

Analytic Properties of Band Functions

J. E. AVRON*

Department of Physics, Princeton University, Princeton, New Jersey 08540

AND

B. SIMON†

Department of Physics, Yeshiva University, New York, New York 10033

Received February 23, 1977

We consider the analytic properties of Bloch Hamiltonians and their energy levels as functions of quasi-momentum k . Among other results are: (1) The only entire band is the trivial free electron parabola. (2) The only isolated singularities allowed are branch points. (3) In one dimension, no natural boundaries or logarithmic branch points occur. (4) The periodic attractive screened Coulomb lattice has a nondegenerate lowest band (i.e., the “direct gap” is strictly positive for all k).

1. INTRODUCTION

In this work, we discuss the analytic properties of the energy bands of Bloch Hamiltonians [13]. One of our main motivations is an attempt to extend Kohn’s elegant analysis [14] of the one-dimensional case to more than one dimension. In one dimension, one has available ordinary differential equation methods which go back to Lyapunov, Hamel [7], Haupt [9], and, in the physics literature, Kramers [16]. These methods lead to a simple implicit expression for the band energies $E(k)$ as solutions of

$$2 \cos k = \Delta(E), \quad (1.1)$$

where Δ is the discriminant for the equation

$$-d^2/dx^2 + V(x) = E; \quad 0 \leq x \leq b. \quad (1.2)$$

These differential equation methods do not seem to extend to the multidimensional case and it seems natural to exploit the operator techniques which have been so useful in the consideration of other aspects of the study of nonrelativistic Hamiltonians [12, 19–21].

* Wigner Fund Fellow, Technische University of Berlin.

† On leave from Departments of Mathematics and Physics, Princeton University; research supported by the National Science Foundation under Grant MPS-75-11864.

The operator theory background for $-\Delta + V$ with V periodic has been developed by several authors [2, 3, 18, 21, 23, 27, 29] and leads to the following formulation. Let $\tilde{\Gamma}$ be a lattice in \mathbb{R}^3 and let V be a potential with

$$V(x + b) = V(x); \quad x \in \mathbb{R}^3; b \in \tilde{\Gamma}.$$

Then $H = -\Delta + V$ is unitarily equivalent to a direct integral described as follows. Let Γ be the lattice reciprocal to $\tilde{\Gamma}$, i.e., those k with $k \cdot a = 2\pi n$ with $n \in \mathbb{Z}$ for any $a \in \tilde{\Gamma}$. For $k \in \mathbb{R}^3$, let $H(k)$ be the operator on $L^2(\Gamma)$

$$H(k) = T(k) + V, \tag{1.3a}$$

$$(T(k)\psi)(p) = (p + k)^2 \psi(p), \quad p \in \Gamma, \quad k \in \mathbb{R}^3, \tag{1.3b}$$

$$(V\psi)(p) = \sum_{q \in \Gamma} \tilde{v}(q) \psi(p - q), \tag{1.3c}$$

$$\tilde{v}(q) = |Q|^{-1} \int_Q \exp(-iq \cdot x) V(x) d^3x, \tag{1.3d}$$

where Q is any basic cell for the direct lattice $\tilde{\Gamma}$. Obviously, for $p_0 \in \Gamma$, $H(k)$ is unitarily equivalent to $H(k + p_0)$. Moreover, H is unitarily equivalent to the direct integral $\int_B^{\oplus} H(k) dk$, where B is the Brillouin zone, i.e., $\{k \mid k \text{ is nerer to } 0 \text{ than to any other point on } \Gamma\}$. Band functions are just analytic eigenvalues $E(k)$ of $H(k)$.

It is clear that $H(k)$ can be continued to complex k just by letting k in (1.3) be complex. We can then study the analytic properties of the energy bands of $H(k)$ by using the powerful tools of operator perturbation theory [12] and some of our results below are transcriptions of results from another problem in analytic perturbation theory; that of the anharmonic oscillator [24].

It is useful for comparison purposes to note what can be proved using differential equation methods in the one-dimensional case (normalized so that $\Gamma = \mathbb{Z}$).

(1) One can analytically continue between any two bands (if some gaps are missing this must be suitably interpreted; see Appendix B).

(2) The only points k where $E(k)$ has a finite limit but is nonanalytic are square root branch points and at such points $\text{Re } k \in \mathbb{Z}$.

(3) The only entire band function is a parabolic onc. (1) and (2) (under unnecessary extra assumptions which we eliminate in Appendix B) are results of Kohn. We discuss (3) in Appendix B.

The O.D.E. methods leave open the following which we prove in Section 6:

(4) Natural boundaries do not occur under analytic continuation.

We have nothing to say about (1) in the general case except for some handwaving in Section 2, but we can prove (3) in general and results related to (2). We establish (4) in the one-dimensional case. We also prove the ground band nondegenerate in some cases. (See Appendix C.)

We generally make two restrictive assumptions in our work below. First, we have little to say about intrinsic multivariable analyticity but rather restrict ourselves to the consideration of $E(k_0 + ak)$, where $k_0, a \in \mathbb{R}^3$ and $k \in \mathbb{C}$. Typically we chose $a \in \tilde{\Gamma}$ the direct lattice. Second, we sometimes suppose that the x -space potential, V , is bounded. The latter hypothesis is known to be a restriction since V locally L^2 is all that is needed for $-\Delta + V$ to be self-adjoint on $D(-\Delta)$ and V locally in $L^{3/2}$ can be accommodated using form methods.

We have studied the problem of analyticity of band functions primarily for its own sake but we close this introduction by mentioning some possible "applications."

(1) Kohn [15] has used analyticity in studying magnetic interactions.

(2) There is an indication of singularities in the complex quasi-momentum plane of some physical quantities (although we should emphasize that branch points can drop out of suitable sums over bands, e.g., $\text{Tr}(H(k)^{-2})$).

(3) A WKB analysis can be developed [1, 32] which suggests that tunneling is enhanced between bands which can be linked by analytic continuation along a short path in the k -plane.

