PHYSICAL REVIEW B VOLUME 37, NUMBER 8 15 MARCH 1988-I

Bargmann's theorem and position-dependent effective mass

Kiyoshi Kawamura* and Ronald A. Brown
School of Mathematics and Physics, Macquarie University, North Ryde, New South Wales, 2113 Australia
(Received 22 May 1987)

The physical significance of Galilean transformations applied to effective-mass (EM) equations for Bloch electrons in Wannier representation is discussed and contrasted with that of Galilean coordinate transformations applied to the free-particle Schrödinger equation. Mass constraints imposed on the latter by Bargmann's (1954) superselection rule do not extend to the EM, and criticisms of the position-dependent EM concept which have invoked Bargmann's theorem are shown to be without foundation. Other criticisms concerning the nonuniqueness and non-Hermiticity of effective Hamiltonians which employ this concept to describe crystals of graded composition are discussed, and it is argued that the problems are associated with the heuristic nature of the virtual-crystal model which is adopted rather than with the position-dependent EM.

I. INTRODUCTION

By means of effective-mass (EM) methods the dynamics of Bloch electrons, subject to externally applied fields or those due to lattice imperfections, may be studied free of complexities due to the periodic lattice potential. In familiar applications of these methods to homogeneous crystals[1-3] the EM parameters are constant over the crystal volume. However, in extending the method to describe crystals which are inhomogeneous due to spatially varying lattice distortions,[4-9] compositions,[10-13] or structure[14,15] one encounters EM parameters which depend on position. The logical self-consistency of this concept has been questioned[16-18] by noting the identical forms of the free-particle Schrödinger equation and the conventional isotropic EM equation appropriate to a homogeneous crystal and then appealing to Bargmann's theorem[19-22] by which the (Newtonian) relativity principle constrains the mass of a free particle to be constant. It is the aim of the present paper to address this criticism and some problems of nonuniqueness and non-Hermiticity of the effective Hamiltonian which have also been noted[15,17,18] in connection with the application of the variable EM concept to semiconductors of graded composition. We are motivated by a belief in both the usefulness and the rigorous justifiability (see Sec. IV) of position-dependent EM formalisms, at least in some instances, and particularly in the contexts of distorted crystals,[4-9] heterostructures,[14,15] and epitaxially laminated superlattices.

The plan of the paper is as follows. In Sec. II we review the principles surrounding the Galilean invariance of the Schrödinger equation (in position representation) to clarify the physical basis and content of Bargmann's theorem, which depends on the equivalence of inertial coordinate systems (having relative speeds $v \ll c$) for the classical and (via Ehrenfest's theorem) the quantum descriptions of nature. It is shown, following Refs. 21 and 22, that this equivalence is inconsistent with a mass spectrum, i.e., that a dependence of the mass parameter on *state*, or wave vector, is not permitted (Bargmann's superselection rule). The derivation presented in Sec. II makes it clear that position independence of mass is essential to the proof of Bargmann's theorem so even if an analogue of Bargmann's theorem were established for the conventional EM equation (with position-independent EM) it would say nothing about the logical consistency of the position-dependent EM concept. It would, however, raise serious problems for band theory as it would imply that *superpositions of states coming from more than one band* were impossible. In view of the doubts about the admissibility of the above concept which were raised by von Roos[16-18] there is a need for clarification and we therefore investigate, in Sec. III, the meaning and implications of applying Galilean transformations to the Wannier-representation EM equation. Following Slater,[2,3] we emphasize there that the spacelike variable appearing therein represents a label which designates the Wannier base states onto which the quantum state of the system may be projected. Although a multitude of (Galilean transformed) labeling schemes may be adopted, there exists no relativity principle by which the form invariance of the EM equations appropriate to different schemes may be asserted. Rather, because the EM equation is derived from the Schrödinger equation, its form is predetermined by the latter and by the specific scheme of continuous labeling adopted for the atomic sites. By investigating the form which the EM equation takes under Galilean label transformations and by modifying the argument of Sec. II we show that there exists no analogue of Bargmann's theorem in the case of effective mass and that therefore neither restrictions on the band dependence of the EM nor, *a fortiori*, on its position dependence, are implied. An alternative view of the label transformation procedure, leading to the same conclusions, is presented in an appendix.

In Sec. IV we consider the questions of uniqueness and Hermiticity of effective Hamiltonians containing position-dependent EM. Emphasis is placed on the need for a physically precise model in achieving the former and on the need to consider the physical interpretation of the "effective wave function" in assessing the latter. These points are illustrated by reference to an effective Hamiltonian appropriate to distorted crystals. Our conclusions are summarized in Sec. V.

II. BARGMANN'S THEOREM

Although the free-particle Schrödinger equation and the effective-mass equation appropriate to a crystalline environment have identical forms (at least for the case of isotropic effective mass) the physical interpretations of their solutions are quite different and this difference is central to our argument. In this section the physical principles underlying Bargmann's theorem are outlined following Refs. 20 and 21, to make the differences between the transformation properties of the quantum theory as encapsulated in Schrödinger's equation and those of effective-mass theory (Sec. III) more evident.

