

Oscillation theory and computational procedures for matrix Sturm-Liouville eigenvalue problems, with an application to the hydrogen molecular ion

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1969 J. Phys. A: Gen. Phys. 2 278

(<http://iopscience.iop.org/0022-3689/2/3/006>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 31/05/2010 at 11:01

Please note that [terms and conditions apply](#).

Oscillation theory and computational procedures for matrix Sturm-Liouville eigenvalue problems, with an application to the hydrogen molecular ion

J. McI. CALVERT and W. D. DAVISON

Department of Applied Mathematics and Theoretical Physics, Queen's University of Belfast

MS. received 18th December 1968

Abstract. A generalized Prüfer transformation is used to derive comparison and oscillation theorems for the matrix Sturm-Liouville eigenvalue problem. It is shown in particular that if the eigenvectors are required to vanish at one (at least) of the end points, then the eigenvalues are uniquely characterized by the number of points (termed *focal points*) at which a certain determinant (of linearly independent solutions each satisfying the other boundary condition) vanishes. Simple but effective computational procedures based on this analysis are suggested, and their use is illustrated by a single-centre calculation on the $3s\sigma$ state of the hydrogen molecular ion.

1. Introduction

The problem of determining the eigenvalues and eigenfunctions of a single Sturm-Liouville equation with prescribed boundary conditions is familiar in many branches of theoretical physics, and numerical procedures for its solution have been extensively studied. In the particularly important case of the radial Schrödinger equation, the method commonly adopted is to integrate several times from the two end points in opposite directions, adjusting the trial eigenvalue (e.g. by the Newton-Raphson method) until the logarithmic derivatives of the inward and outward solutions are equal at some suitable intermediate point of the range (see Cooley 1961, Mayers 1962). The eigenvalue so obtained is uniquely labelled by the number of nodes in the corresponding eigenfunction, and indeed, as Nachamkin (1968) has pointed out, the eigenvalues can often be quite accurately determined by integrating in one direction only and observing the position and number of nodes in the solution for different choices of trial eigenvalue. This latter method appeals directly to the well-known Sturm comparison and oscillation theorems. An elegant proof of these theorems may be given by use of the Prüfer transformation (Prüfer 1926, Brand 1966) and alternative numerical procedures based on the transformed equation have been discussed recently (Bailey 1966, Godart 1966, Banks and Kurowski 1968); these methods have the advantage that the m th eigenvalue is obtained directly by setting the required value of m in the transformed boundary conditions.

The matrix Sturm-Liouville problem (in which one seeks to determine the eigenvalues and eigenvectors of a set of coupled equations) has received rather less attention, although it is assuming an increasing importance in quantum mechanical investigations. We have come to consider the problem in some detail as a result of our interest in numerical single-centre molecular wave functions (see Bishop 1967 for references), but many other applications could be mentioned; for example, the investigation of resonant states of the Feshbach or closed-channel type in which a particle is temporarily bound in the field of an excited system (see, e.g., Levine *et al.* 1968, von Seggern and Toennies 1969). In this paper we first use a generalization of the Prüfer transformation to discuss how the concept of a node, which is so useful in the scalar case, can be extended to a matrix context. Generalized comparison and oscillation theorems are derived for a wide class of boundary conditions. Then we outline some numerical procedures suggested by our analysis which, while simpler than those previously proposed, appear to work very satisfactorily for coupled radial equations arising from a decomposition of the Schrödinger equation. As an illustration, we describe a single-centre calculation on the $3s\sigma$ state of the hydrogen molecular ion.

2. Oscillation theory

The approach that we adopt here has been suggested by the admirable exposition of Atkinson (1964), to which the reader is particularly referred for the derivation of several results which we quote without proof (see also the appendix to this paper).

We consider a set of n coupled Sturm–Liouville equations, written in the general form

$$[\mathbf{R}(x)\mathbf{y}'(x, \lambda)]' + [\lambda \mathbf{P}(x) + \mathbf{Q}(x)]\mathbf{y}(x, \lambda) = 0 \tag{1}$$

where a prime denotes differentiation with respect to x . \mathbf{R} , \mathbf{R}' , \mathbf{P} and \mathbf{Q} are Hermitian $n \times n$ matrices, continuous on a finite closed interval $[a, b]$, \mathbf{y} , \mathbf{y}' are $n \times 1$ column matrices and λ is a real parameter. It will be further assumed that both \mathbf{P} and \mathbf{R} are positive definite. Suppose that we are seeking values of λ for which there exist solutions of (1) satisfying the boundary conditions

$$\cos \frac{1}{2}\alpha \mathbf{y}(a, \lambda) - \sin \frac{1}{2}\alpha \mathbf{v}(a, \lambda) = 0 \tag{2a}$$

$$\cos \frac{1}{2}\beta \mathbf{y}(b, \lambda) - \sin \frac{1}{2}\beta \mathbf{v}(b, \lambda) = 0 \tag{2b}$$

where

$$\mathbf{v}(x, \lambda) = \mathbf{R}(x)\mathbf{y}'(x, \lambda) \tag{3}$$

and $0 \leq \alpha < 2\pi$, $0 < \beta \leq 2\pi$. More general boundary conditions are possible and will be considered later, but for the sake of clarity our initial discussion will be restricted to the simpler conditions (2).