2. COUNTING DIMENSIONS

Many years ago, Wigner and Von Neumann [31] made a precise statement which they proposed gave meaning to the idea that the eigenvalue degeneracy "couldn't" be accidental. They proved:

THEOREM 2.1 [31]. *In the (n^2) -dimensional real vector space of $n \times n$ self-adjoint (complex) matrices, those with a degenerate eigenvalue are a "surface" of real codimension 3.*

This theorem suggests that in three-dimensional solids, accidental degeneracies occur [10]. In this section we examine similar situations with regard to cubic branch points and irreducibility of eigenvalues of $A + \lambda B$ whose A and B are self-adjoint $n \times n$ matrices. Dimension counting can never prove or disprove a result, particularly when one counts in finitely many dimensions and is interested in the infinite-dimensional case. But it can serve as an indicator of what will be true unless some special mechanism is found.

THEOREM 2.2. *In the n^2 -dimensional complex vector space of $n \times n$ matrices, those with a degenerate eigenvalue are a complex variety of (complex) codimension 1 (i.e., they are of real dimension $2n^2 - 2$).*

Proof. It is a fundamental result of Galois theory [11] that any symmetric polynomial of the roots $\lambda_1, \dots, \lambda_n$ of $X^n + a_1 X^{n-1} + \dots + a_n$ is a polynomial in a_1, \dots, a_n . In particular, $\prod_{i \neq j} (\lambda_i - \lambda_j)$ is a polynomial $D(a_1, \dots, a_n)$ (this is a standard function called the discriminant [6]). The coefficients a_1, \dots, a_n of the polynomial $\det(X - T)$

for T a matrix and X a numerical unknown are polynomials in the t_{ij} . Thus $d(t_{ij}) = D(a_k(t_{ij}))$ is a polynomial in the t_{ij} which vanishes if and only if T has a degenerate eigenvalue. Thus, these T 's are the (complex) zeros of a single polynomial and so a variety of codimension 1. ■

Remarks. (1) At first sight, this proof seems to apply also to the case where Wigner and von Neumann say the codimension is 3 not 1. It is true that the set, D , of self-adjoint matrices with a degenerate eigenvalue as a subset of the set, S , of all self-adjoint matrices is given by the vanishing of a single real polynomial. But this does *not* imply that D has codimension 1. To see what is going on, consider the case $n = 2$.

$$T = \begin{pmatrix} a & c + id \\ c - id & b \end{pmatrix}.$$

The condition for T to have a degenerate eigenvalue is $\det(T) = \frac{1}{4}[\text{tr}(T)]^2$ or equivalently,

$$\frac{1}{4}(a - b)^2 + c^2 + d^2 = 0. \quad (2.1)$$

In the self-adjoint case, a, b, c, d real, the single condition (2.1) yields three real conditions $a = b; c = d = 0$. But for complex numbers (2.1) defines a "surface" of complex codimension 1. Theorem 2.2 is easier to prove than Theorem 2.1 because the fundamental theorem of algebra allows one to measure codimensions in terms of the vanishing of polynomials.

(2) Another way of understanding why Theorems 2.1 and 2.2 yield different codimensions is to remark that among those $n \times n$ matrices with a degenerate eigenvalue, those which are diagonalizable are a surface of (complex) codimension 1 (i.e., total real dimension $2n^2 - 4$). These extra two real codimensions explain why there are two fewer dimensions in Theorem 2.1 than indicated by Theorem 2.2.

(3) The point is that according to Remark 2, the "generic" matrix with a degenerate eigenvalue has Jordan "anomalies," i.e., is of the form

$$T = \begin{pmatrix} \lambda_1 & 1 & & 0 \\ & \lambda_2 & & \\ & 0 & \lambda_3 & \\ & & & \ddots \end{pmatrix}, \quad \lambda_1 = \lambda_2.$$

If S has $S_{21} \neq 0$ in the above basis, then $T + \alpha S$ has eigenvalues $\lambda_i(\alpha)$ so that λ_1 and λ_2 are the two branches of a function with a square root singularity at $\alpha = 0$. We have thus proved:

COROLLARY 2.3. *Among the $2n^2$ -dimensional real vector spaces of pairs (T, S) of self-adjoint $n \times n$ complex matrices those with the property that some eigenvalue of $T + \alpha S$ has a square root branch point is a dense open set.*

In fact, the complement is a union of "surfaces" of dimension $2n^2 - 2$ or less. This is a good point to explain what we mean by a "surface" of dimension k . S is such a

surface if there is a surface S_1 of dimension $m \leq k - 1$ (dimension of S_1 defined inductively) so that $S \setminus S_1$ is a manifold of dimension k .

THEOREM 2.4. *In the $2n^2$ -dimensional real vector space of $n \times n$ complex matrices, those with a triply degenerate eigenvalue are a surface of real codimension 4.*

Proof. In terms of the notation of the proof of Theorem 2.2,

$$\sum_{i=1}^n \prod_{\substack{j, k \neq i \\ j \neq k}} (\lambda_j - \lambda_k)$$

is a polynomial \tilde{D} in the coefficients of P . Both $D(a)$ and $\tilde{D}(a)$ vanish when and only when either P has a triple root or two double roots. It follows that the set of matrices with $n - 2$ or fewer distinct eigenvalues is a variety, V , of (complex) codimension 2 so that the set τ of matrices with a triply degenerate eigenvalue is contained in a surface of (real) codimension 4. If $T \in \tau$ has a triply degenerate eigenvalue and all other eigenvalues distinct, then $T \in V$ and so there is a piece of V through T which is a manifold of real dimension $2n^2 - 4$. But no matrices near T have two doubly degenerate eigenvalues, so the piece of V through T is part of τ ; i.e., most of τ is a manifold of dimension $2n^2 - 4$. By continuing this analysis we see that τ is a surface. ■

Remarks. (1) This makes precise the notion that third-order branch points are unlikely. But it says that in general they are not too unlikely. That they do not occur in one-dimensional solids seems to be somewhat special and on the basis of Theorem 2.4, we expect triple points to occur in solids for suitable three-dimensional potentials. We also expect such branch points for anharmonic oscillators with polynomial interactions of a suitable type.

(2) The explicit example $T(\lambda) = A + \lambda B$ with

$$A = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}; \quad B = \frac{1}{2} \begin{pmatrix} 0 & -i & -i \\ i & 0 & -i \\ i & i & 0 \end{pmatrix}$$

is of some interest. A and B are both self-adjoint but $T(+i)$ are triply degenerate and the eigenvalues have a cube root singularity there.