The time evolution of any quantum state $|\Psi\rangle$ of a free particle of mass m is determined by

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = H|\Psi\rangle \ , \tag{2.1}$$

where H is to be understood as the kinetic energy operator $(-\hbar^2\nabla^2/2m)$. (We adopt absolute Newtonian time without further comment, in keeping with the nonrelativistic connotations of the Schrödinger equation and the Galilean transformation). The well-known correspondence (Ehrenfest's theorem) between the quantum and classical descriptions implies that coordinate representations of H and $|\Psi\rangle$ expressed in terms of any inertial frame must yield the same physics, at least for particle speeds and relative frame speeds very much less than that of light (as will henceforth be assumed). Consider a Cartesian reference frame F in which a material particle P has coordinates $(x,y,z)\equiv\mathbf{r}$ and another, F', in which the same particle at the same time has coordinates $(x',y',z')\equiv\mathbf{r}'$, where

$$\mathbf{r}'=\mathbf{r}-\mathbf{v}t+\mathbf{a} \ , \tag{2.2}$$

$\mathbf{a}$ being a constant vector (a_x,a_y,a_z). The state $|\Psi\rangle$ may be described either in terms of its projections $\psi_F(\mathbf{r},t)\equiv\langle\mathbf{r},F|\Psi\rangle$ on the eigenstates $|\mathbf{r},F\rangle$ of position $\mathbf{r}$ as measured in F or in terms of $\psi_{F'}(\mathbf{r},t)\equiv\langle\mathbf{r},F'|\Psi\rangle$ where $|\mathbf{r},F'\rangle$ is an eigenstate of position $\mathbf{r}$ as measured in F'. Note that $\mathbf{r}$ is a dummy variable in these relations and may be replaced by $\mathbf{r}'$ or any other symbol. The equivalence of the two quantum descriptions then demands that

$$\frac{-\hbar^2}{2m}\nabla^2\psi_F(\mathbf{r},t)=i\hbar\frac{\partial}{\partial t}\psi_F(\mathbf{r},t) \tag{2.3a}$$

and

$$\frac{-\hbar^2}{2m}\nabla^2\psi_{F'}(\mathbf{r},t)=i\hbar\frac{\partial}{\partial t}\psi_{F'}(\mathbf{r},t) \ , \tag{2.3b}$$

and furthermore demands that the physical interpretation of the two wave functions be the same. In particular the probability density at P must be the same whether expressed as $|\psi_F(\mathbf{r},t)|^2$ or $|\psi_{F'}(\mathbf{r}-\mathbf{v}t+\mathbf{a},t)|^2$ and this quite generally implies

$$\psi_F(\mathbf{r},t)=e^{ig(\mathbf{r},t)}\psi_{F'}(\mathbf{r}-\mathbf{v}t+\mathbf{a},t) \ , \tag{2.4}$$

where the function g is real. That g must also be *independent* of the state $|\Psi\rangle$ follows from the required invariance of inner products such as $\langle\Phi|\Psi\rangle$ which the theory relates to observable transition probabilities. On substituting (2.4) into (2.3a) it follows[20,21] that the former relationship is consistent with the form invariance of the Schrödinger equation only if

$$g(\mathbf{r},t)=\frac{m}{\hbar}(\mathbf{v}\cdot\mathbf{r}-\tfrac{1}{2}v^2t) \ . \tag{2.5}$$

It should be emphasized that this result is dependent upon the position independence of m.

It is shown in Refs. 21 and 22 that (2.4) and (2.5) generate an accumulated phase change of arbitrary magnitude when a suitable sequence of boost and translation operations is applied, the phase change being proportional to m for a given sequence. This indeterminacy of phase is specifically associated[22] with nonrelativistic aspects, in particular the failure to distinguish between coordinate time and proper time and the artificial separation of energies of excitation from the rest mass. If stationary solutions $\psi^{(j)}(\mathbf{r},t)=\phi_j(\mathbf{r})\exp(-iE_jt/\hbar)$ of (2.3) existed, *in which the parameter m was allowed to assume values m_j dependent on the state of excitation,* then a superposition of such states would exhibit interference effects which depended on the coordinate system chosen. To avoid this manifestly unphysical situation "Bargmann's superselection rule," expressly forbidding such a mass spectrum and restricting the application of the Schrödinger equation to particles of constant mass, is formulated. In view of the questions raised by von Roos[16-18] it is of interest to construct an argument which parallels that of the present section as far as possible, but applying instead to the EM equation. This is done in the following section and it is found to lead to quite different conclusions.

III. THE EFFECTIVE-MASS EQUATION

In the preceding section we emphasized that the derivation of Bargmann's theorem for the free-particle Schrödinger equation depends on two physical principles. Thus the principle of (Newtonian) relativity requires that the wave functions which represent the same free-particle state in coordinate representations corresponding to frames in uniform relative motion should satisfy the same free-particle Schrödinger equation (2.3), while the standard interpretation of the wave function requires in addition that these wave functions should be related as in (2.4). The suggestion that the physical consequence, expressed in the superselection rule for particle mass, should also apply to the effective mass of particles propagating in a crystalline environment has been made by von Roos.[16-18] This suggestion arises because of the identity of form between the usual EM equation (3.14) and the free-particle Schrödinger equation. However its substantiation must rest upon a detailed study of the transformation properties of the EM equation and the physical interpretation of its solutions. In the present section we deduce the forms taken by the EM equation when Galilean transformations are applied to the independent variables and deduce the interpretation which attaches to the "effective wave functions" or "modulating functions." Such transformations correspond to time-dependent labeling of Wannier base states.

To prepare the way for the time-dependent transformation (3.15) we first consider a monatomic lattice whose atoms we suppose to be rigidly displaced from each lattice point by $\boldsymbol{\sigma}$, i.e., they are located at positions

$$\mathbf{R}_n=\mathbf{T}_n+\boldsymbol{\sigma} \ , \tag{3.1}$$

where the lattice translations $\mathbf{T}_n \equiv n_i \mathbf{a}_i$ (summation implied), $\mathbf{a}_i$ being primitive basis vectors and n_i being any integers. For the moment $\boldsymbol{\sigma}$ is a time-independent vector. We define the Wannier functions

$$a_j(\mathbf{r},\mathbf{R})=[\Omega^{1/2}/(2\pi)^{3/2}] \int e^{-i\mathbf{k}\cdot\mathbf{R}} b_j(\mathbf{k},\mathbf{r}) d^3k \ , \tag{3.2}$$

