Bivariational methods applied to Schrodinger's and Dirac's equations

1982 J. Phys. A: Math. Gen. 15 3083

(http://iopscience.iop.org/0305-4470/15/10/014)

J. Phys. A: Math. Gen. 15 (1982) 3083–3107. Printed in Great Britain

Bivariational methods applied to Schrödinger's and Dirac's equations

G Wachutka and H Bross

Sektion Physik der Universität München, and Sonderforschungsbereich 128, D-8000 Munich, Federal Republic of Germany

Received 26 January 1982

Abstract. A class of bivariational functionals is derived whose stationary points are pairs of solutions of the single-particle Schrödinger equation (or Dirac equation, respectively) subject to so-called 'complementary boundary conditions'. The formulation of the boundary value problem is sufficiently general to include matching conditions and Bloch conditions as well as scattering conditions. It is shown how bivariational translational techniques may be applied to problems with three- and two-dimensional translational symmetry (calculation of complex band structures and propagation matrices, scattering problems).

1. Introduction

Many methods for the calculation of solutions of the Schrödinger equation or Dirac equation are based on variational principles. If, for instance, the Hamiltonian H can be considered as a symmetric operator defined on a Hilbert space $(\mathscr{H}, \langle\,|\,\rangle)$, it is well known that the 'energy functional' $\varphi \mapsto \langle\varphi|(H-E)\varphi\rangle$ is stationary for solutions of $(H-E)\varphi = 0$. In practical applications, however, the boundary conditions imposed on φ are often such that the formal differential operator H cannot be assigned a domain $D(H)$ such that it becomes a symmetric operator. But in this situation it is also well known (Courant and Hilbert 1924) that for certain boundary conditions it is possible to modify the energy functional by adding 'boundary functionals' so that variational methods may be applied as in the symmetric case. The essential idea is that the boundary conditions are not included in the definition of the domain of H but arise as part of the Euler equation (so-called 'natural boundary conditions').

Nevertheless, there are problems with boundary conditions which cannot be treated in this way (for example, scattering problems, or the calculation of propagation matrices). Here, as an appropriate technique, 'bivariational methods' may be used; their feature is to construct a functional $(\chi, \varphi) \mapsto J(\chi, \varphi)$ such that the derivative with respect to χ yields the equation $(H-E)\varphi = 0$ and the natural boundary conditions for φ, while the derivative with respect to φ implies a second ('auxiliary') equation $(H-E)\chi = 0$ and 'complementary' natural boundary conditions for χ. Further boundary conditions ('essential' boundary conditions) are imposed on χ and φ by the definition of the domain of J.

For special applications, both boundary functionals of a single variable and bivariational expressions have already been derived (Kohn 1948, 1952, Leigh 1956, Bevensee 1961, Schlosser and Marcus 1963, Marcus 1967, Ferreira *et al* 1974, 1975, Sarker and Taj-ul Islam 1975, Lopez-Aguilar 1979). It seems useful to present a comprehensive class of such bivariational functionals which includes the special cases and provides new variational approaches for problems which have not yet been treated in practice; § 4 deals with some of them. The explicit construction of the bivariational functionals and the formulation of the pertinent variational principle are given in § 3; the required definitions and notations are introduced and illustrated in § 2. Mathematical reasoning is postponed to the appendices as far as possible.

2. Formulation of the boundary value problem

Consider the Hamiltonians

$$H=-\Delta+w \qquad \text{(Schrödinger theory; units: } \hbar=1,\ m=\tfrac{1}{2}) \tag{1_S}$$

or

$$H=-\mathrm{i}\boldsymbol{\alpha}\boldsymbol{\nabla}+\beta+w \qquad \text{(Dirac theory; units: } \hbar=c=m=1) \tag{1_D}$$

respectively†.

It is required to find a pair (χ,φ) of functions which satisfy

$$H\varphi=E\varphi \qquad \text{and} \qquad H\chi=E\chi \qquad \text{(with real energy } E) \tag{2}$$

in a region Ω of three-dimensional space whose boundary $\partial\Omega$ can be represented as the union of a piecewise smooth‡ surface F and pairs of piecewise smooth surfaces $F_{k,0}$ and $F_{k,1}$ $(k=1,\ldots,M)$

$$\partial\Omega=F\cup\bigcup_{k=1}^{M}(F_{k,0}\cup F_{k,1}) \tag{3}$$

with $F_{k,0}$ being parallel to and displaced from $F_{k,1}$ by a displacement vector $\boldsymbol{T}_k$:

$$F_{k,0}=\boldsymbol{T}_k+F_{k,1} \qquad \text{for } k=1,\ldots,M. \tag{4}$$

The union of boundary surfaces $\bigcup_{k=1}^{M}F_{k,1}$ is denoted by S_t.

In the case where the potential w has three-dimensional translational symmetry with lattice G_3, Ω can be identified with the three-dimensional Wigner–Seitz cell WSZ_3; in this case $F=\varnothing$ and $\boldsymbol{T}_k$ belongs to G_3. As a further case, one may think of a potential w with two-dimensional translational symmetry with layer lattice G_2; then Ω will be a 'column' $WSZ_2\times(z_\mathrm{L},z_\mathrm{R})$ whose cross sections parallel to the layer lattice are two-dimensional Wigner–Seitz cells such that $\boldsymbol{T}_k\in G_2$. Here, F consists of those parts of $\partial\Omega$ which are contained in the planes $z=z_\mathrm{L}$ and $z=z_\mathrm{R}$ (see figure 1(*a*)).

In scattering problems with a localised potential (Kohn 1948), Ω can be chosen as a large sphere: then $F_{k,0}=\varnothing$ and $F_{k,1}=\varnothing$ (i.e. $F=\partial\Omega$).

† (i) Schrödinger's and Dirac's theories are treated simultaneously here and in the following sections; subscripts 'S' and 'D' are used to denote corresponding relationships. (ii) The potential w may be local $(w(\boldsymbol{r}))$ or non-local $(w((\boldsymbol{r},\boldsymbol{r}')))$ in principle, but for simplification the notation $w(\boldsymbol{r})$ is used throughout because the generalisation to a non-local w is obvious. (iii) For the Dirac matrices the representation

$$\boldsymbol{\alpha}=\begin{pmatrix}0&1\\1&0\end{pmatrix}\otimes\boldsymbol{\sigma}; \qquad \beta=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\otimes I_2$$

is used ($\boldsymbol{\sigma}$ = Pauli matrices; $I_2=2\times2$ unit matrix).

‡ For the meaning of 'piecewise smooth' refer to appendix 1.

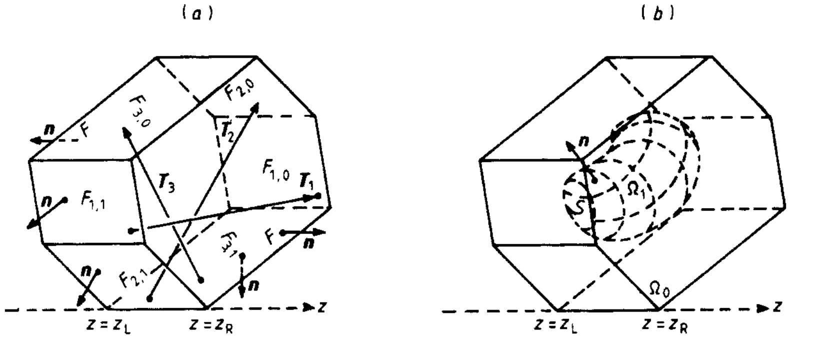

Figure 1. (*a*) Boundary of the region Ω in the case of planar translational symmetry. (*b*) Inner partition of Ω.

It is further supposed that the region Ω is partitioned into the parts $\Omega_0 = \bigcup_{j=1}^{q_0} \Omega_0^{(j)}$ and $\Omega_1 = \bigcup_{j=1}^{q_1} \Omega_1^{(j)}$ where $\Omega_i^{(j)}$ are bounded domains with piecewise smooth boundaries such that $\partial\Omega_i = \bigcup_{j=1}^{q_i} \partial\Omega_i^{(j)}$ holds and Ω_0 and Ω_1 are separated by a piecewise smooth surface $S = \partial\Omega_0 \cap \partial\Omega_1$ (figure 1(*b*)); the potential w is assumed to be continuous within Ω_0 and Ω_1. For example, in the case of a 'warped muffin-tin potential' Ω_1 represents the union of all muffin-tin spheres $K(\boldsymbol{r}_j; s_j)$ within Ω (such that $w(\boldsymbol{r}) = V_{\text{sph}}(|\boldsymbol{r} - \boldsymbol{r}_j|)$ for $|\boldsymbol{r} - \boldsymbol{r}_j| \leqslant s_j$), or S can be thought of as being a plane surface separating two adjacent layers of a layered crystal structure (figure 2).

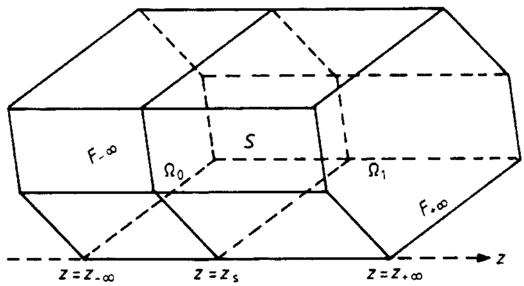

Figure 2. Region Ω in the case of scattering at $z = z_s$.

The solutions of equation (2) have to be elements of an appropriate Hilbert space V and they are subject to certain boundary conditions on $S \cup \partial\Omega$ which can be expressed conveniently using the following notations. For any domain $\tilde{\Omega}$ let $L_2(\tilde{\Omega}, \mathbb{C}^n)$ denote the set of square-integrable functions $\psi : \tilde{\Omega} \to \mathbb{C}^n$ and let $W_2^{(k)}(\tilde{\Omega})$ denote the Sobolev space consisting of those functions $\psi \in L_2(\tilde{\Omega}, \mathbb{C}^n)$ for which all partial derivatives (in the distribution sense) $D^l\psi := \partial^{|l|}\psi / \partial x_1^{l_1} \partial x_2^{l_2} \partial x_3^{l_3}$ with $|l| := l_1 + l_2 + l_3 \leqslant k$ belong to $L_2(\tilde{\Omega}, \mathbb{C}^n)$. Equipped with the inner product $(\cdot|\cdot)_{k,\tilde{\Omega}}$ defined as $(\chi | \varphi)_{k,\tilde{\Omega}} := \Sigma_{|l| \leqslant k} \int_{\tilde{\Omega}} D^l\chi^+ D^l\varphi \, \mathrm{d}^3 r$, $W_2^{(k)}(\tilde{\Omega})$ is a Hilbert space.

The Hilbert space V of which the solutions of equation (2) shall be elements is defined as

$$V := \{\psi : \Omega \to \mathbb{C};\ \psi|_{\Omega_i^{(j)}} \in W_2^{(2)}(\Omega_i^{(j)}) \text{ for } i = 0, 1; j = 1, \ldots, q_i\} \qquad (5_S)$$

with inner product $(\chi|\varphi) := \Sigma_{i=0,1} \Sigma_{j=1}^{q_i} (\chi|\varphi)_{2,\Omega_i^{(j)}}$,

$$V := \{\psi : \Omega \to \mathbb{C}^4;\ \psi|_{\Omega_i^{(j)}} \in W_2^{(1)}(\Omega_i^{(j)}) \text{ for } i = 0, 1; j = 1, \ldots, q_i\} \qquad (5_D)$$

with inner product $(\chi|\varphi) := \Sigma_{i=0,1} \Sigma_{j=1}^{q_i} (\chi|\varphi)_{1,\Omega_i^{(j)}}$ ($\psi|_{\tilde{\Omega}}$ denotes the restriction of ψ to $\tilde{\Omega}$).

Note that the Hilbert space V contains all functions which solve equation (2) in Ω_0 and Ω_1 in the classical sense and have a continuous probability density and probability current in $\bar{\Omega}$.

Following Nečas (1967, ch 2, § 4 and ch 3, § 1) a superficial measure σ can be defined along $\partial\Omega_0 \cup \partial\Omega_1$; so for any measurable part $\Gamma \subset \partial\Omega_0 \cup \partial\Omega_1$ the space $L_2(\Gamma)$ is well defined. The unit normal

$$\boldsymbol{n}(\boldsymbol{r}) := \begin{cases} \text{outward normal to region } \Omega_1 & \text{for } \boldsymbol{r} \in S \\ \text{outward normal to region } \Omega & \text{for } \boldsymbol{r} \in \partial\Omega \backslash S \end{cases} \tag{6}$$

exists almost everywhere along $\partial\Omega_0 \cup \partial\Omega_1$. The subspace $\mathscr{E} := \{\psi \in V;\ \psi|_{\Omega_i^{(j)}} \in \mathscr{C}^\infty(\bar{\Omega}_i^{(j)})$ for $i = 0, 1;\ j = 1, \ldots, q_i\}$ is dense in V (with respect to the norm $\|\cdot\| := \sqrt{(\cdot|\cdot)}$). For $\psi \in \mathscr{E}$ and $\alpha = 0, 1$ it makes sense to define almost everywhere:

$$\partial_n^\alpha \psi_1(\boldsymbol{r}) := \begin{cases} \lim\limits_{\substack{\boldsymbol{r}' \to \boldsymbol{r} \\ \boldsymbol{r}' \in \Omega_1}} (\boldsymbol{n}(\boldsymbol{r})\boldsymbol{\nabla})^\alpha \psi(\boldsymbol{r}') & \text{for } \boldsymbol{r} \in S \\ \lim\limits_{\boldsymbol{r}' \to \boldsymbol{r}} (\boldsymbol{n}(\boldsymbol{r})\boldsymbol{\nabla})^\alpha \psi(\boldsymbol{r}') & \text{for } \boldsymbol{r} \in S_t \end{cases} \tag{7a}$$

$$\partial_n^\alpha \psi_0(\boldsymbol{r}) := \begin{cases} \lim\limits_{\substack{\boldsymbol{r}' \to \boldsymbol{r} \\ \boldsymbol{r}' \in \Omega_0}} (\boldsymbol{n}(\boldsymbol{r})\boldsymbol{\nabla})^\alpha \psi(\boldsymbol{r}') & \text{for } \boldsymbol{r} \in S \\ \lim\limits_{\boldsymbol{r}' \to \boldsymbol{r} + \boldsymbol{T}_k} (\boldsymbol{n}(\boldsymbol{r})\boldsymbol{\nabla})^\alpha \psi(\boldsymbol{r}') & \text{for } \boldsymbol{r} \in F_{k,1} \text{ and } F_{k,0} = \boldsymbol{T}_k + F_{k,1}\dagger \end{cases} \tag{7b}$$

$$\partial_n^\alpha \psi(\boldsymbol{r}) := \lim_{\boldsymbol{r}' \to \boldsymbol{r}} (\boldsymbol{n}(\boldsymbol{r})\boldsymbol{\nabla})^\alpha \psi(\boldsymbol{r}') \qquad \text{for } \boldsymbol{r} \in F. \tag{7c}$$

As shown by Nečas (1967), the mappings $\mathscr{E} \ni \psi \mapsto \partial_n^\alpha \psi_i \in L_2(S \cup S_t)$ and $\mathscr{E} \ni \psi \mapsto \partial_n^\alpha \psi \in L_2(F)$ are continuous; hence there exist unique linear continuous extensions on V (so-called 'trace operators'). Since no confusion arises these extensions are denoted by the same symbols $\partial_n^\alpha \psi_i$ and $\partial_n^\alpha \psi$; it is said '$\psi \in V$ takes on its boundary values in the sense of traces'.

