PHYSICAL REVIEW B VOLUME 32, NUMBER 8 15 OCTOBER 1985

Spatially varying band structures

P. Roblin
Department of Electrical Engineering, The Ohio State University, Columbus, Ohio 43210

M. W. Muller
Department of Electrical Engineering, Washington University, St. Louis, Missouri 63130

(Received 13 May 1985)

Advances in technology have made possible the fabrication of rapidly varying heterostructures which hold the promise of important applications. We develop a set of approximate treatments of electron states in a variety of layered heterostructures. The approximations are all based on the concept of one-band generalized Wannier functions. Following a discussion of the validity of this representation, we apply it to an evaluation of the bound states in a narrow quantum well in GaAs, which clearly demonstrates the mixing of main and satellite valley states as well as the contribution of evanescent states, and of the states of a superlattice in a model structure of up to 20 quantum wells. As a final example we discuss the application of generalized Wannier functions to the matching of electronic states at a heterojunction between two model band structures with different effective masses, and compare the formalism with alternative approaches to this problem.

With the development of special growth techniques such as molecular beam epitaxy (MBE) and metal organic chemical vapor deposition (MOCVD), tailor-made submicron semiconductor heterostructures can now be designed. An entirely new variety of structures, including quantum wells, superlattices, and modulation-doped structures, can be conceived and some have already led to successful device applications. Quantum wells are being investigated in semiconductor lasers[1] and resonant tunneling[2] for submillimeter radiation. Potential applications for superlattices include nonlinear optics in relation to optical switching (bistability),[3] the Bloch oscillator,[4] and the Zener oscillator[5] for submillimeter radiation. An important application of modulation doping[6] is the high electron mobility transistor, a low-noise microwave amplifier,[7] and a candidate for high-speed logic.[8]

These semiconductor heterostructures are made of successive semiconductor crystal layers grown on top of each other. Provided the lattice parameters are closely matched, the lattices of the semiconductors essentially cohere with a minimal perturbation. Due to the special growth techniques used, the spatial variation of the semiconductor materials and/or of the doping can be controlled so as to occur in a few lattice parameters.

The physical properties of the structures of these semiconductors are the object of extensive theoretical and experimental effort with respect to important potential device applications, some of which were mentioned above, but most of which, however, remain unrealized.

A fundamental property on which all these heterostructures rely is the spatial variation of the band gap and the conduction and valence band structures. In this paper, we shall be concerned with the development of a consistent and simplified picture of spatially varying band structures.

Band structures derive from the periodic nature of the crystal potential and are therefore mathematically well defined only for an infinite crystal. However, in practice, the concept of bands holds on a microscopic scale. This is exemplified by the use of band diagrams in classical device theory. The total classical Hamiltonian H can be written as

$$H(\mathbf{k},\mathbf{r})=E(\mathbf{k})-eV(\mathbf{r}) , \tag{1}$$

where $E(\mathbf{k})$ is the conduction or valence band structure, e the electron charge, and $V(\mathbf{r})$ the electrostatic potential. The symbol $\mathbf{r}$ denotes the spatial location and $\mathbf{k}$ the Bloch wave vector or quasimomentum. In a band diagram, usually only the bottom of the conduction band and the top of the valence band are represented, which is equivalent to specifying $\mathbf{k}$.

Equation (1), together with the acceleration theorem,[9] constitute the semiclassical picture on the basis of which one can describe the ballistic motion of an electron or derive (using the Boltzmann transport equation) the standard device equations.[10] In the heterostructures discussed above, the spatial dimensions are made sufficiently small for the quantization of the electron states to be significant. One needs then to revert to the quantum-mechanical form of Eq. (1) where the quasimomentum is replaced by the operator

$$\mathbf{k}=-i\frac{\partial}{\partial\mathbf{r}} , \tag{2}$$

and to solve the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\psi(t,\mathbf{r})=H\psi(t,\mathbf{r}) .$$

Equation (1), together with (2), is often referred to as the Wannier theorem, the derivation of which can be found in textbooks on semiconductor theory.[11] The Wannier theorem relies on two basic assumptions. The first is the one-band approximation which holds if the perturbation potential $V(\mathbf{r})$ is not strong enough for band mixing to occur or interband transition such as tunneling to take place. Secondly, the perturbation potential should vary slowly or smoothly with position so as to constrain the potential overlap matrix to a diagonal representation.

These assumptions are adequate for the classical representation in which the electrons are essentially free and can be represented by a wave packet $\Psi(\mathbf{r})$. For the submicron heterostructures of interest here, these assumptions do not hold, as the potential variations in modulation-doped or in spatially varying band-gap structures are both large and can occur in a few lattice parameters. In this paper we shall propose a concept of generalized band structures for which the Wannier theorem can be generalized and applied to submicron heterostructures.

First, we shall rederive the Wannier theorem in Sec. I, specializing it to one-dimensional structures. Since the submicron heterostructures grown by MBE or MOCVD are layered structures, a one-dimensional picture can be implemented, provided some assumptions are made. Such one-dimensional models are of heuristic interest as their simplicity promotes a better insight, and they are sometimes used in semiquantitative studies.

In Sec. II we discuss the extension of the Wannier theorem, using the generalized Wannier functions, and study in Sec. III two examples of heterostructures. These examples constitute both a test and a demonstration of applications of the generalized Wannier picture. Finally, we present in Sec. IV the application of the generalized Wannier picture to "true" heterostructures and discuss the generalization of the concept of bands to spatially varying band structures.

I. THE ONE-DIMENSIONAL WANNIER PICTURE

In this section we consider the problem of a crystal electron in an external one-dimensional potential U. The total Hamiltonian of the electron is given by

$$H=H_0+U(\hat{\mathbf{d}}\cdot\mathbf{r})\;,$$

where H_0 is the unperturbed crystal Hamiltonian and $U(x)$ is the one-dimensional energy potential varying along an axis (the device axis) represented by the unit vector $\hat{\mathbf{d}}$.