where Ω is the unit-cell volume, $\mathbf{R}$ (for convenience of notation we omit the index n) denotes a member of the set (3.1), $b_j(\mathbf{k},\mathbf{r})$ is a Bloch energy eigenfunction of the crystal Hamiltonian H_c corresponding to band j and wave vector $\mathbf{k}$, and the integral is over the first Brillouin zone. The existence of stationary energy eigenstates implies that the coordinates $(x,y,z)\equiv\mathbf{r}$ are measured in a lattice rest frame. It is also implied, by the constancy (in time) of $\boldsymbol{\sigma}$, that the atomic positions $\mathbf{R}$ are (for the present) measured in such a frame; the latter condition will later be relaxed. When $\boldsymbol{\sigma}=0$, or some other member of the set $\{\mathbf{T}_n\}$, (3.2) are Wannier functions and depend[23] only on the difference $(\mathbf{r}-\mathbf{R})$. For general $\boldsymbol{\sigma}$ this is not the case but the properties of orthogonality

$$\int a_{j'}^*(\mathbf{r},\mathbf{R}') a_j(\mathbf{r},\mathbf{R}) d^3r = \delta_{jj'}\delta_{\mathbf{R}\mathbf{R}'} \ , \tag{3.3}$$

where the integral is over all space, and completeness

$$\sum_{j,\mathbf{R}} a_j^*(\mathbf{r},\mathbf{R}) a_j(\mathbf{r}',\mathbf{R}) = \delta(\mathbf{r}-\mathbf{r}') \tag{3.4}$$

are still readily proved, along the lines of Ref. 23.

In view of (3.4) we may expand an arbitrary one-electron wave function of the crystal as

$$\Psi(\mathbf{r},t)=\sum_{j,\mathbf{R}} C_j(\mathbf{R},t) a_j(\mathbf{r},\mathbf{R}) \ . \tag{3.5}$$

Wave functions (3.5) with components from several bands with different EM are considered since it is for this case (only) that Bargmann's theorem as invoked by von Roos could possibly have interesting, i.e., disturbing, consequences (viz., the existence of an EM-dependent phase factor depending on an arbitrary choice of reference frame) if the theorem applied at all. On substituting in the Schrödinger equation and using (3.3) we find

$$i\hbar\frac{\partial C_j}{\partial t}(\mathbf{R},t)=\sum_{\mathbf{R}'} \varepsilon_j(\mathbf{R},\mathbf{R}') C_j(\mathbf{R}') \ , \tag{3.6}$$

where no sum over j is implied and where

$$\varepsilon_j(\mathbf{R},\mathbf{R}')=\int a_j^*(\mathbf{r},\mathbf{R}) H_c a_j(\mathbf{r},\mathbf{R}') d^3r \ . \tag{3.7}$$

By following a similar procedure to that of Ref. 23, and using (3.1) to relate atomic positions and corresponding lattice translations, we find

$$\varepsilon_j(\mathbf{R},\mathbf{R}')=\varepsilon_j(\mathbf{R}-\mathbf{R}',0)\equiv\varepsilon_j(\mathbf{R}-\mathbf{R}')=\varepsilon_j(\mathbf{T}-\mathbf{T}') \ . \tag{3.8}$$

Now it follows from (3.8) and (3.2) that

$$\varepsilon_j(\mathbf{T})=[\Omega/(2\pi)^3] \int E_j(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{T}} d^3k \tag{3.9}$$

with inverse

$$E_j(\mathbf{k})=\sum_{\mathbf{T}} \varepsilon_j(\mathbf{T}) e^{i\mathbf{T}\cdot\mathbf{k}} \ , \tag{3.10}$$

where $E_j(\mathbf{k})$ is the energy eigenvalue corresponding to $b_j(\mathbf{k},\mathbf{r})$. We may use (3.10) to define the operator $E(-i\nabla_{\mathbf{R}})$ and consider its operation on an infinitely differentiable function $C_j(\mathbf{R})$ of a continuous variable $\mathbf{R}$. By using Taylor's theorem in the form[2,3]

$$\{\exp(\mathbf{T}\cdot\nabla_{\mathbf{R}})\}F(\mathbf{R})=F(\mathbf{R}+\mathbf{T}) \ , \tag{3.11}$$

(3.10) yields

$$\begin{aligned} E_j(-i\nabla_{\mathbf{R}})C_j(\mathbf{R}) &= \sum_{\mathbf{T}} \varepsilon_j(\mathbf{T}) C_j(\mathbf{R}+\mathbf{T}) \\ &\equiv \sum_{\mathbf{R}'} \varepsilon_j(\mathbf{R}'-\mathbf{R}) C_j(\mathbf{R}') \ , \end{aligned} \tag{3.12}$$

where we have used the fact that $(\mathbf{R}+\mathbf{T})$ is a member of the set (3.1). It then follows from (3.12) that solutions of the differential equation

$$i\hbar\frac{\partial}{\partial t} C_j(\mathbf{R},t)=E_j(-i\nabla_{\mathbf{R}})C_j(\mathbf{R},t) \ , \tag{3.13}$$

when evaluated at the discrete points (3.1), yield solutions of the difference equations (3.6). For those solutions $C_j(\mathbf{R})$ which vary slowly on the scale of the lattice constant [this requires that solutions of (3.14) be restricted, *in each band,* to plane waves of small wave vector] we are justified in retaining only the leading order of derivatives on the right-hand side of (3.13) to obtain the EM equation

$$i\hbar\frac{\partial}{\partial t} C_j(\mathbf{R},t)=-\frac{\hbar^2}{2m_j^*}\nabla_{\mathbf{R}}^2 C_j(\mathbf{R},t) \ , \tag{3.14}$$

in which m_j^* is the EM for the jth band and we ignore an unimportant constant term which may be absorbed into the time derivative. *If* the considerations of Sec. II applied to (3.14), unphysical interference effects would be predicted unless the parameters m_j^* were independent not only of wave vector but also of band index—a conclusion which must be ruled out on both theoretical and experimental grounds. In fact, Bargmann's theorem does *not* apply to the EM equation, as we now discuss.