We define an $n \times n$ matrix $\mathbf{Y}(x, \lambda)$ as the ‘forward’ solution of

$$[\mathbf{R}(x)\mathbf{Y}'(x, \lambda)]' + [\lambda \mathbf{P}(x) + \mathbf{Q}(x)]\mathbf{Y}(x, \lambda) = 0 \tag{4}$$

satisfying the conditions

$$\cos \frac{1}{2}\alpha \mathbf{Y}(a, \lambda) - \sin \frac{1}{2}\alpha \mathbf{V}(a, \lambda) = 0 \tag{5a}$$

$$\sin \frac{1}{2}\alpha \mathbf{Y}(a, \lambda) - \cos \frac{1}{2}\alpha \mathbf{V}(a, \lambda) = \mathbf{A} \tag{5b}$$

where

$$\mathbf{V}(x, \lambda) = \mathbf{R}(x)\mathbf{Y}'(x, \lambda) \tag{6}$$

and \mathbf{A} is an $n \times n$ matrix formed from any suitable set of n linearly independent column vectors. The columns of $\mathbf{Y}(x, \lambda)$ form what may be termed a *conjugate system of solutions* of (1); they are linearly independent and any solution of (1) satisfying (2a) may be expressed as a linear combination of them (cf. Bliss and Schoenberg 1931).

To obtain a matrix generalization of the Prüfer transformation, we consider the unitary matrix

$$\mathbf{W}(x, \lambda) = (\mathbf{V} + i\mathbf{Y})(\mathbf{V} - i\mathbf{Y})^{-1}. \tag{7}$$

We shall be interested in the eigenvalues of \mathbf{W} which (since \mathbf{W} is unitary) lie on the unit circle in the complex plane. For suppose that there exists a non-trivial solution $\mathbf{y}(x, \lambda)$ of (1) satisfying the condition (2a) and

$$\cos \frac{1}{2}\beta \mathbf{y}(\xi, \lambda) - \sin \frac{1}{2}\beta \mathbf{v}(\xi, \lambda) = 0 \tag{8}$$

at some $x = \xi > a$. Introducing the vector $\mathbf{z}(\lambda)$ which is such that

$$\mathbf{y}(x, \lambda) = \mathbf{Y}(x, \lambda)\mathbf{z}(\lambda) \tag{9}$$

$$\mathbf{v}(x, \lambda) = \mathbf{V}(x, \lambda)\mathbf{z}(\lambda)$$

it follows that at $x = \xi$

$$\cos \frac{1}{2}\beta \mathbf{Y}\mathbf{z} - \sin \frac{1}{2}\beta \mathbf{V}\mathbf{z} = 0 \tag{10}$$

and hence that

$$\mathbf{W}(\xi, \lambda)\mathbf{w}(\lambda) = e^{i\beta}\mathbf{w}(\lambda) \tag{11}$$

where

$$\mathbf{w}(\lambda) = (\mathbf{V} - i\mathbf{Y})\mathbf{z}. \tag{12}$$

Thus $\mathbf{W}(\xi, \lambda)$ has $e^{i\beta}$ among its eigenvalues. Conversely, if one of the eigenvalues of

$\mathbf{W}(\xi, \lambda)$ is $e^{i\beta}$ and $\mathbf{w}(\lambda)$ is the corresponding eigenvector, then

$$\mathbf{y}(x, \lambda) = \mathbf{Y}(x, \lambda)[\mathbf{V}(\xi, \lambda) - i\mathbf{Y}(\xi, \lambda)]^{-1}\mathbf{w}(\lambda) \tag{13}$$

is a solution of (1) satisfying (2a) and (8).

We define the *right-conjugate points* of the point $x = a$ as those values of ξ in the open interval $a < \xi < b$ for which there exists a non-trivial solution of (1) satisfying the boundary conditions (2a) and (8). At such points (if they exist), \mathbf{W} has $e^{i\beta}$ among its eigenvalues. We shall be particularly interested in the case where $\beta = 2\pi$ in (8); the corresponding right-conjugate points (henceforth denoted by $x_1 \leq x_2 \leq x_3 \dots$) will be termed *focal points* of the conjugate system of solutions forming \mathbf{Y} (Morse 1930). At each focal point x_i , $\mathbf{W}(x_i, \lambda)$ has $+1$ among its eigenvalues and there exists a non-trivial solution of (1) satisfying (2a) and

$$\mathbf{y}(x_i, \lambda) = 0. \tag{14}$$

It follows from (13) that the focal points are just those points at which

$$D(x_i, \lambda) \equiv \det \mathbf{Y}(x_i, \lambda) = 0. \tag{15}$$

We must now consider how the arguments $\theta_i(x, \lambda)$ of the eigenvalues $\omega_i(x, \lambda)$ of $\mathbf{W}(x, \lambda)$ ($i = 1, \dots, n$) vary with x and λ . We first note that $\mathbf{W}(a, \lambda)$ is just $e^{i\alpha}$ times the unit $n \times n$ matrix and thus all its eigenvalues $\omega_i(a, \lambda)$ are $e^{i\alpha}$; it is convenient to choose all their arguments to be the same, viz.

$$\theta_i(a, \lambda) = \alpha \quad i = 1, \dots, n. \tag{16}$$

It may be shown that the $\theta_i(x, \lambda)$ can then be continued uniquely and continuously so that

$$\theta_1(x, \lambda) \leq \theta_2(x, \lambda) \leq \dots \leq \theta_n(x, \lambda) \leq \theta_1(x, \lambda) + 2\pi. \tag{17}$$

From the theorems given in the appendix, we have that the angles $\theta_i(x, \lambda)$ are monotonic increasing functions of x when they are passing through a multiple of 2π (i.e. when $\omega_i(x, \lambda)$ is passing through $+1$ and x is passing through a focal point), and also on any interval for which the matrix $[\lambda\mathbf{P}(x) + \mathbf{Q}(x)]$ is positive definite; furthermore they are monotonic increasing functions of λ . Using these properties we can immediately derive a generalization of Sturm's *comparison theorem*: as λ is increased, the focal points x_i , as given by the roots of (15), move to the left (cf. Bliss and Schoenberg 1931).

