Homotopy and Quantization in Condensed Matter Physics

J. E. Avron[(a)]
Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, California 91125

and

R. Seiler
Institute for Theoretical Physics, Free University of Berlin, D-1000 Berlin 33, West Germany

and

B. Simon
Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, California 91125
(Received 31 May 1983)

It is shown that the integers found by Thouless *et al.* in the quantized Hall effect are the only quantized quantities associated with the energy bands. It is also proved that if two bands touch and then come apart as a parameter is varied, then their individual integers (conductances) may not be preserved but their sum is preserved.

PACS numbers: 72.20.My, 02.40.+m, 71.25.-s

Thouless, Kohmoto, Nightingale, and den Nijs (TKNN), in a remarkable paper,[1] considered the motion in two dimensions of noninteracting electrons in a periodic potential and homogeneous magnetic field with rational flux through the crystalline unit cell. A first main result was that the Kubo-Greenwood formula for the conductance of any filled, isolated band is, in fact, an expression for a topological invariant, and is an integral multiple of e^2/h. Second, in the limits of either a very weak periodic potential or a very weak magnetic field, they calculated the conductance of any given diamagnetic band explicitly in terms of certain diophantine equations. This work is an important step towards an understanding of the quantized Hall effect of von Klitzing, Dorda, and Pepper.[2]

Our primary purpose is to explain a negative result: Namely, that in a sense which we will make precise, the invariants found by TKNN are the *only* topological invariants, not only for the two-dimensional case, but in any dimension including three. We also prove a sum rule for the TKNN integers when two bands collide. We sketch the basic ideas here; fuller details will appear elsewhere.[3]

We shall establish results about the homotopy of periodic (infinite dimensional) Hermitian matrices: They are relevant if the band structure is described by giving to each $\vec{k}$ in the Brillouin zone a Hermitian matrix $H(\vec{k})$ with a discrete spectrum (bands) $\epsilon_n(\vec{k})$, where $\epsilon_n(\vec{k}) \to \infty$ as $n \to \infty$. The matrices $H(\vec{k})$ vary continuously[4] with $\vec{k}$, are periodic because the Brillouin zone is topologically a torus (T^2 in two dimensions), and—as the word matrix implies—act on a fixed Hilbert space.

We stress that it is not true in general that Bloch Hamiltonians (with or without a magnetic field) can be described by periodic matrices. In general, the operator $H(\vec{k})$ acts on a $\vec{k}$-dependent Hilbert space.[5] However, if $\vec{B} \neq 0$ and the flux has suitable values, the Bloch Hamiltonian can indeed be brought into the form of a periodic matrix.

To be able to associate a TKNN integer to each band, one needs to suppose that for *every* k, the $E_n(k)$'s are distinct, i.e., for all n, k, $E_n(k) < E_{n+1}(k)$. Replacing $H(k)$ by $[H(k)+E_0]^{-1}$ for suitable E_0, we have a map from T^2 to the space $\mathfrak{N}$ of compact, self-adjoint, positive operators with nondegenerate eigenvalues. We will indicate the set of all such periodic matrix functions as $[T^2 \to \mathfrak{N}]$. By a quantized invariant on T^2 we mean the association of a value in a discrete set to every element of $[T^2 \to \mathfrak{N}]$ which is continuous (and so constant by the discreteness assumption) under continuous variation of the maps. Our main results are as follow:

Theorem 1.—Every quantized invariant on T^2 is a function of the TKNN integers.

Theorem 2.—Every quantized invariant on T^d is a function of the $\frac{1}{2}d(d-1)$ sets of TKNN integers obtained by slicing[6] T^d by the $\frac{1}{2}d(d-1)$ distinct[6] T^2's.

We claim that it is obvious that these results follow from the assertion that two matrix functions $A(k)$ and $B(k)$ with values in $\mathfrak{N}$ are homotopic (can be deformed continuously one into another through maps with values in $\mathfrak{N}$) if and only if their

TKNN integers are the same.

If one looks at maps of the ν-dimensional sphere S^ν into $\mathcal{H}$, the families of maps which are homotopic are precisely the homotopy group $\pi_\nu(\mathcal{H})$. Think about two maps of T^2 into an arbitrary space X. If we take the two basic loops in T^2, we obtain from each map two elements of $\pi_1(X)$ and the two maps cannot be homotopic unless the corresponding pair of elements of $\pi_1(X)$ are the same. Even if they are the same, there is clearly a leftover map from S^2 into X. In this way, one can classify maps from T^ν into X by ν elements of $\pi_1(X)$, $\frac{1}{2}\nu(\nu-1)$ elements of $\pi_2(X)$, ... For general X the precise structure is complicated (see, e.g., Fox[7]), but if all π's are zero but for one value of j, the map is precisely classified by $\binom{\nu}{j}$ elements of π_j. Thus theorems 1 and 2 follow from the following theorem.