(3) The first remark illustrates the dangers of the Wigner-von Neumann philosophy.

(4) One can show (R. Gunning, private communication) that the "surface" above is a variety in the technical sense [6].

Finally we consider the question of whether the eigenvalues $E(\lambda)$ of $A + \lambda B$ are values of a single analytic function. We first prove the weak but interesting

THEOREM 2.5. *Let \mathcal{M} be the $2n^2$ -dimensional real vector space of pairs (A, B) of self-adjoint matrices. Let \mathcal{S} consist of these pairs (A, B) with the property that the*

eigenvalues of $A + \lambda B$ are all the values of a single analytic function. Then $\mathcal{M} \setminus \mathcal{S}$ is a proper real subvariety and, in particular, \mathcal{S} is a dense open subset of \mathcal{M} .

Proof. Let $P(\lambda, \mu)$ be an arbitrary polynomial of degree n in λ and μ . The set of such P which can be factored as $P_1(\lambda, \mu) P_2(\lambda, \mu)$ in a nontrivial way is a variety in the set of polynomials; i.e., it is the zero set for some family of polynomials Q_1, \dots, Q_k in the coefficients in of P . (This follows from the fact that such P 's can be written as a union of parameterized "curves," the parameters being the coefficients of P_1 and P_2 .) $(A, B) \in \mathcal{M} \setminus \mathcal{S}$ if and only if the polynomial $\det(A + \lambda B - \mu)$ is reducible, and so if and only if certain polynomials Q_1, \dots, Q_k all vanish. This realizes $\mathcal{M} \setminus \mathcal{S}$ as a variety. To see that it is proper we need only find some pair $(A, B) \in \mathcal{S}$. Take

$$A = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ \vdots & & & & \\ 1 & 1 & 1 & \cdots & 0 \end{pmatrix}; \quad B = \begin{pmatrix} 0 & -i & -i & \cdots & -i \\ i & 0 & -i & \cdots & -i \\ \vdots & & & & \\ i & i & i & \cdots & 0 \end{pmatrix}.$$

As in Remark 2 above, $A + \lambda B$ has an eigenvalue with an n th-order branch point at $+i$ so that the n eigenvalues are branches of a single analytic function. Thus \mathcal{S} is non-empty. ■

We have not been able to identify precisely the dimension of $\mathcal{M} \setminus \mathcal{S}$ but we make the following

Conjecture 1. $\mathcal{M} \setminus \mathcal{S}$ has codimension $2n - 2$.

Before explaining our reason for this conjecture, let us see what this suggests about band structure of solids. We claim that the natural restriction of a Bloch Hamiltonian to $n \times n$ matrices is to consideration of families of the form $A + B_0 + \lambda B_1 + \lambda^2$ where B_0, B_1 are fixed and A is an n -parameter family. The B 's come from the quasi-momentum and A from the potential. If the $n \times n$ approximation is made by making x take finitely many values then A is clearly diagonal and so n -parameter; if we get the $n \times n$ approximation by making p take only finitely many values, V is convolution with something depending only on $(k - m) \bmod n$ and so is n -parameter. In general, an n -parameter set will not meet a set of codimension $2n - 2$ and this becomes more so as n increases! This numerology suggests that:

Conjecture 2. For any (bounded) potential the eigenvalues of $H(k)$ are different Riemann sheets of a single multivalued analytic function, with the proviso that obvious geometric symmetries must be taken into account.

Let us now try to explain Conjecture 1. One way that a pair (A, B) can lie in $\mathcal{M} \setminus \mathcal{S}$ is for A and B to have a common invariant subspace S . The only way it can then happen that $(A, B) \in \mathcal{S}$ is for the eigenvalues on S and S^\perp to be identical. The real dimension of \mathcal{M} is $2n^2$. The dimension of all j -dimensional, S , subspaces is easily seen to be $2j[n - j]$ (e.g., one can choose an O.N. basis for S in $(2n - 1) + \cdots + (2n - 2j + 1)$ -dimensional ways, each j -dimensional space has a family of O.N. basis of dimension j^2). Given S , the dimension of self-adjoint A with S as invariant subspace

is $j^2 + (n - j)^2$ and similarly for B . Thus the dimension of (A, B) with a common j -dimensional invariant subspace is

$$2j(n - j) + 2j^2 + 2(n - j)^2 = 2n^2 - 2j(n - j)$$

and so of codimension $2j(n - j)$. This is minimized by $j = 1$ and, in particular, $\mathcal{M} \setminus \mathcal{J}$ has a piece of codimension $2n - 2$; i.e., $\text{codim}(\mathcal{M} \setminus \mathcal{J}) \geq 2n - 2$.

Let us try to explain why it is reasonable that the above case is the smallest codimensional piece of $\mathcal{M} \setminus \mathcal{J}$. At a point of double-crossing, the generic possibility is that there is a square root branch point. Suppose that only double degeneracies occur and that at each such degeneracy there is either a square root branch point or a partly accidental degeneracy, i.e., analytic eigenvalues and *projections* at that point (this can fail to happen; see [12]). In this situation (which is “generic”) if $(A, B) \in \mathcal{M} \setminus \mathcal{J}$, then the spectral projection for $A + \lambda B$ associated to the eigenvalues for a given irreducible part of $\det(A + \lambda B - \mu) = 0$ is an entire function approaching a spectral projection for B as $\lambda \rightarrow \infty$. It follows that it is constant so that (A, B) have a common nontrivial invariant subspace. For this reason, we make Conjecture 1.

3. BEHAVIOR ON THE REAL AXIS

In this section, we briefly describe the analytic behavior of the energy band functions over the real Brillouin zone. There are two kinds of possibilities about which one might worry a priori and such worries appear in the physics literature [30]:

- (a) Can a band have a loose end, i.e., suddenly stop?
- (b) Can “closed bubble” bands occur?

That neither of these things happens is a simple consequence of the Kato–Rellich theory, and it is useful to compare the simplicity and power of these methods with the weakness of results obtained in the attempts (e.g., [30]) to eliminate these phenomena using classical methods of analysis.