Now our discussion indicates that the eigenvalues of the problem (1), (2) are the roots of

$$\theta_i(b, \lambda) = \beta + 2m\pi \tag{18}$$

for $i = 1, \dots, n$ and m any integer. However, from (16) and the fact that θ_i cannot pass through a multiple of 2π from the above, we have that

$$\theta_i(x, \lambda) > 0 \quad a < x \leq b \tag{19}$$

for $i = 1, \dots, n$ and any λ , so (18) has no solution for negative m . Now it may be shown that

$$\lim_{\lambda \rightarrow -\infty} \theta_i(b, \lambda) = 0 \quad i = 1, \dots, n \tag{20}$$

and so, as λ increases from $-\infty$, the angles $\theta_i(b, \lambda)$ each increase from zero, assuming each positive value once only. It follows that the values of λ at which the angles $\theta_i(b, \lambda)$ pass through values $\beta + 2m\pi$ ($m \geq 0$) may be labelled

$$\lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \dots$$

The first n eigenvalues correspond to the passage of the angles $\theta_i(b, \lambda)$ through the value β , the next n eigenvalues to passage through $\beta + 2\pi$, and so on. Hence a particular angle, $\theta_k(b, \lambda)$ say, goes through $\beta + 2m\pi$ when $\lambda = \lambda_p$, where

$$p = (m + 1)n - k. \tag{21}$$

We now ask, how many focal points will there be in $a < x < b$ when $\lambda = \lambda_p$? To begin with, there will be m focal points at values of x for which

$$\theta_k(x, \lambda_p) = 2\pi, 4\pi, \dots, 2m\pi.$$

Then there will be additional focal points at values of x where the other angles $\theta_i(x, \lambda_p)$ ($i \neq k$) pass through multiples of 2π . In determining their number, it is useful to consider separately the cases $\beta = 2\pi$ and $\beta \neq 2\pi$. For convenience we write θ_i for $\theta_i(b, \lambda_p)$ in the following discussion.

Suppose first that $\beta = 2\pi$. Then $\theta_n, \theta_{n-1}, \dots, \theta_{k+1}$ will all lie in the range

$$2(m+1)\pi < \theta < 2(m+2)\pi$$

and will each give rise to $(m+1)$ focal points in $a < x < b$, a total of $(m+1)(n-k)$. The other angles $\theta_{k-1}, \theta_{k-2}, \dots, \theta_1$ will lie in the range

$$2m\pi < \theta < 2(m+1)\pi$$

and will each give rise to m focal points, a total of $m(k-1)$. So the total number of focal points is precisely

$$N(\lambda_p) = m + (m+1)(n-k) + m(k-1) = (m+1)n - k = p. \tag{22}$$

Now consider the case where $0 < \beta < 2\pi$. The total number of focal points will then depend on the distribution of the other θ_i around θ_k . However, we can readily derive upper and lower bounds on $N(\lambda_p)$, using (17). The *maximum* number of focal points will occur when $\theta_n, \theta_{n-1}, \dots, \theta_{k+1}$ all lie in the range

$$2(m+1)\pi < \theta < \beta + 2(m+1)\pi$$

and $\theta_{k-1}, \theta_{k-2}, \dots, \theta_1$ all lie in the range

$$2m\pi < \theta < \beta + 2m\pi.$$

Then it is clear that $N(\lambda_p)$ has the same value as in the case where $\beta = 2\pi$, namely p . The *minimum* number of focal points will occur when $\theta_n, \theta_{n-1}, \dots, \theta_{k+1}$ all lie in the range

$$\beta + 2m\pi < \theta < 2(m+1)\pi$$

and $\theta_{k-1}, \theta_{k-2}, \dots, \theta_1$ all lie in the range

$$\beta + 2(m-1)\pi < \theta < 2m\pi$$

(except when $m = 0$). In this case we have that

$$N(\lambda_p) = m + (n-k)m + (k-1)(m-1) = p - n + 1. \tag{23}$$

For the special case $m = 0$, $\theta_{k-1}, \theta_{k-2}, \dots, \theta_1$ must lie in the range

$$0 < \theta < \beta$$

and the minimum number of focal points is zero.

We therefore conclude that in general $N(\lambda_p)$, the number of focal points when $\lambda = \lambda_p$, satisfies

$$\max(p - n + 1, 0) \leq N(\lambda_p) \leq p \tag{24}$$

the upper equality holding for the case $\beta = 2\pi$. This *oscillation theorem* is a generalization of that proved by Bliss and Schoenberg (1931) for the simple case $\alpha = 0$, $\beta = 2\pi$ and of course it includes the familiar Sturm oscillation theorem for the scalar problem ($n = 1$). We observe that (17) admits the possibility of two (or more) of the eigenvalues of \mathbf{W} being equal for certain values of x and λ ; such an accidental degeneracy may in particular occur at a focal point, more than one of the eigenvalues of \mathbf{W} being then equal to $+1$. In determining $N(\lambda)$, each focal point should therefore be counted according to its order, i.e. the number of linearly independent solutions of (1) satisfying (2a) and vanishing at the

focal point.† In practice, focal points of multiple order are not likely to occur very often, but in any case they should be easy to detect, since they will be observed to split apart as λ is varied. In the remainder of the discussion we shall assume that any such degeneracies have been taken into account in finding $N(\lambda)$.

