Theory of the effective Hamiltonian for degenerate bands in an electric field

REVIEW ARTICLE

# Theory of the effective Hamiltonian for degenerate bands in an electric field

Bradley A Foreman

Department of Physics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

**Abstract.** Recent experiments have generated a renewed interest in the properties of the degenerate valence states in semiconductors under the influence of a uniform external electric field. In response, a number of authors have proposed that the standard Luttinger–Kohn effective-mass Hamiltonian should be modified to include the energy of interaction between the electric field and the dipole matrix elements of the relevant zone-centre Bloch functions. This article examines these proposals by comparing the proposed dipole interaction with rigorous derivations of the field-dependent Hamiltonian in the Bloch and Luttinger–Kohn representations. It is shown that the dipole matrix element of a unit cell is not a suitable foundation for a Hamiltonian because it depends on the choice of unit cell, which is arbitrary. Moreover, the correct Luttinger–Kohn Hamiltonian has no wave-vector-independent matrix elements that are linear in the applied field $\mathcal{E}$, except for small terms of relativistic origin. Therefore, the proposed modifications to the Luttinger–Kohn theory are not valid. The correct form of the Luttinger–Kohn Hamiltonian is derived here through terms of order $k^3$, $k\mathcal{E}$, and $\mathcal{E}^2$, along with the momentum matrix to first order in $k$ and $\mathcal{E}$. In addition, the recent discovery of field-induced mixing at the Brillouin zone boundary in the Luttinger–Kohn representation is studied in detail.

## 1. Introduction

The electronic properties of a periodic crystal in a uniform electric field have been of interest throughout the history of solid-state physics. Although conceptually simple, the problem is subtle mathematically because the scalar potential of the field is unbounded and non-periodic. Consequently, this problem has been studied and debated since the advent of quantum mechanics (Bloch 1928, Jones and Zener 1934, Zener 1934, Houston 1940) through to the present day. Perhaps the best-known controversy is a long-running argument over the existence of Wannier–Stark localization (Wannier 1960, 1962, Zak 1972, Nenciu 1991), which was only resolved in modern times (Mendez *et al* 1988, Voisin *et al* 1988).

Research in this area has proceeded primarily from the standpoint of 'effective-Hamiltonian' theory (Wannier 1960, 1962, Blount 1962, Zak 1972, Nenciu 1991, Callaway 1991), in which the electron is viewed as having a kinetic energy determined by the dispersion relation of the unperturbed crystal, and a potential energy due to the field alone. This is an extremely useful but approximate way of describing the problem. The critical question is, to what extent must this picture be modified to account for the mixing of different Bloch bands caused by the electric field?

The vast majority of research on this topic has focused on the properties of non-degenerate bands. This problem is sufficiently intricate to generate a wealth of interesting physics, and there are undoubtedly many avenues yet to be explored. However, in the 1990s, interest began

to shift towards the additional difficulties posed by the presence of degeneracies in the energy band structure. A major stimulus was the discovery of the quantum-well Pockels effect by Kwok *et al* (1992). In this effect, the optical properties of a (001) GaAs/AlAs quantum well in a [001] electric field were found to exhibit strong in-plane polarization anisotropy (for other material systems see Krebs *et al* (1998) and Platonov *et al* (1999)). The anisotropy arises from a reduction in crystal symmetry caused by the electric field and by the change in material properties at a heterojunction. This broken symmetry leads to a mixing of the degenerate valence-band states that is not included in the effective-mass approximation of Luttinger and Kohn (1955), which is the most widely used theoretical tool for the study of valence electrons in semiconductors.

Much recent work has therefore been devoted to finding ways of incorporating such mixing into effective-mass theory. Valence-band mixing at a heterojunction is now fairly well understood from both the phenomenological (Aleĭner and Ivchenko 1992, Ivchenko *et al* 1993, 1996) and microscopic perspectives (Foreman 1996, 1998, Burt and Foreman 1999). The mechanisms by which this mixing leads to optical anisotropy are also well established (Krebs and Voisin 1996, Magri and Ossicini 1998, Ivchenko *et al* 1998).

However, the contribution from the electric field has proved more troublesome. Although the basic tools needed to solve this problem are already available in the literature, there are many conceptual traps for the unwary. Several different Hamiltonians have been proposed in recent years, some of which contradict the results obtained by earlier authors. The objective of the present article is to examine these proposals at the most fundamental level, and thereby to resolve the current state of confusion in the literature.

An electric field can break the symmetry of a crystal in two ways: by shifting the equilibrium positions of the atomic nuclei, and by polarizing the electrons. The former contribution may be treated using deformation-potential theory (Bir and Pikus 1974, Blacha *et al* 1984), which is not a subject of controversy. The present article therefore focuses exclusively on electronic polarization, assuming throughout that the nuclei are fixed in position at their zero-field equilibrium locations.

Three major lines of approach have been used to derive models for field-induced valence-band mixing. The simplest and (in recent years) most widely used approach is to construct a Hamiltonian from matrix elements of the dipole operator as follows. An electron of charge $-e$ in a uniform electric field $\mathcal{E}$ has potential energy

$$U(\boldsymbol{x}) = e\mathcal{E} \cdot \boldsymbol{x} \tag{1.1}$$

where $\boldsymbol{x}$ is the electron coordinate. Now since $\boldsymbol{k} \cdot \boldsymbol{p}$ theory (Kane 1982) uses zone-centre Bloch functions as basis states, it seems reasonable to expect that the potential (1.1) would give rise to Hamiltonian matrix elements of the form

$$e\mathcal{E} \cdot \boldsymbol{x}_{nn'} \tag{1.2}$$

where $\boldsymbol{x}_{nn'}$ is a matrix element of $\boldsymbol{x}$ connecting two zone-centre Bloch functions (see equation (2.10) below). This procedure has been used by many authors, both for the construction of effective-mass Hamiltonians (Aspnes 1975, Lommer *et al* 1985, 1988, Zhu and Chang 1994, Khurgin 1994, Krebs and Voisin 1996, Khurgin and Voisin 1997, 1998) and for the calculation of interband optical transition rates (Khurgin (1988); Wang (1989), p 603; Weisbuch and Vinter (1991), p 58; Singh (1993), p 154; Haug and Koch (1994), p 71; Chuang (1995), p 355; Szmulowicz (1995); Fiore *et al* (1995); Rosencher *et al* (1996); Khurgin (1999)).

The difficulty with this procedure is that the matrix element $\boldsymbol{x}_{nn'}$ is ill defined (Burt 1993). The nature of the problem is illustrated in figure 1, which shows that the dipole moment of a unit cell is not independent of the choice of unit cell (Ashcroft and Mermin (1976), p 554). In an infinite lattice, this choice is completely arbitrary; hence, the matrix elements $\boldsymbol{x}_{nn'}$ are
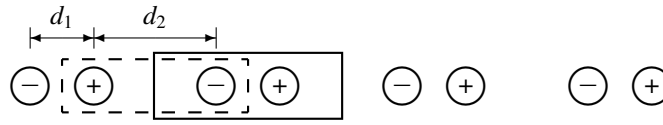
**Figure 1.** A portion of an infinite lattice of point charges, showing two possible choices of unit cell. If the magnitude of each charge is $q$, the dipole moment of the cell indicated by the solid lines is $+qd_1$, whereas that of the cell indicated by the dashed lines is $-qd_2$.

physically meaningless. This problem is well known in the theory of polarization (Martin 1974, Tagantsev 1987, 1991, Resta 1994). A more detailed discussion of the dipole matrix element is presented in section 2.

The source of this problem is that, although zone-centre Bloch functions are complete with respect to *periodic* functions (Luttinger and Kohn 1955), they cannot represent non-periodic functions such as $x$. The situation is identical to that found in Fourier representation theory: although a Fourier integral can represent a non-periodic function, a Fourier series cannot. The solution to this problem is clear: to represent the coordinate $x$, one must use a complete basis such as the Bloch (Adams 1953, Blount 1962) or Luttinger–Kohn (1955) representation.

The properties of $x$ in the Bloch representation are examined in section 3. The main focus of this section is a difficulty that was mentioned frequently in the early literature but is appreciated less widely today. Namely, the matrix elements of $x$ in the Bloch representation are singular functions of $k$ at a degeneracy. This means that the Bloch representation is not well suited for the study of degenerate bands in an electric field.

The Luttinger–Kohn representation, on the other hand, is ideal for this problem because its basis states are analytic in $k$ everywhere, including at degeneracies. The bulk of this paper consists of a study of the Luttinger–Kohn representation in sections 4 and 5. An apparent oversight in Luttinger and Kohn's (1955) analysis of the coordinate operator, recently discovered by Burt (1999), is discussed in section 4. Section 5 presents a derivation of the effective-mass Hamiltonian for degenerate or quasi-degenerate bands in an electric field (Foreman 2000b).

The results derived in this paper are discussed and compared with other approaches in the literature in section 6. Section 7 provides a summary of the main conclusions.

## 2. The dipole matrix element of a unit cell

As mentioned previously, the dipole matrix element of a unit cell does not provide a suitable means for defining a Hamiltonian because it is inherently arbitrary. This topic is explored below, beginning with a definition of the fundamental quantities in $k \cdot p$ theory (Luttinger and Kohn 1955, Blount 1962, Kane 1982).

The Pauli Hamiltonian for an electron in an ideal periodic crystal is

$$H_0 = \frac{p^2}{2m} + V(x) + \frac{\hbar}{4m^2c^2}(\boldsymbol{\sigma} \times \boldsymbol{\nabla}V) \cdot \boldsymbol{p}. \tag{2.1}$$

Here $\boldsymbol{p} = -i\hbar\boldsymbol{\nabla}$ is the canonical momentum operator, $m$ is the mass of a free electron, $V$ is the potential energy ($V(x) = V(x + R)$, where $R$ is a Bravais lattice vector), $c$ is the speed of light, and the components of $\boldsymbol{\sigma}$ are the Pauli spin matrices. In addition to the canonical momentum $\boldsymbol{p}$, it is convenient to define the kinematic (or mechanical) momentum $\boldsymbol{\pi}$ as follows:

$$\boldsymbol{\pi} \equiv \frac{m}{i\hbar}[x, H_0] = \boldsymbol{p} + \frac{\hbar}{4mc^2}(\boldsymbol{\sigma} \times \boldsymbol{\nabla}V). \tag{2.2}$$

The eigenfunctions of $H_0$ are the Bloch functions $\psi_{nk}(x)$:

$$H_0\psi_{nk}(x) = E_n(k)\psi_{nk}(x) \tag{2.3}$$

in which

$$\psi_{nk}(x) = (2\pi)^{-3/2}e^{ik\cdot x}u_{nk}(x). \tag{2.4}$$

Here $k$ is the crystal momentum, $n$ labels the independent solutions of (2.3), $E_n(k)$ is the energy eigenvalue, and $u_{nk}(x) = u_{nk}(x + R)$ is periodic. Both $\psi_{nk}$ and $u_{nk}$ are spinors. The latter are normalized over the volume of any unit cell $\Omega$ (not necessarily a primitive cell):

$$\frac{1}{\Omega}\int_\Omega u_{nk}^\dagger(x)u_{n'k}(x)\,\mathrm{d}^3x = \delta_{nn'} \tag{2.5}$$

and they satisfy the eigenvalue equation (Blount 1962)

$$H_0(k)u_{nk}(x) = E_n(k)u_{nk}(x) \tag{2.6}$$

in which

$$H_0(k) = e^{-ik\cdot x}H_0e^{ik\cdot x} = H_0 + \frac{\hbar}{m}k\cdot\pi + \frac{\hbar^2k^2}{2m} \tag{2.7}$$

is the $k\cdot p$ Hamiltonian. In $k\cdot p$ theory, one defines matrix elements of momentum as follows (Luttinger and Kohn 1955):

$$p_{nn'}(k) = \frac{(2\pi)^3}{\Omega}\int_\Omega \psi_{nk}^\dagger(x)p\psi_{n'k}(x)\,\mathrm{d}^3x = \frac{1}{\Omega}\int_\Omega u_{nk}^\dagger(x)pu_{n'k}(x)\,\mathrm{d}^3x + \hbar k\delta_{nn'} \tag{2.8}$$

$$\pi_{nn'}(k) = \frac{1}{\Omega}\int_\Omega u_{nk}^\dagger(x)\pi u_{n'k}(x)\,\mathrm{d}^3x + \hbar k\delta_{nn'}. \tag{2.9}$$

For consistency, the coordinate matrix element must be defined likewise:

$$x_{nn'}(k) = \frac{1}{\Omega}\int_\Omega u_{nk}^\dagger(x)xu_{n'k}(x)\,\mathrm{d}^3x. \tag{2.10}$$

It should be apparent after a moment's reflection that equation (2.10) depends on the choice of unit cell $\Omega$, because $x$ is not a periodic function. For example, consider the following pseudopotential wave functions (Yu and Cardona 1997, Foreman 1998) for the $\Gamma_{15}$ valence-band states in a zinc-blende crystal:

$$u_{X0} = \sqrt{8}\left(\alpha\sin\frac{2\pi x}{a}\cos\frac{2\pi y}{a}\cos\frac{2\pi z}{a} + \beta\cos\frac{2\pi x}{a}\sin\frac{2\pi y}{a}\sin\frac{2\pi z}{a} + \frac{\gamma}{2}\sin\frac{4\pi x}{a}\right) \tag{2.11}$$

$$u_{Y0} = \sqrt{8}\left(\alpha\cos\frac{2\pi x}{a}\sin\frac{2\pi y}{a}\cos\frac{2\pi z}{a} + \beta\sin\frac{2\pi x}{a}\cos\frac{2\pi y}{a}\sin\frac{2\pi z}{a} + \frac{\gamma}{2}\sin\frac{4\pi y}{a}\right). \tag{2.12}$$

Here $a$ is the lattice parameter, and the coefficients $\alpha$, $\beta$, and $\gamma$ are real, with $\alpha^2 + \beta^2 + \gamma^2 = 1$. These three coefficients are the dominant terms in the $\Gamma_{15}$ valence-band wave function (for example, in GaAs, the complete wave function has $\alpha^2 + \beta^2 + \gamma^2 \simeq 0.96$ (Foreman 1998)). If the unit cell is chosen to be a conventional cube of volume $a^3$ centred on the point $(x_0, y_0, z_0)$, the coordinate matrix element (2.10) is easily shown to be

$$x_{XY}(0) = -\hat{z}\alpha\beta\frac{a}{2\pi}\cos\frac{4\pi z_0}{a}. \tag{2.13}$$

Thus, the matrix element $x_{nn'}(k)$ is meaningless, since it depends essentially on the location of the unit cell, which is arbitrary.