It is clear from Sec. II that Bargmann's theorem is based on the physical requirement that the free-particle Schrödinger equation provide a valid description of appropriate physical phenomena, independent of which of several frames of reference is used to specify particle position coordinates. The formal similarity between (3.14) and the Schrödinger equation has led von Roos[16–18] to conclude that Bargmann's theorem has implications for the band effective mass. To assess this possibility calls for consideration of the physics underlying Galilean transformations in the context of the EM equation. We have been led to examine two different possibilities in this regard. On the one hand we consider a Galilean transformation of all coordinates (including ionic coordinates implied in the one-electron potential) which appear in the one-electron Schrödinger equation. Physically, this corresponds to choosing a reference frame in which the lattice moves uniformly. We then derive the EM equation (A15), by expanding the wave function in terms of Wannier functions centered on the atomic sites of this uniformly moving lattice, and show that its form differs from (3.14). Thus the essential [from the point of view of establishing Bargmann's theorem—cf. (2.3a) and (2.3b)] form invariance of the EM equations corresponding to the two frames is not present and consequently Bargmann's theorem cannot be invoked. The details of the above viewpoint are developed in the Appendix.

In the remainder of this section we take an alternative view which corresponds more closely to that of von Roos,[16–18] who considers Galilean transformations of the variables appearing in (3.14). It is clear from the derivation leading to (3.14) that the variable $\mathbf{R}$ appearing therein represents a set of three numbers which serve to label the atomic sites of the crystal. In the conventional scheme these numbers are time independent and are simply the atomic coordinates as measured in the lattice rest frame (the transition from discrete coordinates to continuous variable having been made by restricting consideration to slowly varying envelope functions). Therefore Galilean transformations applied to a given $\mathbf{R}_{\mathrm{n}}$ yield time-dependent labels

$$\mathbf{R}_{\mathrm{v}}=\mathbf{T}_{\mathrm{n}}-\mathbf{v}t+\boldsymbol{\sigma}-\mathbf{a}\equiv\mathbf{R}_{\mathrm{n}}-\mathbf{v}t+\mathbf{a} \tag{3.15}$$

corresponding to the coordinates of atom $\mathbf{n}$ with respect to a frame which moves uniformly with respect to the lattice rest frame. Although clumsy, such a labeling scheme is legitimate, establishing a one-to-one correspondence between label and atomic site provided the time is also specified; i.e., a clock is indispensable in the interpretation of such a scheme. Bearing in mind the central role of the relativity principle in establishing the invariant form of (2.3a) and (2.3b), which played a crucial part in the proof of Bargmann's theorem, the corresponding physical justification behind von Roos's application of this theorem to the EM rests on showing that the *form of the EM equation is invariant* under different choices of labeling schemes (3.15). We now consider this. We noted earlier that the completeness and orthonormality of the functions (3.2) are retained when $\mathbf{R}$ is replaced by (3.15), allowing the expansion

$$\Psi(\mathbf{r},t)=\sum_{j,\mathbf{R}_{\mathrm{v}}}\tilde{C}_j(\mathbf{R}_{\mathrm{v}},t)a_j(\mathbf{r},R_{\mathrm{v}}) , \tag{3.16}$$

where the sum is over that set $\{\mathbf{R}_{\mathrm{v}}\}$ which is specified at time t by the Galilean transformation (3.15).

If expansions (3.5) and (3.16) are to represent the same state we have, for all t,

$$\sum_{j,\mathbf{R}_{\mathrm{n}}} C_j(\mathbf{R}_{\mathrm{n}},t)a_j(\mathbf{r},\mathbf{R}_{\mathrm{n}}) = \sum_{j,\mathbf{R}_{\mathrm{n}}}\tilde{C}_j(\mathbf{R}_{\mathrm{n}}-\mathbf{v}t+\mathbf{a})a_{\mathrm{j}}(\mathbf{r},\mathbf{R}_{\mathrm{n}}-\mathbf{v}t+\mathbf{a}) . \tag{3.17}$$

On substituting (3.2) and using the orthonormality of the Bloch functions, (3.17) yields, for all t,

$$\sum_{\mathbf{R}_{\mathrm{n}}}\tilde{C}_j(\mathbf{R}_{\mathrm{n}}-\mathbf{v}t+\mathbf{a},t)e^{-ik\cdot(\mathbf{R}_{\mathrm{n}}-\mathbf{v}t+\mathbf{a})} = \sum_{\mathbf{R}_{\mathrm{n}}} C_j(\mathbf{R}_{\mathrm{n}},t)e^{-i\mathbf{k}\cdot\mathbf{R}_{\mathrm{n}}} . \tag{3.18}$$

Further reduction of (3.18) can be achieved for those special combinations of $\mathbf{v}$ and t such that $(\mathbf{v}t-\mathbf{a})$ is some lattice translation $\mathbf{T}_{\mu}$. In these cases one finds, on integration of (3.18) over the first Brillouin zone and using[24]

$$\int d^3k\, e^{i\mathbf{k}\cdot(\mathbf{R}_{\mathrm{n}}-\mathbf{R}_{\mathrm{m}})}=\frac{(2\pi)^3}{\Omega}\delta_{\mathrm{m,n}} , \tag{3.19}$$

that

$$\tilde{C}_j(\mathbf{R}_{\mathrm{m}},t_{\mu})=C_j(\mathbf{R}_{\mathrm{m}},t_{\mu}) \tag{3.20}$$

for all $\mathbf{R}_{\mathrm{m}}$. Here t_{μ} are those (discrete) times for which