The number of focal points thus provides a unique characterization of the eigenvalues λ_i if $\beta = 2\pi$, but not for other values of β (except in the scalar case $n = 1$). Now we could equally well have considered a 'backward' solution of (4), which we denote by $\bar{\mathbf{Y}}(x, \lambda)$, and which satisfies the conditions

$$\begin{aligned}\cos \frac{1}{2}\beta \bar{\mathbf{Y}}(b, \lambda) - \sin \frac{1}{2}\beta \bar{\mathbf{V}}(b, \lambda) &= 0 \\ \sin \frac{1}{2}\beta \bar{\mathbf{Y}}(b, \lambda) - \cos \frac{1}{2}\beta \bar{\mathbf{V}}(b, \lambda) &= \mathbf{A}.\end{aligned}\quad (25)$$

If we form the unitary matrix

$$\bar{\mathbf{W}}(x, \lambda) = (\bar{\mathbf{V}} + i\bar{\mathbf{Y}})(\bar{\mathbf{V}} - i\bar{\mathbf{Y}})^{-1} \quad (26)$$

the arguments $\bar{\theta}_i(x, \lambda)$ of its eigenvalues $\bar{\omega}_i(x, \lambda)$ may be chosen so that

$$\bar{\theta}_i(b, \lambda) = \beta + 2s\pi \quad (27)$$

where s is any fixed integer, and they then satisfy (17) for all x in $[a, b]$. They are monotonic increasing functions of x under the same conditions as are the angles $\theta_i(x, \lambda)$ but they differ in being monotonic *decreasing* functions of λ . Just as in the case of the forward solution, we can define a family of focal points $\{\bar{x}_i\}$ as left-conjugate points of $x = b$ relative to the boundary condition

$$\mathbf{y}(\bar{x}_i, \lambda) = 0. \quad (28)$$

At such points one of the eigenvalues of $\bar{\mathbf{W}}(\bar{x}_i, \lambda)$ is passing through $+1$ and

$$\bar{D}(\bar{x}_i, \lambda) \equiv \det \bar{\mathbf{Y}}(\bar{x}_i, \lambda) = 0. \quad (29)$$

The same comparison theorem may then be derived for the points \bar{x}_i except that 'right' replaces 'left'. Once again we can show that $\bar{N}(\lambda_p)$, the number of focal points \bar{x}_i when $\lambda = \lambda_p$, is precisely p when $\alpha = 0$ and has the bounds (24) when $\alpha > 0$.

The practical importance of the foregoing discussion is that it indicates under what conditions we can investigate the eigenvalue spectrum simply by counting the number of focal points for different values of λ . Suppose for example that $\alpha = 0$. Then if

$$\lambda_{p-1} < \lambda \leq \lambda_p \quad (30)$$

it is clear from our analysis that $\bar{N}(\lambda)$, the number of focal points in $\bar{\mathbf{Y}}(x, \lambda)$, is precisely p ; as λ increases through λ_p a new focal point appears at $x = a$ and moves to the right along with the other p focal points already present. So if for a certain choice of λ we find by counting the roots of (29) that there are p focal points, then we can conclude that λ is an upper bound on λ_{p-1} (provided $p > 0$) and a lower bound on λ_p . For other boundary conditions at $x = a$, $\bar{N}(\lambda)$ may take values between $\max(p-n, 0)$ and p when λ is in the range (30), and in general $\bar{N}(\lambda)$ does not change by 1 as λ goes through an eigenvalue. It is therefore not possible to obtain bounds on the eigenvalues by consideration of $\bar{N}(\lambda)$ alone. However, if in addition we compute the eigenvalues of $\bar{\mathbf{W}}(a, \lambda)$, it is easily shown that we may determine their arguments $\bar{\theta}_i(a, \lambda)$ unambiguously from a knowledge of $\bar{N}(\lambda)$, and so deduce the value of p in (30).

Similar arguments apply if instead we consider $N(\lambda)$, the number of focal points in $\mathbf{Y}(x, \lambda)$. However, in investigating the eigenvalue spectrum of coupled equations resulting from a decomposition of the Schrödinger equation, it is clearly preferable to compute the 'backward' solution $\mathbf{Y}(x, \lambda)$ and to consider $\bar{N}(\lambda)$. For one thing, numerical integration errors here tend to be of less significance for the backward integration. But a more important feature is that the vector \mathbf{y} can generally be defined so that the boundary condition at

† Similarly, in indexing the eigenvalues λ_i , account should be taken of any degeneracy therein corresponding to more than one of the eigenvalues of \mathbf{W} being $e^{i\alpha}$ at $x = b$.

$x = a$ ($= 0$) takes the form (2a) with $\alpha = 0$. It is then possible to establish bounds on the eigenvalues from a knowledge of $\bar{N}(\lambda)$ only, but not of $N(\lambda)$ (unless $\beta = 2\pi$ or $n = 1$). Practical computational procedures for this and other cases are discussed in detail in § 3.