Most authors, however, do not attempt to evaluate $x_{nn'}(k)$ directly from the definition (2.10). Instead, they invoke the well-known textbook relationship (Shankar (1994), p 503) between the matrix elements of coordinate and momentum:

$$x_{nn'}(k) \overset{?}{=} \frac{\hbar}{\mathrm{i}m} \frac{\pi_{nn'}(k)}{E_n(k) - E_{n'}(k)}. \tag{2.14}$$

The difficulty with this approach is that equation (2.14) cannot possibly be valid, since the left-hand side depends on the choice of unit cell $\Omega$, whereas the right-hand side is independent of this choice (Burt 1993). The correct relation for the matrix elements of the cell-periodic functions $u_{nk}(x)$ is (Yafet 1957)

$$\pi_{nn'}(k) = \frac{\mathrm{i}m}{\hbar}[E_n(k) - E_{n'}(k)]x_{nn'}(k) + \frac{m}{\Omega} \int_S x[\hat{n} \cdot j_{nn'}(k, x)] \, \mathrm{d}S \tag{2.15}$$

where $\hat{n}$ is the outward normal to the surface $S$ enclosing the volume $\Omega$, and

$$j_{nn'}(k, x) = \frac{1}{2m}\{u_{nk}^\dagger[(\pi + \hbar k)u_{n'k}] + [(\pi + \hbar k)u_{nk}]^\dagger u_{n'k}\} \tag{2.16}$$

is a generalized current density. Equation (2.15) is derived from Green's theorem and the commutator (2.2). In the textbook derivation, one implicitly assumes that the current goes to zero sufficiently rapidly at infinity for the surface integral to vanish. This assumption does not hold for the periodic wave functions considered here. Note that both terms on the right-hand side of (2.15) are cell-dependent quantities; only their sum is cell independent. Equation (2.15) is thus the direct analogue of the expression given by Martin (1974) for the polarization of a unit cell.

The dipole matrix element of a unit cell is sometimes introduced as the starting point of a calculation, but it is also often introduced as a quantity derived from a more fundamental calculation of the dipole matrix element of a macroscopic crystal. The transition from the macroscopic crystal to the unit cell is achieved via a theorem that is presented in virtually all textbooks on semiconductor physics (see, for example, Bastard (1988), p 54; Weisbuch and Vinter (1991), p 58; Corzine *et al* (1993), p 30; Chuang (1995), p 353). The theorem is applicable to integrals of the form $\int A(x)B(x) \, \mathrm{d}^3x$, where $B(x)$ is a periodic function. It states that if the function $A(x)$ is 'slowly varying'—a vaguely defined concept—then, to some undefined degree of approximation, one can express this integral as a product of two integrals; i.e.,

$$\int A(x)B(x) \, \mathrm{d}^3x = \left[\int A(x) \, \mathrm{d}^3x\right]\left[\frac{1}{\Omega} \int_\Omega B(x) \, \mathrm{d}^3x\right]. \tag{2.17}$$

This theorem may be developed in more precise terms by working with the Fourier representations of $A$ and $B$:

$$A(x) = (2\pi)^{-3/2} \int A(q)\mathrm{e}^{\mathrm{i}q \cdot x} \, \mathrm{d}^3q$$
$$A(q) = (2\pi)^{-3/2} \int A(x)\mathrm{e}^{-\mathrm{i}q \cdot x} \, \mathrm{d}^3x \tag{2.18}$$

$$B(x) = \sum_G B_G \mathrm{e}^{\mathrm{i}G \cdot x}$$
$$B_G = \frac{1}{\Omega} \int_\Omega B(x)\mathrm{e}^{-\mathrm{i}G \cdot x} \, \mathrm{d}^3x \tag{2.19}$$

where $G$ is a reciprocal-lattice vector. The integral on the left-hand side of (2.17) takes a particularly simple form in reciprocal space:

$$\int A(x)B(x)\,\mathrm{d}^3x = \sum_G \int A(x)B_G\mathrm{e}^{\mathrm{i}G\cdot x}\,\mathrm{d}^3x = (2\pi)^{3/2}\sum_G A(-G)B_G \tag{2.20}$$

which is just a special case of Parseval's theorem. As a corollary to (2.20), if $A(x)$ happens to be slowly varying in the sense that $A(G) = 0$ for all $G \neq 0$, then the textbook theorem (2.17) is *exact*.

This derivation of (2.17) emphasizes the fundamental importance of $B(x)$ being periodic (if it is not periodic, it does not possess a Fourier series). The significance of this criterion is sometimes forgotten, and $B(x)$ is viewed as 'rapidly varying' rather than periodic. This has led some authors (Khurgin (1988); Wang (1989), p 603; Weisbuch and Vinter (1991), p 58; Rosencher *et al* (1996); Krebs and Voisin (1996)) to include the potential energy $e\mathcal{E} \cdot x$ in $B(x)$, since if $u_{nk}^\dagger(x)u_{n'k}(x)$ is 'rapidly varying,' $xu_{nk}^\dagger(x)u_{n'k}(x)$ is also. However, it is not periodic; therefore, terms such as $e\mathcal{E} \cdot x$ can only appear in $A(x)$. With this restriction, the meaningless quantity $x_{nn'}(k)$ does not arise in applications of theorem (2.20).

## 3. The Bloch representation

The reason why $x_{nn'}(k)$ is not physically meaningful is that the cell-periodic functions $u_{nk}(x)$ do not form a basis for non-periodic functions such as $x$. For any given $k$, the set $\{u_{nk}(x)\}$ is complete with respect to cell-periodic functions (Luttinger and Kohn (1955); see equation (4.5) below), but $x$ lies outside this function space. The Bloch functions $\psi_{nk}(x)$, on the other hand, form a complete set even for non-periodic functions. This section reviews the properties of $x$ in the Bloch representation, paying particular attention to the behaviour near a degeneracy.

### 3.1. Matrix elements of the coordinate operator

For wave vectors inside the first Brillouin zone, the set $\{\psi_{nk}(x)\}$ is orthonormal:

$$\int \psi_{nk}^\dagger(x)\psi_{n'k'}(x)\,\mathrm{d}^3x = \sum_G \delta(k - k' + G)\delta_{nn'}. \tag{3.1}$$

This may be derived from Parseval's theorem (2.20), the orthogonality relation (2.5), and the periodicity condition $\psi_{nk}(x) = \psi_{n,k+G}(x)$. If one takes the gradient of equation (3.1) with respect to $k$ or $k'$, one obtains the following expression for the matrix elements of $x$ (Jones and Zener 1934, Adams 1953, Blount 1962, Lax 1974, Lifshitz and Pitaevskiǐ 1980):

$$\int \psi_{nk}^\dagger(x)x\psi_{n'k'}(x)\,\mathrm{d}^3x = \sum_G [\mathrm{i}\,\nabla_k\delta(k - k' + G)\delta_{nn'} + \xi_{nn'}(k)\delta(k - k' + G)]. \tag{3.2}$$

Here

$$\xi_{nn'}(k) = \frac{1}{\Omega}\int_\Omega u_{nk}^\dagger(x)[\mathrm{i}\,\nabla_k u_{n'k}(x)]\,\mathrm{d}^3x \tag{3.3}$$

is a matrix element of the 'crystal coordinate' operator $\xi \equiv \mathrm{i}\,\nabla_k$ (Adams 1953) connecting two cell-periodic functions. The sum over $G$ in equation (3.2) is customarily omitted (i.e., only the $G = 0$ term is retained). However, it is shown in section 4.4 that this sum is necessary for a correct treatment of points in the immediate vicinity of the Brillouin zone boundary.

The first term in equation (3.2) is the crystal coordinate operator, which is diagonal in the band index $n$; this is the only term retained in the usual 'effective-mass' approximation. The matrix $\xi_{nn'}(k)$ describes the interband mixing generated by $x$, which is neglected in the

effective-mass approximation. It should be emphasized at the outset that the crystal coordinate matrix (3.3) is qualitatively different from the coordinate matrix (2.10) (Yafet (1957); Blount (1962), p 317; Zak (1991)), since (3.3) is independent of the choice of unit cell. In fact, $\xi_{nn'}(k)$ has many of the properties one might naively expect $x_{nn'}(k)$ to have.

The matrix $\xi_{nn'}(k)$ is Hermitian, as can be seen by applying the operator $\xi$ to the orthogonality relation (2.5), which gives

$$\xi_{nn'}(k) = \xi^*_{n'n}(k). \tag{3.4}$$

It is also periodic: $\xi_{nn'}(k) = \xi_{nn'}(k+G)$. This follows from the periodicity of $\psi_{nk}$, which may be expressed as $u_{n,k+G}(x) = \mathrm{e}^{-\mathrm{i}G\cdot x}u_{nk}(x)$. The relation between $\xi_{nn'}(k)$ and $\pi_{nn'}(k)$ can be found by applying $\xi$ to the equation

$$\frac{1}{\Omega}\int_\Omega u^\dagger_{nk}(x)H_0(k)u_{n'k}(x)\,\mathrm{d}^3x = E_n(k)\delta_{nn'} \tag{3.5}$$

where $H_0(k)$ is defined in (2.7). This gives

$$\pi_{nn'}(k) = \begin{cases} (m/\hbar)\,\nabla_k E_n(k) & n = n' \\ (\mathrm{i}m/\hbar)[E_n(k) - E_{n'}(k)]\xi_{nn'}(k) & n \neq n'. \end{cases} \tag{3.6}$$

Hence, for $E_n(k) \neq E_{n'}(k)$, we have (Adams 1952)

$$\xi_{nn'}(k) = \frac{\hbar}{\mathrm{i}m}\frac{\pi_{nn'}(k)}{E_n(k) - E_{n'}(k)} \tag{3.7}$$

which is identical in form to the invalid relation (2.14). Comparing equations (3.7) and (2.15), we see that $\xi_{nn'}(k) - x_{nn'}(k)$ is proportional to the surface integral in equation (2.15) (Yafet 1957). There is in general no choice of unit cell such that $\xi_{nn'}(k) = x_{nn'}(k)$ (Blount (1962), p 317; Martin (1974)).

## 3.2. Singularities in $\xi_{nn'}(k)$ at a degeneracy

It is now apparent that many of the difficulties encountered in section 2 can be resolved simply by working in the Bloch representation and replacing $x_{nn'}(k)$ with $\xi_{nn'}(k)$. However, there is a disturbing aspect to equation (3.7): in the limit as $k$ approaches a degeneracy, if $\pi_{nn'}(k)$ does not tend to zero, or if $E_n(k)$ approaches $E_{n'}(k)$ faster than $\pi_{nn'}(k)$ tends to zero, then $\xi_{nn'}(k)$ will approach infinity.

The question of whether such singularities actually occur can be answered by returning to the definition (3.3) of $\xi_{nn'}(k)$, which involves $\nabla_k u_{nk}$. Since $\mathrm{e}^{\mathrm{i}k\cdot x}$ is an entire function of $k$, the analytical properties of $u_{nk}$ (as a function of $k$) are the same as those of $\psi_{nk}$. This means that $u_{nk}$ is not analytic in $k$, since it has been proven that Bloch functions are always discontinuous in $k$ at a degeneracy (Zak 1985). To be specific, for a degeneracy at $k = k_0$, the Bloch functions depend critically on the *direction* of $\kappa \equiv k - k_0$, and may differ for arbitrarily small values of $\kappa$ in different directions (Bir and Pikus (1974), p 155).

Thus, from equation (3.3) we see that $\xi_{nn'}(k)$ is always singular at a degeneracy. As shown by Blount (Blount (1962), p 331), the singularity occurs in the transverse ($\xi \perp k$) components of $\xi_{nn'}(k)$, which vary as $\kappa^{-1}$ in the neighbourhood of the degeneracy. The longitudinal components ($\xi \parallel k$) remain finite, however, because Bloch functions are continuous functions of $\kappa$ along any fixed direction of $\kappa$ (Bir and Pikus (1974), pp 157–8).