The boundary conditions to which the solutions $(\chi, \varphi) \in V \times V$ of equation (2) are subject are divided into two groups. The first group consists of the 'natural' boundary conditions defined by

$$\left.\begin{aligned} \partial_n^\alpha \varphi_0(\boldsymbol{r}) &= t(\boldsymbol{r}) \partial_n^\alpha \varphi_1(\boldsymbol{r}) \\ \partial_n^\alpha \chi_1(\boldsymbol{r}) &= t(\boldsymbol{r})^* \partial_n^\alpha \chi_0(\boldsymbol{r}) \end{aligned}\right. \qquad \text{for } \boldsymbol{r} \in S \cup S_t \tag{8}$$

($\alpha = 0, 1$: Schrödinger theory; $\alpha = 0$: Dirac theory). Here, $t(\boldsymbol{r})$ is a given complex-valued 'transfer function', bounded almost everywhere (i.e. $t \in L_\infty(S \cup S_t)$), and chosen appropriately to the problem. For example, $t(\boldsymbol{r}) := 1$ means that χ and φ are continuously differentiable (or continuous, respectively) at $\boldsymbol{r} \in S$, if they belong to the

† i.e. S_t is used as parameter space for the arguments of ψ and its normal derivative on $\bigcup_{k=1}^{M} F_{k,0}$; see figure 1($a$).

subspace $V \cap \mathscr{C}^1(\bar{\Omega}_0) \cap \mathscr{C}^1(\bar{\Omega}_1)$ (or $V \cap \mathscr{C}^0(\bar{\Omega}_0) \cap \mathscr{C}^0(\bar{\Omega}_1)$, respectively); in the case of a periodic potential $t(\boldsymbol{r}) := \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{T}_l)$ for $\boldsymbol{r} \in F_{l,1}$ is equivalent to a Bloch condition†.

The second group comprises the 'essential' boundary conditions compelling χ and φ to take the following form on the surface F. Let f_0 and g_0 be given square-integrable functions $F \to \mathbb{C}^n$ (with $n = 2$ in Schrödinger's theory and $n = 4$ in Dirac's) and let $U_0(F)$ and $W_0(F)$ be linear spaces of square-integrable functions $F \to \mathbb{C}^n$ such that

$$\int_F g^+ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} f \,\mathrm{d}\sigma = \int_F (g^{(1)*} f^{(0)} - g^{(0)*} f^{(1)}) \,\mathrm{d}\sigma = 0 \tag{9$_\mathrm{S}$}$$

for any $f = (f^{(0)}, f^{(1)})^\mathrm{T} \in U_0(F)$ and $g = (g^{(0)}, g^{(1)})^\mathrm{T} \in W_0(F)$ or

$$\int_F g^+ \boldsymbol{\alpha n} f \,\mathrm{d}\sigma = 0 \qquad \text{for any } f \in U_0(F) \text{ and } g \in W_0(F), \tag{9$_\mathrm{D}$}$$

respectively. Then, writing $f_0 = (f_0^{(0)}, f_0^{(1)})^\mathrm{T}$ and $g_0 = (g_0^{(0)}, g_0^{(1)})^\mathrm{T}$ in the case of Schrödinger's theory, χ and φ are subject to

$$\partial_n^\alpha \varphi(\boldsymbol{r}) = f_0^{(\alpha)}(\boldsymbol{r}) + f^{(\alpha)}(\boldsymbol{r}) \qquad \text{for } \boldsymbol{r} \in F \text{ and } \alpha = 0, 1 \tag{10$_\mathrm{S}$}$$

$$\partial_n^\alpha \chi(\boldsymbol{r}) = g_0^{(\alpha)}(\boldsymbol{r}) + g^{(\alpha)}(\boldsymbol{r}) \tag{11$_\mathrm{S}$}$$

or in Dirac's theory

$$\varphi(\boldsymbol{r}) = f_0(\boldsymbol{r}) + f(\boldsymbol{r}) \qquad \text{for } \boldsymbol{r} \in F \tag{10$_\mathrm{D}$}$$

$$\chi(\boldsymbol{r}) = g_0(\boldsymbol{r}) + g(\boldsymbol{r}) \tag{11$_\mathrm{D}$}$$

where $f \in U_0(F)$ and $g \in W_0(F)$ are arbitrary.

Since the trace operators are surjective mappings $W_2^{(k)}(\tilde{\Omega}) \ni \psi \mapsto \partial_n^\alpha \psi \in W_2^{(k-\alpha-1/2)}(\partial\tilde{\Omega})$‡ the essential boundary conditions only make sense if the data f_0, g_0, $U_0(F)$, $W_0(F)$ are such that

$$\begin{aligned} &f_0 + U_0(F) \cap W_2^{(3/2)}(F) \times W_2^{(1/2)}(F) \neq \varnothing, \\ &g_0 + W_0(F) \cap W_2^{(3/2)}(F) \times W_2^{(1/2)}(F) \neq \varnothing, \end{aligned} \tag{12$_\mathrm{S}$}$$

or

$$\begin{aligned} &f_0 + U_0(F) \cap W_2^{(1/2)}(F)^4 \neq \varnothing, \\ &g_0 + W_0(F) \cap W_2^{(1/2)}(F)^4 \neq \varnothing, \end{aligned} \tag{12$_\mathrm{D}$}$$

respectively.

Boundary conditions of the kind (10), (11) appear in scattering problems (for example, see Kohn 1948); here f_0 and g_0 stand for the incoming waves, while f and g represent the scattered waves (and their normal derivatives on F, respectively). The computation of propagation matrices (Marcus and Jepsen 1968, Bross 1977) also leads to boundary conditions of the form (10), (11) with an appropriate definition of f_0, g_0, $U_0(F)$ and $W_0(F)$ as will be shown in § 4.1.

It must be emphasised that it is not sufficient to pose the boundary value problem only for a single function φ, if variational techniques are to be applied, but it is

† See Kohn (1952), for example.

‡ See Nečas (1967, ch 2, theorems 5.5 and 5.8); the definition of the Sobolev spaces $W_2^{(P/2)}(\partial\tilde{\Omega})$ is also given there.

necessary to consider pairs (χ, φ) of functions with 'complementary' boundary conditions (i.e. the transfer functions of χ and φ are related by $t \leftrightarrow (t^*)^{-1}$, and the 'free' parts $f \in U_0(F)$ and $g \in W_0(F)$ satisfy (9), the so-called 'compatibility condition') to ensure the existence of a variational expression $J(\chi, \varphi)$ whose stationary points are solutions of equation (2) with (8) as natural boundary conditions and (10), (11) as essential boundary conditions in the usual sense of the calculus of variations. Except for some special boundary conditions, a functional of a single variable $J(\varphi)$ would not yield these properties (Arthurs 1980, theorem 3.1.1).

3. Variational expressions

We now derive a class of variational expressions $J: W \times U \to \mathbb{C}$, $(\chi, \varphi) \mapsto J(\chi, \varphi)$ which become stationary only for solutions of the boundary value problem posed above. Consider the 'energy functional'

$$I: W \times U \to \mathbb{C}, \qquad I(\chi, \varphi) := \int_{\Omega_0 \cup \Omega_1} \chi^+ (H - E)\varphi \, \mathrm{d}^3 r \tag{13}$$

defined on the affine submanifolds of V

$$\begin{aligned} U &:= \{\varphi \in V; \varphi \text{ satisfies equation (10)}\} = \varphi_0 + U_0, \\ W &:= \{\chi \in V; \chi \text{ satisfies equation (11)}\} = \chi_0 + W_0, \end{aligned} \tag{14}$$

with some (arbitrarily chosen) $\varphi_0 \in U$ and $\chi_0 \in W$ (the existence of which is ensured by condition (12)) and subspaces of V

$$U_0 := \{\varphi \in V; (\varphi|_F, \partial_n \varphi|_F)^{\mathrm{T}} \in U_0(F)\}, \qquad W_0 := \{\chi \in V; (\chi|_F, \partial_n \chi|_F)^{\mathrm{T}} \in W_0(F)\}, \tag{15$_\mathrm{S}$}$$

or

$$U_0 := \{\varphi \in V; \varphi|_F \in U_0(F)\}, \qquad W_0 := \{\chi \in V; \chi|_F \in W_0(F)\}, \tag{15$_\mathrm{D}$}$$

respectively. Here $\psi|_F$ means the restriction of the trace of ψ to F. By the definition (15) the trial functions χ and φ are subject to the 'essential' boundary conditions only (i.e. their variations must be elements of W_0 and U_0), whereas the boundary conditions (8) are required to arise as part of the Euler–Lagrange equations (i.e. they can be derived from $\delta J = 0$). It is well known (Arthurs 1980, after theorem 2.3.1) that this can be accomplished by adding an appropriate 'boundary functional' K to I. To derive the general form of K, the first variation δI is calculated for $\chi \in W$, $\varphi \in U$, $\delta\chi \in W_0$ and $\delta\varphi \in U_0$, assuming that χ and φ are solutions of (2) and (8). Hence

$$\delta I = \int_{S \cup S_t} [\chi_0^* (\partial_n \delta\varphi_0 - t\, \partial_n \delta\varphi_1) - \partial_n \chi_0^* (\delta\varphi_0 - t\delta\varphi_1)] \, \mathrm{d}\sigma + \int_F (g_0^{(1)*} \delta\varphi - g_0^{(0)*} \, \partial_n \delta\varphi) \, \mathrm{d}\sigma, \tag{16$_\mathrm{S}$}$$

$$\delta I = \mathrm{i} \int_{S \cup S_t} \chi_0^+ \boldsymbol{\alpha n} (\delta\varphi_0 - t\delta\varphi_1) \, \mathrm{d}\sigma - \mathrm{i} \int_F g_0^+ \boldsymbol{\alpha n} \delta\varphi \, \mathrm{d}\sigma. \tag{16$_\mathrm{D}$}$$

(Note the correspondence $(\psi_\mathrm{S}, \mathrm{i}\partial_n \psi_\mathrm{S}) \leftrightarrow (\psi_\mathrm{u}, \psi_\mathrm{d}) = \psi_\mathrm{D}$ between the Schrödinger wavefunction ψ_S and its first derivative and the 'large' and 'small' components ψ_u and ψ_d of the Dirac bispinor ψ_D; thus here and in the following expressions the corresponding term in Dirac's theory can be obtained from Schrödinger's theory by simply substituting

$$\mathrm{i}(\chi^*\partial_n\varphi-\partial_n\chi^*\varphi)=(\chi,\mathrm{i}\partial_n\chi)^*\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\varphi\\ \mathrm{i}\partial_n\varphi\end{pmatrix}\mapsto -\chi^+\boldsymbol{\alpha n}\varphi).$$

Since $J:=I+K$ is required to satisfy $\delta J=0$ (i.e. $\delta K=-\delta I$), equation (16) suggests K is a sesquilinear form

$$K(\chi,\varphi)=\int_{S\cup S_t}\int_{S\cup S_t}\sum_{p,q=0,1}\sum_{\alpha,\beta=0,1}\partial_n^\alpha\chi_p(\boldsymbol{r})^*k_{pq}^{\alpha\beta}(\boldsymbol{r},\boldsymbol{r}')\partial_n^\beta\varphi_q(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r})$$
$$+\int_F\int_F\sum_{\alpha,\beta=0,1}\partial_n^\alpha\chi(\boldsymbol{r})^*k^{\alpha\beta}(\boldsymbol{r},\boldsymbol{r}')\partial_n^\beta\varphi(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r}) \qquad (17_\mathrm{S})$$
$$+\int_F[A(\boldsymbol{r})^*\varphi(\boldsymbol{r})+B(\boldsymbol{r})^*\partial_n\varphi(\boldsymbol{r})+\chi(\boldsymbol{r})^*C(\boldsymbol{r})+\partial_n\chi(\boldsymbol{r})^*D(\boldsymbol{r})]\,\mathrm{d}\sigma(\boldsymbol{r}),$$

$$K(\chi,\varphi)=\int_{S\cup S_t}\int_{S\cup S_t}\sum_{p,q=0,1}\chi_p(\boldsymbol{r})^+k_{pq}(\boldsymbol{r},\boldsymbol{r}')\varphi_q(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r})$$
$$+\int_F\int_F\chi(\boldsymbol{r})^+k(\boldsymbol{r},\boldsymbol{r}')\varphi(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r})$$
$$+\int_F[A(\boldsymbol{r})^+\varphi(\boldsymbol{r})+\chi(\boldsymbol{r})^+C(\boldsymbol{r})]\,\mathrm{d}\sigma(\boldsymbol{r}), \qquad (17_\mathrm{D})$$

with some suitable kernels $k_{pq}^{\alpha\beta}$, $k^{\alpha\beta}$, k_{pq}, k, A, B, C, D. Evaluating the first variation δK for solutions $(\chi,\varphi)\in W\times U$ of the equations (2) and (8) and for arbitrary variations $(\delta\chi,\delta\varphi)\in W_0\times U_0$ and setting $\delta I=-\delta K$, we obtain the result

$$K(\chi,\varphi)=\int_{S\cup S_t}\int_{S\cup S_t}\{[\partial_n\chi_0(\boldsymbol{r})^*(\delta(\hat{r},\hat{r}')-t(\boldsymbol{r})\mu(\hat{r},\hat{r}'))$$
$$+\partial_n\chi_1(\boldsymbol{r})^*\mu(\hat{r},\hat{r}')][\varphi_0(\boldsymbol{r}')-t(\boldsymbol{r}')\varphi_1(\boldsymbol{r}')]$$
$$-[\chi_0(\boldsymbol{r})^*(\delta(\hat{r},\hat{r}')-t(\boldsymbol{r})\lambda(\hat{r},\hat{r}'))+\chi_1(\boldsymbol{r})^*\lambda(\hat{r},\hat{r}')][\partial_n\varphi_0(\boldsymbol{r}')-t(\boldsymbol{r}')\partial_n\varphi_1(\boldsymbol{r}')]$$
$$+[\chi_0(\boldsymbol{r})^*t(\boldsymbol{r})-\chi_1(\boldsymbol{r})^*]\rho(\hat{r},\hat{r}')[\varphi_0(\boldsymbol{r}')-t(\boldsymbol{r}')\varphi_1(\boldsymbol{r}')]$$
$$+[\partial_n\chi_0(\boldsymbol{r})^*t(\boldsymbol{r})-\partial_n\chi_1(\boldsymbol{r})^*]\tau(\hat{r},\hat{r}')[\partial_n\varphi_0(\boldsymbol{r}')-t(\boldsymbol{r}')\partial_n\varphi_1(\boldsymbol{r}')]\}\,\mathrm{d}\sigma(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r})$$
$$+\tfrac{1}{2}\int_F[(\chi(\boldsymbol{r})^*+g_0^{(0)}(\boldsymbol{r})^*)(\partial_n\varphi(\boldsymbol{r})-f_0^{(1)}(\boldsymbol{r}))-(\partial_n\chi(\boldsymbol{r})^*+g_0^{(1)}(\boldsymbol{r})^*)$$
$$\times(\varphi(\boldsymbol{r})-f_0^{(0)}(\boldsymbol{r}))]\,\mathrm{d}\sigma(\boldsymbol{r}) \qquad (18_\mathrm{S})$$

or

$$K(\chi,\varphi)=-\mathrm{i}\int_{S\cup S_t}\int_{S\cup S_t}[\chi_0(\boldsymbol{r})^+(\boldsymbol{\alpha n}(\boldsymbol{r})\delta(\hat{r},\hat{r}')-t(\boldsymbol{r})\rho(\hat{r},\hat{r}'))+\chi_1(\boldsymbol{r})^+\rho(\hat{r},\hat{r}')]$$
$$\times[\varphi_0(\boldsymbol{r}')-t(\boldsymbol{r}')\varphi_1(\boldsymbol{r}')]\,\mathrm{d}\sigma(\boldsymbol{r}')\,\mathrm{d}\sigma(\boldsymbol{r})$$
$$+\tfrac{1}{2}\mathrm{i}\int_F[\chi(\boldsymbol{r})^++g_0(\boldsymbol{r})^+]\boldsymbol{\alpha n}(\boldsymbol{r})[\varphi(\boldsymbol{r})-f_0(\boldsymbol{r})]\,\mathrm{d}\sigma(\boldsymbol{r}), \qquad (18_\mathrm{D})$$

respectively, where λ, μ, ρ, τ: $(S \cup S_t) \times (S \cup S_t) \to \mathbb{C}$ (or $\rho: (S \cup S_t) \times (S \cup S_t) \to \mathbb{C}^{(4,4)}$, respectively) are arbitrary† kernels which can be interpreted as weight functions, and $\delta(\hat{r}, \hat{r}')$ means the surface distribution $\int_{S \cup S_t} \delta(\hat{r}, \hat{r}')\psi(\boldsymbol{r}')\,\mathrm{d}\sigma(\hat{r}') := \psi(\boldsymbol{r})$ for $\boldsymbol{r} \in S \cup S_t$ while vanishing elsewhere.