$$\mathbf{v}t_{\mu}-\mathbf{a}=\mathbf{T}_{\mu} , \tag{3.21}$$

some lattice translation. Although (3.20) can only be established for discrete $(\mathbf{v},t_{\mu})$ satisfying (3.21), this restriction becomes less and less important as the lattice constant is decreased to reach the continuum approximation or, alternatively, if the functions (3.20) vary slowly on the scale of the (fixed) lattice constant. In any case the establishment of (3.20) for the above special values of $\mathbf{v}t$ shows that no relation of type (2.4), viz.,

$$\tilde{C}_j(\mathbf{R},t)=e^{iG(\mathbf{R},t)}C_j(\mathbf{R},t) \tag{3.22}$$

can apply unless

$$G(\mathbf{R},t_{\mu})=0 \tag{3.23}$$

and this is sufficient to rule out form (2.5); the latter appears to be essential, according to the discussion following that equation, to the proof of phase indeterminacy underlying Bargmann's theorem.

Although nothing further need be said it is useful to derive the EM equation appropriate to the Galilean-transformed labeling scheme (3.15) in order to show where the arguments of Refs. 16 and 17 (in particular, Sec. II of Ref. 17) break down. Substituting (3.16) in the Schrödinger equation one notes that the time dependence imparted through (3.15) appears not only in the coefficients $\tilde{C}$ but also in the Wannier functions. Methods similar to those which led to (3.14) then yield

$$i\hbar\frac{\partial}{\partial t}\tilde{C}_j(\mathbf{R}-\mathbf{v}t+\mathbf{a},t)=[E_j(-i\nabla_{\mathbf{R}})-i\hbar\mathbf{v}\cdot\nabla_{\mathbf{R}}] \times\tilde{C}_j(\mathbf{R}-\mathbf{v}t+\mathbf{a},t) , \tag{3.24}$$

in which we emphasize that *both* arguments on the left are to be differentiated. The drift term on the right arises from the time dependence of the Wannier functions in (3.16). It cancels with an identical term arising from the first argument on the left of (3.24), yielding (for slowly varying solutions)

$$i\hbar \left[\frac{\partial}{\partial t}\tilde{C}_j(\mathbf{R}',t)\right]_{\mathbf{R}'=\mathbf{R}-\mathbf{v}t+\mathbf{a}} = -\frac{\hbar^2}{2m_j^*}[\nabla^2_{\mathbf{R}'}\tilde{C}_j(\mathbf{R}',t)]_{\mathbf{R}'=\mathbf{R}-\mathbf{v}t+\mathbf{a}} , \tag{3.25}$$

in which the time derivative is with respect to the *second argument only*. Bearing in mind (3.14), it follows that (3.25) has solutions (3.20). On the other hand, if a Galilean transformation is applied to (3.14) directly, the drift term appearing on the right of (3.24) does not appear, leading to a different EM equation and different conclusions.[16,17]

To complete the analogy with Sec. II we consider the direct physical interpretation of the modulating functions. It readily follows from the orthonormality of the functions (3.2) that on integrating the absolute square of (3.5) over a volume ΔV (centered on $\mathbf{r}\approx\mathbf{R}$), whose dimensions are large compared to the lattice constant but small compared to the scale on which $C(\mathbf{R})$ varies significantly, that the probability of finding an electron in ΔV is[25]

$$P_{\Delta v}=\frac{\Delta V}{\Omega}\sum_j |C_j(\mathbf{R},t)|^2 \tag{3.26}$$

and

$$P_{\Delta v}=\frac{\Delta V}{\Omega}\sum_j |\tilde{C}_j(\mathbf{R},t)|^2 , \tag{3.27}$$

when calculated from (3.5) and (3.6), respectively. Therefore the invariance of (averaged) charge density between the two labeling schemes implies

$$\sum_j |\tilde{C}_j(\mathbf{R},t)|^2=\sum_j |C_j(\mathbf{R},t)|^2 . \tag{3.28}$$

To make our point we consider a state $|\Psi\rangle$ made of components from a single band. Equation (3.28) then requires [cf. (2.4)]

$$C(\mathbf{R},t)=e^{iG(\mathbf{R},t)}\tilde{C}(\mathbf{R},t) , \tag{3.29}$$

where, as in Sec. II, the real function $G(\mathbf{R},t)$ is determined by the fact that both C and $\tilde{C}$ must satisfy certain differential equations, in this case (3.14) and (3.25), respectively. These imply $G=0$, agreeing with (3.23). The contrast between (3.23) and the analogous expression (2.5) is obvious and clearly negates the suggestion that the EM should be subject to Bargmann's theorem. It is a corollary of this result that a more general solution may be constructed from states of several bands without leading to logical inconsistencies.

IV. POSITION-DEPENDENT EFFECTIVE MASS

The concept of position-dependent EM has been associated with an evident non-Hermiticity and nonuniqueness of the associated effective Hamiltonians.[10–18] Such theories present problems and have been criticized. The cause of these difficulties is not the variable EM concept itself, but rather the unsatisfactory models (e.g., the virtual-crystal model) from which the EM equation has been derived and the associated difficulties of ascribing a precise physical interpretation to its solutions. In contrast to Sec. III, where effective Hamiltonians were rigorously deduced from the Schrödinger equation, the lack of a precise deductive derivation leads inevitably to some arbitrariness and, although the requirement of Hermiticity may be invoked to rule out certain possibilities, there can be no guarantee that the result is unique, correct, or even meaningful. It must also be emphasized, as we discuss in more detail below, that the Hermiticity of an operator cannot be judged without considering the physical interpretation of the functions on which it operates, since this in turn determines the mathematical structure of the integrals representing inner products [see Eq. (4.3)].