Theorem 3.—$\pi_k(\mathcal{H})=0$ if $k\neq 2$; $\pi_2(\mathcal{H})=Z^\infty$, an infinite set of integers, precisely given by the S^2 analogs of TKNN. Namely, let $\psi_j(k)$ be the normalized eigenfunction associated with the jth band and $k\in S^2$. Then the integer associated with the jth band is

$$n_j=(i/2\pi)\int_{S^2}(d\psi_j,d\psi_j)$$

with the $d\psi=\sum_i(\partial_{k_i}\psi)dk_i$ and $dk_i dk_l=-dk_l dk_i$.

Theorem 3 is actually a rather simple exercise in homotopy theory: Let $\mathfrak{M}$ be the set of elements of $\mathcal{H}$ whose eigenvalues are $1,\frac{1}{2},\frac{1}{3},\ldots$. It is clear that any A in $\mathcal{H}$ can be continuously deformed to an element in $\mathfrak{M}$ and vice versa by deforming the eigenvalues but keeping the eigenvectors fixed, so that $\pi_k(\mathcal{H})=\pi_k(\mathfrak{M})$.

Now any element of $\mathfrak{M}$ can be written as $UA_0U^{-1}(k)$, where A_0 is the diagonal and k-independent matrix $(A_0)_{nm}=(1/n)\delta_{mn}$ and $U(k)\in\mathfrak{U}(\mathcal{K})$, $k\in S^2$, are unitary. $U(k)$ is uniquely determined up to an element of $D\mathfrak{U}(\mathcal{K})$, the unitary diagonal matrices. It follows that $\mathfrak{M}$ is a homogeneous space,[8]

$$\mathfrak{M}=\mathfrak{U}(\mathcal{K})/D\mathfrak{U}(\mathcal{K}).$$

Most spaces whose homotopy groups are of interest to condensed matter physics are homogeneous spaces and their homotopy groups are computable in terms of "the exact sequence of a fibration"; see, e.g., Mermin.[9] Theorem 3 follows from this by knowing first[10] that

$$\pi_k(\mathfrak{U}(\mathcal{K}))=0$$

which implies that $\pi_k(\mathfrak{M})=\pi_{k-1}(D\mathfrak{U}(\mathcal{K}))$ and then by noting that $D\mathfrak{U}(\mathcal{K})$ is just an infinite-dimension torus (given by the set of eigenvalues) so that $\pi_k(D\mathfrak{U}(\mathcal{K}))=0$ if $k\geq 2$ and is Z^∞ if $k=1$. The identification of π_2 with the TKNN integers is easy.[3]

Here are some additional remarks about this proof:

(1) It is remarkable that the only invariants are those associated with individual bands and there are none associated with all the bands. If $\mathcal{H}$ is replaced with $m\times m$ matrices there are additional global invariants[3] produced by the fact that $\pi_k(U(m))\neq 0$.

(2) The proof also shows that given any set of integers $n_1,n_2,\ldots$ one can find a map $A(k)$ whose TKNN numbers are precisely $n_1,n_2,\ldots$. One can construct such $A(k)$ explicitly.[3] In the $m\times m$ case, there is a restriction on the n's, namely $\sum_{i=1}^m n_i=0$ which comes from the fact that $\pi_1(U(m))=Z$.

These homotopy ideas are also useful in studying some questions about the TKNN invariants. Suppose we start with an $A_0(k)$ with given values of n_3 and n_4.[11] We vary an additional parameter,[12] θ, and for some θ_0 and k, the 3 and 4 eigenvalues of $A_\theta(k)$ are equal, but as θ is further increased, $A_\theta(k)$ lies again in $\mathcal{H}$ for all k. Let n_3', n_4' be the TKNN integers for A_θ when $\theta>\theta_0$. It is not true that necessarily $n_3=n_3'$, $n_4=n_4'$ but we claim that