Let us begin by picking a direction k_0 and considering the function $f(\lambda) = \epsilon_n(\lambda k_0)$. Depending on whether a multiple of k_0 lies in Γ or not, $H(\lambda k_0)$ is periodic in λ or not. In any event, one has:

THEOREM 3.1. *Let $V \in L_{\text{c}}^{3/2}(\mathbb{R}^3)$ and fix $\lambda_0 k_0, n$. Then the n th eigenvalue $\epsilon_n(\lambda_0 k_0)$ is the value at λ_0 of an analytic function $f(\lambda)$ in a neighborhood, N , of $(-\infty, \infty)$ so that $f(\lambda)$ is an eigenvalue of $H(\lambda k_0)$ for all $\lambda \in N$. Moreover, for some fixed a and b (independent of n, λ_0, k_0)*

$$|df(\lambda)/d\lambda| \leq a + b |f(\lambda)|; \quad \lambda \in \mathbb{R}, \tag{3.1}$$

and in particular,

$$|f(\lambda)| \leq C e^{b|\lambda|}, \quad \lambda \in \mathbb{R}. \tag{3.2}$$

Remark. This theorem, in particular, eliminates possibilities (a) and (b) above.

Proof. Since $V \in L_{loc}^{3/2}$, we have the bound; $\phi \in L^2(\Gamma)$, $k \in \mathbb{R}^3$,

$$(\phi, |V| \phi) \leq \epsilon(\phi, T(k) \phi) + c_\epsilon(\phi, \phi) \tag{3.3}$$

for since V is periodic, it is uniformly locally $-L^{3/2}$ and thus by a bound of Strichartz [26] (see [19, 21])

$$(\psi, V\psi) \leq \epsilon(\psi, -\Delta\psi) + c_\epsilon(\psi, \psi) \tag{3.4}$$

for $\psi \in L^2(\mathbb{R}^3)$. One obtains (3.3) from (3.4) by passing to the direct integral decomposition, noting thereby that (3.3) holds for a.e. k . By continuity it holds for all $k \in \mathbb{R}^3$.

From (3.3) one finds that for suitable c, d ,

$$(\phi, T(k) \phi) \leq c[(\phi, H(k) \phi) + d] \tag{3.5}$$

(by noting that $H(k) \geq T(k) - |V| \geq (1 - \epsilon) T + c_\epsilon$). Now

$$H(\lambda k_0) = \lambda^2 k_0^2 + 2\lambda k_0 \cdot p + H(0)$$

so

$$\pm dH/d\lambda = 2(\lambda k_0^2 + k_0 \cdot p) \leq 2|k_0| |\lambda k_0 + p| \leq 2|k_0|^2 + T(\lambda k_0).$$

Therefore

$$\pm \langle \phi, (dH/d\lambda) \phi \rangle \leq C(1 + (\phi, H(\lambda) \phi)). \tag{3.6a}$$

Moreover,

$$d^n H/d\lambda^n = \text{const} \quad \text{for } n \geq 2. \tag{3.6b}$$

From (3.6), it follows that $H(\lambda)$ is analytic family (of type B)) in the sense of Kato [12] so that by the general Kato–Rellich theory [12, 21], one has local regularity; i.e., any eigenvalue $\epsilon_n(\lambda_1, k)$ can be continued uniquely to λ near λ_1 (if degeneracies occur, this depends on the fact that $H(\lambda)$ is self-adjoint and λ_1 is real). To be able to continue globally we need to know that (i) $f(\lambda)$ cannot go to infinity as λ approaches a finite point; (ii) $f(\lambda)$ cannot have loose ends, i.e., if $\lambda_n \rightarrow \lambda$ and $f_n(\lambda_n)$ is bounded then $f(\lambda_n)$ has a limit which is an eigenvalue of $H(\lambda k)$. (ii) follows from the fact that $H(\lambda k)$ has compact resolvent (see [24]). We prove more than this compactness in Appendix A; compactness alone is an easy consequence of (3.1), (3.3), and the compactness of $(T(0) + 1)^{-1}$. (i) clearly follows from (3.2), which follows from (3.1). Equation (3.1) follows from (3.6) and the standard formula

$$df(\lambda)/d\lambda = (\eta, (dH/d\lambda) \eta) \tag{3.7}$$

(where $H(\lambda) \eta = f(\lambda) \eta$ and $\|\eta\| = 1$).

Remark. The proof of (3.1) is standard; see [12].

In part, the above arguments extend to complex λ , but only in part. The germane facts are as follows.

(i) $H(\lambda k_0)$ is an entire function of type (b), so local continuation is possible *modulo* (iii) below.

(ii) $H(\lambda k_0)$ still has compact resolvent so “loose ends cannot occur.”

(iii) Since $H(\lambda k_0)$ is no longer self-adjoint, algebraic branch points can occur and indeed, we show they do occur in many cases.

(iv) The proof above that $f(\lambda)$ is bounded no longer works because (3.7) no longer holds with a normalized η , but is replaced by

$$df/d\lambda = (\eta(\bar{\lambda}), (dH/d\lambda) \lambda(\lambda))$$

with $(\eta(\bar{\lambda}), \eta\lambda) = 1$. In fact, at usual square root singularities $df/d\lambda$ will diverge (although only as $|\lambda - \lambda_1|^{-1/2}$. We use the phrase “no natural boundaries occur” as shorthand for the statement that f is bounded as λ varies; of course, such a statement also implies that no poles occur.

We close this section with a few remarks on multivariable analyticity. For a one-variable problem, it is a theorem of Rellich [22] that self-adjointness for real λ implies analyticity of eigenvalues for real λ even at level crossings: This comes from the fact that if f is a function with an algebraic branch point on the real axis so that every branch is real on the real axis, then the algebraic branch point has order 1, i.e., is really a point of analyticity. Rellich [22] gives the example

$$T(\alpha, \beta) = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}$$

with eigenvalues $\pm(\alpha^2 + \beta^2)^{1/2}$ so that one can have singularities on the real axis in the multivariable case. In fact, such singularities are believed to occur in semiconductor energy bands [13].

We do have a conjecture to make about singularities in this case.