Mixed crystals of graded composition present a very difficult problem in which the deficiencies of the virtual-crystal model are to some extent compensated by its heuristic appeal; in any case the absence of alternative strategies makes adoption of the model appear unavoidable. However, to illustrate the points made in the previous paragraph it suffices to consider the example of distorted crystals, in which the concept of position-dependent EM also arises[9,26] (Ref. 9 also discusses earlier formulations[4–8]). In this case the deduction of a unique[24] and Hermitian effective Hamiltonian and the physical interpretation of its solutions both follow rigorously from the Schrödinger equation using methods similar to those of Sec. III. Although all the details need not be presented here, we wish to emphasize the importance of physically interpreting the envelope functions $C(\mathbf{R},t)$ before judging the Hermiticity of the effective Hamiltonian. Thus we observe, following Sec. III of Ref. 9, that we may integrate the absolute square of the total wave function over a small volume ΔV to obtain a probability analogous to (3.26). Now, however, if we denote the elastic dilatation by $\Theta(R)$ the number of atoms in ΔV is inversely proportional to $1+\Theta$, leading to a spatially averaged probability density [see Eq. (19) of Ref. 9]

$$\rho(\mathbf{R},t)=|C(\mathbf{R},t)|^2/[1+\Theta(\mathbf{R})] , \tag{4.1}$$

where, as before, Ω denotes the unit-cell volume of the undeformed crystal. Hence the normalization condition

$$\int |C(\mathbf{R},t)|^2\frac{d^3R}{1+\Theta(\mathbf{R})}=1 \tag{4.2}$$

must be applied to the modulating functions. It follows that the integral defining the inner product of two modulating functions

$$\langle C_\alpha | C_\beta\rangle \equiv \int C_\alpha^*(\mathbf{R},t)C_\beta(\mathbf{R},t)\frac{d^3R}{1+\Theta(\mathbf{R})} \tag{4.3}$$

must contain the same weighting factor and hence that the question of the Hermiticity of the effective Hamiltonian H_e, i.e., the property by which[27]

$$\langle C_\alpha | H_e C_\beta\rangle = \langle C_\beta | H_e C_\alpha\rangle^* \tag{4.4}$$

for all C_α and C_β, cannot be separated from that of physical interpretation of the modulating functions. In the interest of simplicity we have written Eqs. (4.1)–(4.4) in forms appropriate to a one-band model; the many-band extensions are available in Ref. 26.

V. SUMMARY AND CONCLUSIONS

After investigating the theoretical basis of Bargmann's theorem (Sec. II) we pointed out that if this theorem were to apply to the EM equation (3.14) its effect would be to restrict the *band dependence* (not the position dependence) of the EM. We conclude that criticisms[16–18] of the position-dependent EM concept which appeal to Bargmann's theorem cannot be sustained.

In fact, Bargmann's theorem does not apply to the EM equation and no restrictions on the band dependence of the EM are implied. This was established in Sec. III by investigating the physical significance of Galilean transformations in this context and showing that the physical invariance of EM theory when different base-state labeling schemes (related by Galilean transformation) are employed implies that no EM-dependent phase changes, of the type (2.5) which led to Bargmann's theorem, accrue.

In the absence of justifiable criticism the concept of position-dependent EM remains a useful and powerful one with application to a number of problems. However, as emphasized in Sec. IV, it is important that effective Hamiltonians which embody this concept be based upon physically sound models which permit adequate consideration to be given to the direct physical interpretation of the modulating functions which they govern.

ACKNOWLEDGMENT

Both authors wish to acknowledge the assistance of the Australian Academy of Science and the Japan Society for the Promotion of Science and thank their respective host universities [Macquarie University (K.K.) and Hiroshima and Keio Universities (R.A.B)] for their hospitality.

APPENDIX

In the main body of this paper we have attempted to clarify the physical meaning of the Galilean transformations applied by von Roos[16–18] to the spacelike label variable appearing in the EM equation and have shown that no EM version of Bargmann's theorem follows. However, given the invariance of the complete Schrödinger equation under Galilean coordinate (both electronic and ionic) transformations we should also examine what this invariance implies regarding the corresponding EM equations. In this appendix we show, by explicit derivation of the EM equation arising when a coordinate frame other than the crystal rest frame is adopted, that this equation does not possess the form invariance[28] assumed by von Roos[16–18] so that an essential ingredient for the proof of Bargmann's theorem is missing.

We consider the variation of the effective-mass equation due to the Galilean transformation of the coordinate system. We will show that the effective-mass equation is not Galilean invariant.

We begin with the coordinate system F in which ions are at rest. Then, Ψ_F in (2.3a) is expanded in terms of Wannier functions as follows:

$$\Psi_F(\mathbf{r},t;\{\mathbf{R}_\mathbf{m}\})=\sum_{j,\mathbf{n}} C_j(\mathbf{n},t)a_{Fj}(\mathbf{r},\mathbf{R}_\mathbf{n}) . \tag{A1}$$

This is equivalent to (3.5) in the text and the apparent difference comes from the notations which are employed here to facilitate the following discussion. The subscript F is retained on the right-hand side since we wish to also consider another coordinate system F' which moves against F. The same subscript is attached to the Wannier function since this function needs to be defined [see (A9)] according to the coordinate system. In this appendix ions are labeled by n and ionic positions (with respect to F) are denoted by $\mathbf{R}_\mathbf{n}$. The dependence of the one-electron wave function on the position of ions is indicated explicitly on the left-hand side of (A1); this dependence comes through the crystal potential in the Schrödinger equation which is satisfied by Ψ_F.

The definition of the Wannier function (3.2) may be rewritten as

$$a_{Fj}(\mathbf{r},\mathbf{R}_\mathbf{n})=[\Omega^{1/2}/(2\pi)^{3/2}]e^{-i\mathbf{k}\cdot\mathbf{R}_\mathbf{n}}b_{Fj}(\mathbf{k},\mathbf{r};\{\mathbf{R}_\mathbf{m}\})d^3k , \tag{A2}$$

where we again acknowledge that the Bloch function depends on ionic positions through the crystal potential.