Note that the volume integrals of the functional J may be altered using Gauss' integral theorem. For example, J can be represented in a symmetric form as‡

$$\begin{aligned} J(\chi, \varphi) = & \int_{\Omega_0 \cup \Omega_1} [\nabla \chi^* \nabla \varphi + (w - E)\chi^* \varphi]\,\mathrm{d}^3 r \\ & + \int_{S \cup S_t} [(\chi_0^* t - \chi_1^*)\partial_n \varphi_1 + \partial_n \chi_0^* (\varphi_0 - t\varphi_1)]\,\mathrm{d}\sigma \\ & + \int_{S \cup S_t} \int_{S \cup S_t} (\chi_0(\boldsymbol{r})^* t(\boldsymbol{r}) - \chi_1(\boldsymbol{r})^*, \partial_n \chi_0(\boldsymbol{r})^* t(\boldsymbol{r}) - \partial_n \chi_1(\boldsymbol{r})^*) \\ & \times \begin{bmatrix} \rho(\hat{r}, \hat{r}'), \lambda(\hat{r}, \hat{r}') \\ -\mu(\hat{r}, \hat{r}'), \tau(\hat{r}, \hat{r}') \end{bmatrix} \begin{bmatrix} \varphi_0(\boldsymbol{r}') - t(\boldsymbol{r}')\varphi_1(\boldsymbol{r}') \\ \partial_n \varphi_0(\boldsymbol{r}') - t(\boldsymbol{r}')\partial_n \varphi_1(\boldsymbol{r}') \end{bmatrix} \mathrm{d}\sigma'\,\mathrm{d}\sigma \\ & - \tfrac{1}{2} \int_F [(\partial_n \chi^* + g_0^{(1)*})(\varphi - f_0^{(0)}) + (\chi^* - g_0^{(0)*})(\partial_n \varphi + f_0^{(1)})]\,\mathrm{d}\sigma + \text{constant} \end{aligned} \tag{19$_\mathrm{S}$}$$

or

$$\begin{aligned} J(\chi, \varphi) = & \tfrac{1}{2} \int_{\Omega_0 \cup \Omega_1} \chi^+ (H - E)\varphi\,\mathrm{d}^3 r + \tfrac{1}{2} \int_{\Omega_0 \cup \Omega_1} [(H - E)\chi]^+ \varphi\,\mathrm{d}^3 r \\ & - \tfrac{1}{2}\mathrm{i} \int_{S \cup S_t} [(\chi_1^+ - t\chi_0^+)\boldsymbol{\alpha n}\varphi_1 + \chi_0^+ \boldsymbol{\alpha n} (\varphi_0 - t\varphi_1)]\,\mathrm{d}\sigma \\ & - \mathrm{i} \int_{S \cup S_t} \int_{S \cup S_t} [\chi_1(\boldsymbol{r})^+ - t(\boldsymbol{r})\chi_0(\boldsymbol{r})^+] \rho(\hat{r}, \hat{r}') [\varphi_0(\boldsymbol{r}') - t(\boldsymbol{r}')\varphi_1(\boldsymbol{r}')]\,\mathrm{d}\sigma'\,\mathrm{d}\sigma \\ & + \tfrac{1}{2}\mathrm{i} \int_F (g_0^+ \boldsymbol{\alpha n}\varphi - \chi^+ \boldsymbol{\alpha n} f_0)\,\mathrm{d}\sigma + \text{constant}, \end{aligned} \tag{19$_\mathrm{D}$}$$

respectively.

The variational principle for the boundary value problem posed above can be formulated in the following way.

(1) If $(\chi, \varphi) \in W \times U$ is a pair of solutions of equation (2) in $\Omega_0 \cup \Omega_1$ satisfying the natural boundary conditions (8), then $\delta J(\chi, \varphi, \delta\chi, \delta\varphi) = 0$ for any $(\delta\chi, \delta\varphi) \in W_0 \times U_0$.

(2) Conversely, if $(\chi, \varphi) \in W \times U$ satisfies $\delta J(\chi, \varphi, \delta\chi, \delta\varphi) = 0$ for any $(\delta\chi, \delta\varphi) \in W_0 \times U_0$, then (χ, φ) is a pair of solutions of equation (2) subject to the boundary conditions (8).

For the proof, see appendix 1; it is worth mentioning that the second statement is proven under the weaker assumption $\delta J(\chi, \varphi, \delta\chi, \delta\varphi) = 0$ for any $(\delta\chi, \delta\varphi) \in V_0 \times V_0$ with $V_0 := \{\psi \in V: \partial_n^\alpha \psi|_F = 0\}$ (Schrödinger theory: $\alpha = 0, 1$; Dirac theory: $\alpha = 0$).

† More precisely: the kernels must define linear operators $L_2(S \cup S_t) \to L_2(S \cup S_t)$; δ has to be interpreted as identity.

‡ The form (19$_\mathrm{S}$) indicates how the domain of J may be extended to functions $\psi: \Omega \to \mathbb{C}$; $\psi|_{\Omega_i^{(j)}} \in W_2^{(1)}(\Omega_i^{(j)})$ for $i = 0, 1$; $j = 1, \ldots, q_i$ if the boundary conditions are such that all normal derivatives (which are not well defined even in the sense of traces in this situation) can be eliminated from the functional J; see appendix 4.

The weight functions λ, μ, ρ, τ may be chosen arbitrarily in principle; if the variations $(\delta\chi, \delta\varphi)$ are allowed to vary in the whole space $W_0 \times U_0$, the result of $\delta J = 0$ will be independent of the weights. But if $(\delta\chi, \delta\varphi)$ are restricted to some submanifold of $W_0 \times U_0$ (for example, in numerical applications; see Marcus (1967), Ferreira *et al* (1974, 1975), Lopez-Aguilar (1979)), this is not true, and the choice of the weights will have an influence on the quality of approximate solutions. An analysis of the weight dependence of approximate solutions often leads to criteria which show how to choose the weights best.

Some conditions on the weight functions can be obtained by demanding that the Euler–Lagrange equations derived from $\delta J = 0$ should show the same symmetries as the initial boundary value problem. Since equations (2), (8), (9) and (10), (11) are invariant under the transformation

$$\begin{gathered} \varphi \mapsto \chi, \qquad \chi \mapsto \varphi, \qquad t \mapsto (t^*)^{-1}, \\ (U, U_0(F), f_0) \mapsto (W, W_0(F), g_0), \qquad (W, W_0(F), g_0) \mapsto (U, U_0(F), f_0), \end{gathered} \tag{20}$$

the functional J should transform into its complex conjugate under this operation. As shown in appendix 2 this condition is satisfied if and only if

$$\begin{gathered} \lambda(\hat{r}, \hat{r}')t(r') + t(r)^{-1}\mu(\hat{r}', \hat{r})^* = \delta(\hat{r}, \hat{r}'), \qquad \mu(\hat{r}, \hat{r}')t(r') + t(r)^{-1}\lambda(\hat{r}', \hat{r})^* = \delta(\hat{r}, \hat{r}'), \\ \rho(\hat{r}, \hat{r}')t(r') = t(r)^{-1}\rho(\hat{r}', \hat{r})^*, \qquad \tau(\hat{r}, \hat{r}')t(r') = t(r)^{-1}\tau(\hat{r}', \hat{r})^*, \end{gathered} \tag{21$_\mathrm{S}$}$$

or

$$\rho(\hat{r}, \hat{r}')t(r') + t(r)^{-1}\rho(\hat{r}', \hat{r})^+ = \boldsymbol{\alpha n}(\boldsymbol{r})\delta(\hat{r}, \hat{r}'), \tag{21$_\mathrm{D}$}$$

respectively.

Furthermore, if $t^{-1} \neq t^*$, λ and μ are uniquely determined by (21_S):

$$\lambda(\hat{r}, \hat{r}') = \mu(\hat{r}, \hat{r}') = (t(r) + (t(r)^*)^{-1})^{-1}\delta(\hat{r}, \hat{r}'). \tag{22$_\mathrm{S}$}$$

The correspondence between Schrödinger wavefunctions and Dirac bispinors mentioned above suggests that the 4×4 matrix $\rho(\hat{r}, \hat{r}')$ has the form†

$$\rho(\hat{r}, \hat{r}') = \begin{bmatrix} -\mathrm{i}\rho_\mathrm{S}(\hat{r}, \hat{r}')I_2, & \lambda_\mathrm{S}(\hat{r}, \hat{r}')\boldsymbol{\sigma n}(\boldsymbol{r}') \\ \boldsymbol{\sigma n}(\boldsymbol{r})\mu_\mathrm{S}(\hat{r}, \hat{r}'), & -\mathrm{i}\tau_\mathrm{S}(\hat{r}, \hat{r}')I_2 \end{bmatrix} \tag{23$_\mathrm{D}$}$$

where I_2 is the 2×2 unit matrix, $\boldsymbol{\sigma}$ are the Pauli matrices and $\lambda_\mathrm{S}, \mu_\mathrm{S}, \rho_\mathrm{S}, \tau_\mathrm{S}$ are complex valued. Thus a complete analogy between Schrödinger's and Dirac's theory can be achieved, because condition (21_D) becomes equivalent to (21_S) with $\lambda_\mathrm{S}, \mu_\mathrm{S}, \rho_\mathrm{S}, \tau_\mathrm{S}$ substituted for λ, μ, ρ, τ.

By the choice $\tau = 0$ and a change of the division of the boundary conditions into natural and essential ones, the domain of the non-relativistic functional J can be modified so as to include less regular functions of the class $W_2^{(1)}$ (see appendix 4).

4. Applications and comparison with other work

In this section, some examples are given which illustrate how well known methods for solving Schrödinger's equation (or Dirac's equation, respectively) can be derived

† Weights of the form (23_D) are used in the RAPW method; for example, see Loucks (1967, with $\lambda_\mathrm{S} = \mu_\mathrm{S} = \frac{1}{2}$; $\rho_\mathrm{S} = \tau_\mathrm{S} = 0$) or Sarker and Taj-ul Islam (1975, with $\lambda_\mathrm{S} = \rho_\mathrm{S} = 0$; $\mu_\mathrm{S} = 1$; $\tau_\mathrm{S} \in \mathbb{R}$).

from the variational principle stated above. The variational principle can be used to justify these methods as 'optimal' approximation techniques. The way in which these methods can be generalised, or new variational methods can be established, for problems subject to more general boundary conditions of the form (8) and (10), (11), is also shown.

4.1. Calculation of propagation matrices

Let w be a potential with two-dimensional translational symmetry, and let Ω be the 'column' $WSZ_2 \times (z_L, z_R)$ shown in figure 1(a). Set $F := F_L \cup F_R$, where $F_L(F_R)$ means that part of $\partial\Omega$ which is contained in the planes $z = z_L(z_R)$, and look for solutions of the following 'initial boundary value' problem.

Given an energy E and a two-dimensional propagation vector $\boldsymbol{k}_\parallel$ find a pair of functions (χ, φ) which satisfy equation (2) in Ω and equation (8) (with $t(\boldsymbol{r}) := \exp(\mathrm{i}\boldsymbol{k}_\parallel \boldsymbol{T}_l)$ for $\boldsymbol{r} \in F_{l,1}$ and $F_{l,0} = \boldsymbol{T}_l + F_{l,1}$) and the 'initial conditions'

$$\partial_n^\alpha \varphi(\boldsymbol{r}) = f_R^{(\alpha)}(\boldsymbol{r}) \text{ for } \boldsymbol{r} \in F_R \text{ and } \alpha = 0, 1, \qquad \partial_n^\alpha \chi(\boldsymbol{r}) = g_L^{(\alpha)}(\boldsymbol{r}) \text{ for } \boldsymbol{r} \in F_L \text{ and } \alpha = 0, 1, \tag{24_S}$$

or

$$\varphi(\boldsymbol{r}) = f_R(\boldsymbol{r}) \text{ for } \boldsymbol{r} \in F_R, \qquad \chi(\boldsymbol{r}) = g_L(\boldsymbol{r}) \text{ for } \boldsymbol{r} \in F_L, \tag{24_D}$$

respectively.

Of special interest is the value of $\partial_n^\alpha \varphi|_{F_L}$ and $\partial_n^\alpha \chi|_{F_R}$; i.e. it is required to find a pair of linear operators $P(z_L, z_R)$ and $\tilde{P}(z_R, z_L)$ ('propagation matrices') with†

$$\begin{aligned}
&P(z_L, z_R) : W_2^{(3/2)}(F_R) \times W_2^{(1/2)}(F_R) \to W_2^{(3/2)}(F_L) \times W_2^{(1/2)}(F_L), \\
&\begin{pmatrix} \varphi|_{F_L} \\ \partial_n \varphi|_{F_L} \end{pmatrix} = P(z_L, z_R) \begin{pmatrix} f_R^{(0)} \\ f_R^{(1)} \end{pmatrix}, \\
&\tilde{P}(z_R, z_L) : W_2^{(3/2)}(F_L) \times W_2^{(1/2)}(F_L) \to W_2^{(3/2)}(F_R) \times W_2^{(1/2)}(F_R), \\
&\begin{pmatrix} \chi|_{F_R} \\ \partial_n \chi|_{F_R} \end{pmatrix} = \tilde{P}(z_R, z_L) \begin{pmatrix} g_L^{(0)} \\ g_L^{(1)} \end{pmatrix},
\end{aligned} \tag{25_S}$$

or

$$\begin{aligned}
&P(z_L, z_R) : W_2^{(1/2)}(F_R)^4 \to W_2^{(1/2)}(F_L)^4, \qquad \varphi|_{F_L} = P(z_L, z_R) f_R, \\
&\tilde{P}(z_R, z_L) : W_2^{(1/2)}(F_L)^4 \to W_2^{(1/2)}(F_R)^4, \qquad \chi|_{F_R} = \tilde{P}(z_R, z_L) g_L,
\end{aligned} \tag{25_D}$$

respectively.