As a specific example, consider the following effective-mass Hamiltonian for the $\Gamma_{15}$ valence electrons in a zinc-blende crystal (Shockley 1950, Dresselhaus *et al* 1955), which is derived below in section 5:

$$H = \begin{bmatrix} Lk_x^2 + M(k_y^2 + k_z^2) & Nk_xk_y & Nk_xk_z \\ Nk_xk_y & Lk_y^2 + M(k_x^2 + k_z^2) & Nk_yk_z \\ Nk_xk_z & Nk_yk_z & Lk_z^2 + M(k_x^2 + k_y^2) \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & iG(k_x^2 - k_y^2)k_z & iG(k_x^2 - k_z^2)k_y \\ iG(k_y^2 - k_x^2)k_z & 0 & iG(k_y^2 - k_z^2)k_x \\ iG(k_z^2 - k_x^2)k_y & iG(k_z^2 - k_y^2)k_x & 0 \end{bmatrix}. \tag{3.8}$$

Here $L$, $M$, $N$, and $G$ are real constants. The basis functions for this Hamiltonian are 'pseudo-Bloch' functions (Kittel and Mitchell 1954, Dresselhaus *et al* 1955) constructed from linear combinations of zone-centre Bloch functions; the linear combination is chosen such that it diagonalizes the $\boldsymbol{k} \cdot \boldsymbol{p}$ coupling (to first order in $k$) between the given $\Gamma_{15}$ states and all other $\Gamma$ states of different energy. A precise definition of the pseudo-Bloch basis states is given in section 5. At present all one needs to know is that these functions are analytic in $\boldsymbol{k}$ by construction, so they do not contribute to the above-mentioned singularities in $\boldsymbol{\xi}_{nn'}(\boldsymbol{k})$. In the limit as $\boldsymbol{k}$ approaches zero—from any direction—the pseudo-Bloch functions approach the zone-centre Bloch functions $u_{n0}$ given in equations (2.11) and (2.12), where $n = X$, $Y$, or $Z$.

To find the actual Bloch functions $\psi_{nk}$, one must diagonalize the Hamiltonian (3.8). The first example treated here is a calculation of the $z$-component of $\boldsymbol{\xi}_{nn'}(\boldsymbol{k})$ for the transverse case $\boldsymbol{k} = (k_x, 0, 0)$. To calculate the gradient in equation (3.3), one must find the eigenfunctions of (3.8) for $\boldsymbol{k} = (k_x, 0, k_z)$, where $k_z$ is infinitesimal. It is easily verified that to first order in $k_z$, these eigenfunctions are

$$u_{Xk} = \begin{bmatrix} 1 \\ \dfrac{-iGk_x^2k_z}{(L - M)k_x^2} \\ \dfrac{Nk_xk_z}{(L - M)k_x^2} \end{bmatrix} \qquad u_{Yk} = \begin{bmatrix} \dfrac{-iGk_x^2k_z}{(L - M)k_x^2} \\ 1 \\ 0 \end{bmatrix} \qquad u_{Zk} = \begin{bmatrix} \dfrac{-Nk_xk_z}{(L - M)k_x^2} \\ 0 \\ 1 \end{bmatrix}. \tag{3.9}$$

Substituting these expressions into equation (3.3), one finds that

$$\xi_{XY}^z(k_x, 0, 0) = \frac{G}{L - M} + \mathrm{O}(k_x^2) \tag{3.10}$$

$$\xi_{XZ}^z(k_x, 0, 0) = -\frac{iN}{(L - M)k_x} + \mathrm{O}(k_x) \tag{3.11}$$

with all other matrix elements zero. As expected, the transverse part of the crystal coordinate matrix contains elements diverging as $k^{-1}$.

The longitudinal components are somewhat easier to calculate, especially along the $\Delta$ axes. For example, if $\boldsymbol{k} = (0, 0, k_z)$, all off-diagonal terms in the Hamiltonian (3.8) vanish. Indeed, from the symmetry of the wave functions on the $\Delta$ axis (Parmenter 1955), one can see that the off-diagonal elements vanish to any order in $\boldsymbol{k}$. Hence, $\xi_{nn'}^z(0, 0, k_z) = 0$ to within terms of order $k_z^0$. To determine the order of the leading corrections, one can use $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory to find the cell-periodic functions $u_{nk}$. For example, from the symmetry properties of the momentum matrix (Parmenter 1955, Dresselhaus 1955) we know that $u_{Xk}$ and $u_{Yk}$ contain terms proportional to $k_z^2 u_{X'0}$ and $k_z u_{X'0}$, respectively, where $X'$ refers to a $\Gamma_{15}$ conduction-band state. Therefore

$$\xi_{XY}^z(0, 0, k_z) = \mathrm{O}(k_z^2). \tag{3.12}$$

The other off-diagonal matrix elements are forbidden by symmetry to any order in $k_z$ (Parmenter 1955):

$$\xi_{XZ}^z(0, 0, k_z) = \xi_{YZ}^z(0, 0, k_z) = 0 \tag{3.13}$$

but the diagonal elements are not:

$$\xi_{nn}^z(0, 0, k_z) = O(k_z) \tag{3.14}$$

where $n = X$, $Y$, or $Z$. Equation (3.14) assumes that the Bloch functions are given by the standard expressions of $k \cdot p$ perturbation theory, without any extra phase factors. One is of course free to include an arbitrary $k$-dependent phase factor, in which case the diagonal elements (3.14) are likewise arbitrary and may be set equal to any real value (Blount 1962, Lax 1974).

The existence of singularities in $\boldsymbol{\xi}_{nn'}(\boldsymbol{k})$ means that the Bloch representation is ill suited for the analysis of degenerate bands in external fields (Wannier (1960), appendix I; Roth (1962); Blount (1962), p 319; Aspnes and Bottka (1972), p 501). To eliminate the singularities, one must use basis functions that are analytic functions of $\boldsymbol{k}$. The best-known example is the Luttinger–Kohn representation, which was used in the construction of the pseudo-Bloch functions mentioned above. This representation is analysed in detail below.

## 4. The Luttinger–Kohn representation

This section provides a review of the Luttinger–Kohn representation. The basis functions and their properties are considered in section 4.1. Representations of the wave function, momentum operator, and coordinate operator are treated in sections 4.2, 4.3, and 4.4, respectively. Section 4.4 focuses on an effect recently discovered by Burt (1999), in which the coordinate operator gives rise to interband mixing at the surface of the Brillouin zone. An example illustrating this effect is presented in section 4.5. Section 4.6 describes the Luttinger–Kohn Hamiltonian for an electron in a uniform electric field.

### 4.1. Basis functions

The basis functions in the Luttinger–Kohn (LK) representation are defined by (Luttinger and Kohn 1955, Bir and Pikus 1974)

$$\chi_{n\boldsymbol{k}}(\boldsymbol{x}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\psi_{n\boldsymbol{k}_0}(\boldsymbol{x}) = (2\pi)^{-3/2}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{k}_0)\cdot\boldsymbol{x}}U_n(\boldsymbol{x}) \tag{4.1}$$

in which $U_n(\boldsymbol{x}) \equiv u_{n\boldsymbol{k}_0}(\boldsymbol{x})$ is the periodic part of the Bloch function (2.4) at some fixed point $\boldsymbol{k}_0$ (usually $\boldsymbol{k}_0 = \boldsymbol{0}$). Note that the wave vector $\boldsymbol{k}$ in equation (4.1) is not the absolute crystal momentum used in the Bloch functions (2.4); it is the crystal momentum measured relative to $\boldsymbol{k}_0$ (which was referred to as $\boldsymbol{\kappa}$ in section 3).

Since $U_n(\boldsymbol{x})$ is periodic, it can be expanded in a Fourier series:

$$U_n(\boldsymbol{x}) = \sum_{\boldsymbol{G}} U_{n\boldsymbol{G}}\mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{x}} \tag{4.2}$$

$$U_{n\boldsymbol{G}} = \frac{1}{\Omega}\int_{\Omega} U_n(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{x}}\,\mathrm{d}^3x \tag{4.3}$$

where $\boldsymbol{G}$ is a reciprocal-lattice vector. The set $\{U_n\}$ is orthogonal (see equation (2.5)), and it is complete with respect to cell-periodic functions (Luttinger and Kohn 1955, Burt 1988, 1992):

$$\frac{1}{\Omega}\int_{\Omega} U_n^{\dagger}(\boldsymbol{x})U_{n'}(\boldsymbol{x})\,\mathrm{d}^3x = \delta_{nn'} \tag{4.4}$$

$$\frac{1}{\Omega}\sum_n U_n(\boldsymbol{x})U_n^{\dagger}(\boldsymbol{x}') = I\sum_{\boldsymbol{R}}\delta(\boldsymbol{x} - \boldsymbol{x}' - \boldsymbol{R}) \tag{4.5}$$

where $R$ is a Bravais lattice vector and $I$ is the $2 \times 2$ unit matrix. Another way of expressing these properties is

$$\sum_{G} U_{nG}^{\dagger} U_{n'G} = \delta_{nn'} \tag{4.6}$$

$$\sum_{n} U_{nG} U_{nG'}^{\dagger} = I \delta_{GG'}. \tag{4.7}$$

The LK basis functions are written in Dirac notation as

$$\chi_{nk}^{s}(x) = \langle sx|nk \rangle \tag{4.8}$$

where $s$ is a spin index (i.e., (4.8) is one component of the spinor (4.1)), and $|sx\rangle$ is an eigenket of spin and position. The inner product of two basis functions may be calculated from (4.2):

$$\langle nk|n'k' \rangle = \sum_{s} \int \langle nk|sx \rangle \langle sx|n'k' \rangle \, \mathrm{d}^3 x = \sum_{G,G'} U_{nG}^{\dagger} U_{n'G'} \delta(k - k' + G - G')$$

$$= \sum_{G} B_{nn'}(G) \delta(k - k' + G) \tag{4.9}$$

where

$$B_{nn'}(G) = \sum_{G'} U_{n,G'+G}^{\dagger} U_{n'G'} = \frac{1}{\Omega} \int_{\Omega} U_{n}^{\dagger}(x) U_{n'}(x) \mathrm{e}^{\mathrm{i}G \cdot x} \, \mathrm{d}^3 x. \tag{4.10}$$

If $k$ and $k'$ lie inside the same primitive unit cell in reciprocal space, then $k - k'$ is never equal to a non-zero reciprocal-lattice vector. In this case

$$\delta(k - k' + G) = \delta(k - k') \delta_{G,0} \tag{4.11}$$

and equation (4.9) reduces to

$$\langle nk|n'k' \rangle = \delta_{nn'} \delta(k - k'). \tag{4.12}$$

Basis functions from the same unit cell are therefore orthogonal. One can show likewise that they are complete:

$$\sum_{n} \int_{\Omega^*} \langle sx|nk \rangle \langle nk|s'x' \rangle \, \mathrm{d}^3 k = \delta_{ss'} \delta(x - x') \tag{4.13}$$

where $\Omega^* = (2\pi)^3 / \Omega$ is the volume of a primitive unit cell in reciprocal space. The shape and location of this cell are arbitrary (although it is convenient to choose a cell symmetric about $k = 0$).

### 4.2. The wave function

The closure property (4.13) means that one can expand any spinor wave function $\varphi(x)$ in the LK basis:

$$\varphi^{s}(x) = \langle sx|\varphi \rangle = \sum_{n} \int_{\Omega^*} \langle sx|nk \rangle \langle nk|\varphi \rangle \, \mathrm{d}^3 k = \sum_{n} \int_{\Omega^*} \chi_{nk}^{s}(x) \varphi_n(k) \, \mathrm{d}^3 k \tag{4.14}$$

where

$$\varphi_n(k) \equiv \langle nk|\varphi \rangle \tag{4.15}$$

is the wave function in the LK representation. One can calculate $\varphi_n(k)$ from the plane-wave expansion (4.2):

$$\varphi_n(k) = \sum_{s} \int \langle nk|sx \rangle \langle sx|\varphi \rangle \, \mathrm{d}^3 x = \sum_{G} U_{nG}^{\dagger} \varphi(k + k_0 + G) \tag{4.16}$$

where

$$\varphi(\boldsymbol{q}) = (2\pi)^{-3/2} \int \varphi(\boldsymbol{x}) \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}} \, \mathrm{d}^3 x \tag{4.17}$$

is the Fourier transform of $\varphi(\boldsymbol{x})$. Note from (4.15) and (4.16) that $\varphi_n(\boldsymbol{k})$ is defined for all values of $\boldsymbol{k}$. However, equation (4.14) shows that only those values that are within a single unit cell are used in constructing $\varphi(\boldsymbol{x})$. Thus, there is a redundancy of information in $\varphi_n(\boldsymbol{k})$. To emphasize this, $\varphi_n(\boldsymbol{k})$ will be referred to as the 'extended-zone' LK wave function†.

The inverse relation to (4.16) is obtained from (4.7):

$$\varphi(\boldsymbol{k} + \boldsymbol{k}_0 + \boldsymbol{G}) = \sum_n U_{nG} \varphi_n(\boldsymbol{k}). \tag{4.18}$$

The extended-zone wave function $\varphi_n(\boldsymbol{k})$ is not periodic in $\boldsymbol{k}$, since from (4.16), (4.18), and (4.10) we have

$$\varphi_n(\boldsymbol{k} - \boldsymbol{G}) = \sum_{n'} B_{nn'}(\boldsymbol{G}) \varphi_{n'}(\boldsymbol{k}). \tag{4.19}$$

Thus, $B_{nn'}(\boldsymbol{G})$ represents the reciprocal-lattice translation operator for the extended-zone wave function. If equation (4.19) were written in the Bloch representation, $B_{nn'}(\boldsymbol{G})$ would be replaced by $\delta_{nn'}$.