Now we go to another coordinate system F' which moves against the rest system with the velocity $\mathbf{v}$. The coordinate $\mathbf{r}'$ in the new frame is related to $\mathbf{r}$ in the old one by (2.2). The wave function in the new coordinate system is expanded in terms of localized Wannier-type functions which will be appropriately defined below [see (A9)];

$$\Psi_{F'}(\mathbf{r}',t';\{\mathbf{R}'_\mathbf{m}\})=\sum_{j,\mathbf{n}} C'_j(\mathbf{n},t')a_{F'j}(\mathbf{r}',\mathbf{R}'_\mathbf{m}) . \tag{A3}$$

The position $\mathbf{R}'_\mathbf{n}$ of the ion with (frame-independent) label $\mathbf{n}$ is now related to $\mathbf{R}_\mathbf{n}$ by

$$\mathbf{R}'_\mathbf{n}=\mathbf{R}_\mathbf{n}-\mathbf{v}t' . \tag{A4}$$

If the description is not to depend on an arbitrary choice of coordinate frame the functions on the left-hand side of (A1) and (A3) must be related by (2.4) in the text.

To arrive at an appropriate definition for the $a_{F'j}$ functions we consider first the time-dependent Bloch function in the rest frame,

$$b_{Fj}(\mathbf{k},\mathbf{r},\mathrm{t};\{\mathbf{R}_\mathbf{m}\})=\exp[-iE_j(\mathbf{k})t/\hbar]b_{Fj}(\mathbf{k},\mathbf{r};\{\mathbf{R}_\mathbf{n}\}) , \tag{A5}$$

where $E_j(\mathbf{k})$ is the eigenvalue corresponding to $b_{Fj}(\mathbf{k},\mathbf{r};\{\mathbf{R}_\mathbf{n}\})$. Since the latter satisfies Schrödinger's equation it is transformed as follows using (2.4) and (2.5):

$$b_{F'j}(\mathbf{k},\mathbf{r}',t';\{\mathbf{R}'_\mathbf{m}\})=e^{-i[m\mathbf{v}\cdot\mathbf{r}'+(m/2)v^2t']/\hbar} \times b_{Fj}(\mathbf{k},\mathbf{r}'+\mathbf{v}t',t';\{\mathbf{R}_\mathbf{m}\}) . \tag{A6}$$

When we substitute (A5) into (A6), we get a factor showing that the new function oscillates with the frequency

$$E_j'(\mathbf{k})/\hbar = E_j(\mathbf{k})/\hbar + mv^2/2\hbar \ , \tag{A7}$$

meaning that the energy eigenvalue is shifted by the amount of kinetic energy associated with the uniform translatory motion. We choose to drop this factor in defining the Galilean transform of the time-independent Bloch function as

$$b_{F'j}(\mathbf{k},\mathbf{r}';\{\mathbf{R}_\mathbf{m}'\}) = e^{-im\mathbf{v}\cdot\mathbf{r}'/\hbar} b_{Fj}(\mathbf{k},\mathbf{r}'+\mathbf{v}t',t';\{\mathbf{R}_\mathbf{m}\}) \ . \tag{A8}$$

This function still depends on time since, in F', the crystal potential changes with time but quantum-mechanical oscillation has been eliminated. Correspondingly we define the Wannier-like function in the new frame as

$$a_{F'j}(\mathbf{r}',\mathbf{R}_\mathbf{n}') = e^{-i(m/\hbar)\mathbf{v}\cdot\mathbf{r}'} a_{Fj}(\mathbf{r}'+\mathbf{v}t',\mathbf{R}_\mathbf{n}) \tag{A9}$$

and readily establish the orthogonality, completeness, and localization of these functions.

It follows from (2.4), (2.5), (3.5), (A3), and (A9) that

$$\sum_{j,\mathbf{n}} C_j'(\mathbf{n},t') a_{Fj}(\mathbf{r}'+\mathbf{v}t',\mathbf{R}_\mathbf{n}) = e^{-i(m/2\hbar)v^2 t'} \sum_{j,\mathbf{n}} C_j(\mathbf{n},t') a_{Fj}(\mathbf{r}'+\mathbf{v}t',\mathbf{R}_\mathbf{n}) \ . \tag{A10}$$

When the orthonormality of the Wannier function is applied, it follows from (A10) that

$$C_j'(\mathbf{n},t') = \exp\left[-i\frac{m}{2\hbar}v^2 t'\right] C_j(\mathbf{n},t') \ , \tag{A11}$$

i.e., that the frequency of oscillation of C_j' is also shifted by the amount corresponding to the kinetic energy of the uniform motion. Therefore, it is evident from (3.6) that $C_j'(\mathbf{n},t')$ obeys the equation

$$i\hbar\frac{\partial}{\partial t} C_j'(\mathbf{n},t) = \sum_{\mathbf{n}'} \frac{m}{2} v^2 \delta_{\mathbf{n}\mathbf{n}'} + \varepsilon_j(\mathbf{n},\mathbf{n}') C_j'(\mathbf{n}',t) \ , \tag{A12}$$

where $\mathbf{n}'$ denotes a typical member of the invariant label set $\{\mathbf{n}\}$.