We assume the device axis to be parallel to the lattice vector. We only consider face-centered-cubic semiconductor lattices for which a lattice vector $\mathbf{R}(\mathbf{n})$ can be written using the orthonormal basis of the Bravais lattice as $\mathbf{R}(\mathbf{n})=A\mathbf{n}$, with $\mathbf{n}$ a set of integers (n_1,n_2,n_3), a^* the lattice parameter of the cubic lattice, and A the matrix

$$A=\frac{a^*}{2}\begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix}.$$

We shall denote as a the modulus of the smallest lattice vector $\mathbf{a}$ parallel to $\hat{\mathbf{d}}$; a is therefore the effective lattice parameter along the direction $\hat{\mathbf{d}}$, and any lattice vector parallel to $\hat{\mathbf{d}}$ is written $\mathbf{R}=n\mathbf{a}$, with n an integer.

A reciprocal-lattice vector $\mathbf{K}$ is written in the same basis as $\mathbf{K}(\boldsymbol{l})=B\boldsymbol{l}$, with $\boldsymbol{l}$ a set of integers (l_1,l_2,l_3) and B the matrix

$$B=\frac{2\pi}{a^*}\begin{bmatrix}-1&1&1\\1&-1&1\\1&1&-1\end{bmatrix},$$

so that $AB=2\pi$. It is easy to prove that for cubic lattices there always exists a reciprocal-lattice vector parallel to a given direct-lattice vector. This is equivalent to finding the constant c such that $\mathbf{K}=c\mathbf{R}$. A possible solution is $c=8\pi/a^{*2}$, which leads to

$$\boldsymbol{l}=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}\mathbf{n}\;.$$

We shall denote q as the modulus of the smallest reciprocal-lattice vector $\mathbf{q}$ parallel to $\hat{\mathbf{d}}$ so that any reciprocal vector parallel to $\hat{\mathbf{d}}$ can be written $\mathbf{K}=l\mathbf{q}$, with l an integer. Since the lattice vector and the reciprocal-lattice vector satisfy the relation

$$\mathbf{R}(\mathbf{n})\cdot\mathbf{K}(\boldsymbol{l})=2\pi\mathbf{n}\cdot\boldsymbol{l}\;,$$

we have that $\mathbf{q}$ and $\mathbf{a}$ satisfy

$$\mathbf{a}\cdot\mathbf{q}=aq=2\pi p\;,$$

with p an integer. For the $\langle 100\rangle$ direction $[\mathbf{R}=a(1,0,0)]$, we have $\mathbf{n}=(-1,1,1)$, and the smallest reciprocal vector is given by $\boldsymbol{l}=(0,1,1)$, so that $a=a^*$ and $p=2$.

In the extended zone scheme the Bloch-function solution $\phi(\mathbf{k},\mathbf{r})$ of

$$H_0\phi(\mathbf{k})=E(\mathbf{k})\phi(\mathbf{k}) \tag{3}$$

is a periodic function of $\mathbf{k}$ along the device direction $\hat{\mathbf{d}}$. We can introduce a one-dimensional quasimomentum k along the device direction $\hat{\mathbf{d}}$ defined by

$$\mathbf{k}(k)=\mathbf{k}_\perp+k\hat{\mathbf{d}}\;,$$

with $\mathbf{k}_\perp$ the transverse momentum perpendicular to the direction $\hat{\mathbf{d}}$. It follows that a Bloch function along the direction $\hat{\mathbf{d}}$ can be written as $\phi(k\hat{\mathbf{d}}+\mathbf{k}_\perp,\mathbf{r})=\phi(k)$, which is a periodic function of k with period q: $\phi(k+nq)=\phi(k)$ for n, an integer.

We can now define, along the direction $\hat{\mathbf{d}}$, a one-dimensional Wannier function $w(m,\mathbf{r},\mathbf{k}_\perp)$ as the spatial Fourier coefficient of the Bloch function $\phi(k)$,

$$w(m,r,k_\perp)=(1/\sqrt{q})\int_{-q/2}^{q/2}\phi(k\hat{\mathbf{d}}+\mathbf{k}_\perp,\mathbf{r})\times\exp(-ikm\,2\pi/q)dk\;, \tag{4}$$

and reciprocally we have

$$\phi(k\hat{\mathbf{d}}+\mathbf{k}_\perp,\mathbf{r})=(1/\sqrt{q})\sum_m w(m,\mathbf{r},\mathbf{k}_\perp)\exp(ikm\,2\pi/q)\;.$$

From the orthogonality property of the Bloch functions and from Eq. (3), one easily verifies that these Wannier functions (WF's) form an orthogonal set

$$\langle w(n)\,|\,w(m)\rangle=\int_{-\infty}^{\infty}w(n,\mathbf{r},\mathbf{k}_\perp)^*w(m,\mathbf{r},\mathbf{k}_\perp)dx'=\delta_{nm}\;,$$

with $x'=\hat{\mathbf{d}}\cdot\mathbf{r}$, and satisfy the useful property

$$H_0 w(m,\mathbf{r},\mathbf{k}_\perp)=\sum_n E(n-m)w(n,\mathbf{r},\mathbf{k}_\perp)\,,$$

with

$$E(n)=(1/q)\int_{-q/2}^{q/2} E(\mathbf{k}_\perp+k\hat{\mathbf{d}})\exp(-ikn\,2\pi/q)dk\,.$$

The $E(n)$ are recognized as the Fourier coefficients of the band structure along the line $\mathbf{k}(k)$ parallel to the device axis.