If one pulls $U_n(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\boldsymbol{k}_0\cdot\boldsymbol{x}}$ outside the integral in (4.14), the remaining quantity

$$F_n(\boldsymbol{x}) = (2\pi)^{-3/2} \int_{\Omega^*} \varphi_n(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}^3 k \tag{4.20}$$

is referred to as an envelope function. Its Fourier transform is

$$F_n(\boldsymbol{k}) = \begin{cases} \varphi_n(\boldsymbol{k}) & \boldsymbol{k} \in \Omega^* \\ 0 & \boldsymbol{k} \notin \Omega^* . \end{cases} \tag{4.21}$$

The function $F_n(\boldsymbol{k})$ is therefore the 'reduced-zone' LK wave function. In coordinate space, the envelope-function representation for $\varphi$ is (Burt 1988, 1992)

$$\varphi(\boldsymbol{x}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}_0\cdot\boldsymbol{x}} \sum_n F_n(\boldsymbol{x}) U_n(\boldsymbol{x}). \tag{4.22}$$

Since the set $\{F_n\}$ provides complete information about the wave function, it must be possible to express $\varphi_n(\boldsymbol{k})$ in terms of $F_n(\boldsymbol{k})$. The first step is to rewrite equation (4.19) as

$$\varphi_n(\boldsymbol{k}) = \sum_{n'} B_{nn'}(\boldsymbol{G}) \varphi_{n'}(\boldsymbol{k} + \boldsymbol{G}). \tag{4.23}$$

Then, since there is only one $\boldsymbol{G}$ such that $(\boldsymbol{k} + \boldsymbol{G}) \in \Omega^*$, equation (4.23) is equivalent to

$$\varphi_n(\boldsymbol{k}) = \sum_{\boldsymbol{G}} \sum_{n'} B_{nn'}(\boldsymbol{G}) F_{n'}(\boldsymbol{k} + \boldsymbol{G}). \tag{4.24}$$

This reverses the truncation performed in equation (4.21).

---

† This nomenclature was originally developed for the Bloch representation (Ashcroft and Mermin (1976), pp 159–60), where it indicates a non-periodic, non-redundant depiction of the energy band structure covering all values of $\boldsymbol{k}$. The 'repeated-zone' scheme indicates a periodic, redundant depiction. The LK wave function $\varphi_n(\boldsymbol{k})$ is redundant but non-periodic (see equation (4.19)), which fits neither definition strictly. To avoid the connotation of periodicity associated with 'repeated-zone', I have chosen to use the term 'extended-zone'.

### 4.3. The momentum operator

To determine the properties of the momentum operator $p$, it is simplest to start from the momentum representation (4.17):

$$\varphi^s(q) = \langle sq|\varphi\rangle \tag{4.25}$$

in which $|sq\rangle$ is an eigenket of momentum (i.e., $p|sq\rangle = \hbar q|sq\rangle$); hence,

$$\langle sq|p|\varphi\rangle = \hbar q\langle sq|\varphi\rangle. \tag{4.26}$$

The LK basis functions are written in the momentum representation as follows:

$$\langle sq|nk\rangle = \sum_{s'}\int \langle sq|s'x\rangle\langle s'x|nk\rangle \, \mathrm{d}^3x = \sum_G U_{nG}^s\delta(k+k_0+G-q) \tag{4.27}$$

in which $\langle sq|s'x\rangle = (2\pi)^{-3/2}\mathrm{e}^{-\mathrm{i}q\cdot x}\delta_{ss'}$. The action of $p$ in the LK representation is obtained by combining equations (4.26) and (4.27):

$$\langle nk|p|\varphi\rangle = \sum_s\int \langle nk|sq\rangle\langle sq|p|\varphi\rangle \, \mathrm{d}^3q = \sum_G U_{nG}^\dagger\hbar(k+k_0+G)\varphi(k+k_0+G). \tag{4.28}$$

One can then express this result in terms of $\varphi_n(k)$ using equation (4.18):

$$\langle nk|p|\varphi\rangle = \sum_{n'}[p_{nn'}(k_0) + \hbar k\delta_{nn'}]\varphi_{n'}(k) \tag{4.29}$$

where

$$p_{nn'}(k_0) = \sum_G U_{nG}^\dagger U_{n'G}\hbar(k_0+G) \tag{4.30}$$

is the momentum matrix defined previously in equation (2.8). For the special case $|\varphi\rangle = |n'k'\rangle$, equation (4.29) reduces to

$$\langle nk|p|n'k'\rangle = [p_{nn'}(k_0) + \hbar k\delta_{nn'}]\delta(k-k'). \tag{4.31}$$

### 4.4. The coordinate operator

The action of the coordinate operator in the LK representation is simply

$$\langle nk|x|\varphi\rangle = \sum_s\int \langle nk|sx\rangle x\langle sx|\varphi\rangle \, \mathrm{d}^3x = \mathrm{i}\,\nabla_k\sum_s\int \langle nk|sx\rangle\langle sx|\varphi\rangle \, \mathrm{d}^3x$$
$$= \mathrm{i}\,\nabla_k\langle nk|\varphi\rangle = \mathrm{i}\,\nabla_k\varphi_n(k) \tag{4.32}$$

which, apart from the index $n$, is the same as the well-known result for the momentum representation (Merzbacher 1998). One can use equation (4.32) to obtain a variety of related expressions, such as the product

$$x\varphi^s(x) = \langle sx|x|\varphi\rangle = \sum_n\int_{\Omega^*} \langle sx|nk\rangle\langle nk|x|\varphi\rangle \, \mathrm{d}^3k = \sum_n\int_{\Omega^*} \chi_{nk}^s(x)[\mathrm{i}\,\nabla_k\varphi_n(k)] \, \mathrm{d}^3k \tag{4.33}$$

and the matrix element

$$\langle \varphi^{(i)}|x|\varphi^{(j)}\rangle = \sum_n\int_{\Omega^*} \varphi_n^{(i)*}(k)[\mathrm{i}\,\nabla_k\varphi_n^{(j)}(k)] \, \mathrm{d}^3k \tag{4.34}$$

where $i$ and $j$ are labels for different states (not spin indices). Equation (4.34) is equal to the matrix element in the momentum representation:

$$\langle \varphi^{(i)}|x|\varphi^{(j)}\rangle = \int \varphi^{(i)\dagger}(q)[\mathrm{i}\,\nabla_q\varphi^{(j)}(q)] \, \mathrm{d}^3q \tag{4.35}$$

as can be seen by substituting (4.16) into (4.34), then using the completeness relation (4.7).

The matrix element of $x$ connecting two basis functions is just equation (4.32) for the special case $|\varphi\rangle = |n'\boldsymbol{k}'\rangle$:

$$\langle n\boldsymbol{k}|x|n'\boldsymbol{k}'\rangle = \mathrm{i}\,\boldsymbol{\nabla}_k\langle n\boldsymbol{k}|n'\boldsymbol{k}'\rangle. \tag{4.36}$$

It is often assumed, both in the original literature (Luttinger and Kohn 1955, Kjeldaas and Kohn 1957) and in textbooks (Callaway 1991), that equation (4.36) is equivalent to

$$\langle n\boldsymbol{k}|x|n'\boldsymbol{k}'\rangle = \mathrm{i}\,\boldsymbol{\nabla}_k\delta(\boldsymbol{k}-\boldsymbol{k}')\delta_{nn'}. \tag{4.37}$$

However, to evaluate the gradient in (4.36), one must (in general) make use of $\boldsymbol{k}$-values lying outside the cell $\Omega^*$. Consequently, the orthogonality relation (4.12) is not valid in (4.36); one should instead use equation (4.9), which gives

$$\langle n\boldsymbol{k}|x|n'\boldsymbol{k}'\rangle = \mathrm{i}\sum_G B_{nn'}(\boldsymbol{G})\,\boldsymbol{\nabla}_k\delta(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G}). \tag{4.38}$$

Hence, the coordinate operator in the LK representation does give rise to interband mixing, albeit only at the surfaces of the unit cell $\Omega^*$. This was first discovered by Burt (1999)†.

There is a subtlety here, however, because the mixing is not obvious from equations (4.33) and (4.34), which have no terms off-diagonal in $n$. Indeed, equations (4.33) and (4.34) look as if they should be consistent with the matrix element (4.37) and inconsistent with (4.38). To show that the truth is in fact just the opposite, one can insert the matrix elements (4.37) and (4.38) into the integral

$$\langle\varphi^{(i)}|x|\varphi^{(j)}\rangle = \sum_{n,n'}\int_{\Omega^*}\int_{\Omega^*}\langle\varphi^{(i)}|n\boldsymbol{k}\rangle\langle n\boldsymbol{k}|x|n'\boldsymbol{k}'\rangle\langle n'\boldsymbol{k}'|\varphi^{(j)}\rangle\,\mathrm{d}^3k\,\mathrm{d}^3k' \tag{4.39}$$

and compare the results with equation (4.34). In either case, the integral over $\boldsymbol{k}'$ is evaluated using the identity

$$\int_{\Omega^*}\delta(\boldsymbol{k}-\boldsymbol{k}')\varphi_n(\boldsymbol{k}')\,\mathrm{d}^3k' = F_n(\boldsymbol{k}) \tag{4.40}$$

where $F_n(\boldsymbol{k})$ is the reduced-zone wave function. Equation (4.37) therefore yields

$$\langle\varphi^{(i)}|x|\varphi^{(j)}\rangle = \sum_n\int F_n^{(i)*}(\boldsymbol{k})[\mathrm{i}\,\boldsymbol{\nabla}_k F_n^{(j)}(\boldsymbol{k})]\,\mathrm{d}^3k. \tag{4.41}$$

This expression disagrees with the correct result (4.34) because, at the surface of $\Omega^*$, $F_n(\boldsymbol{k})$ drops abruptly to zero, whereas $\varphi_n(\boldsymbol{k})$ does not (cf. equations (4.16) and (4.21)). The gradient in equation (4.41) therefore generates a surface $\delta$-function with a magnitude different from that in (4.34). For example, if $\varphi(\boldsymbol{q})$ is continuous, then $\varphi_n(\boldsymbol{k})$ is also, and the surface $\delta$-function is zero for equation (4.34) but non-zero for equation (4.41). To illustrate further this discrepancy, an explicit numerical comparison between (4.34) and (4.41) is presented below in section 4.5.

If one inserts the matrix element (4.38) into equation (4.39), one finds instead

$$\langle\varphi^{(i)}|x|\varphi^{(j)}\rangle = \sum_G\sum_{n,n'}B_{nn'}(\boldsymbol{G})\int F_n^{(i)*}(\boldsymbol{k})[\mathrm{i}\,\boldsymbol{\nabla}_k F_{n'}^{(j)}(\boldsymbol{k}+\boldsymbol{G})]\,\mathrm{d}^3k. \tag{4.42}$$

The sum in (4.42) covers all $\boldsymbol{G}$, but the only non-zero $\boldsymbol{G}$s that contribute are those that connect two surfaces of the cell $\Omega^*$. The sums over $\boldsymbol{G}$ and $n'$ can be evaluated using equation (4.24); this

---

† Although Burt was the first to give an explicit description of this effect, Kohn was undoubtedly aware of its existence as early as 1959. Consider the following statement: 'Two courses of action suggest themselves: One is to hold fast to the basis of Luttinger and Kohn and consider very carefully what happens when $\boldsymbol{k}$ is *on* the Brillouin zone boundary. This procedure we found not tractable.' (Kohn (1959), p 1462).

yields precisely the result found earlier in equation (4.34). Equation (4.42) is also equivalent to the expression derived by Burt (Burt (1999), equation (6.12)).

Thus, we see that the interband mixing found in equation (4.38) is manifest in expressions such as (4.42) involving the reduced-zone wave function, but hidden in expressions such as (4.34) involving the extended-zone wave function. I say 'hidden' rather than 'non-existent' because calculation of the gradient in (4.34) requires knowledge of $\varphi_n(k)$ for $k$-values lying outside the cell $\Omega^*$. This information is redundant and can be expressed in terms of $k$-values inside $\Omega^*$ through the use of equation (4.19). But this procedure merely replicates the reduced-zone expression (4.42).

In the Bloch representation, the equations corresponding to (4.34) and (4.42) have the same form as that shown here, except that $B_{nn'}(G)$ is replaced by $\delta_{nn'}$, and there is an additional integral containing $\xi_{nn'}(k)$.

Further discussion of the coordinate operator (using the method of Blount (1962)) is presented in appendix A, including a proof that the integrals (4.14) and (4.33) are independent of the choice of unit cell $\Omega^*$.

### 4.5. Example of zone-boundary mixing effects

The validity of equations (4.34), (4.41), and (4.42) is tested here for a simple example, which also serves to emphasize the care that must be taken in dealing with the surface of the cell $\Omega^*$. The reference crystal in this example is a one-dimensional spin-zero 'empty lattice' of period $a$, for which the zone-centre Bloch functions are just the plane waves $U_n(x) = \exp(\mathrm{i}\,2\pi nx/a)$, where $n$ is an integer. The associated Fourier coefficients are $U_{nG} = \delta_{nG}$, where $G$ is measured in units of $2\pi/a$. The wave function in the extended-zone LK representation (equation (4.16)) is therefore $\varphi_n(k) = \varphi(k + 2\pi n/a)$.

The problem to be treated here involves two states $|\varphi^{(0)}\rangle$ and $|\varphi^{(1)}\rangle$, which are given in the coordinate, momentum, extended-zone, and reduced-zone representations by

$$
\begin{aligned}
\varphi^{(0)}(x) &= \sqrt{2\pi}\,\sin(\pi x/a)/\pi x \\
\varphi^{(0)}(q) &= Z(q) \\
\varphi_n^{(0)}(k) &= Z(k + 2\pi n/a) \\
F_n^{(0)}(k) &= \delta_{n0} Z(k)
\end{aligned}
\tag{4.43}
$$

and

$$
\begin{aligned}
\varphi^{(1)}(x) &= \sqrt{2\pi}\,\mathrm{e}^{\mathrm{i}2\pi x/a}\sin(\pi x/a)/\pi x \\
\varphi^{(1)}(q) &= Z(q - 2\pi/a) \\
\varphi_n^{(1)}(k) &= Z(k + 2\pi(n-1)/a) \\
F_n^{(1)}(k) &= \delta_{n1} Z(k)
\end{aligned}
\tag{4.44}
$$

in which

$$
Z(k) = \begin{cases} 1 & |k| < \pi/a \\ 0 & |k| > \pi/a \end{cases}
\tag{4.45}
$$

is a function defining the boundaries of the first Brillouin zone (which is chosen to coincide with $\Omega^*$).