Since $(t^*)^{-1} = t$, the problem is posed symmetrically for χ and φ, and therefore $\tilde{P}(z_R, z_L) = P(z_L, z_R)^{-1}$ holds. The appropriate data f_0, g_0, $U_0(F)$ and $W_0(F)$ are:

$$\begin{aligned}
f_0(\boldsymbol{r}) &:= \begin{cases} (f_R^{(0)}(\boldsymbol{r}), f_R^{(1)}(\boldsymbol{r}))^T & \text{for } \boldsymbol{r} \in F_R, \\ 0 & \text{for } \boldsymbol{r} \in F_L, \end{cases} \\
g_0(\boldsymbol{r}) &:= \begin{cases} 0 & \text{for } \boldsymbol{r} \in F_R, \\ (g_L^{(0)}(\boldsymbol{r}), g_L^{(1)}(\boldsymbol{r}))^T & \text{for } \boldsymbol{r} \in F_L, \end{cases}
\end{aligned} \tag{26_S}$$

† See Marcus and Jepsen (1968) and Bross (1977).

or

$$f_0(\boldsymbol{r}) := \begin{cases} f_{\mathrm{R}}(\boldsymbol{r}) & \text{for } \boldsymbol{r} \in F_{\mathrm{R}}, \\ 0 & \text{for } \boldsymbol{r} \in F_{\mathrm{L}}, \end{cases}$$
$$g_0(\boldsymbol{r}) := \begin{cases} 0 & \text{for } \boldsymbol{r} \in F_{\mathrm{R}}, \\ g_{\mathrm{L}}(\boldsymbol{r}) & \text{for } \boldsymbol{r} \in F_{\mathrm{L}}, \end{cases} \tag{26$_\mathrm{D}$}$$

respectively and

$$\begin{aligned} U_0(F) &:= \{f: F \to \mathbb{C}^n \text{ square integrable}; f|_{F_{\mathrm{R}}} = 0\}, \\ W_0(F) &:= \{g: F \to \mathbb{C}^n \text{ square integrable}; g|_{F_{\mathrm{L}}} = 0\} \end{aligned} \tag{27}$$

(with $n = 2$ in Schrödinger's theory and $n = 4$ in Dirac's). By this choice of the data, condition (24) can be written in the form (10), (11), and the compatibility condition (9) is satisfied. The $\int_F$ term of the boundary functional K is given by

$$\int_{F_{\mathrm{L}}} (g_{\mathrm{L}}^{(0)*}\, \partial_n \varphi - g_{\mathrm{L}}^{(1)*} \varphi)\, \mathrm{d}\sigma, \tag{28$_\mathrm{S}$}$$

$$\mathrm{i} \int_{F_{\mathrm{L}}} g_{\mathrm{L}}^{+} \boldsymbol{\alpha n} \varphi \, \mathrm{d}\sigma. \tag{28$_\mathrm{D}$}$$

If f_{R} and g_{L} are chosen such that

$$\begin{pmatrix} g_{\mathrm{L}}^{(0)} \\ g_{\mathrm{L}}^{(1)} \end{pmatrix} = P(z_{\mathrm{L}}, z_{\mathrm{R}}) \begin{pmatrix} f_{\mathrm{R}}^{(0)} \\ f_{\mathrm{R}}^{(1)} \end{pmatrix} \tag{29$_\mathrm{S}$}$$

or

$$g_{\mathrm{L}} = P(z_{\mathrm{L}}, z_{\mathrm{R}}) f_{\mathrm{R}}, \tag{29$_\mathrm{D}$}$$

respectively, it is accomplished that $g_{\mathrm{L}}^{(\alpha)} = \partial_n^\alpha \varphi|_{F_{\mathrm{L}}}$ (or $g_{\mathrm{L}} = \varphi|_{F_{\mathrm{L}}}$, respectively) holds. It can be recognised from (28) that the physical meaning of the stationary value of $J(\chi, \varphi)$ for $\chi = \varphi$ satisfying (2), (8) and (24) is merely $\mathrm{i}\hbar$ times the probability current passing the boundary surface F_{L}.

The 'propagation matrix method' first proposed by Marcus and Jepsen (1968) can be understood as a special case to which the variational principle is applied. The true wavefunctions (χ, φ) are approximated by a finite planar Fourier series ('mixed representation'):

$$\hat{\varphi}(\boldsymbol{r}) = \sum_{j=1}^{N} u_j(z) \exp[\mathrm{i}(\boldsymbol{k}_\parallel + \boldsymbol{g}_j)\boldsymbol{r}_\parallel], \qquad \hat{\chi}(\boldsymbol{r}) = \sum_{j=1}^{N} v_j(z) \exp[\mathrm{i}(\boldsymbol{k}_\parallel + \boldsymbol{g}_j)\boldsymbol{r}_\parallel]. \tag{30}$$

Here $\boldsymbol{g}_j \in G_2^*$ = two-dimensional reciprocal lattice, $\boldsymbol{r} = (\boldsymbol{r}_\parallel, z)$, and $u_j, v_j \in V(z_{\mathrm{L}}, z_{\mathrm{R}})$ defined as

$$\begin{aligned} V(z_{\mathrm{L}}, z_{\mathrm{R}}) := \{&u: [z_{\mathrm{L}}, z_{\mathrm{R}}] \to \mathbb{C};\ u \text{ is twice continuously differentiable in } (z_{\mathrm{L}}, z_{\mathrm{R}}); \\ &u \text{ and } \mathrm{d}u/\mathrm{d}z \text{ possess a continuous continuation on } [z_{\mathrm{L}}, z_{\mathrm{R}}]; \\ &\mathrm{d}^2u/\mathrm{d}z^2 \text{ is square integrable}\}, \end{aligned} \tag{31$_\mathrm{S}$}$$

$$\begin{aligned} V(z_{\mathrm{L}}, z_{\mathrm{R}}) := \{&u: [z_{\mathrm{L}}, z_{\mathrm{R}}] \to \mathbb{C}^4;\ u \text{ is continuously differentiable in } (z_{\mathrm{L}}, z_{\mathrm{R}}) \\ &\text{and continuous on } [z_{\mathrm{L}}, z_{\mathrm{R}}];\ \mathrm{d}u/\mathrm{d}z \text{ is square integrable}\}. \end{aligned} \tag{31$_\mathrm{D}$}$$

It is convenient to define the column vectors

$$u(z) := \left(u_1(z), \ldots, u_N(z), \frac{\mathrm{d}u_1}{\mathrm{d}z}(z), \ldots, \frac{\mathrm{d}u_N}{\mathrm{d}z}(z)\right)^{\mathrm{T}} \in \mathbb{C}^{2N},$$
$$v(z) := \left(v_1(z), \ldots, v_N(z), \frac{\mathrm{d}v_1}{\mathrm{d}z}(z), \ldots, \frac{\mathrm{d}v_N}{\mathrm{d}z}(z)\right)^{\mathrm{T}} \in \mathbb{C}^{2N}, \qquad (32_{\mathrm{S}})$$

$$u(z) := (u_1(z), \ldots, u_N(z))^{\mathrm{T}} \in \mathbb{C}^{4N}, \qquad v(z) := (v_1(z), \ldots, v_N(z))^{\mathrm{T}} \in \mathbb{C}^{4N}; \qquad (32_{\mathrm{D}})$$

then the essential boundary conditions can be written as

$$u(z_{\mathrm{R}}) = u_{\mathrm{R}} \qquad \text{and} \qquad v(z_{\mathrm{L}}) = v_{\mathrm{L}}. \qquad (33)$$

The 'best' approximations $\hat{\chi}$ and $\hat{\varphi}$ have to satisfy $\delta J = 0$ under the restriction (33); the resulting Euler–Lagrange equations are

$$\mathrm{d}u/\mathrm{d}z = Au \qquad \text{and} \qquad \mathrm{d}v/\mathrm{d}z = Av \qquad (34)$$

with A being defined as

$$A := \begin{pmatrix} 0 & I_N \\ W & 0 \end{pmatrix}, \qquad W_{mn}(z) := \delta_{mn}(|\boldsymbol{k}_\parallel + \boldsymbol{g}_m|^2 - E) + w_{mn}(z), \qquad (35_{\mathrm{S}})$$

$$A := -\mathrm{i}(I_N \otimes \alpha_z) W, \qquad W_{mn}(z) := \delta_{mn}(\boldsymbol{\alpha}(\boldsymbol{k}_\parallel + \boldsymbol{g}_m) + \beta - E) + I_4 w_{mn}(z), \qquad (35_{\mathrm{D}})$$

where I_N denotes the $N \times N$ unit matrix, and $w_{mn}(z)$ are the planar Fourier coefficients

$$w_{mn}(z) = \frac{1}{|F|} \int_F w(\boldsymbol{r}_\parallel, z) \exp[\mathrm{i}(\boldsymbol{g}_n - \boldsymbol{g}_m)\boldsymbol{r}_\parallel] \,\mathrm{d}^2 r_\parallel \qquad (36)$$

($|F|$ means the area of the two-dimensional unit mesh of G_2). Equations (34) and (33) define an initial value problem, which is solved by the fundamental system $\hat{P}(z, z_0)$ satisfying

$$\partial_z \hat{P}(z, z_0) = A(z)\hat{P}(z, z_0), \qquad \hat{P}(z_0, z_0) = I_{2N} \text{ (or } I_{4N}\text{, respectively)}. \qquad (37)$$

Then $u(z) = \hat{P}(z, z_{\mathrm{R}})u_{\mathrm{R}}$ and $v(z) = \hat{P}(z, z_{\mathrm{L}})v_{\mathrm{L}}$, i.e. $\hat{P}(z_{\mathrm{L}}, z_{\mathrm{R}})$ is a finite-dimensional best approximation of the operator $P(z_{\mathrm{L}}, z_{\mathrm{R}})$.

Note that the method of Marcus and Jepsen (1968) is derived here without truncating an infinite system of differential equations, but as a necessary condition for $\delta J = 0$ on a submanifold of $W \times U$ and is hence justified as an approximation procedure. Furthermore, the variational expression can be used to show that the so-called 'matching condition' postulated by Marcus and Jepsen

$$u(z_0 + 0) = u(z_0 - 0) \qquad \text{and} \qquad v(z_0 + 0) = v(z_0 - 0) \qquad (38)$$

for any $z_0 \in (z_{\mathrm{L}}, z_{\mathrm{R}})$ where A is not continuous arises as a natural boundary condition for $\delta J = 0$ (see appendix 3).

It must be emphasised that (38) generally does not hold for the planar Fourier transforms $\tilde{v} = \mathscr{F}[\tilde{\chi}]$ and $\tilde{u} = \mathscr{F}[\tilde{\varphi}]$ of approximate solutions $(\tilde{\chi}, \tilde{\varphi})$ of the true wave-functions (χ, φ) obtained by another approximation method, unless the differential equations (34) are satisfied by $\tilde{v}$ and $\tilde{u}$. Therefore, it is not reasonable to postulate the matching condition (38) in every case, but rather to consider what condition arises from demanding $\delta J=0$. The advantage of a variational approach is confirmed by numerical investigations made by Bross (1982).

4.2. *Calculation of complex band structures*

Let it now be supposed that the potential w has three-dimensional translational symmetry with lattice $G_3=\mathbb{Z}\boldsymbol{a}_1\oplus\mathbb{Z}\boldsymbol{a}_2\oplus\mathbb{Z}\boldsymbol{a}_3$ and that Ω is the oblique column $\Omega=\{\boldsymbol{r}_\parallel+\xi\boldsymbol{a}_3;\ \boldsymbol{r}_\parallel\in WSZ_2,\ 0\leqslant\xi\leqslant 1\}$ shown in figure 3. Replacing the essential boundary conditions given in (24) by natural ones given in (8) for $\boldsymbol{r}\in F_{\mathrm{R}}\cup F_{\mathrm{L}}$ yields the following eigenvalue problem (which is closely connected with the problem considered in § 4.1).

Given a real energy E and a real two-dimensional propagation vector $\boldsymbol{k}_\parallel$, find a complex number $k_\perp$ and a pair of functions (χ,φ) such that equations (2) and (8) are satisfied with the transfer function t chosen such that $t(\boldsymbol{r})=\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{T}_l)$ for $\boldsymbol{r}\in F_{l,1}$; $F_{l,0}=\boldsymbol{T}_l+F_{l,1}$; $\boldsymbol{k}:=\boldsymbol{k}_\parallel+k_\perp\boldsymbol{e}_z$.

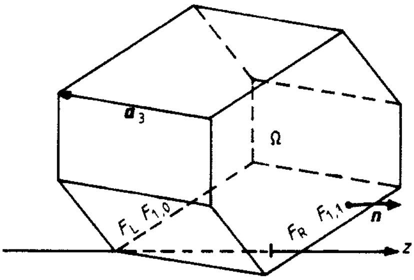

Figure 3. Region Ω in the case of three-dimensional translational symmetry.

Here, there are no essential boundary conditions; instead, one of the 'left' or 'right' boundary surfaces F_{L} and F_{R} is contained in S_t and is support for a natural boundary condition. For example, set $F_{1,1}:=F_{\mathrm{L}}$ and $F_{1,0}=\boldsymbol{a}_3+F_{1,1}$. Since $t_1:=t|_{F_{1,1}}$ is given by $t_1=\exp[\mathrm{i}(\boldsymbol{k}_\parallel+k_\perp\boldsymbol{e}_z)\boldsymbol{a}_3]$, t_1 may be regarded as eigenvalue instead of $k_\perp$.

One possibility to determine t_1 is to calculate the spectrum of the propagation matrix obtained from the initial value problem (24) (Bross 1977), but in practice this method involves difficulties, because the planar Fourier expansion will slowly converge for potentials with singularities. Therefore, it seems reasonable to use the variational principle $\delta J=0$ directly in the following way. Approximate χ and φ by

$$\hat{\varphi}=\sum_{j=1}^{N}C_j\varphi_j,\qquad \hat{\chi}=\sum_{j=1}^{N}D_j\chi_j,\qquad C_j,D_j\in\mathbb{C},\tag{39}$$

where $\{\varphi_j;\,j\in\mathbb{N}\}\subset V$ and $\{\chi_j;\,j\in\mathbb{N}\}\subset V$ are appropriately chosen. The condition $\delta J=0$ yields a linear homogeneous system of equations for C_j and D_j whose matrix elements depend nonlinearly on t_1 in general; but since the weight functions λ,μ,ρ,τ in the functional J can be freely chosen, a *linear* (generalised) eigenvalue problem for t_1 is obtained by setting† $\lambda=\mu=\rho=\tau=0$ in the $F_{1,1}$ integral. Then J separates into the sum

$$J(\chi,\varphi)=\sum_{k,l}D_k^*\tilde{J}(\chi_k,\varphi_l)C_l-t_1D_k^*B(\chi_k,\varphi_l)D_l\tag{40}$$

† Alternatively, to obtain a linear eigenvalue problem for $(t_1)^{-1}$ we could choose $\rho=\tau=0$, $\lambda(\hat{r},\hat{r}')=\mu(\hat{r},\hat{r}')=(t_1)^{-1}\delta(\hat{r},\hat{r}')$.

where the boundary operator B is defined as

$$B(\chi,\varphi):=\int_{F_{1,1}}(\partial_n\chi_0^*\varphi_1-\chi_0^*\partial_n\varphi_1)\,\mathrm{d}^2r_\parallel, \tag{41$_\mathrm{S}$}$$

$$B(\chi,\varphi):=-\mathrm{i}\int_{F_{1,1}}\chi_0^+\boldsymbol{\alpha n}\varphi_1\,\mathrm{d}^2r_\parallel, \tag{41$_\mathrm{D}$}$$

and $\tilde{J}$ means the remaining terms of J. The eigenvalue problem

$$\sum_{l=1}^{N}\tilde{J}(\chi_k,\varphi_l)C_l=t_1\sum_{l=1}^{N}B(\chi_k,\varphi_l)C_l,\qquad \sum_{k=1}^{N}\tilde{J}(\chi_k,\varphi_l)^*D_k=t_1^*\sum_{k=1}^{N}B(\chi_k,\varphi_l)^*D_k, \tag{42}$$

must be handled carefully because $B(\chi_k,\varphi_l)$ will be a singular matrix with, generally, large kernel. Since no physical meaning can be attached to $|t_1|=\infty$, numerical and analytical solutions to (42) must exclude this condition. The formulation of (42) provides a way to calculate complex band structures $(E,\boldsymbol{k}_\parallel)\mapsto k_\perp$ by means of linear eigensystem analysis (for which quite efficient computer programs are available; e.g. see Wilkinson and Reinsch (1971)) instead of determining the zeros of a secular determinant (which is done in KKR methods; for example, see Holzwarth and Lee (1978)).