Conjecture 3. Let $E(\lambda_1, \dots, \lambda_n)$ be an algebraic function defined near $(0, \dots, 0)$; i.e., E solves an equation $\sum_{j=0}^N a_j(\lambda_1, \dots, \lambda_n) E^j$ with a_j analytic near zero. Suppose that all the branches of E are real for $\lambda_1, \dots, \lambda_n$ real. Then, there exists a family of analytic functions $\{f_{jk}^i\}$, $j \leq i$, $0 \leq k \leq m$, such that

$$E(\lambda) = f_{00}^0(\lambda) + \sum_i \pm (f_{0i}^0(\lambda))^{1/2} + \dots + \sum_i \pm (f_{0i}^n(\lambda) + (f_{1i}^n + \dots)^{1/2})^{1/2}$$

where each f_i obeys $f_i(z\lambda_0) = O(z^{2k})$ as $z \rightarrow 0$ (k may depend on λ_0 ; e.g., $f_i(\lambda_1, \lambda_2) = \lambda_1^2 + \lambda_2^4$ is allowed).

Conjecture 3 would imply that only the Rellich type of singularity can occur.

4. BOUNDS ON THE REAL PART

In a study of the analyticity of the anharmonic oscillator energy levels [24], a major role was played by Martin's observation that in that case $\text{Im } E(\beta) \geq 0$ if $\text{Im } \beta > 0$. This is certainly not true for energy bands but a suitable replacement exists, namely:

THEOREM 4.1. *Let $V \in L_{\text{loc}}^{3/2}(\mathbb{R}^3)$ be a periodic potential. Then for some V -dependent constant b and all $k \in \tilde{\Gamma}^3$:*

$$\sigma(H(k)) \subset \{z \mid \text{Re } z \geq b - |\text{Im } k|^2\}. \quad (4.1)$$

Proof. By (1.3) for any ϕ ,

$$\text{Re}(\phi, H(k)\phi) = (\phi, H(\text{Re } k)\phi) - |\text{Im } k|^2 \|\phi\|^2.$$

Because of (3.3) and (3.5), $(\phi, H(\text{Re } k)\phi) \geq b(\phi, \phi)$ so we have

$$\text{Re}(\phi, H(k)\phi) \geq (b - |\text{Im } k|^2) \|\phi\|^2.$$

Equation (4.1) now follows from the well-known result [20]

$$\sigma(A) \subseteq \overline{\text{cvx null}\{(\phi, A\phi) \mid \|\phi\| < 1\}}. \quad (4.2)$$

Actually, for the case at hand (4.2) is easy since $H(k)$ has a discrete spectrum and $A\psi = e\psi$ implies $e = (\psi, A\psi)$ if $\|\psi\| = 1$. ■

COROLLARY 4.2. *The energy band functions $f(\lambda) = \epsilon_n(\lambda k_0)$ cannot have isolated (single-sheeted) singularities.*

Proof. In the neighborhood of any poles or isolated essential singularity there are values with arbitrary large negative real part violating (4.1). ■

COROLLARY 4.3. *If an energy band function $f(\lambda) = \epsilon_n(\lambda k_0)$ has an algebraic singularity (i.e., isolated multisheeted singularity with a finite number of sheets), the Pusieux expansion (expansion in $(\lambda - \lambda_0)^{n/p}$) has no negative terms.*

Proof. Same as for Corollary 4.2. ■

5. ENTIRE BANDS ARE PARABOLIC

As we have emphasized, Theorem 4.1 is an analog of the Herglotz property for the anharmonic oscillator. In this section, we mimic the analysis of isolated singularities at infinity for Herglotz functions [24] to prove that the only entire energy band functions are parabolic. The basic input result is

THEOREM 5.1. *Let $f(z)$ be a function analytic in $D = \{z \mid |z| > R\}$ obeying $\operatorname{Re} f \geq b - c(\operatorname{Im} z)^2$; $f(\bar{z}) = \overline{f(z)}$. Then*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n z^n$$

with $a_2 \leq c$.

Proof. Since f has an isolated singularity, it has a convergent Laurent expansion $f(z) = \sum_{-\infty}^{\infty} a_n z^n$ with a_n real. Thus

$$\operatorname{Re} f(re^{i\theta}) = \sum_{-\infty}^{\infty} a_n r^n \cos n\theta. \tag{5.1}$$

Moreover, we have

$$\operatorname{Re} f(re^{i\theta}) \geq b - cr^2 \sin^2 \theta. \tag{5.2}$$

Multiplying (5.2) by the positive function $(2\pi)^{-1} (1 \pm \cos(n\theta))$ and integrating and then using (5.1) we obtain

$$a_0 \pm \frac{1}{2}(a_n r^n + a_{-n} r^{-n}) \geq -(\operatorname{const}) r^2.$$

Taking the right choice of \pm and $r \rightarrow \infty$, we see that $a_n = 0$ for $n \geq 3$. ■

This theorem clearly implies that if ϵ_n is an energy band function and $f(\lambda) = \epsilon_n(\lambda k_0)$ is entire for some k_0 , then $f(\lambda) = a\lambda^2 + b\lambda + c$ so that f is parabolic. One would expect that only the free parabola is possible and we can prove this for suitable k_0 . Our proof is only one step past Thomas' proof [27] that flat bands are impossible and is modeled on his proof (obviously our result implies his).

THEOREM 5.2. *Let V be a bounded periodic potential. Let $k_0 \in \tilde{\Gamma}$ (the x -space lattice or direct lattice) and suppose that $f(\lambda) = \epsilon_n(\lambda k_0)$ is analytic near infinity for some band. Then*

$$f(\lambda) = (k_1 + \lambda k_0)^2 + O(1)$$

for some $k_1 \in \Gamma$ (the p -space or reciprocal lattice).