In (A12) we have derived the equation of motion for the envelope $C_j'(\mathbf{n},t)$ expressed as a function of the labels. Since atoms are moving in the present case (coordinate frame F') this does not represent the time evolution of the amplitude at a fixed point. To get the time evolution at a fixed point, we should replace the first entry of C' with the position $\mathbf{R}_\mathbf{n}'$ which is time dependent when $\mathbf{n}$ is fixed. The time derivatives of these two functions are related by

$$\frac{\partial C'(\mathbf{n},t)}{\partial t} = -\mathbf{v}\cdot\nabla_\mathbf{R} C'(\mathbf{R},t) + \frac{\partial C'(\mathbf{R},t)}{\partial t} \tag{A13}$$

and the substitution of this into (A12) yields

$$\left[i\hbar\frac{\partial}{\partial t} - i\hbar\mathbf{v}\cdot\nabla_\mathbf{R} - \frac{m}{2}v^2\right] C_j'(\mathbf{R},t) = \sum_{\mathbf{R}'} \varepsilon_j(\mathbf{R},\mathbf{R}') C_j'(\mathbf{R}',t) \ , \tag{A14}$$

in which $\mathbf{R}'$ is now a typical member of the set $\{\mathbf{R}_\mathbf{n}\}$, *not* a Galilean transform of $\mathbf{R}$. In the latter equation it is important to distinguish the free-electron mass m from the band effective masses m_j^* which enter the right-hand side via the Fourier transform (3.10), leading to

$$\left[i\hbar\frac{\partial}{\partial t} - i\hbar\mathbf{v}\cdot\nabla_\mathbf{R} - \frac{m}{2}v^2\right] C_j'(\mathbf{R},t) = -\frac{\hbar^2}{2m_j^*}\nabla_\mathbf{R}^2 C_j'(\mathbf{R},t) \ . \tag{A15}$$

On comparing (A15) with (3.14) we observe that the invariance of form which was postulated (or, rather, which is prescribed by the correspondence principle) in the case of (2.3), and which plays a vital role in the proof[21] of Bargmann's theorem, fails to carry over to the EM equations which derive from Galilean-transformed coordinate frames. As a result Bargmann's theorem cannot be established for the EM (and of course, this applies *a fortiori* if the m_j^* depend on position.) The only exception would be if all the m_j^* were equal, in which case (A15) could be reduced to the form (3.14) by a unitary transformation; then the considerations of Sec. II and Ref. 21 would apply but, with all m_j^* equal, would be without physical consequence.

---

[*] Permanent address: Faculty of Science and Technology, Keio University 14-1, Hiyoshi, 3 Chome, Kohokuku Yokohama, 223 Japan.

[1] G. Weinreich, *Solids: Elementary Theory for Advanced Students* (Wiley, New York, 1965), Chap. 9.

[2] J. C. Slater, Phys. Rev. **76**, 1592 (1949).

[3] J. C. Slater, *Insulators, Semiconductors, and Metals*, Vol. 3 of *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1967), Appendix 1.

[4] J. Bardeen and W. Shockley, Phys. Rev. **80**, 72 (1950).

[5] W. A. Harrison, Phys. Rev. **110**, 14 (1958).

[6] M. C. Gutzwiller and R. O. Wells, J. Phys. Chem. Solids **27**, 349 (1966).

[7] K. Kawamura, Z. Phys. B **29**, 101 (1978).

[8] H. Teichler, Phys. Lett. **87A**, 113 (1981).

[9] R. A. Brown, Aust. J. Phys. **36**, 321 (1983).

[10] T. Gora and F. Williams, Phys. Rev. **177**, 1179 (1969).

[11] L. Leibler, Phys. Rev. B **12**, 4443 (1975).

[12] C. M. van Vliet and A. H. Marshak, Phys. Rev. B **26**, 6734 (1982).

[13] R. A. Morrow, Phys. Rev. B **29**, 5962 (1984).

[14] Q.-G. Zhu and H. Kroemer, Phys. Rev. B **27**, 3519 (1983).

[15] R. A. Morrow and K. R. Brownstein, Phys. Rev. B **30**, 678 (1984).

[16] O. von Roos, Appl. Phys. Commun. **2**, 57 (1982).

[17] O. von Roos, Phys. Rev. B **27**, 7547 (1983).

[18] O. von Roos and H. Mavromatis, Phys. Rev. B **31**, 2294 (1985).

[19] V. Bargmann, Ann. Math. **59**, 1 (1954).

[20] J.-M. Levy-Leblond, J. Math. Phys. **4**, 776 (1963).

[21] F. A. Kaempffer, *Concepts in Quantum Mechanics* (Academic, New York, 1965), Appendix 7.

[22] D. M. Greenberger, Am. J. Phys. **47**, 35 (1979).

[23] J. Callaway, *Quantum Theory of the Solid State* (Academic, New York, 1974), Pts. A and B, Sec. 5.1.3.

[24] The effective Hamiltonian of Eq. (30) of Ref. 9 is Hermitian. It is unique only to the extent that terms proportional to distortion gradients are neglected. The level of approximation used in Ref. 9 is rather trivial in this regard. It may be extended as in Ref. 26 to include distortion gradients or higher derivatives, leading to effective Hamiltonians which are Hermitian and unique to a desired order of approximation. Full details of these extensions are planned to be reported in a future publication.

[25] We assume $\Psi(\mathbf{r},t)$ is normalized to unity. The proof of (3.26) and (3.27) depends on the localization of the Wannier functions within distances of order of a lattice constant. In the case of the functions occurring in (3.16) this localization can only be established for the special $(\mathbf{v},t)$ combinations of (3.21). Although (3.27) is only proven for these cases it may be taken to hold more generally when $\tilde{C}$ is sufficiently slowly varying (or in continuum approximation) as discussed following (3.21).

[26] R. A. Brown, in *Dislocations in Solids,* edited by H. Suzuki, T. Ninomiya, K. Sumino, and S. Takeuchi (University of Tokyo Press, Tokyo, 1985), pp. 105–108.

[27] A. Messiah, *Quantum Mechanics,* (North-Holland, Amsterdam, 1961), Vols. 1 and 2, Sec. 5.3.

[28] Nor can such form invariance be restored (simultaneously, for all bands) by a simple transformation, unitary or otherwise, on the label variable.