Since each eigenvalue t_1 corresponds to a different boundary condition, it is essential that the basis functions χ_j and φ_j have sufficient degrees of freedom on $F_{1,1}$ and $F_{1,0}$ to satisfy condition (8) for a continuum of t_1 values enclosing the spectrum of (42); otherwise the subspaces of V spanned by those linear combinations of χ_j and φ_j which satisfy the boundary condition (8) on $F_{1,1}\cup F_{1,0}$ could be null for some t_1, in which case $\hat{\chi}$ and $\hat{\varphi}$ would not represent the behaviour of the 'true' solutions χ and φ at the boundaries $F_{1,1}\cup F_{1,0}$ sufficiently well.

The usual way to formulate the problem is the following.

Given a real two-dimensional vector $\boldsymbol{k}_\parallel$ and a complex number $k_\perp$, find a real energy E and a pair of functions (χ,φ) such that equations (2) and (8) are satisfied with the transfer function t chosen as above. For *real* $k_\perp$ this is exactly the well known 'single-cell formulation' of the band structure problem; since $t(\boldsymbol{r})^{-1}=t(\boldsymbol{r})^*$ for $\boldsymbol{r}\in S_\mathrm{t}$ in this special case, it is not necessary to distinguish between φ and χ (see equation (20)), and the variational expression J may be regarded as a functional of the single argument $\varphi=\chi$. Specifying the weight functions λ,μ,ρ,τ yields expressions $J(\varphi)$ which have already been derived by many authors for special numerical methods (APW method and related methods; see Leigh (1956), Schlosser and Marcus (1963), Loucks (1967), Marcus (1967), Bross and Hofmann (1969), Bross *et al* (1970), Ferreira *et al* (1974, 1975), Sarker and Taj-ul Islam (1975), Lopez-Aquilar (1979)). For the non-relativistic case, Marcus (1967) has given the most general functional $\varphi\mapsto J(\varphi,\varphi)$ with local weight functions $(\lambda(\hat{r},\hat{r}')=\lambda(\hat{r})\delta(\hat{r},\hat{r}')$ etc), where the trial functions φ may be discontinuous for $\boldsymbol{r}\in S$ but have to satisfy the Bloch condition (8) for $\boldsymbol{r}\in S_\mathrm{t}$. Sarker and Taj-ul Islam (1975) used a special form of the relativistic functional J (see footnote on p 3091) in the RAPW method. Non-local weights (= angular momentum projection operators) were considered by Ferreira *et al* (1975).

Making use of expression (19), all these methods can be generalised for *complex* values of $k_\perp$ in the following way. Introduce the complex propagation vector

$\boldsymbol{k} = (\boldsymbol{k}_{\|}, \boldsymbol{k}_{\perp})$ and define

$$V_{\boldsymbol{k}} := \{\varphi \in V;\, \partial_n^\alpha \varphi_0(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{T}_l)\partial_n^\alpha \varphi_1(\boldsymbol{r}) \text{ for } \boldsymbol{r} \in S_\mathrm{t} \text{ and } \alpha = 0, 1\},$$
$$V_{\boldsymbol{k}^*} := \{\chi \in V;\, \partial_n^\alpha \chi_0(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}^*\boldsymbol{T}_l)\partial_n^\alpha \chi_1(\boldsymbol{r}) \text{ for } \boldsymbol{r} \in S_\mathrm{t} \text{ and } \alpha = 0, 1\}, \tag{43$_\mathrm{S}$}$$

$$V_{\boldsymbol{k}} := \{\varphi \in V;\, \varphi_0(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{T}_l)\varphi_1(\boldsymbol{r}) \text{ for } \boldsymbol{r} \in S_\mathrm{t}\},$$
$$V_{\boldsymbol{k}^*} := \{\chi \in V;\, \chi_0(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}^*\boldsymbol{T}_l)\chi_1(\boldsymbol{r}) \text{ for } \boldsymbol{r} \in S_\mathrm{t}\}, \tag{43$_\mathrm{D}$}$$

respectively (i.e. (χ, φ) satisfy the natural boundary conditions for $\boldsymbol{r} \in S_\mathrm{t}$). Since $V_{\boldsymbol{k}} \subset U$ and $V_{\boldsymbol{k}^*} \subset W$ the functional $V_{\boldsymbol{k}^*} \times V_{\boldsymbol{k}} \ni (\chi, \varphi) \mapsto J(\chi, \varphi) \in \mathbb{C}$ is a variational expression which can be used in the usual way (Rayleigh–Ritz method) to get approximations

$$\hat{\varphi} = \sum_{j=1}^{N} A_j \varphi_j \qquad \text{and} \qquad \hat{\chi} = \sum_{j=1}^{N} B_j \chi_j, \qquad A_j, B_j \in \mathbb{C}, \tag{44}$$

where $\{\varphi_j;\, j \in \mathbb{N}\} \subset V_{\boldsymbol{k}}$ and $\{\chi_j;\, j \in \mathbb{N}\} \subset V_{\boldsymbol{k}^*}$ are suitable linearly independent sets of functions. The condition $\delta J = 0$ leads to a generalised linear algebraic eigenvalue problem of the form

$$\sum_{n=1}^{N} \hat{J}(\chi_m, \varphi_n) A_n = E(\boldsymbol{k}) \sum_{n=1}^{N} S(\chi_m, \varphi_n) A_n,$$
$$\sum_{m=1}^{N} \hat{J}(\chi_m, \varphi_n)^* B_m = E(\boldsymbol{k})^* \sum_{m=1}^{N} S(\chi_m, \varphi_n)^* B_m, \tag{45}$$

with $S(\chi, \varphi) := \int_\Omega \chi^+ \varphi \,\mathrm{d}^3 r$ and $\hat{J}(\chi, \varphi) := J(\chi, \varphi) + E(\boldsymbol{k}) S(\chi, \varphi)$, thus yielding the dependence $\boldsymbol{k} \mapsto E(\boldsymbol{k}) \in \mathbb{C}$. Seeking those values of $k_\perp$ for which $\mathrm{Im}[E(\boldsymbol{k})] = 0$, the 'complex band structure' $\boldsymbol{k} \mapsto E(\boldsymbol{k}) \in \mathbb{R}$ is finally obtained. In particular, if φ_j and χ_j are continuously differentiable or continuous, respectively (e.g. MAPW functions used by Bross and Hoffmann (1969) and by Bross *et al* (1970)), the variational expression J reduces to

$$J(\chi, \varphi) = \int_\Omega [\nabla \chi^* \nabla \varphi + (w - E)\chi^* \varphi] \,\mathrm{d}^3 r, \tag{46$_\mathrm{S}$}$$

$$J(\chi, \varphi) = \tfrac{1}{2} \int_\Omega \chi^+ (H - E)\varphi \,\mathrm{d}^3 r + \tfrac{1}{2} \int_\Omega [(H - E)\chi]^+ \varphi \,\mathrm{d}^3 r. \tag{46$_\mathrm{D}$}$$

This form of the functional proved adequate for numerical calculations of the complex band structures of Al, Cu and Ni using the complex version of the MAPW method (Bross 1976, Bross *et al* 1979).

4.3. *Scattering problems*

An important feature of the variational principle $\delta J = 0$ is the fact that it also holds for scattering problems. As a first example a result derived by Kohn (1948) will be verified.

Let w be a potential, which scatters 'incoming' plane waves $\exp(\mathrm{i}\boldsymbol{k}_1 \boldsymbol{r})$ ($\boldsymbol{k}_1 \in \mathbb{R}^3$; energy $E = \boldsymbol{k}_1^2$) elastically into 'outgoing' waves with asymptotic behaviour:

$$\varphi(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}_1 \boldsymbol{r}) + A(\hat{\boldsymbol{k}}_1, \hat{\boldsymbol{r}}) \frac{\exp(\mathrm{i}kr)}{r} \qquad (r \to \infty) \tag{47}$$

(here $\hat{\boldsymbol{x}}$ means the unit vector in direction of $\boldsymbol{x}$; $k = |\boldsymbol{k}_1|$; $r = |\boldsymbol{r}|$). An appropriate formulation as a variational problem is possible in the following way: Set $\Omega := K_R := \{\boldsymbol{r} \in \mathbb{R}^3;\ |\boldsymbol{r}| \leq R\}$ and $F := \partial\Omega = \partial K_R$ (and let $R \to \infty$ afterwards) while $F_{k,1} = F_{k,0} = S = \varnothing$. Define $E(r) := \exp(\mathrm{i}kr)/r$ and let for $\boldsymbol{k}_1, \boldsymbol{k}_2 \in \mathbb{R}^3$

$$
\begin{aligned}
U_0(F) &:= \{f : F \to \mathbb{C}^2;\ f(\boldsymbol{r}) = A(\hat{k}_1, \hat{r})(E(R), \partial_r E(R))^{\mathrm{T}};\\
&\qquad \hat{r} \mapsto A(\hat{k}_1, \hat{r}) \text{ is square integrable}\},\\
&\qquad f_0(\boldsymbol{r}) := (\exp(\mathrm{i}\boldsymbol{k}_1\boldsymbol{r}), \partial_r[\exp(\mathrm{i}\boldsymbol{k}_1\boldsymbol{r})])^{\mathrm{T}} \in W_2^{(3/2)}(F) \times W_2^{(1/2)}(F),\\
W_0(F) &:= \{g : F \to \mathbb{C}^2;\ g(\boldsymbol{r}) = B(\hat{k}_2, \hat{r})^*(E(R)^*, \partial_r E(R)^*)^{\mathrm{T}};\\
&\qquad \hat{r} \mapsto B(\hat{k}_2, \hat{r}) \text{ is square integrable}\},\\
&\qquad g_0(\boldsymbol{r}) := (\exp(-\mathrm{i}\boldsymbol{k}_2\boldsymbol{r}), \partial_r[\exp(-\mathrm{i}\boldsymbol{k}_2\boldsymbol{r})])^{\mathrm{T}} \in W_2^{(3/2)}(F) \times W_2^{(1/2)}(F).
\end{aligned}
\tag{48}
$$

Then the boundary condition (47) takes the form of equations (10), (11) and the compatibility condition (9) is also satisfied since

$$
\begin{aligned}
&\int_F (g^{(1)*} f^{(0)} - g^{(0)*} f^{(1)})\, \mathrm{d}\sigma\\
&\qquad = \int_{\partial K_1} B(\hat{k}_2, \hat{r}) A(\hat{k}_1, \hat{r})\, \mathrm{d}\sigma(\hat{r}) R^2 [E(R)\partial_r E(R) - E(R)\partial_r E(R)] = 0.
\end{aligned}
\tag{49}
$$

Thus the variational expression is given by

$$
\begin{aligned}
J(\chi, \varphi) &= \int_\Omega \chi^*(H - k^2)\varphi\, \mathrm{d}^3 r\\
&\qquad + \int_{\partial K_1} [\exp(\mathrm{i}\boldsymbol{k}_2 r\hat{r})\partial_r E(r) - \partial_r[\exp(\mathrm{i}\boldsymbol{k}_2 r\hat{r})] E(r)]|_{r=R} A(\hat{k}_1, \hat{r}) R^2\, \mathrm{d}\sigma.
\end{aligned}
\tag{50}
$$

In the limit $R \to \infty$, the factor multiplying $A(\hat{k}_1, \hat{r})$ becomes $-4\pi\delta(-\hat{k}_2, \hat{r})$ (Dirac 1947) and (50) reduces to Kohn's expression

$$
J(\chi, \varphi) = \int_{\mathbb{R}^3} \chi^*(H - k^2)\varphi\, \mathrm{d}^3 r - 4\pi A(\hat{k}_1, -\hat{k}_2)
\tag{51}
$$

which reveals the physical meaning of the functional J in this case.

As another example consider two semi-infinite media $\mathbb{R}^2 \times (-\infty, z_s)$ and $\mathbb{R}^2 \times (z_s, +\infty)$ with common boundary $z = z_s$; suppose each of them to have three-dimensional (either infinitesimal or discrete) translational symmetry when extended to full three-dimensional space†. Further, suppose the complex band structures $E(\boldsymbol{k}_\parallel, k_\perp)$ have been calculated for each semi-infinite medium separately (see § 4.2), so that for a given energy E and a given real two-dimensional propagation vector $\boldsymbol{k}_\parallel$ the sets $\{k_\perp^{(\mu,<)};\ \mu = 1, 2, , \ldots\}$ (for $z < z_s$) and $\{k_\perp^{(\nu,>)};\ \nu = 1, 2, \ldots\}$ (for $z > z_s$) of complex values of $k_\perp$ are known. Adopting the notation of Bross (1977), let $b(k_\perp^{(\mu,<)}, \boldsymbol{r})$ and $b(k_\perp^{(\nu,>)}, \boldsymbol{r})$ denote the corresponding solutions φ of Schrödinger's equation (2) subject to the

† In this section reference is made to Bross (1977) and free use is made of his results.