Problems in which (1) has singular points at $x = a$ or $x = b$, or in which one or both of the end points is at infinity, require special attention, but we shall not discuss them here in any detail. Suffice it to say that, in the majority of such problems, no difficulty should arise if full use is made of the known asymptotic behaviour of the solution (see Bailey 1966 and Banks and Kurowski 1968 for discussions of the scalar case). The Schrödinger problem to which we have just referred is a case in point, and an example is given in § 4.

Finally we consider briefly the extension to include boundary conditions more general than (2). Suppose for example that the angle β in (2) is a function of λ . Then according to theorem IV of the appendix all our conclusions with regard to the focal points \bar{x}_i will certainly continue to hold if the matrix

$$\bar{\mathbf{V}}^*(b, \lambda) \frac{\partial \bar{\mathbf{Y}}}{\partial \lambda}(b, \lambda) - \bar{\mathbf{Y}}^*(b, \lambda) \frac{\partial \bar{\mathbf{V}}}{\partial \lambda}(b, \lambda)$$

is negative semi-definite for all λ . This in turn implies that

$$\frac{d\beta(\lambda)}{d\lambda} \leq 0 \tag{31}$$

for all λ . More generally, suppose that the boundary condition at $x = b$ takes the form

$$\mathbf{B}(\lambda)\mathbf{y}(b, \lambda) - \mathbf{C}(\lambda)\mathbf{v}(b, \lambda) = 0 \tag{32}$$

where \mathbf{B} and \mathbf{C} are non-singular $n \times n$ matrices which may depend on λ . Defining $\bar{\mathbf{Y}}$ and $\bar{\mathbf{V}}$ by an obvious generalization of (25), we find that our conclusions remain unchanged provided the matrix $\mathbf{B}^{-1}\mathbf{C}$ is Hermitian and $(d/d\lambda)(\mathbf{B}^{-1}\mathbf{C})$ is negative semi-definite for all λ .

3. Computational procedures

Given a set of coupled equations of the form (1), together with certain boundary conditions, a comprehensive investigation of the eigenvalue spectrum may be regarded as falling into two distinct parts. The first problem is to determine the precise number and *approximate* location of eigenvalues below a certain value of λ , λ_{\max} say; it is important that we should be able to detect *all* the eigenvalues without an inordinate amount of computation. Then we want to be able to determine any particular eigenvalue with high precision, together with the corresponding eigenvector if required.

To start with, let us again consider the case where the boundary condition at $x = a$ is just $\mathbf{y}(a, \lambda) = 0$, i.e. $\alpha = 0$, with possibly a more complicated boundary condition at $x = b$. We begin by generating the ‘backward’ solution $\mathbf{Y}(x, \lambda)$ for a limited number of well-spaced values of λ , using some suitable step-by-step method, and in each case we determine $\bar{N}(\lambda)$, the number of focal points \bar{x}_i as given by the roots of (29). We then look to see if there are any cases in which $\bar{N}(\lambda)$ jumps by more than one between one value of λ and the next above: if so, we repeat the backward integration for further values of λ until we have filled in all such ‘gaps’, also ensuring that $\bar{N}(\lambda) = 0$ for the lowest value of λ considered. In this way, we establish upper and lower bounds (poor though some may be) on each of the eigenvalues below λ_{\max} , and we can be sure that no eigenvalue has gone undetected.

Then concentrating on a particular eigenvalue, λ_p say, we can proceed to progressively refine the upper and lower bounds on λ_p by a simple ‘halving’ procedure similar to that suggested for the scalar problem by Nachamkin (1968) (see also Blatt 1967). In principle this can be continued until the upper and lower bounds differ by less than some prescribed arbitrarily small quantity. However, since we gain only one binary digit of accuracy at each stage, the convergence of the process is rather slow. Additional difficulties arise when $x = a$ is a singular point; the integration then has to be terminated just before the point $x = a$ is reached, and the effect of increasing numerical errors as the singularity is

approached may be considerable, even if the integration interval is made very small. Once reasonably close bounds on λ_p have been established, there is therefore good reason for changing over to a faster iterative procedure.

When $x = a$ is not a singular point, we can make direct use of the fact that the determinant $\bar{D}(a, \lambda)$ is required to vanish when $\lambda = \lambda_p$. Suppose that we have obtained reasonably close upper and lower bounds on λ_p . Then by two-point (linear) inverse interpolation we obtain a first estimate λ' of the value of λ (i.e. λ_p) at which $\bar{D}(a, \lambda)$ vanishes. Computing $\mathbf{Y}(x, \lambda')$ and hence $\bar{D}(a, \lambda')$ we obtain a second estimate λ'' by three-point interpolation. This process is continued (using all the available values of $\bar{D}(a, \lambda)$ for the inverse interpolation at each stage) until two successive estimates of λ_p differ by less than some prescribed quantity.

However, when $x = a$ is a singular point, then to obtain really high accuracy (even in the eigenvalue only—cf. Nachamkin 1968) it is necessary to integrate (4) also in the forward direction from $x = a$ and to match the forward and backward solutions, together with their derivatives, at some intermediate point x_m . Fox (1960, 1961) has discussed how the Newton-Raphson procedure can be adapted to this problem; a number of alternative and apparently simpler methods can, however, be suggested. The matching condition is that there should exist non-trivial vectors \mathbf{a} and \mathbf{b} such that

$$\mathbf{Y}(x_m, \lambda)\mathbf{a} = \bar{\mathbf{Y}}(x_m, \lambda)\mathbf{b} \quad (33a)$$

$$\mathbf{V}(x_m, \lambda)\mathbf{a} = \bar{\mathbf{V}}(x_m, \lambda)\mathbf{b} \quad (33b)$$

when $\lambda = \lambda_p$.