For a sufficiently smooth and weak one-dimensional potential $U(\hat{\mathbf{d}}\cdot\mathbf{r})$, no band mixing is expected and the transverse momentum is conserved. One can then expand the electron state in Wannier functions of the band considered with an envelope $f(n,t)$,

$$\Psi=\sum_n f(n,t)w(n,\mathbf{r},\mathbf{k}_\perp)\,. \tag{5}$$

The weight function $|f(n,t)|^2$ is the probability of finding the electron at the lattice site n at time t. Replacing Ψ in the Schrödinger equation

$$H\Psi=i\hbar\frac{\partial}{\partial t}\Psi \tag{6}$$

by the expansion of Eq. (5), multiplying Eq. (6) by $w(n)^*$, and integrating over x', one obtains the Wannier recurrence equation

$$\sum_m [E(n-m)+U(n,m)]f(m,t)=i\hbar\frac{\partial}{\partial t}f(n,t)\,, \tag{7}$$

where $U(n,m)$ is the matrix element given by

$$\begin{aligned}U(n,m)&=\langle w(n)\,|\,U\,|\,w(m)\rangle\\&=\int_{-\infty}^{\infty} w(n,\mathbf{r},\mathbf{k}_\perp)^* U(\hat{\mathbf{d}}\cdot\mathbf{r})w(m,\mathbf{r},\mathbf{k}_\perp)dx'\,.\end{aligned} \tag{8}$$

The one-band Wannier recurrence equation derived holds for sufficiently small perturbation potentials U. For Eq. (7) to be solved, both the band-structure Fourier coefficients and the matrix elements $U(n,m)$ are required.

As expressed by Eq. (8), the matrix elements are evaluated from the Wannier functions. Techniques for the evaluation of the Wannier functions have been developed by Kohn.[12] The evaluation of the WF for silicon has been recently reported by Kane and Kane;[13] however, the WF derived is a linear superposition of the WF's of the four valence bands. The evaluation of realistic WF's is a complicated matter, which we would like to avoid in the simple picture we intend to develop.

Some direct approximations can be derived from the properties of the WF's. In the case of simple bands, it is possible to select the phase of the Bloch functions so that the WF's are exponentially localized.[14] From both the orthogonality and the tight localization of the WF's, it follows that the matrix element of a smooth potential is accurately represented by the sampling of the potential at the lattice sites:

$$U(n,m)=U(x'=na/p)\delta_{nm}\,. \tag{9}$$

We shall discuss in Sec. III the extent to which these approximations can be used for sharply varying potentials.

Finally, an important property of the one-dimensional WF's which we did not address so far is their invariance under translation from one lattice site to another. Following the approach of Wannier,[15] we first write

$$\Psi(\mathbf{k},\mathbf{r}-\mathbf{R})=(1/\sqrt{q})\sum_m \exp(ikma/p)w(m,\mathbf{r}-\mathbf{R},\mathbf{k}_\perp)\,, \tag{10}$$

which using the translation property of the Bloch function, can be written

$$\Psi(\mathbf{k},\mathbf{r}-\mathbf{R})=(1/\sqrt{q})\sum_m \exp(-i\mathbf{k}\cdot\mathbf{R})\exp(ikma/p)\times w(m,\mathbf{r},\mathbf{k}_\perp)\,.$$

For $\mathbf{R}=n\mathbf{a}$, we have

$$\mathbf{k}\cdot\mathbf{R}=(k\hat{\mathbf{d}}+\mathbf{k}_\perp)\cdot\mathbf{R}=k\hat{\mathbf{d}}\cdot\mathbf{R}=kna\,,$$

and changing the index of summation we obtain

$$\Psi(\mathbf{k},\mathbf{r}-\mathbf{R})=(1/\sqrt{q})\sum_m \exp(ikma/p)w(m+np,\mathbf{r},\mathbf{k}_\perp)\,. \tag{11}$$

Identifying the coefficients in the series (10) and (11), we finally get

$$w(m,\mathbf{r}-n\mathbf{a},\mathbf{k}_\perp)=w(m+np,\mathbf{r},\mathbf{k}_\perp)\,.$$

Considering the $\langle 100\rangle$ direction for which $p=2$, it follows that there exist two sets of WF's, each of which can be generated from two generic WF's by a lattice translation

$$w(2n,\mathbf{r})=w(0,\mathbf{r}-n\mathbf{a}),\quad w(1+2n,\mathbf{r})=w(1,\mathbf{r}-n\mathbf{a})\,.$$

The number of WF's for a given direction is therefore given by p for the face-centered-cubic lattice considered. For our purpose it is convenient to picture these WF's as located at p sites along the lattice vector $\mathbf{a}$ associated with the direction $\hat{\mathbf{d}}$. We can then write these WF's as

$$w(n,\mathbf{r},\mathbf{k}_\perp)=w(x'-na/p,\mathbf{r}_\perp,\mathbf{k}_\perp)=w(x'-na/p)=w(n)\,,$$

where a is the effective lattice parameter for the direction $\hat{\mathbf{d}}$. The one-dimensional WF we have introduced is in fact a hybrid entity, since it corresponds to the Bloch state $\mathbf{k}_\perp$ for the transverse coordinates $\mathbf{r}_\perp$. The properties of the WF's of one-dimensional lattices cannot be assumed to apply directly to the one-dimensional WF we have defined. Therefore, problems concerning the reality of the WF and the simultaneous convergence of the WF's remain to be addressed. These studies are not trivial matters for three-dimensional crystals,[14] in particular for lattices without a center of inversion.

In our simplified treatment we shall assume the picture developed to be truly one-dimensional. The reality and exponential localization of the WF's in one-dimensional structures are discussed by Kohn.[16] More recent work by Zak[17] raises the one-dimensional Wannier picture to the status of a well-defined representation in quantum mechanics and introduces the canonical Wannier function with minimal position uncertainty and located in the zero cell of the Bravais lattice.

Zak also demonstrates that the position operator has a discrete spectrum. This implies that the applied potentials are transformed by the one-band approximation into steplike potentials. This is due to the inability of the one-band approximation to deal with the perturbation of the inner crystal potential. However, as we shall see in the next section, most of the band mixing can be formally accounted for with the use of the generalized Wannier function. For stronger potentials it is further necessary to account for the polarization of the semiconductor.[18]

II. THE GENERALIZED WANNIER PICTURE

In this section we shall be concerned only with abrupt modulation-doped heterostructures or models of abrupt heterostructures for which the band gap is spatially varying while the band structure is assumed to be conserved.

The one-dimensional Wannier picture developed in the previous section is capable of handling fast spatial variation of the potential since the nondiagonal term of the matrix potential $V(n,m)$ can be accounted for. However, the formalism developed relies on the one-band approximation which strictly holds only for small potential perturbations. Its use for submicron heterostructures seems unjustifiable. Indeed, the band-edge variation in these structures is large and band mixing is expected.