In the coordinate representation, the matrix element of $x$ can be calculated using trigonometric identities and the integral $\int_{-\infty}^{\infty} x^{-1}\sin x\,\mathrm{d}x = \pi$; the result is

$$
\langle \varphi^{(0)}|x|\varphi^{(1)}\rangle = \mathrm{i}/2.
\tag{4.46}
$$

For the momentum representation, one inserts the derivative of $\varphi^{(1)}(q)$:

$$d\varphi^{(1)}/dq = \delta(q - \pi/a) - \delta(q - 3\pi/a) \tag{4.47}$$

into equation (4.35), which yields the same result (4.46). This calculation uses the identity

$$\int f(x)\delta(x)\,dx = \frac{1}{2}\left[\lim_{x\to 0^-} f(x) + \lim_{x\to 0^+} f(x)\right] \tag{4.48}$$

which follows from the definition of the generalized integral (4.48) as the limit of a sequence of ordinary integrals (Gelfand and Shilov 1964). Here $f(x)$ is assumed to be a bounded function with a finite number of discontinuities. Note that (4.48) does not depend on the value of $f(x)$ at $x = 0$.

In the extended-zone LK representation, the only non-zero contribution to equation (4.34) comes from $n = 0$ (since $Z(k)Z(k + G) = \delta_{G,0}Z(k)$). The derivative of $\varphi_0^{(1)}(k)$ is

$$d\varphi_0^{(1)}/dk = \delta(k - \pi/a) - \delta(k - 3\pi/a) \tag{4.49}$$

which, when inserted into (4.34), gives the result (4.46) found above. However, for the case of equation (4.41), the reduced-zone wave functions (4.43) and (4.44) have different values of $n$, yielding a matrix element

$$\langle\varphi^{(0)}|x|\varphi^{(1)}\rangle = 0 \tag{4.50}$$

which is clearly incorrect. The matrix element in the reduced-zone representation should instead be calculated using (4.42), where $B_{nn'}(G) = \delta_{n,n'+G}$. Since the reduced-zone wave functions have $n = 0$ and $n' = 1$, this requires $G = -2\pi/a$. The derivative appearing in (4.42) is then

$$\frac{d}{dk}F_1^{(1)}(k - 2\pi/a) = \delta(k - \pi/a) - \delta(k - 3\pi/a) \tag{4.51}$$

which gives the matrix element (4.46) once again.

Hence, this example reinforces the conclusion that equations (4.34) and (4.42) are the correct expressions to use for calculating the matrix element of $x$ in the LK representation. The calculation is simpler if $x\varphi(x)$ is square integrable (Blount 1962), for this ensures that $\varphi(q)$ and $\varphi_n(k)$ are continuous, and consequently that no $\delta$-functions arise in equation (4.34). In this case one can get the right answer even from equation (4.41), provided that one deliberately omits all zone-boundary effects.

However, as Burt (1999) has emphasized, such an approach only works for the *complete* wave function. Even if $x\varphi(x)$ is square integrable, the individual terms in the envelope-function expansion (4.22) generally do not have this property. Consequently, there is considerable danger of error if one chooses to approximate the wave function by one term in expansion (4.22). This is discussed at length by Burt (1993, 1995a, b, 1999) and Coles *et al* (1999).

## 4.6. The Hamiltonian

With the results derived in sections 4.1–4.4, we are now in a position to evaluate the Luttinger–Kohn Hamiltonian for an electron in a uniform electric field. The Pauli Hamiltonian for this case is

$$H = H_0 + U + W \tag{4.52}$$

where $H_0$ is the periodic Hamiltonian defined in (2.1), $U$ is the potential energy of the field defined in (1.1), and

$$W = \frac{e\hbar}{4m^2c^2}\boldsymbol{\mathcal{E}} \cdot (\boldsymbol{p} \times \boldsymbol{\sigma}). \tag{4.53}$$

Taking note of equations (2.3) and (2.7), the matrix elements of $H_0$ are

$$\langle n\boldsymbol{k}|H_0|n'\boldsymbol{k}'\rangle = \left[\left(E_n + \frac{\hbar^2 k^2}{2m}\right)\delta_{nn'} + \frac{\hbar}{m}\boldsymbol{k}\cdot\boldsymbol{\pi}_{nn'}\right]\delta(\boldsymbol{k}-\boldsymbol{k}') \tag{4.54}$$

in which $E_n \equiv E_n(\boldsymbol{k}_0)$ and $\boldsymbol{\pi}_{nn'} \equiv \boldsymbol{\pi}_{nn'}(\boldsymbol{k}_0)$. The matrix elements of $U$ are obtained from equation (4.38):

$$\langle n\boldsymbol{k}|U|n'\boldsymbol{k}'\rangle = \mathrm{i}e\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\nabla}_k \sum_G B_{nn'}(\boldsymbol{G})\delta(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G}). \tag{4.55}$$

Finally, the matrix elements of $W$ are

$$\langle n\boldsymbol{k}|W|n'\boldsymbol{k}'\rangle = \frac{e\hbar}{4m^2c^2}[\boldsymbol{\mathcal{E}}\cdot(\boldsymbol{p}\times\boldsymbol{\sigma})_{nn'} + (\boldsymbol{\mathcal{E}}\times\hbar\boldsymbol{k})\cdot\boldsymbol{\sigma}_{nn'}]\delta(\boldsymbol{k}-\boldsymbol{k}') \tag{4.56}$$

where the matrix elements $(\boldsymbol{p}\times\boldsymbol{\sigma})_{nn'}$ and $\boldsymbol{\sigma}_{nn'}$ are evaluated at $\boldsymbol{k}_0$ according to the prescription given in equation (2.8).

## 5. Elimination of interband coupling in the Luttinger–Kohn Hamiltonian

In the Hamiltonian derived above, states of different energy are coupled at $\boldsymbol{k}\neq\boldsymbol{0}$ by the $\boldsymbol{k}\cdot\boldsymbol{p}$ interaction. Therefore, this Hamiltonian is directly useful only for numerical work. However, one often needs an accurate solution only within a small region near $\boldsymbol{k}=\boldsymbol{0}$. In such cases, one can derive an approximate 'effective-mass' Hamiltonian by treating the $\boldsymbol{k}\cdot\boldsymbol{p}$ interaction and the field-induced interband coupling as small perturbations. This is done below using the unitary-transformation technique of Luttinger and Kohn (1955).

Since effective-mass theory is not valid near the zone boundary, one is justified in neglecting the $\boldsymbol{G}\neq\boldsymbol{0}$ terms in the potential energy (4.55). This means that all operators $A$ of interest have matrix elements of the form

$$\langle n\boldsymbol{k}|A|n'\boldsymbol{k}'\rangle = \langle n|A|n'\rangle\delta(\boldsymbol{k}-\boldsymbol{k}') \tag{5.1}$$

where $\langle n|A|n'\rangle$ is a $\boldsymbol{k}$-dependent operator†. Thus, for brevity, I shall give only $\langle n|A|n'\rangle$ in the equations that follow.

The physical system of primary interest is a set of states that are degenerate or quasi-degenerate in energy at the point $\boldsymbol{k}_0$. 'Quasi-degenerate' refers to states that lie close together in energy ($E_n \approx E_{n'}$), such as the conduction and valence bands of a narrow-gap semiconductor (Kane 1957). With this in mind, the crystal Hamiltonian (4.54) may be written as $H_0 = H^{(0)} + H^{(1)} + H^{(2)}$, where

$$\langle n|H^{(0)}|n'\rangle = \left(E_n + \frac{\hbar^2 k^2}{2m}\right)\delta_{nn'} \tag{5.2}$$

$$\langle n|H^{(1)}|n'\rangle = \frac{\hbar}{m}k_\alpha \pi_{nn'}^\alpha \Theta_{nn'} \tag{5.3}$$

$$\langle n|H^{(2)}|n'\rangle = \frac{\hbar}{m}k_\alpha \pi_{nn'}^\alpha \Delta_{nn'}. \tag{5.4}$$

Here a summation over the Cartesian components $\alpha$ is implied, and

$$\Delta_{nn'} = \begin{cases} 1 & E_n \approx E_{n'} \\ 0 & \text{otherwise} \end{cases} \tag{5.5}$$

---

† In this expression, $n$ and $n'$ are merely labels, and the quantity $|n\rangle$ has no independent meaning as a Dirac ket vector.

with $\Theta_{nn'} = 1 - \Delta_{nn'}$. The contributions from the electric field are $U$ and $W = W^{(0)} + W^{(1)}$, where

$$\langle n|U|n'\rangle = \mathrm{i}e\boldsymbol{\mathcal{E}} \cdot \boldsymbol{\nabla}_k \delta_{nn'} \tag{5.6}$$

$$\langle n|W^{(0)}|n'\rangle = \frac{e\hbar}{4m^2c^2}\boldsymbol{\mathcal{E}} \cdot (\boldsymbol{p} \times \boldsymbol{\sigma})_{nn'} \tag{5.7}$$

$$\langle n|W^{(1)}|n'\rangle = \frac{e\hbar^2}{4m^2c^2}(\boldsymbol{\mathcal{E}} \times \boldsymbol{k}) \cdot \boldsymbol{\sigma}_{nn'}. \tag{5.8}$$

The idea of the unitary-transformation technique is to replace the LK basis kets $|n\boldsymbol{k}\rangle$ with a new basis, say $|\overline{n\boldsymbol{k}}\rangle$, in which the Hamiltonian for each quasi-degenerate set of states is *decoupled* from all other states to any specified order in the small parameters $k$ and $\mathcal{E}$. This effectively reduces the Hamiltonian to block-diagonal form. In what follows, I shall use this technique to eliminate all† off-block-diagonal terms (i.e., terms proportional to $\Theta_{nn'}$) of the first order in $k$ and $\mathcal{E}$; this yields a decoupled Hamiltonian that is valid through terms of order $k^3$, $k\mathcal{E}$, and $\mathcal{E}^2$ inclusive. Kohn (1959) has shown for the magnetic-field problem that this procedure can be extended to arbitrary order.

The new basis kets $|\overline{n\boldsymbol{k}}\rangle$ are related to the old basis kets $|n\boldsymbol{k}\rangle$ by a unitary operator $\mathrm{e}^S$:

$$|\overline{n\boldsymbol{k}}\rangle = \mathrm{e}^S|n\boldsymbol{k}\rangle \tag{5.9}$$

in which $S = -S^\dagger$ is determined below. The matrix elements of $H$ in the new basis are equal to the matrix elements of

$$\bar{H} = \mathrm{e}^{-S}H\mathrm{e}^S = H + [H, S] + \frac{1}{2!}[[H, S], S] + \frac{1}{3!}[[[H, S], S], S] + \cdots \tag{5.10}$$

in the old basis. Putting in the explicit form of $H$, one has

$$\begin{aligned}
\bar{H} = {}& H^{(0)} + H^{(1)} + H^{(2)} + U + W + [H^{(0)}, S] + [H^{(1)}, S] + [H^{(2)}, S] + [U, S] \\
&+ \tfrac{1}{2}[[H^{(0)}, S], S] + \tfrac{1}{2}[[H^{(1)}, S], S] + \tfrac{1}{2}[[H^{(2)}, S], S] + \tfrac{1}{2}[[U, S], S] \\
&+ \tfrac{1}{6}[[[H^{(0)}, S], S], S]
\end{aligned} \tag{5.11}$$

which includes all terms of order $k^3$, $k\mathcal{E}$, and $\mathcal{E}^2$ (except those arising from $[W, S]$), but neglects higher-order terms such as $k^2\mathcal{E}$.

The first task is to eliminate the off-block-diagonal coupling $H^{(1)}$. This is done by choosing $S$ such that $H^{(1)} + [H^{(0)}, S] = 0$, which implies that (Luttinger and Kohn 1955)

$$\langle n|S|n'\rangle = -\frac{k_\alpha \pi_{nn'}^\alpha}{m\omega_{nn'}}\Theta_{nn'} = -\mathrm{i}k_\alpha \xi_{nn'}^\alpha \Theta_{nn'} \tag{5.12}$$

where $\hbar\omega_{nn'} = E_n - E_{n'}$, and equation (3.7) was used in the last step. With this choice of $S$, the Hamiltonian $\bar{H}$ becomes

$$\begin{aligned}
\bar{H} = {}& H^{(0)} + H^{(2)} + \tfrac{1}{2}[H^{(1)}, S] + \tfrac{1}{3}[[H^{(1)}, S], S] + \tfrac{1}{2}[[H^{(2)}, S], S] \\
&+ U + W + [U, S] + \tfrac{1}{2}[[U, S], S].
\end{aligned} \tag{5.13}$$

The term $[H^{(2)}, S]$ appearing in equation (5.11) has been discarded because its matrix elements are all off-block-diagonal terms of order $k^2$. The term $\tfrac{1}{2}[H^{(1)}, S]$ is just the usual effective-mass correction arising from the $\boldsymbol{k} \cdot \boldsymbol{p}$ interaction:

$$\tfrac{1}{2}\langle n|[H^{(1)}, S]|n'\rangle = \frac{\hbar k_\alpha k_\beta}{2m^2}\sideset{}{'}\sum_i \pi_{ni}^\alpha \pi_{in'}^\beta(\omega_{ni}^{-1} + \omega_{n'i}^{-1}). \tag{5.14}$$

---

† The term $W$ represents a special case. It is relativistic in origin, so its contribution is considerably smaller than the other (non-relativistic) terms proportional to $\mathcal{E}$. For this reason, the off-block-diagonal matrix elements of $W$ are ignored here.

For example, for the triply degenerate $\Gamma_{15}$ valence-band states in a zinc-blende crystal, the matrix (5.14) has the form shown previously in equation (3.8). Note that the 'pseudo-Bloch' basis states discussed in section 3 are just the states $|\overline{nk}\rangle$ defined in equations (5.9) and (5.12).

