Anomaly-free singularities in the generalized Kohn variational method

J N Cooper[1], E A G Armour[1] and M Plummer[2]

[1] School of Mathematical Sciences, University Park, Nottingham NG7 2RD, UK
[2] STFC Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, UK

E-mail: james.cooper@maths.nottingham.ac.uk

Abstract

We have carried out an analysis of singularities in Kohn variational calculations
for low-energy (e^+–H_2) elastic scattering. Provided that a sufficiently accurate
trial wavefunction is used, we argue that our implementation of the Kohn
variational principle necessarily gives rise to singularities which are not
spurious. We propose two approaches for optimizing a free parameter of the
trial wavefunction in order to avoid anomalous behavior in scattering phase shift
calculations, the first of which is based on the existence of such singularities.
The second approach is a more conventional optimization of the generalized
Kohn method. Close agreement is observed between the results of the two
optimization schemes; further, they give results which are seen to be effectively
equivalent to those obtained with the complex Kohn method. The advantage
of the first optimization scheme is that it does not require an explicit solution
of the Kohn equations to be found. We give examples of anomalies which
cannot be avoided using either optimization scheme but show that it is possible
to avoid these anomalies by considering variations in the nonlinear parameters
of the trial function.

PACS numbers: 02.10.Yn, 34.80.Uv

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Despite the absence of an explicit minimization principle, variational methods have been used
successfully in many problems of quantum scattering theory. Such calculations typically
exploit a stationary principle in order to obtain an accurate description of scattering processes.
The Kohn variational method [1] has been applied extensively to problems in electron–atom [2]
and electron–molecule [3, 4] scattering, as well as to the scattering of positrons, e^+, by atoms

[5–7] and molecules [8–10]. It has been widely documented, however, that matrix equations derived from the Kohn variational principle are inherently susceptible to spurious singularities. These singularities were discussed first by Schwartz [11, 12] and have subsequently attracted considerable attention [13–16]. In the region of these singularities, results of Kohn calculations can be anomalous.

Although sharing characteristics similar to those exhibited by scattering resonances [15], Schwartz singularities are nonphysical and arise only because the trial wavefunction, used in Kohn calculations to represent scattering, is inexact [13]. For projectiles of a given incident energy, anomalous results are confined to particular formulations of the trial wavefunction and can, in principle, be mitigated by a small change in boundary conditions or some other parameter. It has also been shown [4, 17, 18] that the use of a complex-valued trial wavefunction avoids anomalous behavior except in exceptional circumstances. Alternative versions of the Kohn method have been developed in terms of a Feshbach projection operator formalism [19] and have been found [20] to give anomaly-free results.

In this paper, we will discuss our investigations of Schwartz-type anomalies for generalized Kohn calculations involving the elastic scattering of positrons by molecular hydrogen, H_2. We will find that our choice of trial wavefunction contains a free parameter that can be varied in such a way as to produce singularities which are legitimate in the context of the scattering theory and do not give rise to anomalous results. Indeed, these singularities can be used to formulate an optimization scheme for choosing the free parameter so as to automatically avoid anomalous behavior in calculations of the scattering phase shift. The novelty of determining the phase shift in this way is that an explicit solution of the linear system of Kohn equations is not required. We will also develop an alternative optimization and show that the two schemes give results in close agreement. Further, the results obtained will be seen to be in excellent agreement at all positron energies with those determined via the complex Kohn method.

We will give examples of anomalous behavior which cannot be avoided with either optimization and show that the same anomalies appear in our application of the complex Kohn method. We will discuss circumstances under which these anomalies might occur. We will also show that such results are nonphysical by considering small changes in the nonlinear parameters of the trial wavefunction.

Our investigations of singular behavior have been carried out as part of a wider study on $(e^+–H_2)$ scattering and annihilation using extremely flexible wavefunctions. Our ability to recognize clearly and analyze the anomalous behavior is as good for this system as it would be for a simpler model system, with the advantage that our calculations can be used to provide meaningful and physically relevant results [10].

2. Theory

2.1. The generalized Kohn variational method

The Kohn variational method is used to calculate approximations to exact scattering wavefunctions. Determining an approximation, Ψ_t, allows a variational estimate, η_v, of the scattering phase shift to be calculated, the error in which is of second order in the error of Ψ_t from the exact scattering wavefunction, Ψ [21]. The standard approach in Kohn calculations is to assume an overall form for Ψ_t that depends linearly on a set of unknown parameters, optimal values for which are then determined by the application of a stationary principle.

In our investigations of anomalous behavior in Kohn calculations for $(e^+–H_2)$ scattering, we have studied the lowest partial wave of Σ_g^+ symmetry. This partial wave has been shown [8] to be the only significant contributor to scattering processes for incident positron energies

below 2 eV. The first significant inelastic channel is positronium formation which has a
threshold at 8.63 eV [21]. Although we will consider here positron energies higher than these
thresholds, it is not our intention to provide a comprehensive physical treatment of the scattering
problem taking higher partial waves and inelastic processes into account. The purpose of the
present study is to give a correct and, as far as possible, anomaly-free treatment of the lowest
partial wave. It is important to examine the single-channel case as accurately as possible
as a preliminary for more sophisticated calculations. By not taking into account additional
channels, it is possible that anomalous behavior could occur due to physical inaccuracies in the
trial wavefunction at higher energies. However, we will demonstrate that all of the anomalies
in our results ultimately can be attributed to purely numerical effects.

We have initially used a trial wavefunction having the same general form as described in
our earlier calculations [9],

$$\Psi_t = (\bar{S} + a_t \bar{C} + p_0 \chi_0)\psi_G + \sum_{i=1}^{M} p_i \chi_i, \tag{1}$$

where

$$\begin{bmatrix} \bar{S} \\ \bar{C} \end{bmatrix} = \begin{bmatrix} \cos(\tau) & \sin(\tau) \\ -\sin(\tau) & \cos(\tau) \end{bmatrix} \begin{bmatrix} S \\ C \end{bmatrix}, \tag{2}$$

for some phase parameter, $\tau \in [0, \pi)$, with

$$S = \frac{N}{\lambda_3 - 1} \sin[c(\lambda_3 - 1)] \tag{3}$$

and

$$C = \frac{N}{\lambda_3 - 1} \cos[c(\lambda_3 - 1)]\{1 - \exp[-\gamma(\lambda_3 - 1)]\}. \tag{4}$$