One should then resort to a multiband expansion, and the envelope equation (7) now reads

$$i\hbar\frac{\partial}{\partial t}f(n,b)=\sum_{m}\left[E(n-m,b)f(m,b)+\sum_{b'}\langle w(n,b)\,|\,U(x)\,|\,w(m,b')\rangle f(m,b')\right],$$

where b is the band index.

It is, however, possible to avoid a multiband expansion if one introduces the concept of generalized bands. As was remarked by Zak,[17] "a band is a surprisingly stable entity under external perturbations." This might seem to be a paradoxical statement coming from Zak, who has been very critical of the use of the one-band approximation in general,[19] and in particular for the Stark effect.[20] In fact, the existence of bands under external perturbations does not rely on the one-band approximations. A band of energy is essentially a quasicontinuum of energy states. Such an entity seems to survive numerous mistreatments such as high impurity concentrations, alloying with random distribution of two atoms on a given sublattice in III-V compounds, even for abrupt heterojunctions. However, surface states or band tailing or even alloy and roughness scattering effects occur, and it might be more accurate to refer to these bands as generalized bands. Such bands in most cases can only be represented mathematically as resulting from the mixing of the bands of the unperturbed crystals. This is similar to the expansion of the pseudo-wave-function in terms of plane waves in the calculation of band structures using the pseudopotential method; the bands evaluated result from the mixing of the bands of the empty lattice.

A one-band formalism can then be developed if one uses the exact Wannier functions instead of the Wannier functions of the unperturbed lattice. It remains to be demonstrated that such Wannier functions exist. The Wannier functions as we introduced them [see Eq. (4)] are defined as Fourier coefficients of the Bloch functions which are periodic in k space. The quasimomentum is defined by the translation operator[15] only for periodical potentials. At a heterojunction, the periodicity is broken and no quasimomentum, Bloch functions, or Wannier functions can be defined, but one can attempt to generalize the entities, as the periodicity is essentially broken only at the junction. The most obvious approach is to generalize the Wannier functions which are strongly localized on each lattice site and therefore will be strongly perturbed only at the junction itself.

The existence of generalized Wannier functions (GWF's) has been theoretically established by Kohn and Onffroy[21] for one-dimensional structures. These Wannier functions can be labeled by their generalized band index; they account, however, for the presence of interface states in the forbidden band gap.[22] We shall denote these generalized Wannier functions as $a(n,x)$. They span the same function space as the eigenfunctions of the perturbed lattice for the same generalized band.

The index n will be taken here as the site index, although depending on the potential well there may be more than one or even no GWF's associated with the interface sites. The GWF's are localized about the lattice site (or the periodic potential minima). Rehr and Kohn have shown[23] that they can be exponentially localized. An essential property of the GWF's is that they form an orthonormal set

$$\langle a(n)\,|\,a(m)\rangle=\int_{-\infty}^{\infty}a(n,x)a(m,x)dx=\delta_{nm}\ .$$

The eigenfunctions are linear combinations of $a(n,x)$,

$$\Psi(x,E)=\sum_{n}f(n,E)a(n,x)\ ,$$

with $f(n,E)$ the solution of

$$Ef(n,E)=\sum_{m}H(n,m)f(m,E)\ , \tag{12}$$

where $H(n,m)$ is the Hamiltonian matrix element given by

$$H(n,m)=\langle a(n)\,|\,H\,|\,a(m)\rangle\ .$$

Here H is the total Hamiltonian of the heterojunction given by

$$H=H_0+U(x)\ ,$$

with H_0 the unperturbed crystal potential and $U(x)$ the heterojunction effective potential.

Equation (12) is similar to Eq. (7), except that we use the GWF instead of the unperturbed Wannier functions. The knowledge of the GWF weight $f(n,E)$ is sufficient in most applications and Eq. (12) can be solved if the matrix element $H(n,m)$ is known.

The derivation of the matrix element is the critical part of this method. One approach consists of first deriving the GWF's at each site. An important property of the GWF's demonstrated by Kohn and Onffroy states that for lattice sites far from the interface the GWF's approach the unperturbed WF exponentially in n,

$$a(n,x) \simeq w(x-na) \text{ for } |n| \gg 0 ,$$

assuming the interface at $n=0$. Only a few GWF's are found to be perturbed or to differ from the WF, as was reported by Gay and Smith.[22] It follows that a few lattice parameters away from the interface the Wannier picture developed in Sec. I holds and the only remaining unknown terms are the interface Hamiltonian matrix terms.

A variational method for the derivation of the GWF has been proposed by Kohn and Onffroy for one-dimensional potentials and was tested by Gay and Smith.[22] A three-dimensional treatment is required for the heterostructures of interest here. The derivation of the Wannier functions in three dimensions being quite involved,[13] the derivation of the more complicated GWF will not be considered as a practical approach. For our present purposes, we shall use for the Hamiltonian matrix simple approximations derived from the general properties of the Wannier functions discussed in Secs. I and II.

The essential result brought about by the GWF formalism is that as long as a generalized band of energy can be assumed, it is possible to lump all the band-mixing effects into the interface Hamiltonian matrix elements, whereas the wave function $f(n)$ remains given by an equivalent Hamiltonian recurrence equation.