Bloch condition (8). The numeration of the values of $k_\perp$ will be such that

$$\begin{aligned}
&k^{(\mu,<)} \in \mathbb{R} \Leftrightarrow \mu = 1, \ldots, 2\sigma_<, \qquad k^{(\nu,>)} \in \mathbb{R} \Leftrightarrow \nu = 1, \ldots, 2\sigma_> \\
&\frac{\partial E(\boldsymbol{k}_\parallel, k_\perp^{(\mu,<)})}{\partial k^{(\mu,<)}} \begin{cases} \geqslant 0 \Leftarrow \mu = 1, \ldots, \sigma_< \\ \leqslant 0 \Leftarrow \mu = \sigma_< + 1, \ldots, 2\sigma_<, \end{cases} \\
&\frac{\partial E(\boldsymbol{k}_\parallel, k_\perp^{(\nu,>)})}{\partial k^{(\nu,>)}} \begin{cases} \geqslant 0 \Leftarrow \nu = 1, \ldots, \sigma_>, \\ \leqslant 0 \Leftarrow \nu = \sigma_> + 1, \ldots, 2\sigma_>. \end{cases}
\end{aligned} \tag{52}$$

The scattering problem can be formulated as follows. Define Ω as column $WSZ_2 \times (z_{-\infty}, z_{+\infty})$ (with $z_{+\infty} \to +\infty$ and $z_{-\infty} \to -\infty$ afterwards) which is separated into $\Omega_0 := WSZ_2 \times (z_{-\infty}, z_s)$ and $\Omega_1 := WSZ_2 \times (z_s, z_{+\infty})$ by the surface S(= intersection of the plane $z = z_s$ with Ω). Analogously to § 4.1, set $F := F_{+\infty} \cup F_{-\infty}$ where $F_{\pm\infty}$ means that part of $\partial\Omega$ which is contained in the planes $z = z_{\pm\infty}$; the remaining part of $\partial\Omega$ is of the form $S_t \cup \bigcup_{l=1}^{M} F_{l,0}$ (see figure 2). Given an energy E and a wavevector $\boldsymbol{k}_\parallel$, find a pair (χ, φ), of solutions of the Schrödinger equation (2) in $\Omega_0 \cup \Omega_1$ satisfying the Bloch condition (8) or matching condition, respectively, on $S_t \cup S(t(\boldsymbol{r}) = 1$ for $\boldsymbol{r} \in S$ and $t(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}_\parallel \boldsymbol{T}_l)$ for $\boldsymbol{r} \in F_{l,1})$ with asymptotic behaviour

$$\left.\begin{aligned}
\partial_z^\alpha \varphi(\boldsymbol{r}) &= \partial_z^\alpha b(k_\perp^{(\kappa,<)}, \boldsymbol{r}) + \sum_{\mu=1+\sigma_<}^{2\sigma_<} A_\mu \partial_z^\alpha b(k_\perp^{(\mu,<)}, \boldsymbol{r}) \\
\partial_z^\alpha \chi(\boldsymbol{r}) &= \partial_z^\alpha b(k_\perp^{(\lambda,<)}, \boldsymbol{r}) + \sum_{\mu=1}^{\sigma_<} B_\mu \partial_z^\alpha b(k^{(\mu,<)}, \boldsymbol{r})
\end{aligned}\right\} \text{for } \boldsymbol{r} \in F_{-\infty} \text{ and } \alpha = 0, 1$$

$$\left.\begin{aligned}
\partial_z^\alpha \varphi(\boldsymbol{r}) &= \sum_{\nu=1}^{\sigma_>} C_\nu \partial_z^\alpha b(k_\perp^{(\nu,>)}, \boldsymbol{r}) \\
\partial_z^\alpha \chi(\boldsymbol{r}) &= \sum_{\nu=1+\sigma_>}^{2\sigma_>} D_\nu \partial_z^\alpha b(k_\perp^{(\nu,>)}, \boldsymbol{r})
\end{aligned}\right\} \text{for } \boldsymbol{r} \in F_{+\infty} \text{ and } \alpha = 0, 1 \tag{53}$$

where $\kappa \in \{1, \ldots, \sigma_<\}$ and $\lambda \in \{1 + \sigma_<, \ldots, 2\sigma_<\}$ are fixed.

To make equation (53) correspond to equations (10), (11), choose

$$\begin{aligned}
f_0(\boldsymbol{r}) &:= \begin{cases} (b(k_\perp^{(\kappa,<)}, \boldsymbol{r}), \partial_z b(k_\perp^{(\kappa,<)}, \boldsymbol{r}))^{\mathrm{T}} & \text{for } \boldsymbol{r} \in F_{-\infty}, \\ 0 & \text{for } \boldsymbol{r} \in F_{+\infty}, \end{cases} \\
g_0(\boldsymbol{r}) &:= \begin{cases} (b(k_\perp^{(\lambda,<)}, \boldsymbol{r}), \partial_z b(k_\perp^{(\lambda,<)}, \boldsymbol{r}))^{\mathrm{T}} & \text{for } \boldsymbol{r} \in F_{-\infty}, \\ 0 & \text{for } \boldsymbol{r} \in F_{+\infty}, \end{cases}
\end{aligned} \tag{54}$$

and

$$\begin{aligned}
U_0(F) := \Big\{ f : F \to \mathbb{C}^2;\ f|_{F_{-\infty}} &= \sum_{\mu=1+\sigma_<}^{2\sigma_<} A_\mu (b(k_\perp^{(\mu,<)}, \boldsymbol{r}), \partial_z b(k_\perp^{(\mu,<)}, \boldsymbol{r}))^{\mathrm{T}} \\
\text{and } f|_{F_{+\infty}} &= \sum_{\nu=1}^{\sigma_>} C_\nu (b(k_\perp^{(\nu,>)}, \boldsymbol{r}), \partial_z b(k_\perp^{(\nu,>)}, \boldsymbol{r}))^{\mathrm{T}};\ A_\mu, C_\nu \in \mathbb{C} \Big\},
\end{aligned} \tag{55}$$

$$\begin{aligned}
W_0(F) := \Big\{ g : F \to \mathbb{C}^2;\ g|_{F_{-\infty}} &= \sum_{\mu=1}^{\sigma_<} B_\mu (b(k_\perp^{(\mu,<)}, \boldsymbol{r}), \partial_z b(k_\perp^{(\mu,<)}, \boldsymbol{r}))^{\mathrm{T}} \\
\text{and } g|_{F_{+\infty}} &= \sum_{\nu=1+\sigma_>}^{2\sigma_>} D_\nu (b(k_\perp^{(\nu,>)}, \boldsymbol{r}), \partial_z b(k_\perp^{(\nu,>)}, \boldsymbol{r}))^{\mathrm{T}};\ B_\mu, D_\nu \in \mathbb{C} \Big\}.
\end{aligned}$$

The compatibility condition (9) is also satisfied because

$$\int_{F_{+\infty}} [\partial_z b(k_\perp^{(\nu,>)}, r)^* b(k_\perp^{(\rho,>)}, r) - b(k_\perp^{(\nu,>)}, r)^* \partial_z b(k_\perp^{(\rho,>)}, r)]\, d^2 r_\parallel = 0$$

for $\nu \in \{1+\sigma_>, \ldots, 2\sigma_>\}$ and $\rho \in \{1, \ldots, \sigma_>\}$,

$$\int_{F_{-\infty}} [\partial_z b(k_\perp^{(\mu,<)}, r)^* b(k_\perp^{(\tau,<)}, r) - b(k_\perp^{(\mu,<)}, r)^* \partial_z b(k_\perp^{(\tau,<)}, r)]\, d^2 r_\parallel = 0 \tag{56}$$

for $\mu \in \{1, \ldots, \sigma_<\}$ and $\tau \in \{1+\sigma_<, \ldots, 2\sigma_<\}$ (see Bross 1977, equation (2.31)).

Note that not χ but χ^* is to be interpreted as a solution of a real scattering problem (see figure 4). As above in Kohn's expression, the $\int_F$ term of the boundary functional K has a simple physical meaning because

$$\tfrac{1}{2}\int_F [(\chi^* + g_0^{(0)*})(\partial_n \varphi - f_0^{(1)}) - (\partial_n \chi^* + g_0^{(1)*})(\varphi - f_0^{(0)})]\, d\sigma$$

$$= A_\lambda \int_{F_{-\infty}} [\partial_z b(k_\perp^{(\lambda,<)}, r)^* b(k_\perp^{(\lambda,<)}, r) - b(k^{(\lambda,<)}, r)^* \partial_z b(k_\perp^{(\lambda,<)}, r)]\, d^2 r_\parallel,$$

i.e. the factor multiplying the scattering amplitude A_λ is 'i times the probability current of $b(k_\perp^{(\lambda,<)}, r)$ passing $F_{-\infty}$'.

The practical application of the variational formulation of the scattering problem to the case of a stacking fault was successfully demonstrated by Bross (1981, 1982).

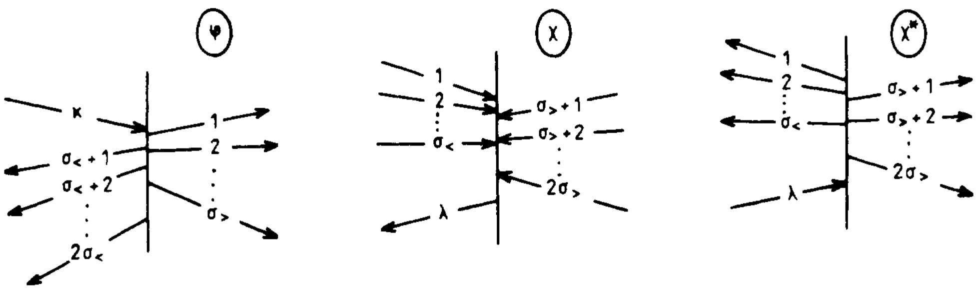

Figure 4. Schematic representation of the incoming and scattered waves contained in the wavefunctions φ, χ and χ^*.

5. Summary

Considering a class of boundary value problems arising in Schrödinger's or Dirac's theory of single-particle states, a family of variational expressions has been derived whose stationary points are solutions of the corresponding boundary value problems. These variational expressions represent a generalisation of previous results obtained by other authors for special applications. It is shown that the given formulation of the boundary conditions is sufficiently general to include scattering conditions, generalised Bloch conditions, or initial value conditions connected with the propagation operator method. The variational principle may be used to derive necessary conditions for approximate solutions of the boundary value problems and thus enables the application of the Ritz method and other numerical techniques. Because the given formulation contains some weight functions controlling the behaviour of the trial functions at the boundaries, it is possible to stress or to relax the natural boundary conditions for approximate solutions by specifying these weights appropriately and to examine the quality of the approximation by varying them.

Appendix 1. Proof of the variational prinxiple

The first part of the proof is based on results presented by Nečas (1967) whose notation will also be used here. Let $\mathcal{M}^{(k),1}$ denote the class of bounded domains $\tilde{\Omega}$ such that the boundary $\partial\tilde{\Omega}$ is described by local maps which have Lipschitz-continuous derivatives up to order k (Nečas 1967, ch 2, § 1.1). It is supposed that any bounded domain $\Omega_i^{(j)}$ of the partition $\Omega = \bigcup_{j=1}^{q_0} \Omega_0^{(j)} \cup \bigcup_{j=1}^{q_1} \Omega_1^{(j)}$ has a 'piecewise regular' boundary in the following sense (Nečas 1967, ch 6, § 1.3). Let $\hat{\Omega}$ be one of the domains $\Omega_i^{(j)}$; then

(i) there exist $\hat{\Omega}_l \in \mathcal{M}^{(2),1}$ such that $\hat{\Omega} = \bigcap_{l=1}^{\hat{p}} \hat{\Omega}_l$;

(ii) $\hat{\Omega} \in \mathcal{M}^{(0),1}$;

(iii) if $\Lambda_l :=$ interior of $\partial\hat{\Omega}_l \cap \partial\hat{\Omega}$ (with respect to $\partial\hat{\Omega}$), then $\Lambda_l \cap \Lambda_k = \emptyset$ for $l \neq k$, and the superficial measure of $\partial\hat{\Omega} \backslash \bigcup_{l=1}^{p} \Lambda_1$ is zero;

(iv) the sets $F \cap \partial\hat{\Omega}$, $F_{k,0} \cap \partial\hat{\Omega}$, $F_{k,1} \cap \partial\hat{\Omega}$, $S \cap \partial\hat{\Omega}$ are of the form $\bigcup_{l\in L} \Lambda_l$ with $L \subset \{1, \ldots, \hat{p}\}$.

If Γ is a measurable subset of $\partial\Omega_0 \cup \partial\Omega_1$ and ψ belongs to V (refer to definition (5)), $\partial_n^\alpha \psi_i|_\Gamma$ means the restriction of $\partial_n^\alpha \psi_i$ (= boundary values in the sense of traces) to Γ (understood as element of $L_2(\Gamma)$).

The above assumptions are sufficient to ensure that Gauss' integral theorem holds (Nečas 1967, ch 3, § 1.2) and that the following lemma is true.

Lemma. Let V_0 be defined as

$$V_0 := \{\psi \in V;\ \psi|_F = \partial_n \psi|_F = 0\}, \tag{A1.1$_S$}$$

$$V_0 := \{\psi \in V;\ \psi|_F = 0\}, \tag{A1.1$_D$}$$

and let the operator Tr be defined as

$$V \ni \psi \mapsto \mathrm{Tr}\,\psi := (\psi_0|_{S\cup S_t}, \psi_1|_{S\cup S_t}, \partial_n\psi_0|_{S\cup S_t}, \partial_n\psi_1|_{S\cup S_t}) \in L_2(S\cup S_t)^4, \tag{A1.2$_S$}$$

$$V \ni \psi \mapsto \mathrm{Tr}\,\psi := (\psi_0|_{S\cup S_t}, \psi_1|_{S\cup S_t}) \in L_2(S\cup S_t)^8. \tag{A1.2$_D$}$$

Then $\mathrm{Tr}(V_0)$ is dense in $L_2(S\cup S_t)^4$ (or $L_2(S\cup S_t)^8$, respectively).

Proof. Let $\hat{\Omega}$ be one of the domains $\Omega_i^{(j)}$, let

$$V_0(\hat{\Omega}) := \{\psi \in W_2^{(2)}(\hat{\Omega});\ \psi|_{F\cap\partial\hat{\Omega}} = \partial_n\psi|_{F\cap\partial\hat{\Omega}} = 0\}, \tag{A1.3$_S$}$$

$$V_0(\hat{\Omega}) := \{\psi \in W_2^{(1)}(\hat{\Omega});\ \psi|_{F\cap\partial\hat{\Omega}} = 0\} \tag{A1.3$_D$}$$

and let $\Gamma := (\partial\Omega_0 \cup \partial\Omega_1 \backslash F) \cap \partial\hat{\Omega}$. It is sufficient to show that $V_0(\hat{\Omega}) \ni \psi \mapsto (\psi|_\Gamma, \partial_n\psi|_\Gamma) \in L_2(\Gamma)^2$ (or $V_0(\hat{\Omega}) \ni \psi \mapsto \psi|_\Gamma \in L_2(\Gamma)^4$, respectively) maps $V_0(\hat{\Omega})$ on a dense subset of $L_2(\Gamma)^2$ (*or* $L_2(\Gamma)^4$).

In the case of Schrödinger's theory, let $(h^{(0)}, h^{(1)})$ be an element of $L_2(\Gamma)^2$. There exist $L, M \subset \{1, \ldots, \hat{p}\}$ such that $\partial\hat{\Omega} \cap F \backslash \bigcup_{l\in L} \Lambda_l$ and $\Gamma \backslash \bigcup_{l\in M} \Lambda_l$ are sets of superficial measure zero. Let χ_l denote the characteristic function of Λ_l; then $h^{(\alpha)} = \Sigma_{l\in M} h_l^{(\alpha)}$

with $h_l^{(\alpha)} := h^{(\alpha)}\chi_l$ (which is considered as an element of $L_2(\partial\hat{\Omega})$). Following Nečas (1967, ch 6, § 1.3) there exists a sequence $(\tilde{h}_{lk}^{(0)}, \tilde{h}_{lk}^{(1)})_{k\in\mathbb{N}} \subset W_2^{(1)}(\partial\hat{\Omega}) \times W_2^{(1)}(\partial\hat{\Omega})$ with $\operatorname{supp}(\tilde{h}_{lk}^{(\alpha)}) \subset \Lambda_l$ which converges to $(h_l^{(0)}, h_l^{(1)})$ in the L_2 norm. Since $W_2^{(3/2)}(\partial\hat{\Omega})$ is dense in $W_2^{(1)}(\partial\hat{\Omega})$ (Nečas 1967, ch 5, § 4.5), $\tilde{h}_{lk}^{(0)}$ can be approximated by $h_{lk}^{(0)} \in W_2^{(3/2)}(\partial\hat{\Omega})$ with $\operatorname{supp}(h_{lk}^{(0)}) \subset \Lambda_l$; so $(h_l^{(0)}, h_l^{(1)})$ is approximated by $(h_{lk}^{(0)}, h_{lk}^{(1)})_{k\in\mathbb{N}} \subset W_2^{(3/2)}(\partial\hat{\Omega}) \times W_2^{(1/2)}(\partial\hat{\Omega})$ in the L_2 norm. Now, there exists a sequence $(\psi_{lk})_{k\in\mathbb{N}} \subset W_2^{(2)}(\hat{\Omega})$ such that $\partial_n^\alpha \psi_{lk} = h_{lk}^{(\alpha)}$ on $\partial\hat{\Omega}$ (Nečas 1967, ch 2, § 5.7). Set $\psi_k := \Sigma_{l\in M}\psi_{lk} \in W_2^{(2)}(\hat{\Omega})$; then $\operatorname{supp}(\partial_n^\alpha \psi_k) \subset \bigcup_{l\in M}\Lambda_l \subset \Gamma$, i.e. $\psi_k \in V_0(\hat{\Omega})$ and $(\psi_k|_\Gamma, \partial_n\psi_k|_\Gamma)_{k\in\mathbb{N}}$ converges to $(h^{(0)}, h^{(1)})$ in $L_2(\Gamma)^2$.