As before [9], we have carried out calculations using the fixed-nuclei approximation [22, 23],
taking the internuclear separation to be at the equilibrium value, $R = 1.4$ au. We have labeled
the electrons as particles 1 and 2, taking the positron to be particle 3. The position vector, $\mathbf{r}_j$, of
each lepton is described by the prolate spheroidal coordinates [24] $(\lambda_j, \mu_j, \phi_j)$, $j \in \{1, 2, 3\}$.
These coordinates are defined implicitly in terms of the Cartesian coordinates, (x_j, y_j, z_j), as

$$x_j = \tfrac{1}{2}R\left[\left(\lambda_j^2 - 1\right)\left(1 - \mu_j^2\right)\right]^{\frac{1}{2}} \cos(\phi_j), \tag{5}$$

$$y_j = \tfrac{1}{2}R\left[\left(\lambda_j^2 - 1\right)\left(1 - \mu_j^2\right)\right]^{\frac{1}{2}} \sin(\phi_j), \tag{6}$$

$$z_j = \tfrac{1}{2}R\lambda_j\mu_j. \tag{7}$$

The functions S and C represent, respectively, the incident and scattered positrons
asymptotically far from the target. The shielding parameter, γ, ensures the regularity of C at
the origin and is taken to have the value $\gamma = 0.75$. The constant, c, is defined to be $c = kR/2$,
with k being the magnitude of the positron momentum in atomic units. N is a normalization
constant and can here be regarded as arbitrary. The unknowns, a_t and $\{p_0, \ldots, p_M\}$, are
constants to be determined. The inclusion of the parameter, τ, in Ψ_t is a generalization of the
Kohn method due to Kato [25, 26]. This parameter is of only minor physical significance,
playing the role of an additive phase factor in the part of the wavefunction representing the
incident and scattered positrons asymptotically far from the target. However, at each value of
k, the value of τ can be varied to avoid spurious singularities in the Kohn calculations. Away
from the spurious singularities, for an accurate trial wavefunction we can expect the variation

in the calculated values of η_v over τ to be small. In the original application of the Kohn method [1], only wavefunctions corresponding to $\tau = 0$ were considered.

The function, ψ_G, is an approximation to the ground-state wavefunction of the unperturbed hydrogen molecule and is determined by the Rayleigh–Ritz variational method [27]. In the calculations presented here, we have taken ψ_G to be the target wavefunction described in detail in another of our previous calculations [10], accounting for 96.8% of the correlation energy [27] of H_2. The function,

$$\chi_0 = \frac{N}{\lambda_3 - 1} \cos[c(\lambda_3 - 1)]\{1 - \exp[-\gamma(\lambda_3 - 1)]\} \exp[-\gamma(\lambda_3 - 1)], \qquad (8)$$

is the same as has been used in our earlier calculations [9, 10] and was introduced first by Massey and Ridley [3]. The remaining short-range correlation functions, $\Omega = \{\chi_1, \ldots, \chi_M\}$, allow for the description of direct electron–positron and electron–electron interactions. Here, we have used the same set of $M = 279$ correlation functions described in detail in equations (5)–(8) of [10]. The general form of each function, χ_i, is

$$\chi_i = f_i(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \exp[-\beta(\lambda_1 + \lambda_2) - \alpha\lambda_3] \qquad (1 \leqslant i \leqslant M), \qquad (9)$$

where each $f_i(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ is symmetric in the coordinates of the electrons. They are a mixture of separable correlation functions and Hylleraas-type functions [28] containing the electron–positron distance as a linear factor. As discussed previously [10], the Hylleraas-type functions in particular allow for high accuracy of results away from anomalous singularities. Unless otherwise noted, we have here chosen values of $\alpha = 0.6$ and $\beta = 1.0$ rather than the values of $\alpha = 0.3$ and $\beta = 0.7$ used earlier [10]. This choice of nonlinear parameters highlights the interesting aspects of Schwartz-type anomalies more clearly.

In our application of the Kohn variational principle [9, 21], the functional

$$\mathcal{J}[\Psi_t] = \tan(\eta_v - \tau + c) = a_t - \frac{2}{\pi N^2 R^2 k} \langle \Psi_t, \Psi_t \rangle \qquad (10)$$

is made stationary with respect to variations in a_t and $\{p_0, \ldots, p_M\}$. Here, we have denoted the integral $\langle \Psi_t | (\hat{H} - E) | \Psi_t \rangle$ by $\langle \Psi_t, \Psi_t \rangle$, where $\hat{H}$ is the nonrelativistic Hamiltonian for the scattering system and E is the sum of the positron kinetic energy and the ground-state energy expectation value of ψ_G. The integral is evaluated over the configuration space of the positron and the two electrons. We will, henceforth, use this notation more generally to denote integrals of the form $\langle X | (\hat{H} - E) | Y \rangle$ by $\langle X, Y \rangle$.

The stationary principle imposed upon (10) leads to the linear system of equations

$$Ax = -b, \qquad (11)$$

where

$$A = \begin{bmatrix} \langle \bar{C}\psi_G, \bar{C}\psi_G \rangle & \langle \bar{C}\psi_G, \chi_0\psi_G \rangle & \cdots & \langle \bar{C}\psi_G, \chi_M \rangle \\ \langle \chi_0\psi_G, \bar{C}\psi_G \rangle & \langle \chi_0\psi_G, \chi_0\psi_G \rangle & \cdots & \langle \chi_0\psi_G, \chi_M \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \chi_M, \bar{C}\psi_G \rangle & \langle \chi_M, \chi_0\psi_G \rangle & \cdots & \langle \chi_M, \chi_M \rangle \end{bmatrix}, \qquad (12)$$

$$b = \begin{bmatrix} \langle \bar{C}\psi_G, \bar{S}\psi_G \rangle \\ \langle \chi_0\psi_G, \bar{S}\psi_G \rangle \\ \vdots \\ \langle \chi_M, \bar{S}\psi_G \rangle \end{bmatrix}, \qquad (13)$$

$$x = \begin{bmatrix} a_t \\ p_0 \\ \vdots \\ p_M \end{bmatrix}. \tag{14}$$

Solving (11) determines the values of a_t and $\{p_0, \dots, p_M\}$, allowing $\langle \Psi_t, \Psi_t \rangle$ and, hence, η_v to be calculated via (10). However, as has been discussed extensively (see, for example, [9, 29–31]), the particular form of the functions, $f_i(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$, used in our calculations does not, in general, permit analytic evaluation of the integrals comprising the matrix elements of A and b. Sophisticated methods to determine these integrals numerically have been developed [29–31]. However, the numerical approaches can give only accurate approximations to the exact values of the integrals, so that small errors in determining the elements of A and b are unavoidable.

2.2. Singularities

Singularities in our generalized Kohn calculations arise from zeros of $\det(A)$, the determinant of (12). Under these circumstances, the linear system (11) has no unique solution. Close to these singularities, it is well known [11–15] that values of η_v obtained by solving (11) can be anomalous; small errors in the elements of A or b can correspond to large errors in the solution, x, particularly when A is close to singularity in a sense that can be defined formally in terms of the condition number [32, 33] of A. A more detailed discussion of the condition number will be given in section 3.2.