This in turn implies that

$$d_1(\lambda_p) \equiv \det \begin{pmatrix} \mathbf{Y}(x_m, \lambda_p) & \bar{\mathbf{Y}}(x_m, \lambda_p) \\ \mathbf{V}(x_m, \lambda_p) & \bar{\mathbf{V}}(x_m, \lambda_p) \end{pmatrix} = 0 \quad (34)$$

and one possible procedure (which we shall term method I) is to adjust λ by an interpolative scheme similar to that already described so that $d_1(\lambda)$ vanishes. Alternatively, elimination of the vector \mathbf{b} in (33) leads to the condition

$$d_2(\lambda) \equiv \det(\bar{\mathbf{V}}\bar{\mathbf{Y}}^{-1}\mathbf{Y} - \mathbf{V}) = 0 \quad (35)$$

which has the advantage that d_2 is of order n only, in contrast to d_1 which is of order $2n$; an interpolation procedure based on (35) will be termed method II.

A third procedure (method III), which is particularly appropriate when both eigenvalue and eigenvector are required, is suggested by the technique developed by Burke and Smith (1962) for the solution of scattering equations with closed channels. We suppress one of the equations (33a), e.g.

$$\sum_{j=1}^n Y_{ij}a_j = \sum_{j=1}^n \bar{Y}_{ij}b_j \quad (36)$$

and replace it by the 'normalization' condition

$$\sum_{j=1}^n Y_{ij}a_j = 1. \quad (37)$$

We then solve (37) and the remaining $2n-1$ equations in (33) for the vectors \mathbf{a} and \mathbf{b} , adjusting λ by an interpolative procedure until

$$d_3(\lambda) \equiv \sum_{j=1}^n \bar{Y}_{ij}b_j - 1 = 0. \quad (38)$$

Often one component of the eigenvector $\mathbf{y}(x, \lambda_p)$ will be known on physical grounds to be dominant and the choice of i will be obvious. This method also has the advantage that the vectors \mathbf{a} and \mathbf{b} are determined in the course of the calculation, so that the eigenvector

$$\mathbf{y}(x, \lambda_p) = \mathbf{Y}(x, \lambda_p)\mathbf{a} = \bar{\mathbf{Y}}(x, \lambda_p)\mathbf{b} \quad (39)$$

can be immediately deduced when the iteration is complete. It is to be noted that all three methods require inverse interpolation in one dimension only.

We have considered the case $\alpha = 0$ in some detail since it is that which generally arises in solving the Schrödinger equation. However, the procedure to be adopted in other situations should by now be clear. For example, if $\beta = 0$, we establish bounds on the eigenvalues by studying $N(\lambda)$, finally determining any particular eigenvalue λ_p either through the requirement that $D(b, \lambda) = 0$ or by matching at $x = x_m$ in one of the ways just described. On the other hand, if neither of the boundary conditions is of the simple form $y = 0$, then we must supplement $N(\lambda)$ (or $\bar{N}(\lambda)$) by the eigenvalues of $\mathbf{W}(b, \lambda)$ (or of $\mathbf{W}(a, \lambda)$) in order to locate the eigenvalues λ_i , subsequently determining any particular eigenvalue by matching at $x = x_m$. Fox (1960, 1961) found that x_m had to be rather carefully chosen if erratic convergence was to be avoided; although further experience is clearly desirable, we feel that the additional information available from our analysis should result in the convergence of the calculation being less critically dependent on the choice of x_m .

4. Example: single-centre expansion for H_2^+

To illustrate the theory and practical techniques developed in the previous sections, we consider the calculation of the eigenvalues and eigenfunctions of the σ_g states of the hydrogen molecular ion in a numerical single-centre formulation, with particular reference to the $3s\sigma$ state. Details of the analysis will be found in previous studies of the problem by Cohen and Coulson (1961) and Temkin (1963). If the internuclear separation is taken to be 2.0 (all quantities being expressed in atomic units), separation of the angular variables from the Schrödinger equation for the system leads to an infinite set of coupled radial equations which are of the form (2) with

$$R_{ij}(x) = \frac{1}{2}P_{ij}(x) = \delta_{ij} \tag{40}$$

and

$$Q_{ij}(x) = -\frac{(2i-1)(2i-2)}{x^2} \delta_{ij} + 4 \sum_k (2i-2, 2k, 0, 0 | 2j-2, 0) (2j-2, 2k, 0, 0 | 2i-2, 0) \times \frac{\{\min(x, 1)\}^{2k}}{\{\max(x, 1)\}^{2k+1}} \tag{41}$$

where $(a, b, \alpha, \beta | c, \gamma)$ is a vector-coupling coefficient. In practice, of course, we solve a truncated set of n equations only, and we can expect that, if l (which must be even) is the azimuthal quantum number in the united atom designation of a particular state, then $y_i(x)$ with $i = \frac{1}{2}l + 1$ will be the dominant component of the eigenvector $\mathbf{y}(x)$, particularly for higher values of the principal quantum number and large x (cf. Mulliken 1964).