In the case of Dirac's theory, $h \in L_2(\Gamma)^4$ is approximated by $(\Sigma_{l\in M} h_{lk})_{k\in\mathbb{N}} \subset W_2^{(1/2)}(\Gamma)^4$ with $\operatorname{supp}(h_{lk}) \subset \Lambda_l$. There exists $\psi_{lk} \in W_2^{(1)}(\hat{\Omega})$ with $\psi_{lk} = h_{lk}$ on $\partial\hat{\Omega}$; so the trace of $\psi_k := \Sigma_{l\in M}\psi_{lk} \in W_2^{(1)}(\hat{\Omega})$ has its support within Γ (i.e. $(\psi_k)_{k\in\mathbb{N}} \subset V_0(\hat{\Omega})$ and it converges to h in $L_2(\Gamma)^4$).

For the following part, it is convenient to define the inner products

$$\langle\ |\ \rangle : V \times V \to \mathbb{C}, \qquad \langle \chi | \varphi \rangle := \int_{\Omega_0 \cup \Omega_1} \chi^+ \varphi \, \mathrm{d}^3 r \tag{A1.4}$$

and

$$[\ |\ ] : L_2(S \cup S_t)^4 \times L_2(S \cup S_t)^4 \to \mathbb{C}, \qquad [\chi|\varphi] := \sum_{j=1}^{4} \int_{S\cup S_t} \chi_j^* \varphi_j \, \mathrm{d}\sigma, \tag{A1.5$_\mathrm{S}$}$$

$$[\ |\ ] : L_2(S \cup S_t)^8 \times L_2(S \cup S_t)^8 \to \mathbb{C}, \qquad [(\chi_1, \chi_2)|(\varphi_1, \varphi_2)] := \sum_{j=1}^{2} \int_{S\cup S_t} \chi_j^+ \varphi_j \, \mathrm{d}\sigma, \tag{A1.5$_\mathrm{D}$}$$

and the linear operators $L_2(S \cup S_t) \to L_2(S \cup S_t)$:

$$\kappa\varphi(\boldsymbol{r}) := \int_{S\cup S_t} \kappa(\hat{r}, \hat{r}')\varphi(\boldsymbol{r}') \, \mathrm{d}\sigma(\hat{r}')$$

for $\kappa(\hat{r}, \hat{r}') \in \{\lambda(\hat{r}, \hat{r}'), \mu(\hat{r}, \hat{r}'), \rho(\hat{r}, \hat{r}'), \tau(\hat{r}, \hat{r}')\}$,

$$t\varphi(\boldsymbol{r}) := t(\boldsymbol{r})\varphi(\boldsymbol{r}), \qquad \boldsymbol{\alpha n}\varphi(\boldsymbol{r}) := \boldsymbol{\alpha n}(\boldsymbol{r})\varphi(\boldsymbol{r}). \tag{A1.6}$$

Then δJ can be written for $(\chi, \varphi) \in W \times U$ and $(\delta\chi, \delta\varphi) \in W_0 \times U_0$ as

$$\begin{aligned}
\delta J = {} & \langle \delta\chi | (H-E)\varphi \rangle + \langle (H-E)\chi | \delta\varphi \rangle + \left[(\delta\chi_0, \delta\chi_1, \partial_n\delta\chi_0, \partial_n\partial\chi_1) \,\middle|\, \left(C \begin{pmatrix} \varphi_0 - t\varphi_1 \\ \partial_n\varphi_0 - t\partial_n\varphi_1 \end{pmatrix} \right)^{\mathrm{T}} \right] \\
& + \left[\left(D \begin{pmatrix} \chi_1 - t^+\chi_0 \\ \partial_n\chi_1 - t^+\partial_n\chi_0 \end{pmatrix} \right)^{\mathrm{T}} \,\middle|\, (\delta\varphi_1, \delta\varphi_0, \partial_n\delta\varphi_1, \partial_n\delta\varphi_0) \right] \\
& + \tfrac{1}{2}\int_F [\delta\chi^*(\partial_n\varphi - f_0^{(1)}) - \partial_n\delta\chi^*(\varphi - f_0^{(0)})] \, \mathrm{d}\sigma \\
& + \tfrac{1}{2}\int_F [(\partial_n\chi^* - g_0^{(1)*})\delta\varphi - (\chi^* - g_0^{(0)*})\partial_n\delta\varphi] \, \mathrm{d}\sigma
\end{aligned} \tag{A1.7$_\mathrm{S}$}$$

where

$$C := \begin{pmatrix} t\rho & t\lambda - 1 \\ -\rho & -\lambda \\ 1 - t\mu & t\tau \\ \mu & -\tau \end{pmatrix}, \qquad D := \begin{pmatrix} t^+\rho^+ & 1 - t^+\mu^+ \\ -\rho^+ & \mu^+ \\ t^+\lambda^+ - 1 & t^+\tau^+ \\ -\lambda^+ & -\tau^+ \end{pmatrix}$$

or

$$\delta J = \langle \delta\chi | (H-E)\varphi \rangle + \langle (H-E)\chi | \delta\varphi \rangle + \mathrm{i}[(\delta\chi_0, \delta\chi_1) | (C(\varphi_0 - t\varphi_1))^{\mathrm{T}}]$$
$$+\mathrm{i}[(D(\chi_1 - t^+\chi_0))^{\mathrm{T}} | (\delta\varphi_1, \delta\varphi_0)]$$
$$+\tfrac{1}{2}\mathrm{i}\int_F \delta\chi^+ \boldsymbol{\alpha n}(\varphi - f_0)\,\mathrm{d}\sigma - \tfrac{1}{2}\mathrm{i}\int_F (\chi^+ - g_0^+)\boldsymbol{\alpha n}\delta\varphi\,\mathrm{d}\sigma \qquad (\mathrm{A1.7_D})$$

where

$$C := \begin{pmatrix} -\boldsymbol{\alpha n} + t\rho \\ -\rho \end{pmatrix}, \qquad D := \begin{pmatrix} -\boldsymbol{\alpha n} + t^+\rho^+ \\ -\rho^+ \end{pmatrix},$$

respectively.

Since

$$\begin{aligned} &(\varphi|_F - f_0^{(0)}, \partial_n\varphi|_F - f_0^{(1)})^{\mathrm{T}} \in U_0(F), \qquad (\delta\varphi|_F, \partial_n\delta\varphi|_F)^{\mathrm{T}} \in U_0(F),\\ &(\chi|_F - g_0^{(0)}, \partial_n\chi|_F - g_0^{(1)})^{\mathrm{T}} \in W_0(F), \qquad (\delta\chi|_F, \partial_n\delta\chi|_F)^{\mathrm{T}} \in W_0(F), \end{aligned} \qquad (\mathrm{A1.8_S})$$

$$\varphi|_F - f_0 \in U_0(F), \qquad \delta\varphi|_F \in U_0(F), \qquad \chi|_F - g_0 \in W_0(F), \qquad \delta\chi|_F \in W_0(F), \qquad (\mathrm{A1.8_D})$$

the $\int_F$ term in equation (A1.7) vanishes due to condition (9). Hence, if it is assumed that (χ, φ) satisfies (2) and (8), then $\delta J = 0$.

Conversely, if $\delta J = 0$ for any $(\delta\chi, \delta\varphi) \in V_0 \times V_0 \subset W_0 \times U_0$, consider the restricted variations $(\delta\chi, \delta\varphi) \in V_r := \{(\delta\chi, \delta\varphi) \in V_0 \times V_0;\ (\delta\chi, \delta\varphi)$ satisfies (8)$\}$. Then $\delta J = \langle \delta\chi | (H-E)\varphi \rangle + \langle (H-E)\chi | \delta\varphi \rangle$ for $(\delta\chi, \delta\varphi) \in V_r$, and since V_r is dense† in $L_2(\Omega_0 \cup \Omega_1)$ one obtains $(H-E)\varphi = (H-E)\chi = 0$ in $\Omega_0 \cup \Omega_1$. Substituting this result in (A1.7) and choosing $(\delta\chi, \delta\varphi) \in V_0 \times V_0$, the condition $\delta J = 0$ yields

$$\begin{aligned} &\left[(\delta\chi_0, \delta\chi_1, \partial_n\delta\chi_0, \partial_n\delta\chi_1) \middle| \left(C\begin{pmatrix} \varphi_0 - t\varphi_1 \\ \partial_n\varphi_0 - t\partial_n\varphi_1 \end{pmatrix}\right)^{\mathrm{T}}\right] = 0,\\ &\left[(\delta\varphi_1, \delta\varphi_0, \partial_n\delta\varphi_1, \partial_n\delta\varphi_0) \middle| \left(D\begin{pmatrix} \chi_1 - t^+\chi_0 \\ \partial_n\chi_1 - t^+\partial_n\chi_0 \end{pmatrix}\right)^{\mathrm{T}}\right] = 0, \end{aligned} \qquad (\mathrm{A1.9_S})$$

$$[(\delta\chi_0, \delta\chi_1) | (C(\varphi_0 - t\varphi_1))^{\mathrm{T}}] = 0, \qquad [(\delta\varphi_1, \delta\varphi_0) | (D(\chi_1 - t^+\chi_0))^{\mathrm{T}}] = 0. \qquad (\mathrm{A1.9_D})$$

According to the lemma given above the boundary values $\partial_n^\alpha\delta\varphi_i$ and $\partial_n^\alpha\delta\chi_i$ have a range which is dense in $L_2(S \cup S_t)$; hence

$$C\begin{pmatrix} \varphi_0 - t\varphi_1 \\ \partial_n\varphi_0 - t\partial_n\varphi_1 \end{pmatrix} = 0, \qquad D\begin{pmatrix} \chi_1 - t^+\chi_0 \\ \partial_n\chi_1 - t^+\partial_n\chi_0 \end{pmatrix} = 0, \qquad (\mathrm{A1.10_S})$$

$$C(\varphi_0 - t\varphi_1) = 0, \qquad D(\chi_1 - t^+\chi_0) = 0. \qquad (\mathrm{A1.10_D})$$

Since the operators

$$A := \begin{pmatrix} 0 & 0 & 1 & t \\ -1 & -t & 0 & 0 \end{pmatrix}, \qquad B := \begin{pmatrix} 0 & 0 & -1 & -t^+ \\ 1 & t^+ & 0 & 0 \end{pmatrix} \qquad (\mathrm{A1.11_S})$$

$$A := (-\boldsymbol{\alpha n}, -\boldsymbol{\alpha n}t), \qquad B := (-\boldsymbol{\alpha n}, -\boldsymbol{\alpha n}t^+) \qquad (\mathrm{A1.11_D})$$

† Note that $\mathscr{C}_0^\infty(\mathring{\Omega}_0 \cup \mathring{\Omega}_1) \times \mathscr{C}_0^\infty(\mathring{\Omega}_0 \cup \mathring{\Omega}_1) \subset V_r \subset L_2(\Omega_0 \cup \Omega_1) \times L_2(\Omega_0 \cup \Omega_1)$; $\mathscr{C}_0^\infty$ is dense in L_2 (see Mikhlin 1970).

have the property

$$AC = BD = 1 \tag{A.12}$$

it can be concluded from (A1.10) that (χ, φ) satisfy the natural boundary conditions (8).

Appendix 2. Condition for symmetric Euler–Lagrange equations

Using the definitions (A1.4), (A1.5) and (A1.6) J can be written as

$$J = \int_{\Omega_0 \cup \Omega_1} [\nabla\chi^* \nabla\varphi + (w - E)\chi^*\varphi]\, \mathrm{d}^3 r$$
$$+ [(\chi_0, \chi_1, \partial_n\chi_0, \partial_n\chi_1) | (T(t)(\varphi_0, \varphi_1, \partial_n\varphi_0, \partial_n\varphi_1)^{\mathrm{T}})^{\mathrm{T}}]$$
$$- \tfrac{1}{2} \int_F [(\partial_n\chi^* + g_0^{(1)*})(\varphi - f_0^{(0)}) + (\chi^* - g_0^{(0)*})(\partial_n\varphi + f_0^{(1)})]\, \mathrm{d}\sigma \tag{A2.1$_\mathrm{S}$}$$

where

$$T(t) := \begin{pmatrix} t\rho & -t\rho t & t\lambda & t(1-\lambda t) \\ -\rho & \rho t & -\lambda & \lambda t - 1 \\ 1 - t\mu & t(\mu t - 1) & t\tau & -t\tau t \\ \mu & -\mu t & -\tau & \tau t \end{pmatrix}$$

or

$$J = \tfrac{1}{2}\langle\chi | (H - E)\varphi\rangle + \tfrac{1}{2}\langle (H - E)\chi | \varphi\rangle + [(\chi_0, \chi_1) | (T(t)(\varphi_0, \varphi_1)^{\mathrm{T}})^{\mathrm{T}}]$$
$$+ \tfrac{1}{2}\mathrm{i} \int_F (g_0^+ \boldsymbol{\alpha n}\varphi - \chi^+ \boldsymbol{\alpha n} f_0)\, \mathrm{d}\sigma \tag{A2.1$_\mathrm{D}$}$$

where

$$T(t) := \mathrm{i}\begin{pmatrix} t\rho - \tfrac{1}{2}\boldsymbol{\alpha n} & \boldsymbol{\alpha n}t - t\rho t \\ -\rho & \rho t - \tfrac{1}{2}\boldsymbol{\alpha n} \end{pmatrix},$$

respectively. Under the operation (20) J is transformed into

$$\tilde{J} = \int_{\Omega_0 \cup \Omega_1} [\nabla\varphi^* \nabla\chi + (w - E)\varphi^*\chi]\, \mathrm{d}^3 r$$
$$+ [(\varphi_0, \varphi_1, \partial_n\varphi_0, \partial_n\varphi_1) | (T((t^+)^{-1})(\chi_0, \chi_1, \partial_n\chi_0, \partial_n\chi_1)^{\mathrm{T}})^{\mathrm{T}}]$$
$$- \tfrac{1}{2} \int_F [(\partial_n\varphi^* + f_0^{(1)*})(\chi - g_0^{(0)}) + (\varphi^* - f_0^{(0)*})(\partial_n\chi + g_0^{(1)})]\, \mathrm{d}\sigma, \tag{A2.2$_\mathrm{S}$}$$

$$\tilde{J} = \tfrac{1}{2}\langle\varphi | (H - E)\chi\rangle + \tfrac{1}{2}\langle (H - E)\varphi | \chi\rangle + [(\varphi_0, \varphi_1) | (T((t^+)^{-1})(\chi_0, \chi_1)^{\mathrm{T}})^{\mathrm{T}}]$$
$$+ \tfrac{1}{2}\mathrm{i} \int_F (f_0^+ \boldsymbol{\alpha n}\chi - \varphi^+ \boldsymbol{\alpha n} g_0)\, \mathrm{d}\sigma. \tag{A2.2$_\mathrm{D}$}$$