Proof. Without loss, we can suppose that k_0 lies in an integral basis for $\tilde{\Gamma}$ since any k_0 is a multiple of such a basis vector. Thus $\{k \cdot k_0 \mid k \in \Gamma\} = 2\pi\mathbb{Z}$. By Theorem 5.1, $f(\lambda)$ has an expansion $a\lambda^2 + b\lambda + O(1)$. The content of this theorem is that $a = k_0^2$ and $b = 4\pi n$, for some $n \in \mathbb{Z}$. Thus suppose that either $a \neq k_0^2$ or $b \neq 4\pi n$. We show below that in that case

$$\|(T(\lambda k_0) - f(\lambda))^{-1}\| \rightarrow 0; \quad \text{as } \lambda \rightarrow \infty \text{ along a suitable path.} \tag{5.3}$$

Equation (5.3) contradicts the fact that $f(\lambda)$ is an eigenfunction of $T(\lambda k_0) + V$, for

$$(T(\lambda k_0) + V) \phi = f(\lambda) \phi$$

implies that

$$(T(\lambda k_0) - f(\lambda))^{-1} V\phi = \phi$$

or $\|V\| \| (T(\lambda k_0) - f(\lambda))^{-1} \| \geq 1$, which implies that (5.3) cannot hold. Thus a proof of (5.3) under the hypothesis $a \neq k_0^2$ or $b \neq 4\pi n$ shows by contradiction that this hypothesis is false and proves the theorem.

$T(\lambda k_0) - f(\lambda)$ has an orthonormal set of eigenvectors with eigenvalues $E_p(\lambda) = (p + \lambda k_0)^2 - a\lambda^2 - b\lambda - r - O(1/\lambda)$; $p \in \Gamma$ (r real). Clearly (5.3) follows from

$$\inf_p |\operatorname{Im} E_p(\lambda)| \rightarrow \infty \quad \text{as } \lambda \rightarrow \infty \text{ suitably.} \quad (5.4)$$

Now letting $p \cdot k_0 = 2\pi n$,

$$|\operatorname{Im} E_1(\lambda)| = |\operatorname{Im} \lambda| |(4\pi n - \beta) + (2k_0^2 - 2a)(\operatorname{Re} \lambda)|.$$

If either $a \neq k_0^2$ or $\beta \notin 4\pi\mathbb{Z}$, we can arrange for the inf in (5.4) to grow like $c |\operatorname{Im} \lambda|$, $c \neq 0$, by choosing $\operatorname{Re} \lambda$ suitably. ■

We conclude this section with a series of remarks about these results.

(i) In particular, if V is bounded and $\epsilon_n(k)$ is an entire function on \mathbb{C}^3 , $\epsilon_n(k)$ must be one of the free bands. We expect that if this occurs, V must be zero although we only know how to prove this in the case of one dimension.

(ii) For very special potentials in one dimension (namely, those with only finitely many “bands of instability”; see [3]), it can happen that a band function is analytic near infinity. The properties of $\Delta(\lambda)$ tell us that such bands must look “free” near infinity. We presume that many such potentials also exist in three dimensions: again they must look “free” near infinity.

(iii) If infinity is not an isolated singularity, it must be a limit point of singularities which must be branch points in some circumstances (see Sect. 6). This infinite number of singularities should not be surprising since in the typical case $f(\lambda)$ is periodic so that one singularity in the basic period region yields an infinite number of singularities.

(iv) The fact that no band functions are entire other than the trivial band function explains why no “elementary” exactly solvable band models have been found.

6. NATURAL BOUNDARIES

It is our expectation that energy band functions must remain bounded when continued along bounded paths in the complex plane. As we have already described, this would imply that their only singularities are algebraic. Unfortunately, we are only able to prove this expectation in one dimension. Our proof is motivated in part by the ideas of Loeffel and Martin [17] in their study of anharmonic oscillators. We first note:

PROPOSITION 6.1. *Let $A(k)$ be an entire analytic family of operators with compact resolvents and let B be a bounded operator. Suppose that for each $k_0 \in \mathbb{C}$, there is a neighborhood N of k_0 and for each $E_0 \in \mathbb{C}$ a simple closed curve Γ surrounding E_0 so that $\Gamma \subset \mathbb{C} \setminus \sigma(A(k))$ for all $k \in N$ and*

$$\sup_{k \in N} \sup_{E \in \Gamma} \|B(A(k) - E)^{-1}\| < 1. \tag{6.1}$$

Then, any eigenvalue e_0 of $A(k) + B$ can be analytically continued along any curve with the only possible singularities being finite algebraic branch points.

Proof. By general arguments [24], it suffices to show that the eigenvalue cannot diverge as it is continued along a curve $\gamma(t)$. Suppose the contrary so that as $\gamma(t) \rightarrow k_0$, some eigenvalue $e(k)$ diverges. Pick t_0 so that for $t = t_0$, $\gamma(t)$ is in the neighborhood N of the hypothesis and let $E_0 = e(\gamma(t_0))$. Then, since (6.1) holds, $(A + B - E) = (1 + B(A - E)^{-1})(A - E)$ is invertible on Γ , so $e(\gamma(t))$ must lie with Γ for $t \geq t_0$. It follows that $e(\gamma(t))$ cannot diverge as $\gamma(t) \rightarrow k_0$. ■

Remark. It is critical in the above that N be independent of E .

THEOREM 6.2. *One-dimensional periodic systems with bounded potentials have band functions without natural boundaries.*

Proof. By periodicity we need only consider the region $|\operatorname{Re} k| \leq \frac{1}{2}$. We use Proposition 6.1 with $A(k) = T(k) - ck^2$ and $B = V$. Pick the period and c so that $A(k)$ has eigenvalues

$$e_n^{(0)}(k) = n^2 + 2kn.$$

It is easy to see that the eigenvalue with real part closest to $e_n^{(0)}(k)$ is one of $e_{\pm(n \pm 1)}^{(0)}(k)$ and that at most one of them is within n of $\operatorname{Re} e_n^{(0)}(k)$. Thus the real parts have gaps in the spectrum of order n . Choosing N a small neighborhood of k_0 and using the fact that $|e_n(k) - e_n(k_0)| = 2(k - k_0)n$ we see that given any E , there is a curve Γ to go with E in Γ and (within the gaps)

$$\sup_{t \in N} \sup_{E \in \Gamma} \|(A(k) - E)^2\| < 1/n.$$

Thus (6.1) holds. ■

Remarks. (1) V bounded is not necessary. Since the gaps are of order $(e^{(0)}(k))^{1/2}$, it suffices that V be $(T(0))^{1/2}$ form bounded with relative bound zero. For example, $V \in L_{\text{loc}}^{1+\epsilon}(\mathbb{R})$ and periodic will do.