The appropriate boundary conditions are that

$$\mathbf{y}(0, \lambda) = 0 \tag{42}$$

and that $\lim_{x \rightarrow \infty} \mathbf{y}(x, \lambda)$ should be finite. In actual computation this second condition must be replaced by a condition at some large value of x (which we call x_∞) derived from a consideration of the asymptotic form of the coupled equations. The most refined condition of this kind would be obtained either by an asymptotic series expansion of $\mathbf{y}(x, \lambda)$ (Burke and Schey 1962) or by an extension of the WKB method to coupled equations (Smith *et al.* 1968, private communication). However, for sufficiently large x_∞ and $\lambda < 0$, the condition obtained by either of these procedures will differ negligibly from the simpler condition

$$\mathbf{v}(x_\infty, \lambda) \equiv \mathbf{y}'(x_\infty, \lambda) = -(-2\lambda)^{1/2} \mathbf{y}(x_\infty, \lambda) \tag{43}$$

to which we restrict ourselves in the present calculation. Thus, in the notation of § 2, we

have

$$\begin{aligned}\alpha &= 0 \\ \beta &= 2 \cot^{-1}\{-(-2\lambda)^{1/2}\}.\end{aligned}\quad (44)$$

It will be noted that β is an explicit function of λ , but, since $d\beta/d\lambda < 0$, the oscillation theory of § 2 applies without modification, as does the subsequent discussion in § 3 which indicates that, since $\alpha = 0$, our initial study should be based on the focal points \bar{x}_i of the 'backward' solution $\bar{\mathbf{Y}}(x, \lambda)$.

The integration of the coupled equations was carried out by the method of de Vogelaere (1955) (see also Lester 1968) using various fixed intervals h ; Aitken's method of interpolation was used in determining the eigenvalues, which were subsequently extrapolated to zero interval by Richardson's method (see e.g. Kopal 1961). In all the calculations x_∞ was given the value 35.0, and it was verified that an increase in x_∞ produced no significant change in the computed eigenvalues. In certain cases we also computed the eigenvalues of the matrix $\bar{\mathbf{W}}(x, \lambda)$, using the triangularization method discussed by Greenstadt (1955) and Lotkin (1956). These calculations serve to further illustrate the theory of § 2, although they are not actually required for the determination of the eigenvalues λ_i .

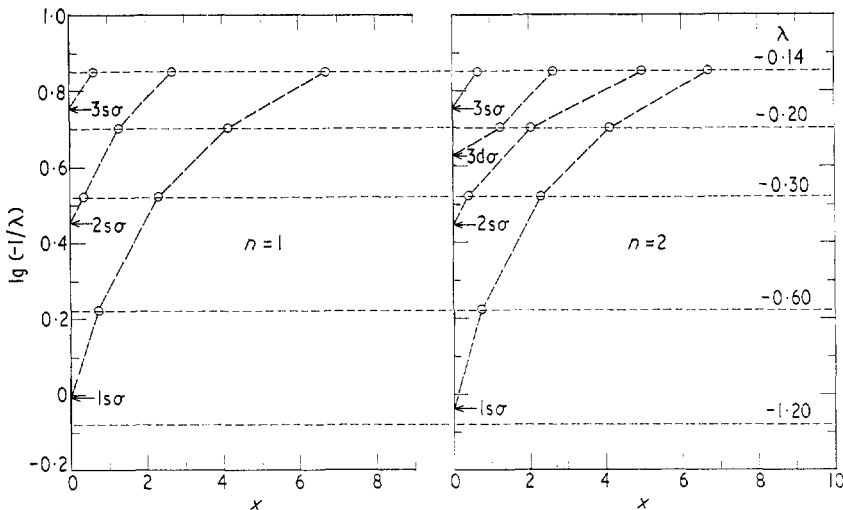


Figure 1. Position of the focal points \bar{x}_i for various values of λ , in the scalar case $n = 1$ and the simplest matrix case $n = 2$. The arrows indicate the accurate eigenvalues in each case.

Figure 1 shows the position of the focal points \bar{x}_i for various trial values of λ in the scalar cases $n = 1$ and the simplest matrix case $n = 2$. The actual eigenvalues for each value of n are also indicated, and the graphs show how a count of the number of focal points enables bounds to be put on the eigenvalues. Corresponding points have been connected by broken lines so as to exhibit clearly how each point moves to the right with increasing λ , in accordance with the comparison theorem of § 2. It will be noted that the 3σ state does not appear in the case $n = 1$, since this corresponds to a spherically symmetric approximation to the electronic wave function (the 'molecular puff' of Hauk and Parr 1965) which can accommodate $s\sigma$ states only.

In figure 2 the angles θ_i in the case $n = 2$ have been plotted as functions of x for two values of λ bounding the 3σ eigenvalue, and the points at which they pass through multiples of 2π have been projected out in order to illustrate how the position and number of focal points are related to the variation of the angles θ_i with x and λ . The occurrence of extrema in figure 2 will be particularly noted, and is related to the fact that when the matrix

$$\mathbf{F}(x, \lambda) \equiv \lambda \mathbf{P}(x) + \mathbf{Q}(x) \quad (45)$$