A simple estimate of the effective-mass parameters $L$, $M$, and $N$ in equation (3.8) may be obtained from a three-band model (Cardona *et al* 1988, Zhu and Chang 1994) that neglects all states outside the $\Gamma_{15}$ valence ($X$, $Y$, $Z$), $\Gamma_1$ conduction ($S$), and $\Gamma_{15}$ conduction ($X'$, $Y'$, $Z'$) bands. The momentum matrix elements connecting these states are $P = (\hbar/im)p_{SX}^x$, $P' = (\hbar/im)p_{SX'}^x$, and $Q = (\hbar/im)p_{X'Z}^y$, and the corresponding energy gaps are $E_0 = E_S - E_X$ and $E_0' = E_{X'} - E_X$. Within this model one has

$$L = \hbar^2/2m - P^2/E_0$$

$$M = \hbar^2/2m - Q^2/E_0'$$

$$N = -(P^2/E_0 + Q^2/E_0').$$

The operator $\frac{1}{3}[[H^{(1)}, S], S]$ gives rise to terms of order $k^3$ in the Hamiltonian:

$$\tfrac{1}{3}\langle n|[[H^{(1)}, S], S]|n'\rangle = -\frac{\hbar}{3m}k_\alpha k_\beta k_\gamma \sum_{i,j}{}' (\pi_{ni}^\alpha \xi_{ij}^\beta \xi_{jn'}^\gamma - 2\xi_{ni}^\alpha \pi_{ij}^\beta \xi_{jn'}^\gamma + \xi_{ni}^\alpha \xi_{ij}^\beta \pi_{jn'}^\gamma). \tag{5.15}$$

The matrix representing this operator for the $\Gamma_{15}$ valence states $X$, $Y$, and $Z$ was shown above in equation (3.8). In the three-band model, the constant $G$ in (3.8) is given by

$$G = 2PP'Q/E_0E_0'.$$

The operator $\frac{1}{2}[[H^{(2)}, S], S]$ is also of order $k^3$. However, in the absence of spin–orbit coupling, $H^{(2)}$ vanishes at the $\Gamma$ point of a zinc-blende crystal (Dresselhaus 1955). This term is therefore relativistically small in comparison to $\frac{1}{3}[[H^{(1)}, S], S]$, and it will not be considered further here.

In conventional effective-mass theory, the field-dependent terms $[U, S]$ and $\frac{1}{2}[[U, S], S]$ are neglected (Luttinger and Kohn 1955). Here they are retained, since they represent the lowest-order band-mixing effects arising from an electric field. From (5.6) and (5.12), one has

$$\langle n|[U, S]|n'\rangle = e\mathcal{E}_\alpha \xi_{nn'}^\alpha \Theta_{nn'}. \tag{5.16}$$

Equation (5.16) has the form of a dipole interaction, but the factor $\Theta_{nn'}$ prohibits the occurrence of any such interaction between degenerate or quasi-degenerate states. Hence, the singularities discussed in section 3.2 do not arise in the Luttinger–Kohn effective-mass Hamiltonian (because the pseudo-Bloch basis states $|\overline{nk}\rangle$ are analytic functions of $k$).

The last term in the Hamiltonian (5.13) is calculated from equations (5.12) and (5.16):

$$\tfrac{1}{2}\langle n|[[U, S], S]|n'\rangle = \frac{ie}{2m^2}(\mathcal{E}_\alpha k_\beta - k_\alpha \mathcal{E}_\beta)\sum_i{}' \frac{\pi_{ni}^\alpha \pi_{in'}^\beta}{\omega_{ni}\omega_{in'}}. \tag{5.17}$$

Note that the momentum matrix elements in (5.17) have the same form as those in the effective-mass Hamiltonian (5.14). Thus, for the previous example of $\Gamma_{15}$ valence electrons, the matrix representing (5.17) has the same form as equation (3.8), but without the diagonal terms:

$$\tfrac{1}{2}[[U, S], S] = \begin{bmatrix} 0 & in(\mathcal{E}_x k_y - k_x \mathcal{E}_y) & in(\mathcal{E}_x k_z - k_x \mathcal{E}_z) \\ in(\mathcal{E}_y k_x - k_y \mathcal{E}_x) & 0 & in(\mathcal{E}_y k_z - k_y \mathcal{E}_z) \\ in(\mathcal{E}_z k_x - k_z \mathcal{E}_x) & in(\mathcal{E}_z k_y - k_z \mathcal{E}_y) & 0 \end{bmatrix}. \tag{5.18}$$

Here $n$ is a real constant that is given in the three-band model by

$$n = -(e/2)(P^2/E_0^2 - Q^2/E_0'^2).$$

The change of basis (5.9) has eliminated all off-block-diagonal terms linear in $k$ in the Hamiltonian (5.13), but as a by-product it has generated the linear-in-$\mathcal{E}$ terms shown in equation (5.16). The next task is therefore to eliminate these terms. This is done by applying another unitary transformation:

$$|\widehat{n\boldsymbol{k}}\rangle = \mathrm{e}^S \mathrm{e}^R |n\boldsymbol{k}\rangle \tag{5.19}$$

yielding a new Hamiltonian $\hat{H} = \mathrm{e}^{-R} \bar{H} \mathrm{e}^R$. This approach was used in the Bloch representation (for non-degenerate bands) by Adams (1953) and Kane (1959). The operator $R$ in equation (5.19) is chosen such that $[U, S] + [H^{(0)}, R] = 0$ (Adams 1953, Kane 1959):

$$\langle n|R|n'\rangle = \frac{\mathrm{i}e\mathcal{E}_\alpha \pi^\alpha_{nn'}}{\hbar m \omega^2_{nn'}} \Theta_{nn'}. \tag{5.20}$$

The new Hamiltonian is therefore

$$\hat{H} = H^{(0)} + H^{(2)} + \tfrac{1}{2}[H^{(1)}, S] + \tfrac{1}{3}[[H^{(1)}, S], S] + \tfrac{1}{2}[[H^{(2)}, S], S]$$
$$+ U + W + \tfrac{1}{2}[[U, S], S] + \tfrac{1}{2}[[U, S], R] \tag{5.21}$$

where the identity $[U, R] = 0$ has been used. In addition, the term $[H^{(2)}, R]$ was discarded because its matrix elements are all off-block-diagonal terms of order $k\mathcal{E}$. The only unknown quantity in (5.21) is $\tfrac{1}{2}[[U, S], R]$:

$$\tfrac{1}{2}\langle n|[[U, S], R]|n'\rangle = -\frac{e^2 \mathcal{E}_\alpha \mathcal{E}_\beta}{2\hbar m^2} \sum_i{}' \frac{\pi^\alpha_{ni}\pi^\beta_{in'}}{\omega_{ni}\omega_{in'}}(\omega^{-1}_{ni} + \omega^{-1}_{n'i}). \tag{5.22}$$

From the structure of this equation, we see that the matrix representing (5.22) for $\Gamma_{15}$ valence electrons is again similar to the effective-mass Hamiltonian (3.8):

$$\tfrac{1}{2}[[U, S], R] = \begin{bmatrix} \lambda\mathcal{E}_x^2 + \mu(\mathcal{E}_y^2 + \mathcal{E}_z^2) & \nu\mathcal{E}_x\mathcal{E}_y & \nu\mathcal{E}_x\mathcal{E}_z \\ \nu\mathcal{E}_x\mathcal{E}_y & \lambda\mathcal{E}_y^2 + \mu(\mathcal{E}_x^2 + \mathcal{E}_z^2) & \nu\mathcal{E}_y\mathcal{E}_z \\ \nu\mathcal{E}_x\mathcal{E}_z & \nu\mathcal{E}_y\mathcal{E}_z & \lambda\mathcal{E}_z^2 + \mu(\mathcal{E}_x^2 + \mathcal{E}_y^2) \end{bmatrix} \tag{5.23}$$

in which $\lambda$, $\mu$, and $\nu$ are real constants. In the three-band model one has

$$\lambda = -e^2 P^2/E_0^3$$
$$\mu = -e^2 Q^2/E_0'^3$$
$$\nu = -e^2(P^2/E_0^3 + Q^2/E_0'^3).$$

The Hamiltonian (5.21) is the Luttinger–Kohn Hamiltonian for degenerate or quasi-degenerate bands in an electric field. It is valid through terms of order $k^3$, $k\mathcal{E}$, and $\mathcal{E}^2$. Notice that there are no dipole terms in this Hamiltonian. The only place that such terms appear is in equation (5.16), which has no matrix elements connecting two states in the same quasi-degenerate set. This conclusion is completely general; it holds for crystals of any symmetry and for any value of the expansion point $\boldsymbol{k}_0$.

There is, however, another possible source of matrix elements linear in $\mathcal{E}$ and independent of $k$: the relativistic term $W^{(0)}$ shown in equation (5.7). The operator $\boldsymbol{p} \times \boldsymbol{\sigma}$ transforms as a vector under rotations, has odd parity, and is invariant with respect to time reversal and translations. Hence, $(\boldsymbol{p} \times \boldsymbol{\sigma})_{nn'}$ has the same symmetry properties as the crystal coordinate matrix $\boldsymbol{\xi}_{nn'}$, and one may consider $W^{(0)}$ to be an 'effective' dipole interaction even though its physical origin is different from that of a true dipole coupling. In practical terms, the main difference is quantitative: since $W^{(0)}$ is relativistic in origin, its magnitude is much smaller than the non-relativistic coupling in equation (5.16). Since the latter is already a small perturbation, one should be safe in neglecting $W^{(0)}$.

Therefore, in the Luttinger–Kohn Hamiltonian, there is no non-relativistic dipole interaction between degenerate or quasi-degenerate states. The only mechanism that can lead to a non-relativistic coupling of order $\mathcal{E}^1 k^0$ is a lattice deformation, which gives rise to deformation-potential terms in the Hamiltonian (Bir and Pikus 1974, Blacha *et al* 1984). The theory of such effects is well established and will not be discussed further here.

In addition to the Hamiltonian, one is often interested in knowing the form of the momentum operator, which is needed for calculating optical transition rates in the $\boldsymbol{A} \cdot \boldsymbol{p}$ gauge. The matrix elements of $\boldsymbol{p}$ in the pseudo-Bloch bases $\{|\overline{n\boldsymbol{k}}\rangle\}$ and $\{|\widehat{n\boldsymbol{k}}\rangle\}$ are the same as the matrix elements of $\bar{\boldsymbol{p}} = \mathrm{e}^{-S}\boldsymbol{p}\,\mathrm{e}^{S}$ and $\hat{\boldsymbol{p}} = \mathrm{e}^{-R}\bar{\boldsymbol{p}}\,\mathrm{e}^{R}$ in the original basis $\{|n\boldsymbol{k}\rangle\}$. To first order in $k$ and $\mathcal{E}$, one has

$$\hat{\boldsymbol{p}} = \boldsymbol{p} + [\boldsymbol{p}, S] + [\boldsymbol{p}, R]. \tag{5.24}$$

The matrix elements of $\boldsymbol{p}$ were given already in equation (4.31):

$$\langle n|p^\alpha|n'\rangle = p_{nn'}^\alpha + \hbar k_\alpha \delta_{nn'} \tag{5.25}$$

while those of $[\boldsymbol{p}, S]$ and $[\boldsymbol{p}, R]$ may be calculated from equations (5.12) and (5.20):

$$\langle n|[p^\alpha, S]|n'\rangle = \frac{k_\beta}{m}\sideset{}{'}\sum_i\left(\frac{p_{ni}^\alpha \pi_{in'}^\beta}{\omega_{n'i}} + \frac{\pi_{ni}^\beta p_{in'}^\alpha}{\omega_{ni}}\right) \tag{5.26}$$

$$\langle n|[p^\alpha, R]|n'\rangle = \frac{\mathrm{i}e\mathcal{E}_\beta}{\hbar m}\sideset{}{'}\sum_i\left(\frac{p_{ni}^\alpha \pi_{in'}^\beta}{\omega_{n'i}^2} - \frac{\pi_{ni}^\beta p_{in'}^\alpha}{\omega_{ni}^2}\right). \tag{5.27}$$

The matrix elements of primary interest are those coupling the $\Gamma_1$ conduction ($n = S$) and $\Gamma_{15}$ valence ($n' = X$, $Y$, or $Z$) states of a zinc-blende semiconductor with no spin–orbit coupling. The symmetry properties of these states (Parmenter 1955, Dresselhaus 1955) can be used to write $\langle S|\hat{\boldsymbol{p}}|X\rangle$ as

$$\frac{\hbar}{\mathrm{i}m}\langle S|\hat{\boldsymbol{p}}|X\rangle = \hat{\boldsymbol{x}}P + \hat{\boldsymbol{y}}(D\mathcal{E}_z - \mathrm{i}Bk_z) + \hat{\boldsymbol{z}}(D\mathcal{E}_y - \mathrm{i}Bk_y) \tag{5.28}$$

with $\langle S|\hat{\boldsymbol{p}}|Y\rangle$ and $\langle S|\hat{\boldsymbol{p}}|Z\rangle$ obtained from cyclic permutations of $x$, $y$, and $z$. The constants in equation (5.28) are real, and are defined by $P = (\hbar/\mathrm{i}m)p_{SX}^x$ and

$$B = \frac{\hbar}{m^2}\sideset{}{'}\sum_i p_{Si}^y p_{iX}^z(\omega_{Xi}^{-1} + \omega_{Si}^{-1}) \tag{5.29}$$

$$D = \frac{e}{m^2}\sideset{}{'}\sum_i p_{Si}^y p_{iX}^z(\omega_{Xi}^{-2} - \omega_{Si}^{-2}) \tag{5.30}$$

in which the summation covers $\Gamma_{15}$ states. In the three-band model one has

$$B = P'Q(1/E_0' + 1/E_1)$$
$$D = eP'Q(1/E_1^2 - 1/E_0'^2)$$

where $E_1 = E_0' - E_0$.