It is appropriate at this point to define a convention that we will adopt in our discussion of singularities in the generalized Kohn method. The type of spurious singularities mentioned by Schwartz [11, 12] here corresponds to zeros in $\det(A)$ for the particular case when $\tau = 0$. We will, however, find it convenient to label as Schwartz singularities those zeros of $\det(A)$ occurring at any $\tau_s \in [0, \pi)$ which give rise to anomalous behavior in the calculations of $\eta_v(\tau)$ when τ is near τ_s. This is an important clarification for the following reason: we claim that, because of our inclusion of τ in Ψ_t, there exist zeros of $\det(A)$ which are not spurious and do not correspond to anomalous behavior in the values of η_v. We will refer to such singularities as anomaly-free singularities.

To understand how anomaly-free singularities might arise, it is helpful to consider the component, Ψ_0, of the exact scattering wavefunction, Ψ, corresponding to the lowest partial wave. Ψ_0 can be expanded as

$$\Psi_0 = (\bar{S} + a\bar{C})\psi + \sum_{i=1}^{\infty} p_i \zeta_i, \tag{15}$$

where ψ is the exact ground-state target wavefunction and the complete set of correlation functions, $\{\zeta_i\}$, describes exactly the leptonic interactions at short range. As noted by Takatsuka and Fueno [16] in their Kohn calculations for single-channel scattering, the exact phase shift, η_0, determined by Ψ_0 is independent of the choice of τ in (15). As a result, there is precisely one value, $\tau_0 \in [0, \pi)$, at each positron energy such that

$$\eta_0 - \tau_0 + c = \pm n\frac{\pi}{2}, \tag{16}$$

for some odd value of $n > 0$, where either $+n$ or $-n$ is chosen to keep $\eta_0 \in (-\pi/2, \pi/2]$. The value of $\cot(\eta_0 - \tau + c)$ will then pass continuously through zero as τ passes through τ_0.

Returning to the trial wavefunction (1), for nonsingular A, it can be shown using (10) and (11) that

$$\cot(\eta_v - \tau + c) = \frac{\pi N^2 R^2 k}{2} \left(\frac{\det(A)}{\hat{x} \cdot b - \det(A)\langle \bar{S}\psi_G, \bar{S}\psi_G \rangle} \right), \tag{17}$$

where we have defined

$$\hat{x} = [\text{adj}(A)]b, \tag{18}$$

with $\text{adj}(A)$ being the adjugate matrix of (12). We note that $\text{adj}(A)$ exists even if A is singular. In section 3.1, we will provide numerical evidence that the limit of $\cot(\eta_v - \tau + c)$ as $\det(A) \to 0$ exists and is equal to zero. We can, therefore, assert a correspondence between zeros of $\cot(\eta_v - \tau + c)$ and zeros of $\det(A)$.

Suppose that, at each k, there are m_k values of τ making A singular, which we will denote by $\Upsilon_k = \left\{ \tau_s^{(k,1)}, \tau_s^{(k,2)}, \ldots, \tau_s^{(k,m_k)} \right\}$. If our previous assertion is correct, each element of Υ_k will correspond to a zero of $\cot(\eta_v - \tau + c)$. We can reasonably claim that if Ψ_t is, in some sense, sufficiently accurate, precisely one of these zeros will correspond to the zero of $\cot(\eta_0 - \tau + c)$ necessarily found at τ_0 for Ψ_0. Assuming that this is the case, we will denote the element of Υ_k corresponding to τ_0 by $\hat{\tau}_s$. Values of η_v determined in the generalized Kohn method should then vary slowly and smoothly with τ as it passes through $\hat{\tau}_s$. Indeed, as $\tau \to \hat{\tau}_s$ from either side, we would expect the values of η_v calculated by solving (11) to converge to the value, $\hat{\eta}_v$, determined directly from

$$\hat{\eta}_v - \hat{\tau}_s + c = \pm n\frac{\pi}{2}, \tag{19}$$

where n is again chosen so that $\hat{\eta}_v \in (-\pi/2, \pi/2]$.

In the following section, we will present results of generalized Kohn calculations exhibiting anomalous behavior due to Schwartz singularities and, further, demonstrate empirically that the anomaly-free singularities do exist and that values of $\hat{\tau}_s$ can be found. At each k, choosing $\tau = \hat{\tau}_s(k)$ then defines an optimization of τ that will be seen to avoid anomalies in η_v due to Schwartz singularities.

3. Results

3.1. Calculations of phase shift

In our generalized Kohn calculations, we have obtained values of $\eta_v \in (-\pi/2, \pi/2]$ using (10) and (11), for a range of positron momenta. Spurious singularities have been accounted for by performing calculations over p different values of τ equidistant in the range $\tau \in [0, \pi)$. For the results presented here, we have taken $p = 1001$. Calculations for a large number of τ values can be carried out with minimal additional computational effort, as it can be shown that the matrix elements of A and b for any τ are readily available from the elements of $A(\tau = 0)$ and $b(\tau = 0)$ via an orthogonal transformation. It is helpful to carry out calculations for large p, as it allows a detailed examination of the behavior of η_v very close to Schwartz singularities to be made. Values of η_v over the p values of τ are given in figure 1 for $k = 0.2$, corresponding to a positron energy of 0.54 eV. Anomalous results due to a Schwartz singularity are clearly evident around $\tau \sim 2.87$. We have indicated the value of τ giving rise to the singularity by a dashed line. Away from this value of τ, the variation in η_v is small.

In figure 2, we have studied the behavior of $\cot(\eta_v - \tau + c)$ close to the singularity at $\tau \sim 2.87$ by calculating $\cot(\eta_v - \tau + c)$ for 101 values of τ equidistant in the range $\tau \in [2.868, 2.872]$. We have again indicated the position of the singularity in this figure by a
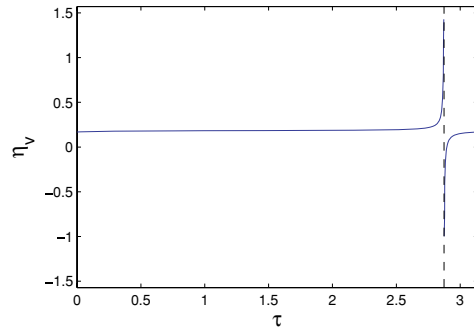
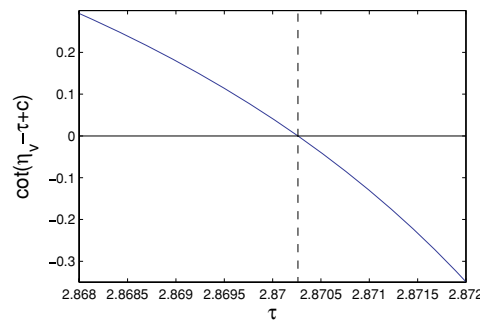**Figure 1.** Values of $\eta_v(\tau)$ at $k = 0.2$.

