators we denote by $H_0^{L,\#}$. We claim the following:

Lemma 2.3. u is an eigenfunction of $H_0^{L,N}$ with eigenvalue e, if and only if the extension of u to Z^ν obtained by reflecting successively in hyperplanes $n_i = (\alpha_i + \frac{1}{2})L$ obeys $H_0 u = Eu$. A similar result holds for $H_0^{L,D}$ if we flip signs upon each reflection.

Proof.

$$(H_0 u)(n) = \sum_{|m-n|=1} [u(n) - u(m)] \quad \text{and} \quad (H_0^{L,N} u)(n) = \sum_{\substack{|m-n|=1 \\ m \in S_0}} [u(n) - u(m)]$$

so the assertion holds if we note that the reflection condition sets $u(n) - u(m) = 0$ if $|n - m| = 1$ and $n \in S_0$, $m \notin S_0$. Similarly

$$(H_0^{L,D} u)(n) = (H_0^{N,D} u)(n) + \sum_{\substack{|m-n|=1 \\ m \notin S_\alpha}} 2u(n)$$

Since the reflection plus sign flip sets $u(n) - u(m) = 2u(n)$ if $|n - m| = 1$ and $n \in S_0$, $m \notin S_0$ we obtain the final statement in the theorem. ∎

Since we can write down all solutions of $H_0 u = eu$ in closed form (plane waves) and check the boundary conditions in Lemma 2.3 by straightforward manipulation, we find the following:

Theorem 2.4. (a) The eigenvalues of $H_0^{L,D}$ are the set of numbers of the form

$$\sum_{i=1}^{\nu} 2 - 2\cos(k_i \pi/L) \tag{2.8}$$

where each k_i is one of $1, 2,..., L$ and all such ν-tuples occur once. In particular, the lowest eigenvalue is

$$e_0^{L,D} = \nu[2 - 2\cos(\pi/L)] \tag{2.9}$$

(b) The eigenvalues of $H_0^{L,N}$ have the form (2.8) where each k_i is one of $0, 1,..., L-1$ and all such ν-tuples occur once. In particular, the two lowest eigenvalues (e_1 is ν-fold degenerate) are

$$e_0^{L,N} = 0, \qquad e_1^{L,N} = 2 - 2\cos(\pi/L) \tag{2.10}$$

Now, we can use these objects to bound $k(E)$ by something involving eigenvalues of finite matrices. Let $H_\omega^{B,\#} = H_0^{B,\#} + V_\omega$, $H_\omega^{L,\#} = H_0^{L,\#} + V_\omega$, etc. Define

$$k_L^\#(E) = \text{Exp}(\#\text{ of e.v. of } H_\omega^{L,\#} \leqslant E)/L^\nu \tag{2.11}$$

where Exp means expectation value (e.v.) over the ensemble of potentials. The following is just a discrete analog of an idea used by Kirsch–Martinelli in several places[10,11,12]:

Theorem 2.5. For any L, $k_L^D(E) \leqslant k(E) \leqslant k_L^N(E)$.

Remark. The inequalities reverse from H to k since a larger operator has *fewer* eigenvalues less than a fixed E.

Proof. By general principles (e.g., Ref. 1), $k(E)$ can be computed by looking at $H_\omega^{nL,\#}$ for a typical ω and computing the number of eigenvalues smaller than E for it, dividing by $(nL)^{-\nu}$, and taking n to ∞. By an obvious extension of Proposition 2.1,

$$H_\omega^{nL,N} \geqslant \bigoplus_{\alpha_i = 1,2,\ldots,n} H_{\omega,\alpha}^{L,N} \tag{2.12}$$

The $H_{\omega,\alpha}^{L,N}$ are identically distributed, independent random operators, so as $n \to \infty$,

$$(nL)^{-\nu} \sum_{1 \leqslant \alpha_i \leqslant n} (\#\text{ of e.v. of } H_{\omega,\alpha}^{L,N} \leqslant E) \to k_L^N(E)$$

by the law of large numbers. By (2.12), we obtain $k(E) \leqslant k_L^N(E)$. The other assertion has a virtually identical proof. ∎

Remark. For general processes V_ω there are two "couplings" between regions. The one in the operators is removed by the boundary conditions. The ones in the process are not present in our case by independence, and are not important so long as a law of large numbers holds for local functions of the process.