To appreciate the significance of equation (5.28), it is helpful to rewrite this expression in terms of the rotated coordinates $x' = (x + y)/\sqrt{2}$, $y' = (-x + y)/\sqrt{2}$, and $z' = z$:

$$\frac{\hbar}{\mathrm{i}m}\langle S|\hat{\boldsymbol{p}}|X'\rangle = \hat{\boldsymbol{x}}'(P + D\mathcal{E}_z - \mathrm{i}Bk_z) + \hat{\boldsymbol{z}}(D\mathcal{E}_{x'} - \mathrm{i}Bk_{x'})$$
$$\frac{\hbar}{\mathrm{i}m}\langle S|\hat{\boldsymbol{p}}|Y'\rangle = \hat{\boldsymbol{y}}'(P - D\mathcal{E}_z + \mathrm{i}Bk_z) - \hat{\boldsymbol{z}}(D\mathcal{E}_{y'} - \mathrm{i}Bk_{y'}) \tag{5.31}$$
$$\frac{\hbar}{\mathrm{i}m}\langle S|\hat{\boldsymbol{p}}|Z\rangle = \hat{\boldsymbol{z}}P + \hat{\boldsymbol{x}}'(D\mathcal{E}_{x'} - \mathrm{i}Bk_{x'}) - \hat{\boldsymbol{y}}'(D\mathcal{E}_{y'} - \mathrm{i}Bk_{y'}).$$

These equations exhibit clearly the field-induced optical anisotropy (i.e., the Pockels effect) mentioned in the introduction. For example, if $\mathcal{E} \parallel [001]$, the optical transition rate is obviously different for light polarized along the $x'$- and $y'$-axes. To lowest order, the anisotropy is linear in $\mathcal{E}$. Both $B$ and $D$ contribute to the Pockels effect, but the contribution from $B$ is several orders of magnitude larger than that from $D$ (Foreman 2000b).

## 6. Discussion of other models in the literature

From the discussion in section 2 it should be clear that any Hamiltonian based on the dipole matrix element of a unit cell is fundamentally incorrect. However, in practice such a model may or may not yield valid results, depending on the details of how it is constructed and used. The possibility of obtaining valid results from an invalid Hamiltonian arises from a serendipitous cancellation of errors: if one uses the invalid identity (2.14) for the matrix elements of the coordinate operator, one obtains (accidentally) the correct relation (3.7) for the matrix elements of the crystal coordinate in the Bloch representation. Thus, if one works in the Bloch representation, one could in principle obtain valid predictions from a Hamiltonian that was derived incorrectly.

However, in practice most work on semiconductors is performed using some variant of the Luttinger–Kohn $\boldsymbol{k} \cdot \boldsymbol{p}$ or effective-mass theory. As discussed in section 5 (see equations (5.9) and (5.19)), this formalism is based on a *pseudo*-Bloch representation. Although the pseudo-Bloch functions are the same as Bloch functions for non-degenerate bands (to within a $\boldsymbol{k}$-dependent phase factor), they are quite different for degenerate or quasi-degenerate bands. Consequently, one cannot expect to obtain correct results by using a Bloch coordinate matrix in the Luttinger–Kohn theory.

For example, a number of authors have proposed $\boldsymbol{k} \cdot \boldsymbol{p}$ models in which states of different energy are coupled simultaneously by the $\boldsymbol{k} \cdot \boldsymbol{p}$ interaction and a dipole matrix element (Aspnes 1975, Lommer *et al* 1985, 1988, Zhu and Chang 1994). That this is incorrect may be seen immediately from equations (5.3), (5.6), and (5.16). In the original Luttinger–Kohn basis $\{|n\boldsymbol{k}\rangle\}$, the potential energy of the field (5.6) is strictly diagonal and there are no dipole terms. After one transforms to the pseudo-Bloch basis $\{|\overline{n\boldsymbol{k}}\rangle\}$, the $\boldsymbol{k} \cdot \boldsymbol{p}$ coupling (5.3) is eliminated and replaced by the dipole coupling (5.16). There is no basis in which (5.3) and (5.16) appear simultaneously.

More recently, it has been proposed that the Luttinger–Kohn Hamiltonian for degenerate valence electrons in zinc-blende semiconductors should contain a dipole mixing term (Khurgin 1994, Krebs and Voisin 1996, Khurgin and Voisin 1997, 1998). The most detailed description of this theory is given by Khurgin and Voisin (1997, 1998), who provide two versions of their Hamiltonian: one that is to be used in the Luttinger–Kohn representation†, and one that is to be used in the Bloch representation‡. The latter is obtained from the former by applying Kane's (1957) direction-dependent unitary transformation.

This model is derived from the dipole matrix element of a unit cell; therefore, the derivation cannot be valid. Furthermore, as shown in section 5 (see especially the discussion below equation (5.23)), there are no circumstances in which a non-relativistic dipole interaction between degenerate bands exists in the Luttinger–Kohn representation (Foreman 1999, 2000a, b). Hence, the Khurgin–Voisin model is fundamentally inconsistent with the theory of Luttinger and Kohn.

The Khurgin–Voisin model also disagrees with the known properties of the coordinate

---

† See equation (9) of Khurgin and Voisin (1997) and the paragraph above equation (6) of Khurgin and Voisin (1998).
‡ See equation (30) of Khurgin and Voisin (1997) and equation (6) of Khurgin and Voisin (1998).

operator in the Bloch representation (Blount 1962). As shown in equations (3.12) and (3.13), if $\mathcal{E}$ and $k$ both lie in the $z$ direction, all off-diagonal matrix elements $e\mathcal{E} \cdot \xi_{nn'}(k)$ vanish at the degeneracy point $k = 0$, since

$$\lim_{k_z \to 0} \xi_{nn'}^z(0, 0, k_z) = 0 \qquad (n \neq n') \tag{6.1}$$

where $n$ and $n'$ are equal to $X$, $Y$, or $Z$. However, the mixing in the Bloch version of the Khurgin–Voisin Hamiltonian does not vanish in this case. Therefore, their Hamiltonian does not correctly describe the angular dependence of the field-induced mixing in the Bloch representation. In addition, the Bloch version of their Hamiltonian omits the singular matrix elements shown in equation (3.11). For both of these reasons, the Khurgin–Voisin model is invalid in the Bloch representation.

In response to this criticism (Foreman 2000a), Khurgin and Voisin (2000) have argued that the theory of Luttinger and Kohn is 'irrelevant' to the question of whether a dipole mixing term should appear in the Luttinger–Kohn Hamiltonian, because the original paper (Luttinger and Kohn 1955) was applied to crystals with inversion symmetry—hence lacking the Pockels effect—and did not proceed beyond terms of order $k^2$ in the Hamiltonian. This response is unsatisfactory because the work of Luttinger and Kohn (1955) establishes clearly a general procedure that may be extended to arbitrary order (Kohn 1959) and applied to crystals of any symmetry (Dresselhaus 1955, Kane 1957). When this extension is performed (Foreman 2000b), one finds unequivocally that the proposed dipole mixing effect does not exist.

This does not imply that there is no electronic mechanism for the Pockels effect in the Luttinger–Kohn representation; it merely implies that the actual mechanism is different from the one proposed by Khurgin, Krebs, and Voisin (Khurgin 1994, Krebs and Voisin 1996, Khurgin and Voisin 1997, 1998). The results derived in section 5 provide evidence for three mechanisms related to the inversion asymmetry of the zinc-blende structure: the terms of order $k^3$ in the Hamiltonian (3.8), and the coefficients $B$ and $D$ in the momentum matrix (5.31). None of these mechanisms was considered in the work of Khurgin, Krebs, and Voisin.

In their response, Khurgin and Voisin (2000) provide an alternative derivation of their model that is based on $k \cdot p$ theory†. The derivation consists of a calculation of the coordinate matrix element (2.10) using the identity (2.14). Since the matrix element (2.10) is physically meaningless and the identity (2.14) is invalid, their calculation is obviously on very shaky ground. However, due to the cancellation of errors mentioned above, the result that they obtain does have some significance—provided that one interprets it as the crystal coordinate matrix element (3.7). This interpretation is adopted in what follows.

Khurgin and Voisin (2000) do not mention what representation they are using, but equation (3.7) is valid only in the Bloch representation. Hence, their derivation has no bearing at all on the Hamiltonian in the Luttinger–Kohn representation‡. One cannot construct a Hamiltonian by grafting together pieces from different representations; one must use a single representation self-consistently, as was done in sections 4 and 5 for the Luttinger–Kohn Hamiltonian.

Khurgin and Voisin's (2000) derivation is also of very limited relevance in the Bloch representation. What they show is that the matrix element $\xi_{XY}^z(k_x, 0, 0)$ is non-zero near the degeneracy point $k = 0$:

$$\lim_{k_x \to 0} \xi_{XY}^z(k_x, 0, 0) \neq 0 \tag{6.2}$$

---

† Strangely enough, in their original articles (Khurgin and Voisin 1997, 1998) these authors stated quite emphatically that $k \cdot p$ theory cannot be used to describe the Pockels effect!

‡ It is important to remember that the zero-field eigenkets of the Luttinger–Kohn Hamiltonian do not form a basis for the Luttinger–Kohn representation; they form a basis for the *Bloch* representation.

which was shown above in equation (3.10). This result is valid, but their interpretation of it is not. They interpret this limiting value as 'the' matrix element of $z$, implying that the value obtained is unique and unambiguous. It is not. If one approaches the degeneracy point along the $k_z$-axis rather than the $k_x$-axis, one obtains a limit of zero:

$$\lim_{k_z \to 0} \xi_{XY}^z (0, 0, k_z) = 0 \qquad (6.3)$$

as shown in equation (3.12). Hence, the crystal coordinate matrix is a non-analytic, direction-dependent function that does not possess a unique limit at $k = 0$. It is incorrect and misleading to treat this quantity as a simple constant.

Thus, although the property (6.2) is certainly suggestive, this property alone is not sufficient to justify the Khurgin–Voisin model. As shown above, most other features of their model are at variance with the known behaviour of the Bloch representation near $k = 0$.

## 7. Conclusions

During the course of this review, no theme has featured more prominently than the importance of careful attention to the basic definitions and theorems of quantum-mechanical representation theory. The representations of primary interest have been the Bloch and Luttinger–Kohn representations. In both formalisms there is a clearly defined and unambiguous procedure for handling the coordinate and momentum operators. Neither theory makes any reference to the coordinate matrix element of a unit cell. Thus, in spite of their obvious intuitive appeal, such matrix elements have no place in either Hamiltonian. Indeed, since they depend on the arbitrary choice of unit cell, they have no place in *any* well-defined Hamiltonian.

The Bloch representation has a long history of success for problems involving non-degenerate bands in external fields. However, for degenerate bands, the existence of singularities in the coordinate matrix has led most authors to abandon the Bloch representation and search for other ways of treating such problems (Wannier (1960), appendix I; Roth (1962); Blount (1962), p 319; Aspnes and Bottka (1972), p 501). These singularities lead directly to a poor localization of Wannier functions (Blount 1962, Zak 1985), and much effort has gone into finding alternative definitions for Wannier functions with better localization properties (des Cloizeaux 1963, 1964a, b, Kohn 1973, Zak 1985). There is of course nothing preventing one from working in the Bloch representation and tackling the singularities by brute force, but to the author's knowledge this approach has never been used in the literature.

The Luttinger–Kohn formalism eliminates these singularities by choosing basis states that are analytic functions of $k$. Consequently, virtually all theoretical work on degenerate bands in external fields is performed using some variant of the Luttinger–Kohn representation. Although the theory has been in use for nearly half a century, some of its features are only just now coming to light; for example, the existence of electric-field-induced mixing at the Brillouin zone boundary was unknown a year ago (Burt 1999). At present there are no clear physical examples in which this effect is important, but it could play a role in the theoretical description of $\Gamma$–$X$ mixing experiments for valence electrons.

However, the primary utility of the Luttinger–Kohn representation is for experiments that sample only a small region in $k$-space. In this case, zone-boundary mixing may be neglected, and the leading non-relativistic mixing terms arising from an electric field are of order $k\mathcal{E}$ and $\mathcal{E}^2$. The only terms of order $\mathcal{E}$ are small terms of relativistic origin, which would be negligible under most circumstances. These conclusions are valid for crystals of any symmetry and at any point in the Brillouin zone.

Thus, even though the concept of a dipole interaction between degenerate states seems plausible at first glance, it cannot be reconciled with the Luttinger–Kohn representation. This

does not rule out the existence of such an interaction in some other representation; however, at present there exists no alternative to the Luttinger–Kohn theory for the analytical description of degenerate bands in semiconductors.