III. TWO SIMPLE EXAMPLES

We shall now consider two simple applications of the Wannier picture. The first consists of the evaluation of the resonant states in a square quantum well 0.5 eV deep using GaAs band structure. One expects to observe the interaction of the upper-valley and central-valley states of GaAs for such a deep well.[24]

These band-structure data, kindly furnished by Professor Karl Hess of the University of Illinois, were given at a set of sampled points in the Brillouin zone. Using the sampling theorem, we reconstituted the band structure. The band structure was in addition fitted at Γ to $0.068m_0$. This consequently slightly altered the smoothness of the upper valley. The resulting band structure of GaAs along Δ is shown in Fig. 1. The amplitudes of the Fourier coefficients of the GaAs band structure along Δ are plotted in Fig. 2. We present in Fig. 3 the locus of the amplitude of the roots $\exp[jk_i(E)a]$ that solve the equation $E(k_i)=E$, where $E(k)$ is the band structure along Δ. One obtains 20 roots. Roots 1 and 2 correspond to Bloch wave functions of the central valley at Γ. Roots 3 and 4 correspond to the Bloch waves of the upper valley centered on X. The unexpected behavior at X of roots 5, 6, 7, and 8 is due to the fact that the upper-valley minimum does not occur exactly at X. The shift of the X minimum

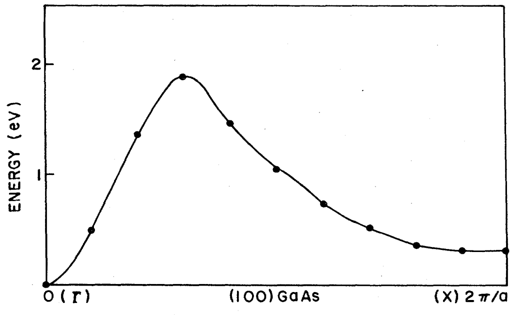

FIG. 1. GaAs band structure along Δ (dots: original sampled data; solid curve: reconstructed band structure with effective mass fitted at Γ to $0.068m$).

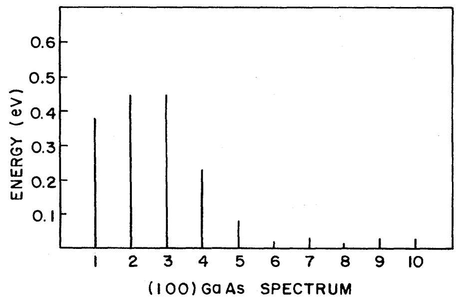

FIG. 2. Fourier coefficients of the Γ to X band structure of Fig. 1.

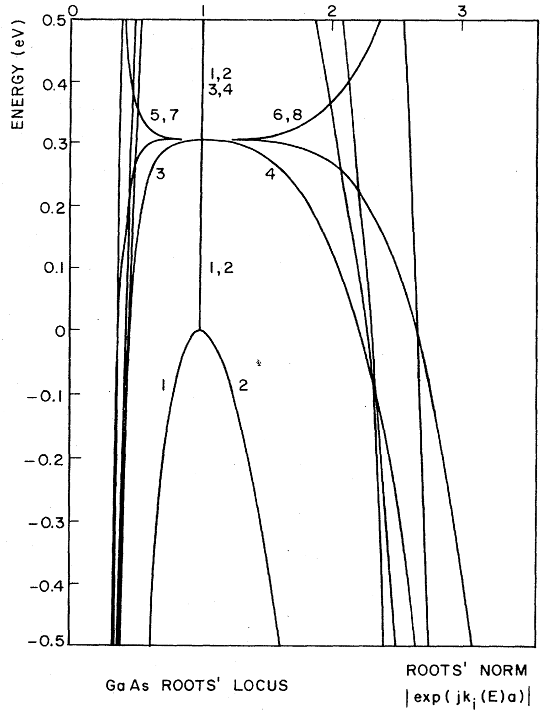

FIG. 3. Locus of the roots of GaAs band structure of Fig. 1 as a function of the energy.

is an artifact of the spectrum truncation. Since the band structure is slightly distorted, we do not expect these roots to contribute much. There are finally 12 remaining roots with a very high damping rate. These roots contribute only within the matching area. One observes in Fig. 3 the evanescent Bloch waves being transformed into propagating Bloch waves at Γ (0 eV) and X (0.3 eV).

The Wannier picture calls for the matrix element of the quantum well potential. For trial purposes we evaluated numerically (with the Gauss-Legendre integration technique) its matrix element using the WF's, the weakly localized WF's of the empty lattice

$$w(x-ma/p)=\left[\frac{p}{a}\right]^{1/2}\frac{\sin[(px/a-m)\pi]}{[(px/a-m)\pi]}\ .$$

It is most enlightening to consider the diagonal terms plotted in Fig. 4 (with an arbitrary energy scale). One observes that these diagonal terms approximately follow the assumed square potential, even for such a narrow well. The energy difference (a few percent) is distributed in the nondiagonal terms and contributes somewhat to the matching. Using the complete matrix element, we found that sufficiently accurate results could be generated by directly reducing the matrix element to diagonal terms corresponding to the sampling of the assumed potential at the lattice site [see Eq. (9)].

We have developed an algorithm to solve the eigenvalue problem and have evaluated the locus of the resonant energy level of the square quantum well of depth 0.5 eV, as a function of the well width L. The result is plotted in Fig. 5 (open circles) and compared with the energy-level locus (solid curve) of the effective-mass approximation. It should be pointed out that in our picture the well width can only be varied by increments of a half-lattice parameter for the 100 direction.

We label the first, resonant state 1. For a wide potential well, both the WF and the continuum methods yield the same results, since the resonant level lies at low energy where the effective mass is quite accurate. For a narrower well, one observes a divergence of the two energy plots as the effective-mass approximation becomes inadequate at higher energy.

For a well smaller than five lattice parameters, the first resonant state occurs above 0.3 eV. This resonant state results, then, from a superposition of the central-valley Bloch wave and the upper-valley Bloch wave. Indeed, one observes the interference of the energy locus (labeled a) of the first resonant level of the upper valley with the energy locus 1 of the central valley. This destroys the smooth variation of these energy loci as they interact. The same phenomenon occurs also as the second and third resonant-level loci (denoted, respectively, 2 and 3) cross the X relative minimum (0.3 eV). A more thorough discussion of this interference effect has been recently reported by Chang and Ting.[24]

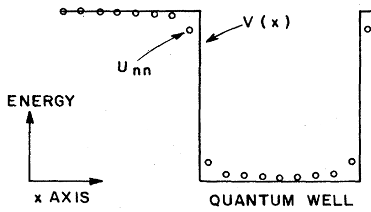

FIG. 4. Diagonal matrix element of a quantum well $U(x)$, nine lattice parameters wide.