Following Lifschitz's intuition, we will get upper and lower bounds on $k(E)$ as $E \downarrow 0$ by using Theorem 2.5 with $L = (\text{const})\, E^{-1/2}$. The details follow in the next two sections.

3. DIRICHLET BOUNDARY CONDITIONS: THE LOWER BOUND

We obtain a lower bound on $k(E)$ by using Dirichlet boundary conditions to get upper bounds on eigenvalues. We look at the contribution of boxes in which $V_\omega(n)$ is small for all n. Fix E_0 small. Choose L so that $L = (\frac{1}{2}E_0/\pi^2 \nu)^{1/2}$. Actually, since L must be an integer, we choose it to be the smallest integer greater than this number so

$$L \leqslant (E_0/2\pi^2 \nu)^{1/2} + 1 \tag{3.1}$$

Moreover, since $2 - 2\cos x \leqslant x^2$, we have by (2.9) that

$$e_0^{L,D} \leqslant \nu\pi^2/L^2 \leqslant \tfrac{1}{2}E_0 \tag{3.2}$$

Now suppose that all $V_\omega(n)$, $n \in S_0^{(L)}$ have $V(n) \leqslant \frac{1}{2}E_0$. Then trivially, $H_\omega^{L,D}$ has at least one eigenvalue smaller than E_0. Thus, by the definition (2.11)

$$\begin{aligned} k_L^D(E_0) &\geqslant L^{-\nu}\, \text{Prob}(V_\omega(n) \leqslant \tfrac{1}{2}E_0, \text{all } n \in S_0^L) \\ &= L^{-\nu}\, \text{Prob}(V_\omega(0) \leqslant \tfrac{1}{2}E_0)^{L^\nu} \\ &\geqslant L^{-\nu} C^{L^\nu} \exp[-lL^\nu \ln(E_0^{-1})] \end{aligned}$$

where we have used hypothesis (2) in the last step. Using the inequality (3.1) and the inequality in Theorem 2.5, we see that

$$\varliminf_{E \downarrow 0} \ln[-\ln k(E)]/\ln E \geqslant -\nu/2$$

which proves one half of Theorem 1.

4. NEUMANN BOUNDARY CONDITIONS: THE UPPER BOUND

Since $d\mu$, the distribution of $V_\omega(0)$ is not a δ function (hypothesis 1), we can find $\varepsilon_0 > 0$ and $f_0 > 0$ so that

$$f_0 = \text{Prob}(V_\omega(0) \geqslant \varepsilon_0) \tag{4.1}$$

The idea of the upper bound on $k(E)$ will be to show not many Neumann boxes can have eigenvalues below E_0 if we pick $L(E_0)=\beta E_0^{-1/2}$. With this choice, we can arrange (since $V \geqslant 0$) that no more than one eigenvalue in a box can be below E_0. The law of large numbers says that typically $f_0 L^\nu$ sites will have $V_\omega(n) \geqslant \varepsilon_0$. We will show that if at least $\frac{1}{2} f_0 L^\nu$ have $V_\omega(0) > \varepsilon_0$ then that box has a ground state larger than E_0 so $k_L^N(E) \leqslant$ Prob(fewer than $\frac{1}{2} f_0 L^\nu$ sites have $V_\omega(n) \geqslant \varepsilon_0$), and by an elementary large deviations estimate this will be $O(\exp(-cL^\nu))$.

To be precise, the two basic inputs of the above argument are the following, whose proofs we defer:

Theorem 4.1. There exists constants L_0 and α_0 so that if $L > L_0$ and if $\#\{n \in S_0^{(L)} \mid V_\omega(n) \geqslant \varepsilon\} L^{-\nu} \geqslant \frac{1}{2} f_0$, then

$$e_{0,\omega}^{L,N} \geqslant \alpha_0 L^{-2}$$

Theorem 4.2. $\operatorname{Prob}(\#\{n \in S_0^{(L)} \mid V_\omega(n) \geqslant \varepsilon\} L^{-\nu} < \frac{1}{2} f_0) \leqslant \exp(-\frac{1}{2} f_0^2 L^\nu)$.