Figure 2. The behavior of $\cot(\eta_v - \tau + c)$ at $k = 0.2$ for values of τ either side of a singularity.

dashed line. The results shown in the figure suggest that $\cot(\eta_v - \tau + c)$ converges smoothly to zero as $\tau \to \tau_s$ from either side, supporting the assertion made in section 2.2 regarding the correspondence of the zeros in $\cot(\eta_v - \tau + c)$ and $\det(A)$. We have found that behavior of the type shown in figure 2 is a general feature of the calculation.

Using (2) and (12), it is straightforward to show that

$$\det(A) = \mathcal{A}(k)\sin^2(\tau) + \mathcal{B}(k)\sin(\tau)\cos(\tau) + \mathcal{C}(k)\cos^2(\tau), \tag{20}$$

where $\mathcal{A}(k)$, $\mathcal{B}(k)$ and $\mathcal{C}(k)$ are constants with respect to variations in τ. For a given positron momentum, the constants, $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$, can be determined by calculating $\det(A)$ directly from (12) at particular values of τ. Strictly speaking, in our calculations we have evaluated $\det(\tilde{A})$, with $\tilde{A}$ being the approximation to A whose elements have been determined using numerical integration. We will assume that the values of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are not unduly sensitive to small changes in the elements of A and henceforth take $\det(A)$ and $\det(\tilde{A})$ to be essentially equivalent.

At each k, provided that $\mathcal{A} \neq 0$, the values, τ_s, making A singular can be found by solving the quadratic equation in $\tan(\tau_s)$,

$$\mathcal{A}\tan^2(\tau_s) + \mathcal{B}\tan(\tau_s) + \mathcal{C} = 0. \tag{21}$$

If only τ is varied, and unless $\mathcal{A} = \mathcal{B} = \mathcal{C} = 0$, there will be no more than two zeros of $\det(A)$ in the range $\tau \in [0, \pi)$. Figure 3 shows $\det(A)$ as a function of τ at $k = 0.2$. The scale on the vertical axis is unimportant, since the value of $\det(A)$ at each k can be made arbitrarily
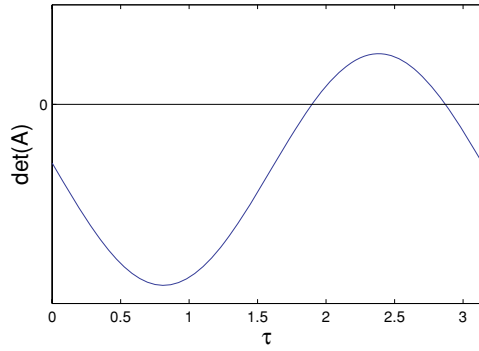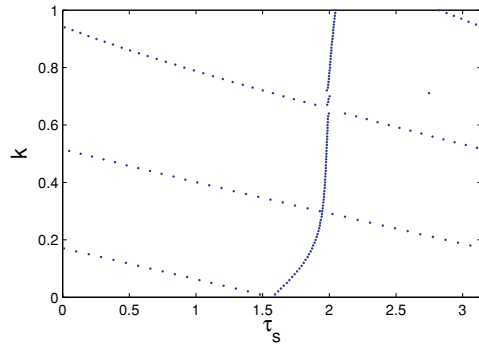
Figure 3. Values of det(A) at $k = 0.2$ for $0 \leqslant \tau < \pi$.

Figure 4. The zeros of det(A) for $0.01 \leqslant k \leqslant 1$.

large or small by a choice of the normalization constant, N. The result of interest in the figure is that it indicates two values of τ at which A is singular. The anomalous behavior in figure 1 corresponds directly to the singularity observed at $\tau \sim 2.87$ in figure 3. However, there are no anomalies in figure 1 corresponding to the singularity at $\tau \sim 1.90$ in figure 3, suggesting that this singularity is of the anomaly-free type described in the previous section.

We have examined this phenomenon at other values of the positron momentum. Figure 4 indicates the roots of (21) for 100 different positron momenta equidistant in the range $0.01 \leqslant k \leqslant 1$, corresponding to a positron energy range from 1.36 meV to 13.6 eV. For the majority of positron momenta considered here, $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are such that there are two values of τ_s at each k, the exceptions being $k = 0.65$ and $k = 0.66$, for which we have found no real-valued solutions of (21). It is apparent from figure 4 that the roots of (21) lie in two families of curves. The first family spans the entire range, $\tau_s \in [0, \pi)$, for $0.01 \leqslant k \leqslant 1$. The second family is confined to values of τ_s in the range $\tau_s \in [1.5, 2.1]$ for all positron momenta considered here. For almost every k where real roots exist, there is precisely one root corresponding to each family. The exception is the result at $k = 0.71$, where there is an irregularity in the otherwise smooth behavior of τ_s over k. We will discuss this phenomenon in more detail in section 3.2.

To illustrate anomalies in η_v at different values of k, it is convenient to define the function,

$$\Delta(k, \tau) = |\eta_v(k, \tau) - \langle \eta_v \rangle(k)|, \tag{22}$$
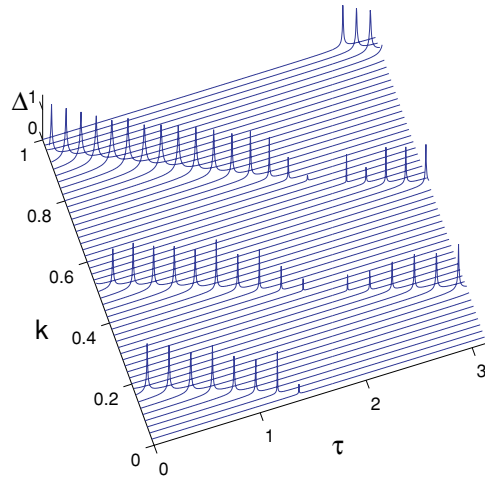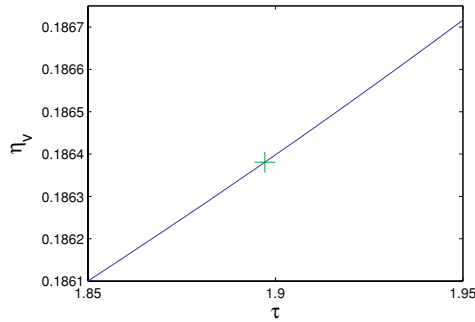
Figure 5. Values of $\Delta(k, \tau)$ for $0.02 \leqslant k \leqslant 1$.