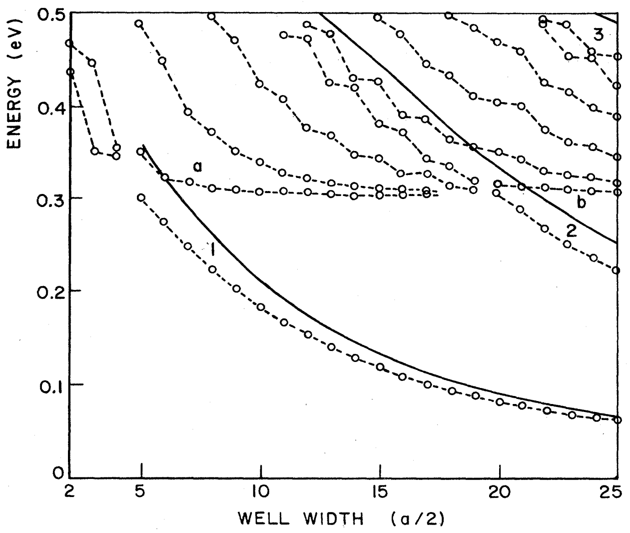

FIG. 5. Energy levels of a GaAs quantum well, as a function of the well width. The well depth is 0.5 eV.

The phenomenon just described constitutes an example of what may be called a full band-structure effect which can be handled by the generalized Wannier picture. The algorithm developed can be used to study any type of variation of the band gap. As a second example, we consider the simple case of a superlattice made of square wells such as in the Kronig-Penney problem.[25] We assume a simple tight-binding band structure $E(k)=A(1-\cos ka)$ with the effective mass at $k=0$ selected to be $0.12m_0$. The well depth is 0.2 eV, the width and separation of the wells is chosen to be $10a$ so that the superlattice parameter is $20a$. This leads to two resonant levels in an individual well. In Fig. 6 we present the result for, successively, 1, 2, and 20 wells. The discrete band structure of the 20-well superlattice is plotted in Fig. 7 (open circles). This is to be compared to the band structure (solid curve) of the Kronig-Penney model in the effective-mass approximation.

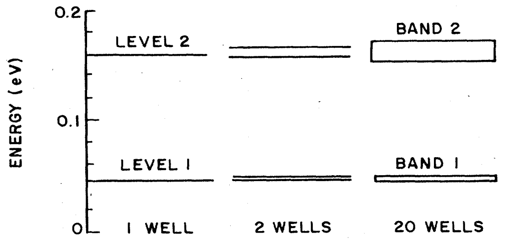

FIG. 6. Superlattice formation.

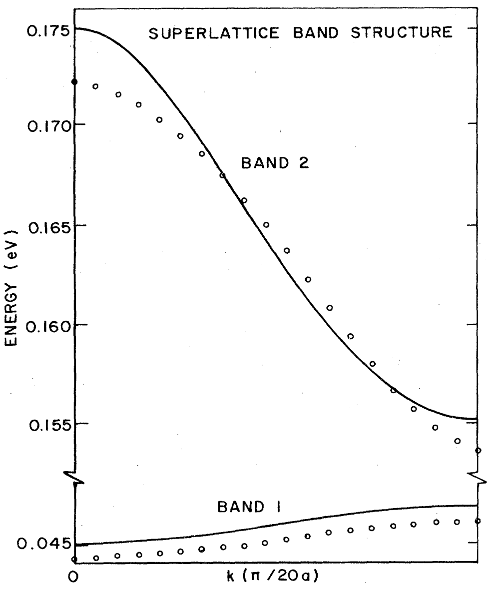

FIG. 7. Superlattice band structure (open circles: tight-binding superlattice of 20 quantum wells; solid curve: Kronig-Penney model).

The low-lying bands (band 1) have a similar shape but are separated by a shift of 1 meV. This small shift originates from the use of the approximate diagonal matrix element for the superlattice potential.

The upper bands (band 2), however, differ appreciably. Part of the difference is due to the deviation from parabolicity of the cosine band and results in a much higher conduction-band maximum for the effective-mass approximation. The remaining difference is essentially due to the finite character of the superlattice composed of only 20 wells.

More sophisticated structures can be studied with the algorithm developed, but this was not our purpose, as we intended to test the generalized Wannier picture using simple examples for which closed-form solutions are available.

IV. SPATIALLY VARYING BAND STRUCTURES

The Wannier picture developed above and its generalization so far do not apply to band structures varying with space. Let us point out right away that the idea of spatially varying band structures is an ill-defined concept, which can be made more rigorous within the framework of the generalized Wannier picture.

A one-dimensional approach will break down for the proper description of a spatially varying band structure, since the variation in the band structure occurs in the three-dimensional k space and therefore requires a three-dimensional matching of the Wannier envelope.

We shall, however, for simplicity partially retain the one-dimensional picture by assuming the spatially varying band structure to be given by

$$E(\mathbf{k},x)=E(k,x)+\frac{\hbar^2 k_\perp^2}{2m^*} ,$$

where m^* is a transverse effective mass which is not varying in space. Such a band structure does not correspond to any realizable interface but constitutes a convenient model. In this model we use the one-dimensional GWF formalism to solve for the eigenstates of this structure. A set of orthogonal GWF's noted $a(n,x)$ can be defined as in Sec. II. These GWF's can be exponentially localized at each lattice site n. For the case of an abrupt heterojunction, the GWF's $a(n,x)$ approach exponentially with n the WF's of semiconductors 1 and 2, for lattice sites far from the interface (at $n=0$). A few lattice sites away from the interface, inside semiconductor 1, a GWF closely resembles the WF of the unperturbed semiconductor 1; the orthogonality with the GWF's of semiconductor 2 across the interface essentially affects only the exponentially decaying tail of this GWF. Finally, one is again led to the generalized one-band formalism of Sec. II, with the GWF envelope given by the one-dimensional Wannier recurrence equation (7).

The derivation of the Hamiltonian matrix elements $H(n,m)$ is again critical to the analysis. We favor a direct evaluation of these elements, which remains to be devised. In its absence, alternative approaches can be conceived, and we shall demonstrate such an approach for a simple system.