Accepting this for the moment, we will show the following:

Proposition 4.3. $\overline{\lim}_{E \downarrow 0} \ln[-\ln k(E)]/\ln E \leqslant -\nu/2$.

Proof. For E_0 small, let L be the largest integer with

$$\alpha_0 L^{-2} > E_0, \qquad 2 - 2\cos(\pi/L) > E_0 \tag{4.2}$$

It is easy to see that as E_0 goes to zero,

$$L/aE_0^{-1/2} \to 1 \tag{4.3}$$

for suitable $a \neq 0$. Since (4.2), (2.10) and the fact that $V \geqslant 0$ imply that for almost every ω, $e_{1,\omega}^{L,N} > E_0$, we see that

$$k_L^N(E) = \operatorname{Prob}(e_{0,\omega}^{L,N} \leqslant E_0)/L^\nu$$

on account of the definition (2.11). But, by Theorem 4.1, this probability is dominated by the probability estimated in Theorem 4.2, and, by that theorem, we have

$$k_L^N(E) \leqslant L^{-\nu} \exp(-\tfrac{1}{2} f_0^2 L^\nu)$$

By (4.3) and Theorem 2.5, the desired $\overline{\lim}$ statement is true. ∎

We have thus reduced the proof of Theorem 1 to the proof of Theorems 4.1 and 4.2. We do the large deviations result first:

Proof of Theorem 4.2. Let $f_\omega(n)=1$ (resp. 0) if $V_\omega(n)\geqslant \varepsilon$ (resp. $<\varepsilon$). Then $\mathrm{Exp}(f_\omega)=f_0$ by (4.1). Define

$$F(y)=\ln \mathrm{Exp}(e^{-y(f_\omega-f_0)})$$

so $F(0)=F'(0)=0$. Moreover, $F''(y)=\langle f_\omega^2\rangle_y-\langle f_\omega\rangle_y^2$ where $\langle\cdot\rangle_y=\mathrm{Exp}(\cdot\, e^{-yf_\omega})/\mathrm{Exp}(e^{-yf_\omega})$. Since $|f_\omega-\frac{1}{2}|\leqslant\frac{1}{2}$, we see that $F''(y)\leqslant\frac{1}{4}$ and thus

$$F(y)\leqslant\tfrac{1}{8}y^2 \tag{4.4}$$

Now, $\sum_{n\in S_0^{(L)}}f_\omega(n)<\frac{1}{2}f_0L^\nu$ if and only if $\#\{n\in S_0^{(L)}\mid V_\omega(n)\geqslant\varepsilon\}<\frac{1}{2}f_0L^\nu$. Thus, for any $y>0$:

$$\begin{aligned}
&\mathrm{Prob}(\#\{n\in S_0^{(L)}\mid V_\omega(n)\geqslant\varepsilon\}\,L^{-\nu}<\tfrac{1}{2}f_0)\\
&\quad\leqslant \mathrm{Exp}\left(\exp\left[+y\sum_\omega(\tfrac{1}{2}f_0-f_\omega)\right]\right)=\exp(-\tfrac{1}{2}f_0L^\nu y)\exp[L^\nu F(y)]\\
&\quad\leqslant\exp(\tfrac{1}{8}L^\nu y^2-\tfrac{1}{2}L^\nu f_0 y)\\
&\quad=\exp(-\tfrac{1}{2}f_0^2L^\nu)
\end{aligned}$$

if we choose $y=2f_0$. ∎

Remark. The use of $F(y)$ above is typical of large deviations. Indeed, Cramer's theorem, the original large deviation result, says that

$$\lim_{L\to\infty}-L^{-\nu}\ln\mathrm{Prob}(\#\{n\in S_0^{(L)}\mid V_\omega(n)\geqslant\varepsilon\}\,L^{-\nu}<\tfrac{1}{2}f_0)=\sup_y(\tfrac{1}{2}yf_0-F(y))$$