Figure 6. Convergence of [——] $\eta_v(\tau)$ to [+] $\hat{\eta}_v$ as $\tau \to \hat{\tau}_s$ at $k = 0.2$.

where, at each k, $\langle \eta_v \rangle$ is the median value of $\eta_v(\tau)$ evaluated over the p values of τ. $\Delta(k, \tau)$ measures the degree to which a given $\eta_v(k, \tau)$ can be considered anomalous. The values of $\Delta(k, \tau)$ are shown in figure 5. For clarity, we have included results only for 50 values of k equidistant in the range $0.02 \leqslant k \leqslant 1$ rather than the 100 values used for figure 4. The omission of the results for $k = 0.71$ in this figure also allows us to delay until section 3.2, the discussion of the atypical singularity observed at this value of k in figure 4.

It is clear from figure 5 that anomalies are observed corresponding to only the first of the two families of curves identified from figure 4. This is strong evidence that the curve for which no anomalies are observed comprises legitimately occurring singularities. It is interesting to note that the size of the anomalies due to the Schwartz singularities becomes noticeably smaller in figure 5 as they tend to coincide with the apparently anomaly-free singularities. Denoting by $\tau_s^{(1)}(k)$ the values of τ describing the anomaly-free curve, at each k we expect that, as $\tau \to \tau_s^{(1)}$, the values of η_v calculated over τ will converge smoothly to the value, $\hat{\eta}_v$, determined directly from (19) and taking $\hat{\tau}_s = \tau_s^{(1)}$. We have found in our calculations that this is indeed the case, and an example is shown in figure 6 for $k = 0.2$.
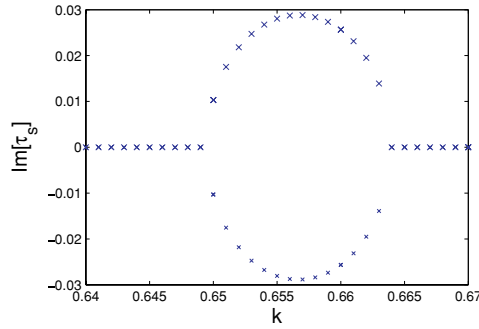
Figure 7. Values of $\text{Im}[\tau_s]$ for $0.64 \leqslant k \leqslant 0.67$.

We can reasonably conclude that we have encountered anomaly-free singularities. Before we can develop an optimization for τ based on these singularities, however, there are two outstanding issues to be addressed. Firstly, we have already noted that no real-valued solutions of (21) were found at $k = 0.65$ and 0.66. This is potentially problematic as, for a sufficiently accurate trial function, we should expect at least one real root of (21) at each k, corresponding to an anomaly-free singularity. However, inspection of figure 4 shows that the solutions of (21) are close together in the regions either side of $k = 0.65$ and 0.66. Were the two roots to coincide at some k, then $\mathcal{B}^2 = 4\mathcal{A}\mathcal{C}$. Near a point of coincidence, $\mathcal{B}^2 \sim 4\mathcal{A}\mathcal{C}$ and small errors in the values of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ could erroneously give rise to $(\mathcal{B}^2 - 4\mathcal{A}\mathcal{C}) < 0$. The values of τ solving (21) at $k = 0.65$ and 0.66 were found to be, respectively, $\tau_s \sim 2.02 \pm 0.01\text{i}$ and $\tau_s \sim 1.98 \pm 0.03\text{i}$. In both cases, the fact that $\text{Im}[\tau_s] \ll \text{Re}[\tau_s]$ suggests that singularities do genuinely exist for $\tau \in \mathbb{R}$ at these values of k, but small errors in our calculations of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ due to inexact numerical integration have prevented us from finding them. Having investigated this problem in more detail, in figure 7 we show the calculated values of $\text{Im}[\tau_s]$ for 31 values of k equidistant in the range $0.64 \leqslant k \leqslant 0.67$. There is a clearly defined region of k where no real roots of (21) have been found. The smoothness of $\text{Im}[\tau_s]$ over k in this region does not necessarily preclude the notion that the failure to find real-valued solutions is due to small numerical errors in our calculations. It is conceivable that inaccuracies in the calculated values of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ could also arise from systematic errors in the algorithm [34] used to calculate the determinants. Nevertheless, the results illustrated in figure 7 are interesting; their exact origin may be speculated upon and will remain a subject of our ongoing investigations.

The second difficulty concerns the choice of $\hat{\tau}_s$ from the two available solutions of (21). A method is needed for identifying at each k, the root of (21) corresponding to a legitimate singularity. This can easily be achieved by inspecting values of η_v at values of τ either side of each singularity, although this approach is not ideal as it requires solutions of (11) to be found. In practice, at each k, it should be possible to determine by inspection which of the two phase shifts is anomaly-free by examining corresponding results at singularities for nearby values of k. For example, figure 4 clearly shows that only one curve in the (τ, k) plane corresponds to a physically acceptable variation of phase shift over k.

With these considerations in mind, we claim that choosing $\tau = \hat{\tau}_s$ at each k defines a consistent optimization that can be used to avoid anomalies due to Schwartz singularities appearing at other values of τ. To evaluate the success of this approach, we have found it helpful to consider an alternative optimization of τ. For the calculations of η_v carried out with $p = 1001$, choosing the value of τ at each k giving rise to the median phase shift, $\langle \eta_v \rangle$, should
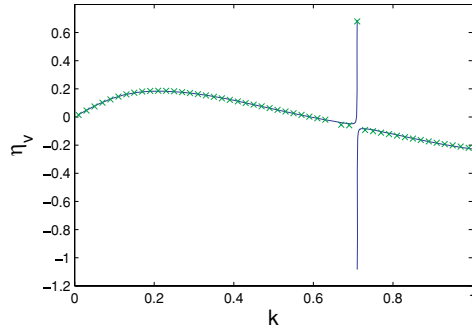
Figure 8. A comparison of optimization schemes for τ, [——] $\langle\eta_v\rangle(k)$ and [×] $\hat{\eta}_v(k)$.

also mitigate anomalous behavior. In figure 8, we have compared results for $\langle\eta_v\rangle(k)$ and $\hat{\eta}_v(k)$ for momenta in the range $0.01 \leqslant k \leqslant 1$. Here, of the two candidates for $\hat{\eta}_v$ corresponding to the two singularities, at each k we have chosen the one whose absolute value is closest to $|\langle\eta_v\rangle|$. For clarity, we have included in the figure values of $\hat{\eta}_v(k)$ for only 50 values of k equidistant in the range $0.01 \leqslant k \leqslant 0.99$.