First consider the trial Hamiltonian

$$H(n,m)=E(n-m,n)+U(n,m) ,$$

where $E(n-m,n)$ is the band-structure Fourier coefficient which is assumed to vary in space with the lattice site n. Such a Hamiltonian, which might be proposed for a smoothly spatially-varying band structure, is not correct as the Hamiltonian is no longer Hermitian:

$$H(n,m)\neq H^*(m,n) .$$

This arises as a consequence of using the intrinsically ill-defined concept of spatially dependent band structures. The GWF formalism is capable of accounting for the breakdown of the band structure. The Hermiticity is indeed compatible with

$$E(m-n,n)\neq E(n-m,n) ;$$

it follows that the related band structure at the lattice site n,

$$E(k,n)=\sum_l E(l,n)e^{ikla} ,$$

is now imaginary and therefore, as expected, has lost its usual physical meaning.

For an abrupt heterojunction, the use of GWF formalism is reduced to the derivation of matching rules. Indeed, away from the interface the Wannier picture of Sec. I holds and the only unknown terms are the interface Hamiltonian element $H(n,m)$.

Consider the tight-binding band structures given by

$$E_1(k,x)=A-A\cos(ka),\quad E_2(k,x)=B-B\cos(ka) ,$$

where the amplitudes A and B are, respectively, related to their effective mass by

$$A=\frac{\hbar^2}{m_1a^2}\ \text{and}\ B=\frac{\hbar^2}{m_2a^2}\ .$$

For simplicity we evaluate the matrix element $H(n,m)$ using the unperturbed WF's on both sides of the junction. In the tight-binding approximation, the only remaining unknown is then the matrix element $H(i,i+1)=-C/2$, assuming the interface is located somewhere between the lattice site i and $i+1$. The resulting Hamiltonian is shown around the interface in Fig. 8 for $U(n,m)=0$.

The evaluation of C depends now on the particular matching theory upon which we choose to rely. There has been much published recently on the matching of wave functions across a heterojunction. The GWF picture proposes a general method which enables us to evaluate the wave function in heterostructures but calls for the knowledge of $H(n,m)$. In its absence we intend first to relate this method to other matching theories.

Popular techniques are the effective-mass matching (see Ref. 26 and the discussion by Kroemer and Qi-Gao Zhu,[27,28] matching developed from the ($\mathbf{k}\cdot\mathbf{p}$) Kane model by Bastard,[29] and White, Margues, and Sham,[30,31] and from tight-binding arguments by White, Margues, and Sham[30] and Ando and Mori.[32] A review of the different methods is briefly given in White, Margues, and Sham.[30] We intend to give a more complete discussion of the matching problem in a forthcoming paper.

It follows from the work of Zhu and Kroemer[28] that the effective-mass matching is related for small energies to the geometrical average for the "ideal" heterostructure (type I, Ref. 28) considered here,

$$C=(AB)^{1/2}\ .$$

From Bastard's work,[29] we have C given by the average

$$C=(A+B)/2\ .$$

Both expressions actually lead to the same results when the variation of effective mass is small across the heterojunction. In Fig. 9 we compare the effective-mass matching with the geometrical average matching for a quantum well. This quantum well is made of a layer of $Ga_{0.47}In_{0.53}As$ sandwiched between $Al_{0.48}In_{0.52}As$. The effective masses are, respectively, taken as 0.041 and $0.075m_0$ for GaInAs and AlInAs. These effective masses should be corrected for narrow wells, as the nonparabolicity is no longer negligible (see Welch, Wicks, and Eastman[33]). The discrete-energy locus is seen to agree with the continuous-energy locus based on the effective-mass formalism.

$\xrightarrow{\quad} n'$ $\quad H_{nn'}$

$n\downarrow$

-A/2	A	-A/2	0	0	0	0	0
0	-A/2	A	-A/2	0	0	0	0
0	0	-A/2	A	-C/2	0	0	0
0	0	0	-C/2	$B-V_0$	-B/2	0	0
0	0	0	0	-B/2	$B-V_0$	-B/2	0
0	0	0	0	0	-B/2	$B-V_0$	-B/2

FIG. 8. Tight-binding Hamiltonian.

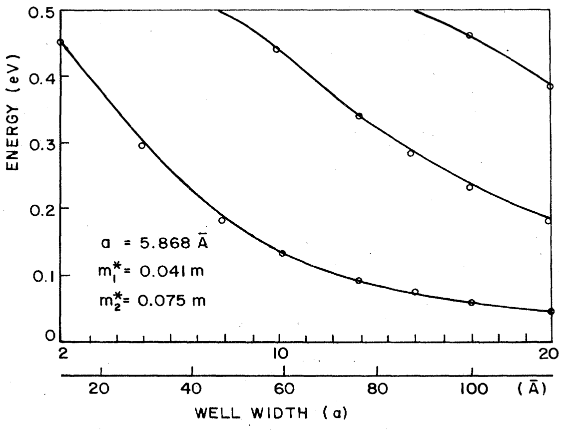

FIG. 9. Energy locus as a function of well width for $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.5}As$ quantum well (open circles: results using Wannier algorithm with the geometrical average matching; solid curve: results using the effective mass matching). The well depth is 0.5 eV.

This result was established within the tight-binding approximation. A more refined scheme will extend the connection rules to include higher-order overlaps.

V. PERSPECTIVE

We have introduced a generalized one-dimensional Wannier picture. This formalism enables us to handle sharp variations in potential as long as the potential strength is sufficiently moderate for the concept of a generalized band to hold. The associated band-mixing effects are then lumped into a generalized one-band Hamiltonian matrix. In the absence of a knowledge of the GWF, one can often rely in practice on the direct sampling of the potential at each lattice site.

The generalized one-band picture applies directly to modulation-doped structures and can be used as a model of heterostructures by treating the band-gap variation as an effective potential. Another application not covered here and for which this model was in fact developed, concerned the study of devices such as the Zener superlattice oscillator, for which the periodicity in $\mathbf{k}$ space of the subband together with its spatial variation are fundamental to the device operation.[5] Finally, practical derivation of the energy levels and eigenfunctions can be performed for arbitrary potentials or heterostructure configurations using the same algorithm.