Our proof of Theorem 4.1 is modeled after an argument we gave in Ref. 24 to prove the Fefferman–Phong theorem in a similar context. It exploits Temple's inequality,[25] which we prove for the sake of completeness:

Lemma 4.4 (Temple's inequality). Let A be a self-adjoint operator with eigenvalues $E_0<E_1$ at the bottom of its spectrum. Let $\varphi\in D(A)$ be such that $\langle\varphi,A\varphi\rangle<E_1$. Then

$$E_0\geqslant\langle\varphi,A\varphi\rangle-[E_1-\langle\varphi,A\varphi\rangle]^{-1}\{\langle A\varphi,A\varphi\rangle-\langle A\varphi,\varphi\rangle^2\} \tag{4.5}$$

Proof. By the spectral theorem, $(A-E_1)(A-E_0)\geqslant 0$. Thus, $(\varphi,(A-E_1)(A-E_0)\varphi)\geqslant 0$ or

$$\begin{aligned}
E_0[(\varphi,A\varphi)-E_1]&\leqslant(A\varphi,A\varphi)-E_1(\varphi,A\varphi)\\
&=[(A\varphi,A\varphi)-(A\varphi,\varphi)^2]+[(\varphi,A\varphi)-E_1](\varphi,A\varphi)
\end{aligned}$$

Since $(\varphi,A\varphi)-E_1<0$, when we divide by it, (4.4) results. ∎

Since $\langle A\varphi, \varphi\rangle^2 > 0$, we see that

$$\langle \varphi, A\varphi\rangle < E_1^* \leqslant E_1 \Rightarrow E_0 \geqslant \langle \varphi, A\varphi\rangle - [E_1^* - (\varphi, A\varphi)]^{-1} \|A\varphi\|^2 \quad (4.6)$$

With this preliminary, we have the following:

Proof of Theorem 4.1. We will prove the result with $\alpha_0 = \frac{1}{3} f_0$ and $L_0 = \varepsilon^{-1/2}$. The theorem is deterministic, i.e., we are given a fixed V_ω with $\#\{n \in S_0^{(L)} \mid V_\omega(n) \geqslant \varepsilon\} L^{-\nu} \equiv \gamma \geqslant \frac{1}{2} f_0$. Define W to be the function

$$\begin{aligned} W(n) &= 0, \qquad &\text{if} \quad V_\omega(n) < \varepsilon \\ &= L^{-2}, \qquad &\text{if} \quad V_\omega(n) \geqslant \varepsilon \end{aligned}$$

Since $L > L_0$, $W(n) \leqslant V_\omega(n)$, so $e_0(A) \leqslant e_{0,\omega}^{L,N}$ if $A = H_0^{L,N} + W$. Let φ be the normalized vector, all of whose components are $L^{-\nu/2}$. Thus, since $H_0^{L,N}\varphi = 0$, we have that

$$(\varphi, A\varphi) = \gamma L^{-2} \qquad (\varphi, A^2\varphi) = \gamma L^{-4}$$

Moreover, since $2 - 2\cos \pi x \geqslant 4x^2$ if $0 < x < 1$, we have that $e_1(A) \geqslant e_1^{L,N} \geqslant 4L^{-2}$, so

$$e_1(A) - (\varphi, A\varphi) \geqslant 3L^{-2}$$

Thus, by (4.5)

$$e_0(A) \geqslant \gamma L^{-2} - (3L^{-2})^{-1} \gamma L^{-4} = \tfrac{2}{3}\gamma L^{-2}$$

Since $e_{0,\omega}^{L,N} \geqslant e_0(A)$ and $\gamma \geqslant \frac{1}{2} f_0$, we have the required result. ∎

NOTE ADDED IN PROOF

The idea of using Dirichlet–Neumann bracketing in analyzing the question of Lifschitz tails appears to have appeared first in a paper by A. B. Harris, *Phys. Rev.* **B8**, 3661 (1973). A discussion closely related to that in this paper appears in a recent preprint of G. A. Mezincescu of the Institute for Physics and Technology of Materials in Bucharest, Rumania.