Both optimization schemes successfully avoid anomalous behavior at most values of k, and there is good agreement between the two sets of results at all momenta. However, the intriguing feature of the figure is the anomaly appearing in both sets of results at $k \sim 0.71$. For $\langle\eta_v\rangle(k)$, we have shown this anomaly in greater detail by including in the figure results of a further 100 Kohn calculations for momenta equidistant in the range $0.7 \leqslant k \leqslant 0.72$, although it is practical to consider henceforth only the anomalous behavior occurring precisely at $k = 0.71$. We believe that the anomaly shown in figure 8 is of a different type to those shown in figure 5, which are due to singularities found by varying only τ at a given k. In the following section, we will examine the circumstances under which persistent anomalies of the kind shown in figure 8 could appear before going on to discuss methods designed to avoid them.

3.2. Persistent anomalous behavior

It is often claimed that anomalous results observed in the region of singularities arise from A having a determinant close to zero. Statements of this kind can be misleading, as the determinant of any nonsingular A can be made arbitrarily close to zero by an appropriate scalar multiplication, without altering the sensitivity of the solution, x, to small errors in the elements of A or b. A better measure for identifying regions where anomalies may occur is the condition number, $\kappa(A)$, defined for nonsingular A. The condition number is independent of the normalization constant, N, and for the linear system (11) is defined as the maximum ratio of the relative error in x and the relative error in b. Formally, it can be shown that

$$\kappa(A) = \|A\|\|A^{-1}\|, \tag{23}$$

with respect to some matrix norm, $\|A\|$ [32]. The value of $\kappa(A)$ is dependent upon the choice of norm. In our calculations, we have considered the matrix 1-norm [32] of A,

$$\|A\|_1 = \max_{1 \leqslant j \leqslant (M+2)} \sum_{i=1}^{M+2} |a_{ij}|, \tag{24}$$

where a_{ij} is the element in the ith row and jth column of A. In what follows, the particular choice of the 1-norm in our calculations will implicitly be assumed.
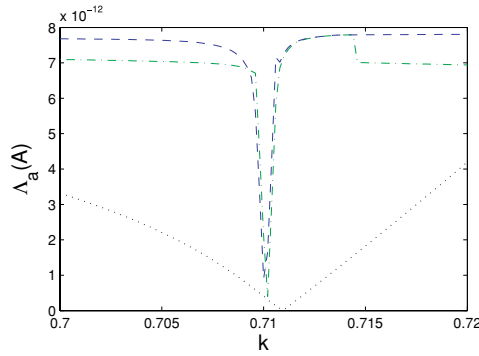
Figure 9. Values of $\Lambda_a(A)$ over k for [- - - -] $\tau = 0$, [— $\cdot$ —] $\tau = \frac{\pi}{4}$ and [$\cdots\cdots$] $\tau = \frac{\pi}{2}$.

A matrix with a large condition number is said to be ill conditioned, and the solution of the corresponding linear system may not be reliable if the elements of A and b are not known exactly. For any invertible A, the condition number can be used to formalize the definition of closeness to singularity in the following way. If ΔA is defined to be any matrix such that $A + \Delta A$ is singular, then the relative distance to singularity, $\Lambda(A)$, for A, is defined [32] to be

$$\Lambda(A) = \min\left(\frac{\|\Delta A\|}{\|A\|} : \quad \det(A + \Delta A) = 0\right). \tag{25}$$

This definition holds for any consistent norm. Further, if $\Lambda(A)$ and $\kappa(A)$ are evaluated using the same choice of norm, it can be shown [32] that

$$\Lambda(A) = [\kappa(A)]^{-1}. \tag{26}$$

In section 3.1 we noted that (21) has, in general, no more than two zeros if only variations of $\tau \in [0, \pi)$ are considered. However, if $\mathcal{A} = \mathcal{B} = \mathcal{C} = 0$ then $\det(A)$ is identically zero independently of τ and no consistent value of phase shift can be calculated, either by solving (11) or directly from (19). There is no obvious physical reason why this circumstance should arise at any $k \in \mathbb{R}$. However, it is conceivable that $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ could coincidentally be close to zero, in some sense, over a narrow range of k. Small errors in the evaluation of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ could then give rise to both unreliable solutions of (21) and persistent anomalies in the calculation of η_v due to ill conditioning in the Kohn equations (11). To see how the latter case arises, using (20) we note that

$$\mathcal{C} = \det[A(\tau = 0)], \tag{27}$$

$$\mathcal{A} = \det\left[A\left(\tau = \frac{\pi}{2}\right)\right], \tag{28}$$

$$\mathcal{B} = 2\det\left[A\left(\tau = \frac{\pi}{4}\right)\right] - \mathcal{A} - \mathcal{C}, \tag{29}$$

so that the notion of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ being close to zero is immediately formalized in terms of $\Lambda(A)$ at $\tau = 0$, $\tau = \pi/4$ and $\tau = \pi/2$.

In our calculations, we have used a numerical algorithm [35] to calculate $\Lambda_a(A)$, an estimate of $\Lambda(A)$. Values of $\Lambda_a(A)$ at $\tau = 0$, $\tau = \pi/4$ and $\tau = \pi/2$ are shown in figure 9 for 100 values of k in the range $0.7 \leqslant k \leqslant 0.72$. Values of $\Lambda_a(A)$ for $\tau = 0$ and $\tau = \pi/4$ are anomalously small at $k \sim 0.71$, with the values for $\Lambda_a(A)$ at $\tau = \pi/2$ also passing through a clear minimum at $k \sim 0.711$. Ordinarily, we would not expect ill conditioning to

occur over a very broad range of τ at a given k. At $k = 0.71$, the small values of $\Lambda_a(A)$ at $\tau = 0, \tau = \pi/4$ and $\tau = \pi/2$ therefore point to a manifestation of ill conditioning which is unusually widespread in τ. In fact, in our calculations we have failed to find any value of $\tau \in [0, \pi)$ such that A is sufficiently well conditioned to avoid anomalous results at $k = 0.71$. We have also confirmed that $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ all pass through zero at least once between $k = 0.71$ and 0.7104.

We can conclude that both optimization schemes developed in section 3.1 successfully avoid anomalies due to Schwartz singularities whose existence depends only upon the choice of τ. However, at certain values of k, A can become close to singularities whose existence is independent of τ, resulting in anomalies in the calculation of η_v that persist even after τ has been optimized. In attempting to address this problem, a number of methods are available. In principle, anomalous behavior can be reduced dramatically by allowing the Kohn trial wavefunction to be complex-valued [4, 17]. An alternative approach for avoiding anomalies persistent in τ is to make a small change in some other parameter of the trial wavefunction. We will explore both of these techniques in the following section.