$$n_3+n_4=n_3'+n_4'. \qquad (1)$$

To see this, let $\mathcal{H}_{34}$ be those matrices whose eigenvalues are nondegenerate except perhaps for the third and fourth. We want to know about deformations in $[T^2\to\mathcal{H}_{34}]$ and (1) follows from a homotopy calculation[3] that $\pi_k(\mathcal{H}_{34})=0$ if $k\neq 0$ and $\pi_2(\mathcal{H}_{34})$ is Z^∞ with integers $m_1,m_2,m_{34},m_5,\ldots$ and that under the natural embedding of $\mathcal{H}$ into $\mathcal{H}_{34}$, $\pi_2(\mathcal{H})$ is mapped to $\pi_2(\mathcal{H}_{34})$ by

$$m_j=n_j,\quad j=1,2,5,\ldots;\quad m_{34}=n_3+n_4.$$

This both proves the sum rule and shows that any time the sum rule is obeyed one can find a deformation with a 3-4 collision producing precisely that change.

In the context of the quantized Hall effect, this conservation law is what one expects on the basis of Streda's formula[13] which relates the sum of the conductances to the density of states.

We have also found a "bare"-hands proof of (1) which we sketch because of certain formulas of independent interest which enter along the way. The TKNN integer for the jth band can be decribed as follows. Let $\psi_j(k)$ be an explicit choice of the jth eigenvector of $A(k)$. Then

$$n_j=(i/2\pi)\int_{T^2}(d\psi_j,d\psi_j). \qquad (2)$$

While n_j is independent of phase transformations, i.e., $\psi_j(k) \to e^{i\theta(k)}\psi_j(k)$, it is not manifestly so. We have found the following manifest phase-invariant formula: Let $P_j(k) = |\psi_j(k)\rangle\langle\psi_j(k)|$ be the (phase invariant) projection of the jth eigenvector. Then[3]

$$n_j = (i/2\pi)\int \mathrm{Tr}(dP_j\, P_j\, dP_j). \quad (3)$$

One can also show that if $P_{jl} = P_j + P_l$, then[3]

$$n_j + n_l = (i/2\pi)\int \mathrm{Tr}(dP_{jl}\, P_{jl}\, dP_{jl}). \quad (4)$$

Since P_{34} varies smoothly as θ varies in the above situation, (4) proves the sum rule.

We would like to thank P. Deift, R. Feynman, B. Fuller, R. Johnson, L. Romans, T. Spencer, B. Souillard, D. Thouless, G. Toulouse, J. Zak, and especially S. Cappell for useful discussions. We thank J. Sokoloff for drawing our attention to Ref. 13. This work was supported part in part through National Science Foundation Grant No. MCS-81-20833.

[a]Address after 1 October 1983: Department of Physics, Technion, 32 000 Haifa, Israel.

[1]D. Thouless, M. Kohmoto, M. Nightingale, and M. den Nijs, Phys. Rev. Lett. 49, 405 (1982).

[2]K. von Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. 45, 494 (1980).

[3]J. Avron, R. Seiler, and B. Simon, to be published.

[4]The correct topology to use is to demand that $[H(k) + i]^{-1}$ be continuous in the norm.

[5]In the most natural formulation, $H(k)$ acts on a k-dependent space and presumably some kind of fibered homotopy should be used, but if the flux through the unit cell is $1/q$, q an integer, one can easily transform to a periodic problem on a fixed space.

[6]The topological invariance of the TKNN integers says that they will be independent of precisely which slice in a given direction is chosen.

[7]R. Fox, Ann. Math. 4a, 471 (1968).

[8]To get the proper norm convergence of $[H(k) + E_0]^{-1}$ one puts the *strong* topology on $\mathcal{U}(\mathcal{H})$.

[9]D. Mermin, Rev. Mod. Phys. 51, 591–648 (1979), and J. Math. Phys. (N.Y.) 19, 457 (1978).

[10]The result for $\mathcal{U}(\mathcal{H})$ in the strong topology follows from its contractibility; see J. Dixmier and A. Douady, Bull. Soc. Math. France 91, 251 (1963). The more subtle contractibility in the norm topology is due to N. Kuiper, Topology 3, 19 (1965). In Ref. 3, we present an easy argument that $\pi_j(\mathcal{U}(\mathcal{H})) = 0$.

[11]3 and 4 are chosen just for definiteness sake. A pair of neighboring integers can be chosen. If three bands come together at a point, the sum of all three TKNN integers is preserved, etc.

[12]For example, in the quantized Hall effect, θ could be the strength of the potential.

[13]P. Streda, J. Phys. C 15, L717–L721 (1982).