(2) If one looks at $T^{(0)}$ for a two-dimensional cubic lattice, one finds arbitrarily large gaps for the simple number theoretic reason that the sums of two squares have density $(\ln n)^{-1}$ [8]. Thus at first sight our argument will extend to this case also. Alas, variations with k are still of order $(E_n)^{1/2}$, overwhelming the $\ln E_n$ gaps!

(3) Since $E(k)$ is given by solutions of the equation $\Delta(E(k)) = 2 \cos k$, one would think it easy to prove that no natural boundaries occur, for “surely” inverse functions of entire functions cannot have natural boundaries. This “surely” statement appears to be false (J. Fornæss, private communication).

APPENDIX A: TRACE IDEAL PROPERTIES OF RESOLVENTS OF BAND HAMILTONIANS

Recall [25]:

DEFINITION. The singular values $\mu_n(B)$ of a bounded operator, B , are the eigenvalues of $|B| = (B^*B)^{1/2}$ ordered $\mu_1 \geq \mu_2 \geq \dots$. B is said to lie in the weak trace ideal $\mathcal{S}_W^p (p > 1)$ if and only if

$$\|B\|_{p,W} \equiv \sup \left[n^{-1+1/p} \sum_{j=1}^n \mu_j(B) \right] < \infty.$$

Here we note [2]:

THEOREM A1. If $V \in L_{loc}^{3/2}(\mathbb{R}^3)$ is periodic, the Bloch Hamiltonians, $H(k)$, have resolvents in $\mathcal{S}_W^{3/2}$ and in no smaller \mathcal{S}_W^p or \mathcal{S}^p space.

Proof. $T(0)$ has a resolvent in $\mathcal{S}_W^{3/2}$ since its n th eigenvalue goes as $n^{2/3}$. Thus $T(0) + V$ has the same property, since by (3.3)

$$(T(0) + V + \lambda)^{-1} = (T(0) + \lambda)^{-1/2} (1 + (T(0) + \lambda)^{-1/2} V (T(0) + \lambda)^{-1/2}) (T(0) + \lambda)^{-1/2}$$

and $(T(0) + \lambda)^{-1/2} \in \mathcal{S}_W^3$ and we have weak Hölder inequalities [25]. Since $T(k) - T(0)$ is $T(0)$ form bounded, the same argument shows that $(T(k) + V + \lambda)^{-1} \in \mathcal{S}_W^{3/2}$. If this resolvent were in some smaller \mathcal{S}_W^p space, we could turn the above around and prove that $T(0)$ is in this smaller space. ■

APPENDIX B: THE DISCRIMINANT AND THE ONE-DIMENSIONAL CASE
(FOLLOWING KOHN [14])

We want to provide here a quick proof of Kohn’s result [14] that in the one-dimensional case only square root branch points occur and that, if all the energy gaps are present, the bands are the value of a single multivalued function. We do this partly for the readers’ convenience, and partly to present an alternative proof (following a suggestion of Trubowitz) of one of Kohn’s main input lemma:

LEMMA B1. The discriminant, $\Delta(\lambda)$, for any one-dimensional problem with periodic L_{loc}^1 potential has the property that $\Delta'(\lambda)$ only vanishes for λ real with $|\Delta(\lambda)| \geq 2$ and all such zeroes are simple.

Proof. $\Delta(\lambda)$ has only real zero and $\lambda_0 < \lambda_1 < \dots$ (since $\Delta(\lambda) = 0$ implies that λ is an eigenvalue of $H(\pi/2)$) and is entire of order $\frac{1}{2}$. Thus by Hadamard's theorem [28],

$$\Delta(\lambda) = c \prod_{i=0}^{\infty} (1 - \lambda/\lambda_i)$$

(if some $\lambda_i = 0$, $1 - \lambda/\lambda_i$ is replaced by λ) with $\sum \lambda_i^{-1} < \infty$. Thus

$$\Delta'(\lambda)/\Delta(\lambda) = \sum_{i=0}^{\infty} (\lambda - \lambda_i)^{-1}$$

so

$$\text{Im}[\Delta'(\lambda)/\Delta(\lambda)] = (\text{Im } \lambda) \sum_{i=0}^{\infty} |\lambda - \lambda_i|^{-1}.$$

It follows that $\Delta'(\lambda)/\Delta(\lambda)$ has all its zeros on the real axis and all these zeros are simple. That $|\Delta(\lambda)| \geq 2$ is required if $\Delta'(\lambda) = 0$ is a consequence of the fact that the zeros are simple and that $\Delta(\lambda)$ is strictly monotone in intervals with $|\Delta(\lambda)| < 2$ (e.g., [21]). ■

Given this lemma, we see that locally Δ^{-1} can only have square root branch points. By Theorem 6.2, it has no other singularities. Thus $2 \cos k$ real and $|\cos k| > 1$ imply $\text{Re } k = 0, \pm\pi, \dots$

THEOREM B2. *If $V \in L^1_{\text{loc}}(\mathbb{R})$ is periodic, the band functions are continuable along any path with only square root branch points occurring. These can only occur at points k with $\text{Im } k \neq 0, \text{Re } k = 0, \pm\pi, \dots$*

Suppose all gaps are present, i.e., $\Delta'(\lambda) = 0$ implies $|\Delta(\lambda)| > 2$, not just ≥ 2 . Then at such points $2 \cos k$ has a nonvanishing derivative so that at such λ_0 's, $e(k) = \lambda_0$ has a square root branch point with nonzero fractional terms in the Puiseux series. If we follow a path around such a branch point we go from one root of $\Delta(\lambda) = c$ to the other. Thus following a path in k so that $2 \cos k$ runs along the curve $2 \cos k = \Delta(\lambda)$ ($\lambda \in (-\infty, \infty)$) with excursions into complex $\cos k$ only to get around the branch points, we run through all energy bands. (see [14, Fig. 4]).