3.3. The complex Kohn method

The complex Kohn method is an extension of the original variational approach in which the boundary conditions of the trial wavefunction are complex. It was originally believed [4, 17] that this method was anomaly-free, although anomalies were subsequently reported by Lucchese [18]. For our complex Kohn calculations on $(e^+$–$H_2)$ scattering, we have used a trial wavefunction, Ψ'_t, of the form

$$\Psi'_t = (\bar{S} + a'_t \bar{T} + p'_0 \chi_0)\psi_G + \sum_{i=1}^{M} p'_i \chi_i, \tag{30}$$

where

$$\bar{T} = \bar{S} + i\bar{C}, \tag{31}$$

with the functions ψ_G and $\Omega = \{\chi_1, \ldots, \chi_M\}$ being the same as in (1). The unknowns a'_t and $\{p'_0, \ldots, p'_M\}$ will not, in general, be real. Application of the variational principle to (30) leads to a matrix equation analogous to (11),

$$A'x' = -b', \tag{32}$$

where A' and b' are identical to A and b, but for the function, $\bar{T}$, replacing $\bar{C}$ in (12) and (13). The determinant, $\det(A')$, conveniently reduces to

$$\det(A') = \mathcal{D}(k)\,e^{-2i\tau}, \tag{33}$$

where

$$\mathcal{D} = (\mathcal{A} - \mathcal{C}) - i\mathcal{B} \tag{34}$$

is a complex constant with respect to variations in τ. The values of $\det(A')$ then describe a circle of radius $|\mathcal{D}|$ in the complex plane for variations of $\tau \in [0, \pi)$. Hence, singularities are obtained only if both the real and imaginary parts of $\mathcal{D}$ are zero; they can neither be located nor avoided by varying only τ. We would therefore expect anomalous results due to singularities arising from the choice of τ at a given k to be eliminated in the complex Kohn method. However, from (34) we would also expect $\mathcal{D}$ to be close to zero at $k = 0.71$, in the same sense that $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ have already been seen to be close to zero at this value of k. It is therefore likely that the anomalies already seen to occur due to relationships between $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ rather than the choice of τ will persist even in the complex Kohn method.
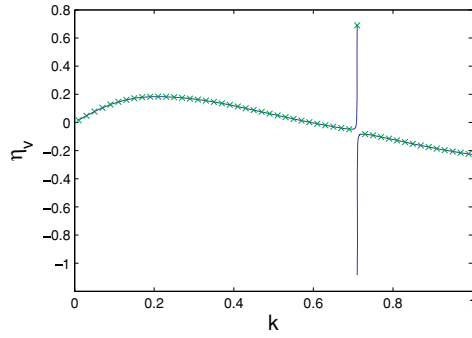
13

Figure 10. A comparison of [——] $\langle \eta_v \rangle(k)$, determined using the trial function, Ψ_t, with [×] $\eta_v(k, \tau = 0)$, evaluated using the complex Kohn method with the trial function, Ψ_t'.
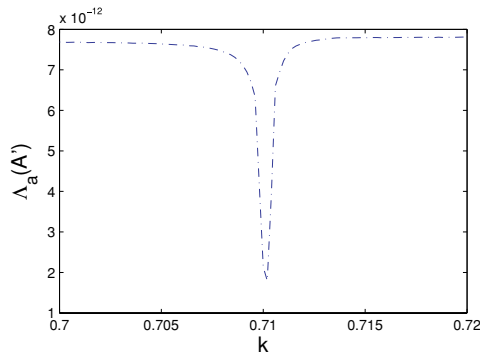

Figure 11. Values of $\Lambda_a(A')$ over k for $\tau = 0$.

We have obtained values of η_v using the trial function, Ψ_t'. We have found that the differences in the calculated values of η_v at different values of $\tau \in [0, \pi)$ are negligible, for all positron momenta considered here. Without loss of generality, we can regard the complex Kohn calculation as effectively independent of τ and choose $\tau = 0$ for simplicity. In figure 10, we have compared results for $\eta_v(k, \tau = 0)$ obtained with the trial function, Ψ_t', with the results for $\langle \eta_v \rangle(k)$ obtained in section 3.1 with Ψ_t. The results of the two methods are essentially equivalent at all positron momenta, with the optimized results for Ψ_t differing from the results for Ψ_t' by no more than 0.1% at each k. As expected, the complex Kohn method automatically avoids anomalies at most values of k without the need for an optimization of τ. Nevertheless, the use of the complex trial function has predictably failed to remove the persistent anomaly at $k = 0.71$. In figure 11, we verify that $\Lambda_a(A')$ is anomalously small at $k \sim 0.71$ for $\tau = 0$. We have found that differences between the results shown in figure 11 and values of $\Lambda_a(A')$ calculated at other values of τ are negligible.

Having failed to find a systematic remedy for the persistent anomalous behavior, we consider a more ad hoc approach. It should be possible to avoid any Schwartz-type anomaly by some variation of parameters in the trial wavefunction. We have found that variations in τ are not always successful, but other candidates exist. In our complex Kohn calculations, we have varied the values of α and β in (9), fixing $\tau = 0$. Recall that the values of the parameters have so far remained fixed at $\alpha = 0.6$ and $\beta = 1.0$. We now consider the results of Kohn
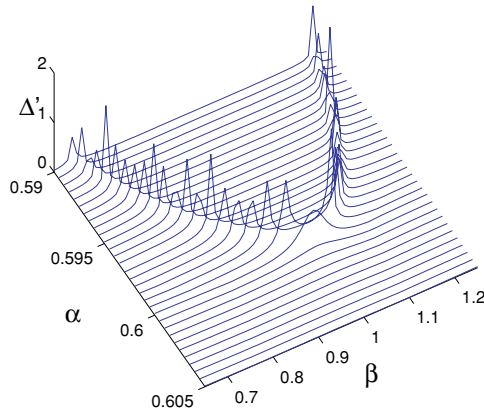
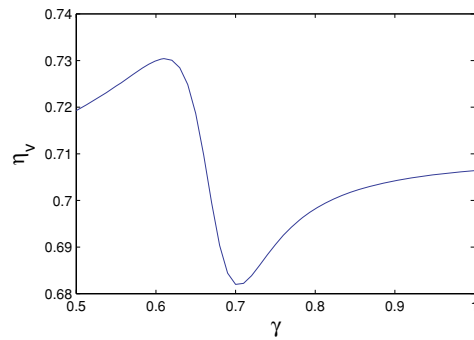**Figure 12.** Values of $\Delta'(\alpha, \beta)$ at $k = 0.71$.