REFERENCES

1. J. Avron and B. Simon, Almost periodic Schrödinger operators, II. The integrated density of states, *Duke Math. J.* **50**:369–391 (1983).
2. M. Benderskii and L. Pastur, On the spectrum of the one dimensional Schrödinger equation with a random potential, *Mat. Sb.* **82**:245–256 (1970).
3. F. Delyon and B. Souillard, The rotation number for finite difference operators and its properties, *Commun. Math. Phys.* **89**:415 (1983).
4. R. Friedberg and J. Luttinger, Density of electronic energy levels in disordered systems, *Phys. Rev. B* **12**:4460 (1975).

5. M. Fukushima, On the spectral distribution of a disordered system and the range of a random walk, *Osaka J. Math.* **11**:73–85 (1974).
6. M. Fukushima, H. Nagai, and S. Nakao, On an asymptotic property of spectra of a random difference operator, *Proc. Japan Acad.* **51**:100–102 (1975).
7. F. Guerra, L. Rosen, and B. Simon, Boundary conditions for the $P(\phi)_2$ Euclidean field, *Ann. Inst. H. Poincaré* **25A**:231–334 (1978).
8. I. M. Lifschitz, Energy spectrum structure and quantum states of disordered condensed systems, *Sov. Phys.-Usp.* **7**:549 (1965).
9. J. M. Luttinger, New variational method with applications to disordered systems, *Phys. Rev. Lett.* **37**:609 (1976).
10. W. Kirsch, Random Schrödinger operators and the density of states, to appear in the Proceedings of the 2nd meeting of French and West German mathematicians and mathematical physicists, Marseille 1983; Springer Lecture Notes in Math.
11. W. Kirsch and F. Martinelli, On the density of states of Schrödinger operators with a random potential, *J. Phys. A* **15**:2139–2156 (1982).
12. W. Kirsch and F. Martinelli, Large deviations and Lifschitz singularity of the integrated density of states of random Hamiltonians, *Commun. Math. Phys.* **89**:27–40 (1983).
13. W. Kirsch and B. Simon, Lifschitz tails for periodic plus random potentials, Caltech Preprint.
14. H. Kunz and B. Souillard, Sur le spectre des opérateurs aux différences finies aléatoires, *Commun. Math. Phys.* **78**:201–246 (1980).
15. H. Nagai, On an exponential character of the spectral distribution function of a random difference operator, *Osaka J. Math.* **14**:111–116 (1977).
16. S. Nakao, On the spectral distribution of the Schrödinger operator with random potential, *Japan J. Math.* **3**:111–139 (1977).
17. L. Pastur, Spectral properties of disordered systems in one-body approximation, *Commun. Math. Phys.* **75**:179 (1980).
18. L. Pastur, On the distribution of the eigenvalues of the Schrödinger equation with a random potential, *Funk. Anal. Pril. (Russian)* **6**:93–94 (1972).
19. L. Pastur, Behavior of some Wiener integrals as $t \to +\infty$ and the density of states of Schrödinger equation with a random potential, *Teo. Mat. Fiz. (Russian)* **32**:88–95 (1977).
20. M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. IV. Analysis of Operators* (Academic Press, New York, 1977).
21. M. Romerio and W. Wreszinksi, On the Lifschitz singularity and the tailing in the density of states for random lattice systems, *J. Stat. Phys.* **21**:169–179 (1979).
22. H. Schmidt, Disordered one-dimensional crystals, *Phys. Rev.* **105**:425–441 (1957).
23. B. Simon, Classical boundary conditions as a tool in quantum physics, *Adv. Math.* **30**:268–281 (1978).
24. B. Simon, Some quantum operators with discrete spectrum but classically continuous spectrum, *Ann. Phys. (N.Y.)* **146**:209–220 (1983).
25. G. Temple, The theory of Rayleigh's principle as applied to continuous systems, *Proc. R. Soc.* **119A**:276–293 (1928).