The one-dimensional generalized Wannier formalism can be used to model true heterostructures, although a rigorous treatment requires a three-dimensional GWF analysis. In the one-dimensional picture, the matching of the wave function across the interface is determined by the interface Hamiltonian matrix elements.

The derivation of Wannier functions is in general a complex problem due to their orthogonality and the fact that they are not eigenfunctions of the Hamiltonian. The best procedure would be to evaluate the generalized Fourier coefficients of the generalized band structure directly, just as the band structure is evaluated directly in pseudopotential calculations using variational techniques.

An alternative approach is to derive the interface Hamiltonian matrix directly from the enforcement of physical constraints. Such an approach was undertaken by Zhu and Kroemer[28] and Ando and Mori[32] for maximized matching (ideal heterojunction) using current continuity. We shall report, in a forthcoming paper, on the extension of these matching theories to a higher energy range (non-effective-mass case) and for the inclusion of higher-order overlap (non-tight-binding cases) for both type I and type II semiconductors (see Ref. 28).

The concept of generalized bands described here is largely intuitive and has been used so far as such. The introduction of the generalized Wannier functions brings support to the concept. One interesting feature of the generalized Wannier picture is that it accounts consistently for the concept of spatially varying band structure. The generalized Wannier picture appears then as an intermediate method situated between the direct study at the lattice level [e.g., linear combination of atomic orbitals (LCAO) techniques] and the techniques based on the effective-mass approximation.

We believe that useful insights can be drawn from this ability of the GWF formalism to handle the concept of band structure, through the use of the band-structure Fourier coefficients together with their spatial dependence on the lattice sites. The band structure appears as a physical entity established in a few lattice parameters. We found that five to ten lattice sites were sufficient for a close representation of a band structure along $\langle 100 \rangle$. This is in qualitative agreement with the Wannier functions evaluated by Kane and Kane.[13] Energy bands calculated from these Wannier functions were accurate to 0.1 eV when the Wannier-function overlaps were evaluated to twelfth neighbors (459 individual bonds). The spatial establishment of a small effective mass requires a larger number of sites since a larger number of Wannier-function overlaps or band-structure Fourier coefficients is required to fit a small effective mass while accurately reproducing the remainder of the band structure.

ACKNOWLEDGMENTS

A portion of this research was carried out at Washington University (St. Louis, MO) and was sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, U.S. Air Force, under Grant No. AFOSR82-0231.

---

[1] N. Holonyak, Jr., R. M. Kolbas, R. D. Dupuis, and P. D. Dapkus, IEEE J. Quantum Electron. **QE-16**, 170 (1980).

[2] T. C. L. Sollner, W. D. Goodhue, P. E. Tannenwa, and C. D. Parker, Appl. Phys. Lett. **43**, 588 (1983).

[3] D. A. B. Miller, D. S. Chemla, D. J. Eilenber, and P. W. Smith, Appl. Phys. Lett. **41**, 679 (1982).

[4] L. Esaki and R. Tsu, IBM J. Res. Dev. Jan., p. 61 (1970).

[5] P. Roblin, Ph.D. thesis, Washington University, 1984.

[6] H. L. Stormer, J. Phys. Soc. Jpn. **49**, 1013 (1980).

[7] N. T. Linh, M. Laviron, P. Delescluse, P. N. Tung, D. Delagebeaudeuf, F. Diamond, and J. Chevrier, in Proceedings of the 9th IEEE Cornell Biennial Conference, 1983 (unpublished).

[8] P. M. Solomon and Hadis Morkoç, IEEE Trans. Electron Devices **ED-31**, No. 8 (1984).

[9] J. Callaway, *Quantum Theory of the Solid State B* (Academic, New York, 1974).

[10] J. E. Carroll, *Hot Electron Microwave Generators* (American Elsevier, New York, 1970).

[11] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, Mass., 1979).

[12] W. Kohn, Phys. Rev. B **7**, 4388 (1972).

[13] E. O. Kane and A. B. Kane, Phys. Rev. B **17**, 2691 (1978).

[14] J. Des Cloizeaux, Phys. Rev. **135**, A698 (1964).

[15] G. H. Wannier, Rev. Mod. Phys. **34**, 645 (1962).

[16] W. Kohn, Phys. Rev. **115**, 809 (1959).

[17] J. Zak, Phys. Rev. B **20**, 2228 (1979).

[18] P. Roblin and M. W. Muller, J. Phys. C **16**, 4547 (1983).

[19] J. Zak, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1972).

[20] J. Avron, L. Gunther, and J. Zak, Solid State Commun. **17**, 189 (1975).

[21] W. Kohn and J. R. Onffroy, Phys. Rev. B **8**, 2485 (1973).

[22] J. G. Gay and J. R. Smith, Phys. Rev. B **11**, 4906 (1975).

[23] J. J. Rehr and W. Kohn, Phys. Rev. B **10**, 448 (1974).

[24] Yia-Chung Chang and D. Z.-Y. Ting, J. Vac. Sci. Technol. B **1**, 435 (1983).

[25] R. de L. Kronig and W. G. Penney, Proc. R. Soc. London **130**, 499 (1930).

[26] R. A. Smith, *Wave Mechanics of Crystalline Solids* (Wiley, New York, 1961), Chap. 4.

[27] H. Kroemer and Qi-Gao Zhu, J. Vac. Sci. Technol. **21**, 551 (1982).

[28] Qi-Gao Zhu and Herbert Kroemer, Phys. Rev. B **27**, 3519 (1983).

[29] G. Bastard, Phys. Rev. B **24**, 5693 (1981).

[30] S. R. White, G. E. Margues, and L. J. Sham, J. Vac. Sci. Technol. **21**, 544 (1982).

[31] S. R. White and L. J. Sham, Phys. Rev. Lett. **47**, 879 (1981).

[32] T. Ando and S. Mori, Surf. Sci. **113**, 124 (1982).

[33] D. F. Welch, G. W. Wicks, and L. F. Eastman, J. Appl. Phys. **55**, 3176 (1984).