Figure 13. Values of $\eta_v(\gamma)$ at $k = 0.71$.

calculations carried out for 31 different values of α in the range $\alpha \in [0.59, 0.605]$ and 61 different values of β in the range $\beta \in [0.65, 1.25]$.

To illustrate persistent anomalous behavior, it is helpful to define a function analogous to (22),

$$\Delta'(\alpha, \beta) = |\eta_v(\alpha, \beta) - \overline{\eta_v}(\alpha)|, \tag{35}$$

where, for each of the values of α considered, $\overline{\eta_v}(\alpha)$ is the median value of η_v evaluated across the range of values of β. Values of $\Delta'(\alpha, \beta)$ are shown in figure 12, from which it is clear that persistent anomalies appear distributed about a curve in the (α, β) plane. For values of α and β away from this curve, the calculations are free of anomalies. Hence, a small change in the values of α or β can indeed be shown to successfully avoid persistent anomalous behavior.

Finally, we consider briefly that the shielding parameter, γ, in (4) and (8) might also be varied in an effort to avoid anomalous behavior. Values of η_v at $k = 0.71$, $\alpha = 0.6$ and $\beta = 1.0$ for $0.5 \leqslant \gamma \leqslant 1.0$ are shown in figure 13. It is apparent that small changes in the value of γ have relatively little effect on the persistent anomaly at $k = 0.71$. This is not unexpected, being consistent with the findings of Lucchese [18], who investigated the effect of varying a parameter analogous to γ in his model potential calculations. He noted that singularities due to the choice of γ occurred when the values of γ and α were not similar, most typically when

$\gamma \ll \alpha$. With this in mind and from inspection of figures 12 and 13, we can conclude that the anomaly observed in figure 10 is due primarily to the choices of α and β rather than the choice of γ.

4. Concluding remarks

We have carried out a thorough examination of singularities and related anomalous behavior in generalized Kohn calculations for (e^+–H_2) scattering. We have argued that singularities do not always occur spuriously and that variational calculations of the scattering phase shift can be anomaly-free at these singularities. Subsequently, we have developed an optimization scheme for choosing a free parameter of the trial wavefunction allowing anomaly-free values of the phase shift to be determined without the need to solve the linear system of equations derived from the Kohn variational principle. This approach has been seen to be largely successful, giving phase shifts in close agreement with those determined by a conventional generalization of the Kohn method, as well as those obtained with the complex Kohn method.

Persistent anomalies in both sets of calculations have been identified and attributed to singularities that cannot be avoided with any choice of the parameter, τ. Further, we have found that our implementation of the complex Kohn method is susceptible to the same behavior. We have demonstrated, however, that persistent anomalies can be avoided by small changes in the nonlinear parameters of the short-range correlation functions. Hence, by studying the behavior of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ over k, we can predict the appearance of persistent anomalous behavior quantitatively and avoid it by an appropriate change in α or β.

References

[1] Kohn W 1948 *Phys. Rev.* **74** 1763–72
[2] Nesbet R K 1980 *Variational Methods in Electron–Atom Scattering Theory* (New York: Plenum)
[3] Massey H S W and Ridley R O 1956 *Proc. Phys. Soc.* A **69** 659–67
[4] Schneider B I and Rescigno T N 1988 *Phys. Rev.* A **37** 3749–54
[5] Van Reeth P and Humberston J W 1995 *J. Phys. B: At. Mol. Opt. Phys.* **28** L511–7
[6] Van Reeth P, Humberston J W, Iwata K, Greaves R G and Surko C M 1996 *J. Phys. B: At. Mol. Opt. Phys.* **29** L465–71
[7] Van Reeth P and Humberston J W 1999 *J. Phys. B: At. Mol. Opt. Phys.* **32** 3651–67
[8] Armour E A G, Baker D J and Plummer M 1990 *J. Phys. B: At. Mol. Opt. Phys.* **23** 3057–74
[9] Cooper J N and Armour E A G 2008 *Nucl. Instrum. Methods* B **266** 452–7
[10] Cooper J N, Armour E A G and Plummer M 2008 *J. Phys. B: At. Mol. Opt. Phys.* **41** 245201
[11] Schwartz C 1961 *Ann. Phys., Lpz.* **16** 36–50
[12] Schwartz C 1961 *Phys. Rev.* **124** 1468–71
[13] Nesbet R K 1968 *Phys. Rev.* **175** 134–42
[14] Brownstein K R and McKinley W A 1968 *Phys. Rev.* **170** 1255–66
[15] Shimamura I 1971 *J. Phys. Soc. Japan* **31** 852–70
[16] Takatsuka K and Fueno T 1979 *Phys. Rev.* A **19** 1011–7
[17] McCurdy C W, Rescigno T N and Schneider B I 1987 *Phys. Rev.* A **36** 2061–6
[18] Lucchese R R 1989 *Phys. Rev.* A **40** 6879–85
[19] Feshbach H 1962 *Ann. Phys., NY* **19** 287–313
[20] Chung K T and Chen J C Y 1971 *Phys. Rev. Lett.* **27** 1112–4

[21] Charlton M and Humberston J W 2005 *Positron Physics (Cambridge Monographs on Atomic, Molecular and Chemical Physics* vol 11*)* ed A Dalgarno *et al* (Cambridge: Cambridge University Press)
[22] Temkin A and Vasavada K V 1967 *Phys. Rev.* **160** 109–17
[23] Temkin A, Vasavada K V, Chang E S and Silver A 1969 *Phys. Rev.* **186** 57–66
[24] Flammer C 1957 *Spheroidal Wave Functions* (Stanford: Stanford University Press)
[25] Kato T 1950 *Phys. Rev.* **80** 475
[26] Kato T 1951 *Prog. Theor. Phys.* **6** 394–407
[27] Bransden B H and Joachain C J 2003 *Physics of Atoms and Molecules* (Harlow: Prentice Hall)
[28] Hylleraas E A 1929 *Z. Phys.* **54** 347–66
[29] Armour E A G and Baker D J 1987 *J. Phys. B: At. Mol. Phys.* **20** 6105–19
[30] Armour E A G 1988 *Phys. Rep.* **169** 1–98
[31] Armour E A G, Todd A C, Jonsell S, Liu Y, Gregory M R and Plummer M 2008 *Nucl. Instrum. Methods* B **266** 363–8
[32] Higham N J 2002 *Accuracy and Stability of Numerical Algorithms* (Philadelphia: SIAM)
[33] Armour E A G and Humberston J W 1991 *Phys. Rep.* **204** 165–251
[34] http://www.nag.co.uk/numeric/Fl/manual20/pdf/F03/f03aaf.pdf
[35] http://www.nag.co.uk/numeric/Fl/manual20/pdf/F07/f07agf.pdf