## Acknowledgments

## Appendix A. Blount's (1962) approach to the coordinate operator

A useful alternative form of the wave-function expansion (4.14) is

$$\varphi(\boldsymbol{x}) = \int_{\Omega^*} \alpha(\boldsymbol{k}, \boldsymbol{x}) \, \mathrm{d}^3 k \tag{A.1}$$

where the function $\alpha$ is defined by

$$\alpha(\boldsymbol{k}, \boldsymbol{x}) = \sum_n \varphi_n(\boldsymbol{k}) \chi_{nk}(\boldsymbol{x}). \tag{A.2}$$

One can use equations (4.2) and (4.18) to express $\alpha$ as

$$\alpha(\boldsymbol{k}, \boldsymbol{x}) = (2\pi)^{-3/2} \sum_G \varphi(\boldsymbol{k} + \boldsymbol{k}_0 + \boldsymbol{G}) \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{k}_0+\boldsymbol{G})\cdot\boldsymbol{x}}. \tag{A.3}$$

This shows that $\alpha$ is periodic in $\boldsymbol{k}$:

$$\alpha(\boldsymbol{k} + \boldsymbol{G}, \boldsymbol{x}) = \alpha(\boldsymbol{k}, \boldsymbol{x}). \tag{A.4}$$

The integrals (A.1) and (4.14) are therefore independent of the choice of unit cell. Since $\alpha$ is periodic, one can expand it in a Fourier series:

$$\alpha(\boldsymbol{k}, \boldsymbol{x}) = \sum_R \alpha_R(\boldsymbol{x}) \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{k}_0)\cdot\boldsymbol{R}}$$

$$\alpha_R(\boldsymbol{x}) = (\Omega^*)^{-1} \int_{\Omega^*} \alpha(\boldsymbol{k}, \boldsymbol{x}) \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}+\boldsymbol{k}_0)\cdot\boldsymbol{R}} \, \mathrm{d}^3 k \tag{A.5}$$

where $\boldsymbol{R}$ is a Bravais lattice vector in coordinate space. The Fourier coefficients $\alpha_R(\boldsymbol{x})$ are calculated by inserting (A.3) into (A.5), which yields

$$\alpha_R(\boldsymbol{x}) = (\Omega^*)^{-1} \varphi(\boldsymbol{x} - \boldsymbol{R}) \tag{A.6}$$

and

$$\alpha(\boldsymbol{k}, \boldsymbol{x}) = (\Omega^*)^{-1} \sum_R \varphi(\boldsymbol{x} - \boldsymbol{R}) \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{k}_0)\cdot\boldsymbol{R}}. \tag{A.7}$$

Substitution of (A.7) into the integrand of (A.1) then yields the expected result $\varphi(\boldsymbol{x})$.

One can now evaluate the action of the coordinate operator in the Luttinger–Kohn representation by taking the gradient of (A.2):

$$\nabla_k \alpha(\boldsymbol{k}, \boldsymbol{x}) = \mathrm{i}\boldsymbol{x}\alpha(\boldsymbol{k}, \boldsymbol{x}) + \sum_n \chi_{nk}(\boldsymbol{x}) \, \nabla_k \varphi_n(\boldsymbol{k}). \tag{A.8}$$

The gradient of a periodic function is just another periodic function, but with an average value of zero. Hence, the integral of (A.8) over a unit cell is zero:

$$\int_{\Omega^*} \nabla_k \alpha(\boldsymbol{k}, \boldsymbol{x}) \, \mathrm{d}^3 k = 0. \tag{A.9}$$

To demonstrate this more explicitly, one can use equation (A.7) to obtain

$$\nabla_k \alpha(k, x) = i(\Omega^*)^{-1} \sum_R R\varphi(x - R)e^{i(k+k_0)\cdot R} \tag{A.10}$$

which implies

$$\int_{\Omega^*} \nabla_k \alpha(k, x) \, d^3k = i \sum_R R\varphi(x - R)\delta_{R,0} = 0. \tag{A.11}$$

This result may also be obtained by using Gauss's theorem (Arfken (1985), p 59):

$$\int_{\Omega^*} \nabla_k \alpha(k, x) \, d^3k = \int_{\Omega^*} \alpha(k, x) \, dS = 0 \tag{A.12}$$

where $dS$ is an element of area normal to the surface enclosing the unit cell $\Omega^*$. The surface integral vanishes due to the periodicity property (A.4). Thus, the integral of equation (A.8) over a unit cell gives

$$x\varphi(x) = \sum_n \int_{\Omega^*} \chi_{nk}(x)[i\,\nabla_k\varphi_n(k)] \, d^3k \tag{A.13}$$

which is the result obtained previously in equation (4.33). The right-hand side of (A.13) may be evaluated using equations (A.2) and (A.7):

$$\sum_n \int_{\Omega^*} \chi_{nk}(x)[i\,\nabla_k\varphi_n(k)] \, d^3k = \int_{\Omega^*} e^{ik\cdot x} i\,\nabla_k[e^{-ik\cdot x}\alpha(k, x)] \, d^3k$$

$$= (\Omega^*)^{-1} \int_{\Omega^*} \sum_R (x - R)\varphi(x - R)e^{i(k+k_0)\cdot R} \, d^3k$$

$$= \sum_R (x - R)\varphi(x - R)\delta_{R,0} = x\varphi(x) \tag{A.14}$$

thus reproducing equation (A.13). The second line shows clearly that the integrand is periodic in $k$, so (A.13), like (4.14), does not depend on the choice of unit cell.

## References

Adams E N 1952 *Phys. Rev.* **85** 41–50
Adams E N 1953 *J. Chem. Phys.* **21** 2013–17
Aleĭner I L and Ivchenko E L 1992 *Pis. Zh. Eksp. Teor. Fiz.* **55** 662–4 (Engl. Transl. 1992 *JETP Lett.* **55** 692–5)
Arfken G 1985 *Mathematical Methods for Physicists* 3rd edn (San Diego, CA: Academic)
Ashcroft N W and Mermin N D 1976 *Solid State Physics* (Philadelphia, PA: Saunders)
Aspnes D E 1975 *Phys. Rev.* B **12** 2297–310
Aspnes D E and Bottka N 1972 *Semiconductors and Semimetals* vol 9, ed R K Willardson and A C Beer (New York: Academic) pp 457–543
Bastard G 1988 *Wave Mechanics Applied to Semiconductor Heterostructures* (New York: Wiley)
Bir G L and Pikus G E 1974 *Symmetry and Strain-Induced Effects in Semiconductors* (New York: Wiley)
Blacha A, Presting H and Cardona M 1984 *Phys. Status Solidi* b **126** 11–36
Bloch F 1928 *Z. Phys.* **52** 555–600
Blount E I 1962 *Solid State Physics* vol 13, ed F Seitz and D Turnbull (New York: Academic) pp 305–73
Burt M G 1988 *Semicond. Sci. Technol.* **3** 739–53
Burt M G 1992 *J. Phys.: Condens. Matter* **4** 6651–90
Burt M G 1993 *J. Phys.: Condens. Matter* **5** 4091–8
Burt M G 1995a *Superlatt. Microstruct.* **17** 335–8
Burt M G 1995b *Semicond. Sci. Technol.* **10** 412–15
Burt M G 1999 *J. Phys.: Condens. Matter* **11** R53–83
Burt M G and Foreman B A 1999 *Proc. 24th Int. Conf. on the Physics of Semiconductors* ed D Gershoni (Singapore: World Scientific) CD-ROM paper 119

Callaway J 1991 *Quantum Theory of the Solid State* 2nd edn (San Diego, CA: Academic)

Cardona M, Christensen N E and Fasol G 1988 *Phys. Rev.* B **38** 1806–27

Chuang S L 1995 *Physics of Optoelectronic Devices* (New York: Wiley)

Coles R A, Abram R A, Brand S and Burt M G 1999 *Phys. Rev.* B **60** 13 306–9

Corzine S W, Yan R H and Coldren L A 1993 *Quantum Well Lasers* ed P S Zory (San Diego, CA: Academic) pp 17–96

des Cloizeaux J 1963 *Phys. Rev.* **129** 554–66

des Cloizeaux J 1964a *Phys. Rev.* **135** A685–97

des Cloizeaux J 1964b *Phys. Rev.* **135** A698–707

Dresselhaus G 1955 *Phys. Rev.* **100** 580–6

Dresselhaus G, Kip A F and Kittel C 1955 *Phys. Rev.* **98** 368–84

Fiore A, Rosencher E, Vinter B, Weill D and Berger V 1995 *Phys. Rev.* B **51** 13 192–7

Foreman B A 1996 *Phys. Rev.* B **54** 1909–21

Foreman B A 1998 *Phys. Rev. Lett.* **81** 425–8

Foreman B A 1999 *Phys. Rev. Lett.* **82** 1339

Foreman B A 2000a *Phys. Rev. Lett.* **84** 4513

Foreman B A 2000b *Phys. Rev. Lett.* **84** 2505–8

Gelfand I M and Shilov G E 1964 *Generalized Functions* (New York: Academic)

Haug H and Koch S W 1994 *Quantum Theory of the Optical and Electronic Properties of Semiconductors* 3rd edn (Singapore: World Scientific)

Houston W V 1940 *Phys. Rev.* **57** 184–6

Ivchenko E L, Kaminskiĭ A Yu and Aleĭner I L 1993 *Zh. Eksp. Teor. Fiz.* **104** 3401–15 (Engl. Transl. 1993 *JETP* **77** 609–16)

Ivchenko E L, Kaminski A Yu and Rössler U 1996 *Phys. Rev.* B **54** 5852–8

Ivchenko E L, Toropov A A and Voisin P 1998 *Fiz. Tverd. Tela* **40** 1925–31 (Engl. Transl. 1998 *Phys. Solid State* **40** 1748–53)

Jones H and Zener C 1934 *Proc. R. Soc.* A **144** 101–17

Kane E O 1957 *J. Phys. Chem. Solids* **1** 249–61

Kane E O 1959 *J. Phys. Chem. Solids* **12** 181–8

Kane E O 1982 *Handbook on Semiconductors* vol 1, ed W Paul (Amsterdam: North-Holland) pp 193–217

Khurgin J B 1988 *Phys. Rev.* B **38** 4056–66

Khurgin J B 1994 *J. Opt. Soc. Am.* B **11** 2492–501

Khurgin J B 1999 *Semiconductors and Semimetals* vol 59, ed E Garmire and A Kost (San Diego, CA: Academic) pp 1–82

Khurgin J B and Voisin P 1997 *Semicond. Sci. Technol.* **12** 1378–87

Khurgin J B and Voisin P 1998 *Phys. Rev. Lett.* **81** 3777–80

Khurgin J B and Voisin P 2000 *Phys. Rev. Lett.* **84** 4514

Kittel C and Mitchell A H 1954 *Phys. Rev.* **96** 1488–93

Kjeldaas T and Kohn W 1957 *Phys. Rev.* **105** 806–13

Kohn W 1959 *Phys. Rev.* **115** 1460–78

Kohn W 1973 *Phys. Rev.* B **7** 4388–98

Krebs O, Rondi D, Gentner J L, Goldstein L and Voisin P 1998 *Phys. Rev. Lett.* **80** 5770–3

Krebs O and Voisin P 1996 *Phys. Rev. Lett.* **77** 1829–32

Kwok S H, Grahn H T, Ploog K and Merlin R 1992 *Phys. Rev. Lett.* **69** 973–6

Lax M 1974 *Symmetry Principles in Solid State and Molecular Physics* (New York: Wiley) section 7.8

Lifshitz E M and Pitaevskiĭ L P 1980 *Statistical Physics* Part 2 (Oxford: Pergamon) section 55

Lommer G, Malcher F and Rössler U 1985 *Phys. Rev.* B **32** 6965–7

Lommer G, Malcher F and Rössler U 1988 *Phys. Rev. Lett.* **60** 728–31

Luttinger J M and Kohn W 1955 *Phys. Rev.* **97** 869–83

Magri R and Ossicini S 1998 *Phys. Rev.* B **58** R1742–5

Martin R M 1974 *Phys. Rev.* B **9** 1998–9

Mendez E E, Agulló-Rueda F and Hong J M 1988 *Phys. Rev. Lett.* **60** 2426–9

Merzbacher E 1998 *Quantum Mechanics* 3rd edn (New York: Wiley)

Nenciu G 1991 *Rev. Mod. Phys.* **63** 91–127

Parmenter R H 1955 *Phys. Rev.* **100** 573–9

Platonov A V, Kochereshko V P, Ivchenko E L, Mikhailov G V, Yakovlev D R, Keim M, Ossau W, Waag A and Landwehr G 1999 *Phys. Rev. Lett.* **83** 3546–9

Resta R 1994 *Rev. Mod. Phys.* **66** 899–915

Rosencher E, Fiore A, Vinter B, Berger V, Bois P and Nagle J 1996 *Science* **271** 168–73

Roth L M 1962 *J. Phys. Chem. Solids* **23** 433–46

Shankar R 1994 *Principles of Quantum Mechanics* 2nd edn (New York: Plenum)

Shockley W 1950 *Phys. Rev.* **78** 173–4

Singh J 1993 *Physics of Semiconductors and their Heterostructures* (New York: McGraw-Hill)

Szmulowicz F 1995 *Phys. Rev.* B **51** 1613–23

Tagantsev A K 1987 *Usp. Fiz. Nauk* **152** 423–48 (Engl. Transl. 1987 *Sov. Phys.–Usp.* **30** 588–603)

Tagantsev A K 1991 *Phase Transitions* **35** 119–203

Voisin P, Bleuse J, Bouche C, Gaillard S, Alibert C and Regreny A 1988 *Phys. Rev. Lett.* **61** 1639–42

Wang S 1989 *Fundamentals of Semiconductor Theory and Device Physics* (Englewood Cliffs, NJ: Prentice-Hall)

Wannier G H 1960 *Phys. Rev.* **117** 432–9

Wannier G H 1962 *Rev. Mod. Phys.* **34** 645–55

Weisbuch C and Vinter B 1991 *Quantum Semiconductor Structures* (San Diego, CA: Academic)

Yafet Y 1957 *Phys. Rev.* **106** 679–84

Yu P Y and Cardona M 1997 *Fundamentals of Semiconductors* (Berlin: Springer)

Zak J 1972 *Solid State Physics* vol 27, ed H Ehrenreich *et al* (New York: Academic) pp 1–62

Zak J 1985 *Phys. Rev. Lett.* **54** 1075–8

Zak J 1991 *Phys. Rev.* B **43** 4519–20

Zener C 1934 *Proc. R. Soc.* A **145** 523–9

Zhu B F and Chang Y C 1994 *Phys. Rev.* B **50** 11 932–48