The requirement $J = \tilde{J}^*$ yields

$$T((t^+)^{-1}) = T(t)^+ \tag{A2.3}$$

which is equivalent to

$$\lambda t + t^{-1}\mu^{+} = 1, \qquad \mu t + t^{-1}\lambda^{+} = 1, \qquad \rho t = t^{-1}\rho^{+}; \qquad \tau t = t^{-1}\tau^{+}, \tag{A2.4$_\mathrm{S}$}$$

$$\rho t + t^{-1}\rho^{+} = \boldsymbol{\alpha n}. \tag{A2.4$_\mathrm{D}$}$$

If $t^{+} \neq t^{-1}$, λ can be eliminated from (A2.4$_\mathrm{S}$):

$$\begin{aligned}
&t\mu t - (t^{-1})^{+}\mu (t^{-1})^{+} = t - (t^{-1})^{+}\\
&\Rightarrow t(\boldsymbol{r})\mu(\hat{r}, \hat{r}')t(\boldsymbol{r}') - (t(\boldsymbol{r})^{*})^{-1}\mu(\hat{r}, \hat{r}')(t(\boldsymbol{r}')^{*})^{-1} = [t(\boldsymbol{r}) - (t(\boldsymbol{r})^{*})^{-1}]\delta(\hat{r}, \hat{r}')\\
&\Rightarrow \mu(\hat{r}, \hat{r}') = \frac{t(\boldsymbol{r}) - (t(\boldsymbol{r})^{*})^{-1}}{t(\boldsymbol{r})t(\boldsymbol{r}') - (t(\boldsymbol{r})^{*})^{-1}(t(\boldsymbol{r}')^{*})^{-1}}\delta(\hat{r}, \hat{r}')\\
&= (t(\boldsymbol{r}) + (t(\boldsymbol{r})^{*})^{-1})^{-1}\delta(\hat{r}, \hat{r}'),\\
&\text{i.e. } \mu = (t + (t^{-1})^{+})^{-1}.
\end{aligned} \tag{A2.5$_\mathrm{S}$}$$

Because $\lambda = (1 - t^{-1}\mu^{+})t^{-1}$, it can be concluded that $\lambda = \mu$.

Appendix 3. Derivation of the 'matching condition'

Let S be the intersection of a plane $z = z_0 (z_\mathrm{L} < z_0 < z_\mathrm{R})$ with Ω and let $\hat{\chi}$ and $\hat{\varphi}$ be of the form (30) but now with planar Fourier coefficients u_j and v_j belonging to

$$\hat{V}(z_\mathrm{L}, z_0, z_\mathrm{R}) := \{u : [z_\mathrm{L}, z_\mathrm{R}] \to \mathbb{C}^n;\\
u|_{[z_\mathrm{L}, z_0]} \in V(z_\mathrm{L}, z_0),\ u|_{[z_0, z_\mathrm{R}]} \in V(z_0, z_\mathrm{R})\}.$$

Then $\delta J = 0$ implies equation (34) for $z_\mathrm{L} < z < z_0$ and $z_0 < z < z_\mathrm{R}$ with initial values (33), if only such variations $\delta\hat{\chi}$ and $\delta\hat{\varphi}$ are chosen which are continuously differentiable (or continuous, respectively) at $z = z_0$ (i.e. $\delta v_j, \delta u_j \in V(z_\mathrm{L}, z_\mathrm{R}) \subset \hat{V}(z_\mathrm{L}, z_0, z_\mathrm{R})$). So (34) may be inserted into δJ. With δv_j and $\delta u_j \in \hat{V}(z_\mathrm{L}, z_0, z_\mathrm{R}) \backslash V(z_\mathrm{L}, z_\mathrm{R})$, then only the $\delta \int_S$ term is left, yielding:

$$\begin{pmatrix} \hat{\rho} & \hat{\lambda} - 1\\ -\hat{\rho} & -\hat{\lambda}\\ 1 - \hat{\mu} & \hat{\tau}\\ \hat{\mu} & -\hat{\tau} \end{pmatrix}(u(z_0+0) - u(z_0-0)) = 0,$$

$$\begin{pmatrix} \hat{\rho}^{+} & -\hat{\mu}^{+}\\ -\hat{\rho}^{+} & \hat{\mu}^{+} - 1\\ \hat{\lambda}^{+} & \hat{\tau}^{+}\\ 1 - \hat{\lambda}^{+} & -\hat{\tau}^{+} \end{pmatrix}(v(z_0+0) - v(z_0-0)) = 0 \tag{A3.1$_\mathrm{S}$}$$

or

$$\begin{pmatrix} I_N \otimes \alpha_z - \hat{\rho}\\ \hat{\rho} \end{pmatrix}(u(z_0+0) - u(z_0-0)) = 0,$$

$$\begin{pmatrix} \hat{\rho}^{+}\\ I_N \otimes \alpha_z - \hat{\rho}^{+} \end{pmatrix}(v(z_0+0) - v(z_0-0)) = 0, \tag{A3.1$_\mathrm{D}$}$$

respectively. Here $\hat{\lambda}$, $\hat{\mu}$, $\hat{\rho}$, $\hat{\tau}$ are of the form $\hat{\kappa}$ with matrix elements

$$\hat{\kappa}_{mn} := \frac{1}{|F|}\int_S\int_S \exp[-\mathrm{i}(\boldsymbol{k}_\| + \boldsymbol{g}_m)\boldsymbol{r}_\|]\kappa(\boldsymbol{r}_\|, \boldsymbol{r}'_\|)\exp[\mathrm{i}(\boldsymbol{k}_\| + \boldsymbol{g}_n)\boldsymbol{r}'_\|]\,\mathrm{d}^2 r_\|\,\mathrm{d}^2 r'_\|. \tag{A3.2}$$

Obviously the ranks of (A3.1$_\mathrm{S}$) and (A3.1$_\mathrm{D}$) are $2N$ and $4N$ independently of the weight functions λ, μ, ρ, τ; so (A3.1) is equivalent to the matching condition (38).

Appendix 4. Extension of the non-relativistic variational expression to functions of class $W_2^{(1)}$

The variational principle can be extended to a larger class of functions in the following way. Let $\tilde{V}$ be defined as function space

$$\tilde{V} := \{\psi : \Omega \to \mathbb{C};\ \psi|_{\Omega_i^{(j)}} \in W_2^{(1)}(\Omega_i^{(j)}) \text{ for } i = 0, 1; j = 1, \ldots, q_i\}. \tag{A4.1}$$

The boundary conditions '$\partial_n^\alpha \varphi_0 = t\partial_n^\alpha \varphi_1$ and $\partial_n^\alpha \chi_1 = t^*\partial_n^\alpha \chi_0$ on $S \cup S_t$' are regarded as essential boundary conditions for $\alpha = 0$ and as natural boundary conditions for $\alpha = 1$. Thus, choosing $\tau = 0$, the $\int_{S\cup S_t}$-terms in equation (19$_\mathrm{S}$) vanish.

For the boundary conditions on the surface F, three special cases are considered.

(i) $U_0(F) = \{0\}$ and $W_0(F) = L_2(F)^2$. Then $\varphi|_F = f_0^{(0)}$ is an essential boundary condition and $\partial_n\varphi|_F = f_0^{(1)}$ is a natural boundary condition; $\partial_n^\alpha\chi|_F$ is not subject to any boundary condition. The variational expression J is given by

$$J = \int_{\Omega_0\cup\Omega_1} [\nabla\chi^*\nabla\varphi + (w - E)\chi^*\varphi]\,\mathrm{d}^3 r - \int_F \chi^* f_0^{(1)}\,\mathrm{d}\sigma \tag{A4.2}$$

with domain $\{(\chi, \varphi) \in \tilde{V}\times\tilde{V};\ \varphi|_F = f_0^{(0)},\ \varphi_0 = t\varphi_1 \text{ and } \chi_1 = t^*\chi_0 \text{ on } S \cup S_t\}$.

(ii) $U_0(F) = W_0(F) = \{0\}\times L_2(F)$. Then $\varphi|_F = f_0^{(0)}$ and $\chi|_F = g_0^{(0)}$ are essential boundary conditions. The variational expression is

$$J = \int_{\Omega_0\cup\Omega_1} [\nabla\chi^*\nabla\varphi + (w - E)\chi^*\varphi]\,\mathrm{d}^3 r \tag{A4.3}$$

with domain

$\{(\chi, \varphi) \in \tilde{V}\times\tilde{V};\ \varphi|_F = f_0^{(0)},\ \chi|_F = g_0^{(0)},\ \varphi_0 = t\varphi_1 \text{ and } \chi_1 = t^*\chi_0 \text{ on } S \cup S_t\}$.

(iii) $U_0(F) = \{(f^{(0)}, f^{(1)})^\mathrm{T} \in L_2(F)^2;\ f^{(1)} = u(f^{(0)}),\ f^{(0)} \in U_{00}(F)\}$, $W_0(F) = \{(g^{(0)}, g^{(1)})^\mathrm{T} \in L_2(F)^2;\ g^{(1)} = w(g^{(0)}),\ g^{(0)} \in W_{00}(F)\}$ where $U_{00}(F)$ and $W_{00}(F)$ are linear spaces such that $W_2^{(1/2)}(F) \subset U_{00}(F) \subset L_2(F)$ and $W_2^{(1/2)}(F) \subset W_{00}(F) \subset L_2(F)$ and u, w are injective linear mappings $u: U_{00}(F) \to L_2(F)$, $w: W_{00}(F) \to L_2(F)$ such that $\int_F [w(g^{(0)})^* f^{(0)} - g^{(0)*}u(f^{(0)})]\,\mathrm{d}\sigma = 0$ for $f^{(0)} \in U_{00}(F)$ and $g^{(0)} \in W_{00}(F)$. Then the boundary conditions (10), (11) can be written as

$$\partial_n\varphi = f_0^{(1)} + u(\varphi - f_0^{(0)}), \qquad \partial_n\chi = g_0^{(1)} + w(\chi - g_0^{(0)}); \tag{A4.4}$$

these have to be regarded as natural boundary conditions for the variational expression

$$J = \int_{\Omega_0\cup\Omega_1} [\nabla\chi^*\nabla\varphi + (w - E)\chi^*\varphi]\,\mathrm{d}^3 r - \int_F [g_0^{(1)*}\varphi + \chi^* f_0^{(1)}]\,\mathrm{d}\sigma$$
$$-\tfrac{1}{2}\int_F [w(\chi - g_0^{(0)})^*(\varphi - f_0^{(0)}) + (\chi^* - g_0^{(0)*})u(\varphi - f_0^{(0)})]\,\mathrm{d}\sigma \tag{A4.5}$$

with domain $\{(\chi, \varphi) \in \tilde{V}\times\tilde{V};\ \varphi_0 = t\varphi_1 \text{ and } \chi_1 = t^*\chi_0 \text{ on } S \cup S_t\}$.

For example, the scattering condition (47) can be rewritten in the form (A4.4) with $U_{00}(F) := W_{00}(F) := L_2(\partial K_R)$:

$$u(f^{(0)}) := \mathrm{i}kf^{(0)} \qquad \text{and} \qquad w(g^{(0)}) := -\mathrm{i}kg^{(0)} \tag{A4.6}$$

(i.e. in the form of Sommerfeld's 'radiation condition' (Sommerfeld 1947)). The variational expression is

$$J = \int_{K_R} [\nabla\chi^*\nabla\varphi + (w - E)\chi^*\varphi]\,\mathrm{d}^3r - \int_{\partial K_R} \{\partial_r[\exp(\mathrm{i}\boldsymbol{k}_2\boldsymbol{r})]\varphi + \chi^*\partial_r[\exp(\mathrm{i}\boldsymbol{k}_1\boldsymbol{r})]\}\,\mathrm{d}\sigma$$

$$-\mathrm{i}k\int_{\partial K_R} [\chi^* - \exp(\mathrm{i}\boldsymbol{k}_2\boldsymbol{r})][\varphi - \exp(\mathrm{i}\boldsymbol{k}_1\boldsymbol{r})]\,\mathrm{d}\sigma \tag{A4.7}$$

defined on $W_2^{(1)}(K_R) \times W_2^{(1)}(K_R)$.

References

Arthurs A M 1980 *Complementary Variational Principles* (Oxford: Clarendon)
Bevensee R M 1961 *Ann. Phys.* **12** 222–63
Bross H 1976 *Verhandl. DPG(VI)* **11** 621
—— 1977 *Z. Phys.* B **28** 173–82
—— 1981 *Verhandl. DPG(VI)* **16** 348
—— 1982 *J. Phys. C: Solid State Phys.* to be published
Bross H, Bohn G, Meister G, Schubö W and Stöhr H 1970 *Phys. Rev.* B **2** 3098–103
Bross H and Hofmann I 1969 *Z. Phys.* **229** 123–30
Bross H, Marschall E and Stöhr H 1979 *Verhandl. DPG(VI)* **14** 253
Courant R and Hilbert D 1924 *Methoden der Mathematischen Physik* I (Berlin: Springer) pp 181–3
Dirac P A M 1947 *The Principles of Quantum Mechanics* (Oxford: Clarendon) p 191
Ferreira L G, Oliveira M J, Silva E Z, Graca G C and Franco B J O 1975 *J. Phys. C: Solid State Phys.* **8** 992–9
Ferreira L G, Pereira J R, Lida D and Agostinho A 1974 *J. Phys. C: Solid State Phys.* **7** 1076–84
Holzwarth N A W and Lee M J G 1978 *Phys. Rev.* B **18** 5350–64
Kohn W 1948 *Phys. Rev.* **74** 1763–72
—— 1952 *Phys. Rev.* **87** 472–81
Leigh R S 1956 *Proc. Phys. Soc.* A **69** 388–400
Lopez-Aguilar F 1979 *Phys. Status Solidi* b **95** 171–8
Loucks T 1967 *Augmented Plane Wave Method* (New York: Benjamin)
Marcus P M 1967 *Int. J. Quantum Chem.* **1S** 567–88
Marcus P M and Jepsen D W 1968 *Phys. Rev. Lett.* **20** 925–9
Mikhlin S G 1970 *Mathematical Physics, an Advanced Course* (Amsterdam: North-Holland) p 16
Nečas J 1967 *Les Méthodes Directes en Théorie des Equations Elliptiques* (Paris: Masson)
Sarker A Q and Taj-ul Islam 1975 *Phys. Status Solidi* b **71** 641–50
Schlosser H and Marcus P M 1963 *Phys. Rev.* **131** 2529–46
Sommerfeld A 1947 *Vorlesungen über Theoretische Physik* (Leipzig: Akad. Verlagsges.) vol VI, p 192
Wilkinson J H and Reinsch C 1971 *Linear Algebra* (Berlin: Springer)