APPENDIX C: NONDEGENERACY OF THE GROUND-STATE BAND

The modern approach to proving the nondegeneracy of the ground state is based on theorems of Perron–Frobenius type. (This idea was first developed by Glimm and Jaffe [5]; see, e.g., [21] for a full discussion.) The ground state of H is the largest eigenvalue of e^{-H} so that one tries to prove that this is an operator with a strictly positive kernel, i.e., if $\psi(x) \geq 0$, then $(e^{-H}\psi)(x) > 0$. Since this approach works for general Schrödinger operators, it certainly applies to those with periodic potentials. The result is that $H(0)$ has a nondegenerate lowest eigenvalue. In general, not all $H(k)$'s will

have that property, e.g., if $V = 0$ in one dimension and $k = \pm\pi$. But we wish to point out that for an interesting class of V 's each $H(k)$ has a nondegenerate ground state.

THEOREM C1. *Let V be a periodic potential whose Fourier "transform" $\tilde{v}(q)$ given by (1.3d) obeys:*

- (a) $\tilde{v}(q) \leq 0$ all q .
- (b) $\{q \mid \tilde{v}(q) \neq 0\}$ contains a basis of the reciprocal lattice, Γ .

Then, for any $k \in \tilde{\Gamma}$, in the Brillouin zone, the smallest eigenvalue of $H(k)$ is nondegenerate.

Proof. Think of $e^{-H(k)}$ as an operator on $l^2(\Gamma)$ in the form (1.3). It suffices to show that it is positivity improving, i.e., $e^{-H(k)}(p, p') > 0$ for all p, p' . Now by a standard argument [4, 21] it suffices to prove that $(e^{-V})(p, p') > 0$ since $e^{-T(k)}$ is a positive multiplication operator. Now, expanding the exponential, it suffices to show that $(-V)^n(p - p') \geq 0$ and > 0 for some n . The ≥ 0 follows from (a) and the > 0 from (b).

EXAMPLE. The periodized attractive Yukawa (screened Coulomb) potential $V(x) = -\sum_{y \in \tilde{\Gamma}} e^{-\mu|x-y|}|x-y|^{-1}$ obeys the hypothesis of Theorem C1.

ACKNOWLEDGMENTS

We are pleased to thank E. Trubowitz for permission to use an argument of his in Appendix B, F. Dyson for a discussion of some number theoretic questions related to Section 6, A. Grossman for his keen interest and for useful discussions, and A. Wightman for advocating the usefulness of the ideas developed in Section 2.

REFERENCES

1. J. AVRON, unpublished.
2. J. AVRON, A. GROSSMAN, AND R. RODRIGUEZ, *Rep. Math. Phys.* **5** (1974), 113; *Ann. Phys.* **102** (1976), 555.
3. M. S. P. EASTHAM, "The Spectral Theory of Periodic Differential Equations," Scott, New York, 1973.
4. W. FARIS AND B. SIMON, *Duke Math. J.* **42** (1975), 559-567.
5. J. GLIMM AND A. JAFFE, *Ann. of Math.* **91** (1970), 362-401.
6. R. GUNNING AND H. ROSSI, "Analytic Functions of Several Complex" Variables," Prentice-Hall, Englewood Cliffs, N.J., 1965.
7. G. HAMEL, *Math. Ann.* **73** (1913), 371-412.
8. G. H. HARDY AND E. M. WRIGHT, "An Introduction to the Theory of Numbers," Oxford Univ. Press, London/New York, 1938.
9. O. HAUPT, *Math. Ann.* **79** (1919), 278-285.
10. C. HERRING, *Phys. Rev.* **52**, 365, 1937; also in "Symmetry in the Solid State" (R. S. Knox and A. Gold, Eds.), p. 240, Benjamin, New York, 1964.
11. N. HERSTEIN, "Topics in Algebra," Xerox, 1964.

12. T. KATO, "Perturbation Theory for Linear Operators," Springer, New York, 1966.
13. C. KITTEL, "Quantum Theory of Solids," Wiley, New York, 1963.
14. W. KOHN, *Phys. Rev.* **115** (1959), 809–821.
15. W. KOHN, *Proc. Phys. Soc. London* **72** (1959), 1147.
16. H. A. KRAMERS, *Physica* **2** (1935), 483–490.
17. J. J. LOEFFEL AND A. MARTIN, Proc. R.C.P. No. 25, 1970; CERN Preprint, TH. 1167-CERN.
18. F. ODEH AND J. B. KELLER, *J. Math. and Phys.* **5** (1964), 1499–1504.
19. M. REED AND B. SIMON, "Methods of Modern Mathematical Physics," Vol. II, "Fourier Transform, Self-Adjointness," Academic Press, New York/London, 1975.
20. M. REED AND B. SIMON, "Methods of Modern Mathematical Physics," Vol. III, "Scattering Theory," Academic Press, New York/London, to appear.
21. M. REED AND B. SIMON, "Methods of Modern Mathematical Physics," Vol. IV, "Analysis of Operators," Academic Press, New York/London, to appear.
22. F. RELICH, *Math. Ann.* **113** (1937), 600–819.
23. G. SCHARF, *Helv. Phys. Acta* **39** (1966), 556–560.
24. B. SIMON, *Ann. Phys.* **58** (1970), 76–136.
25. B. SIMON, *Trans. Amer. Math. Soc.* **224** (1976), 367.
26. R. STRICHARTZ, *J. Math. Mech.* **16** (1967), 1031–1060.
27. L. THOMAS, *Comm. Math. Phys.* **33** (1973), 335–343.
28. E. TITCHMARSH, "Function Theory," Cambridge Univ. Press, London/New York, 1939.
29. G. TROIANIELLO, *J. Math. Phys.* **15** (1974), 2048–2052.
30. G. H. WANNIER, "Elements of Solid State Theory," Cambridge, London/New York, 1957.
31. E. P. WIGNER AND J. VON NEUMANN, *Phys. Z.* **30** (1927), 467; English transl. in "Symmetry in the Solid State" (R. S. Knox and A. Gold, Eds.), pp. 167–172, Benjamin, New York, 1964.
32. C. ZENER, *Proc. Roy. Soc. London* **145** (1934), 523; also in "Mathematical Physics in One Dimension" (E. H. Lieb and D. Mattis, Eds.), Academic Press, New York/London, 1966.