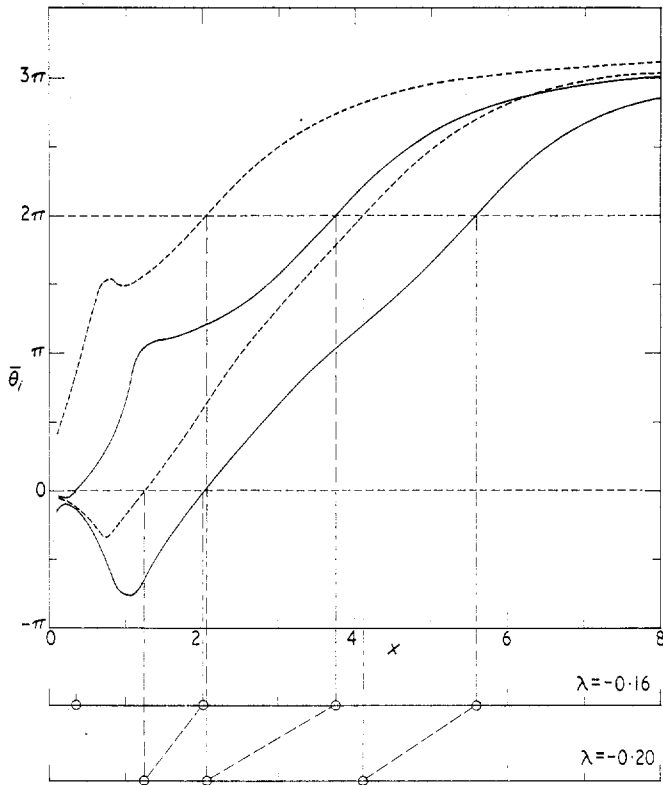


Figure 2. Angles $\bar{\theta}_i$ and focal points \bar{x}_i in the case $n = 2$ for two values of λ bounding the 3σ eigenvalue. Broken curves, $\lambda = -0.20$; full curves, $\lambda = -0.16$.

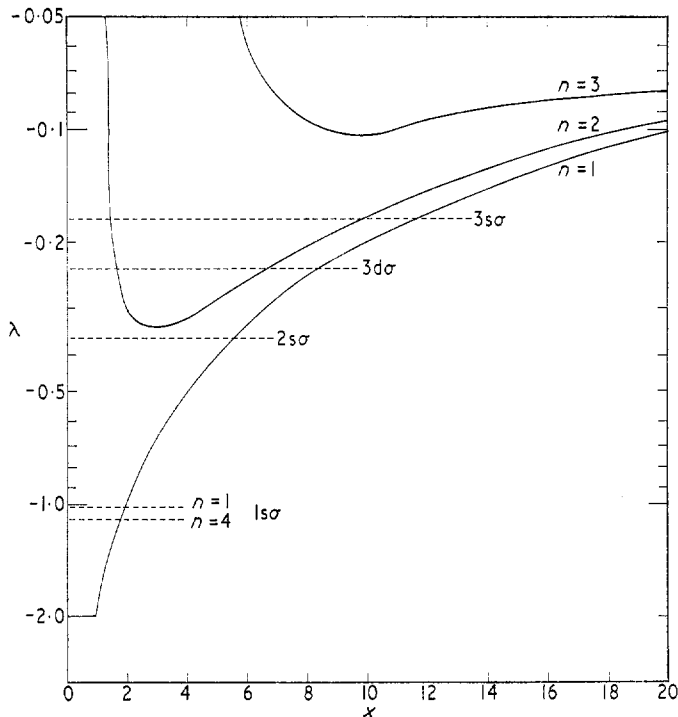


Figure 3. Curves, for various dimensions n , bounding the region of x and λ for which the matrix $\mathbf{F}(x, \lambda)$ defined by (45) is positive definite. The curves for $n \geq 4$ lie outside the figure.

is not positive definite, the angles $\bar{\theta}_i$ are not necessarily all monotonic increasing functions of x . For given values of n and λ , the range of x for which \mathbf{F} is positive definite may be read off from figure 3. In the vicinity of the 3σ eigenvalue, \mathbf{F} is positive definite for $1.5 < x < 10$ if $n = 2$, and as expected the extrema in figure 2 occur outside this range of

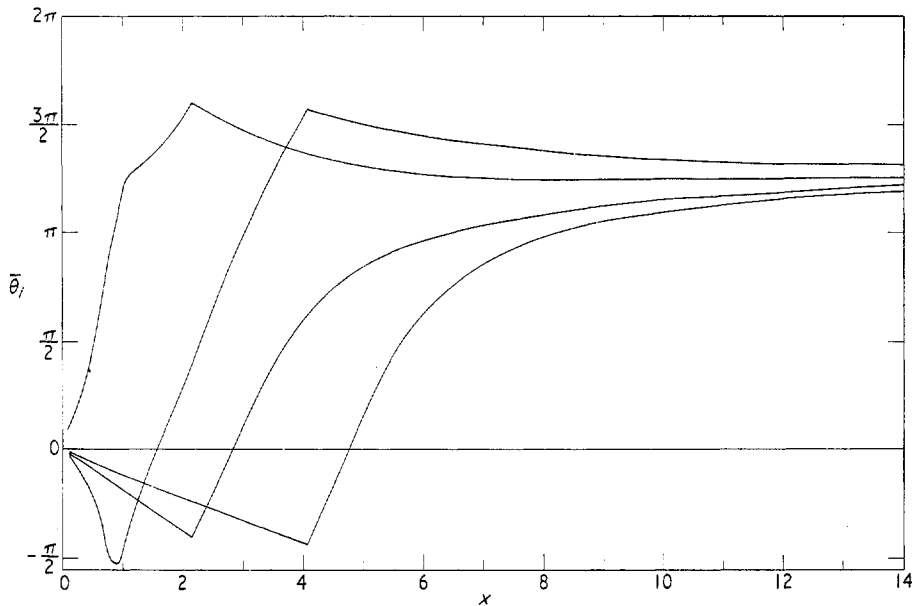


Figure 4. Angles $\bar{\theta}_i$ in the case $n = 4$ at the 3σ eigenvalue ($\lambda = -0.17757$).

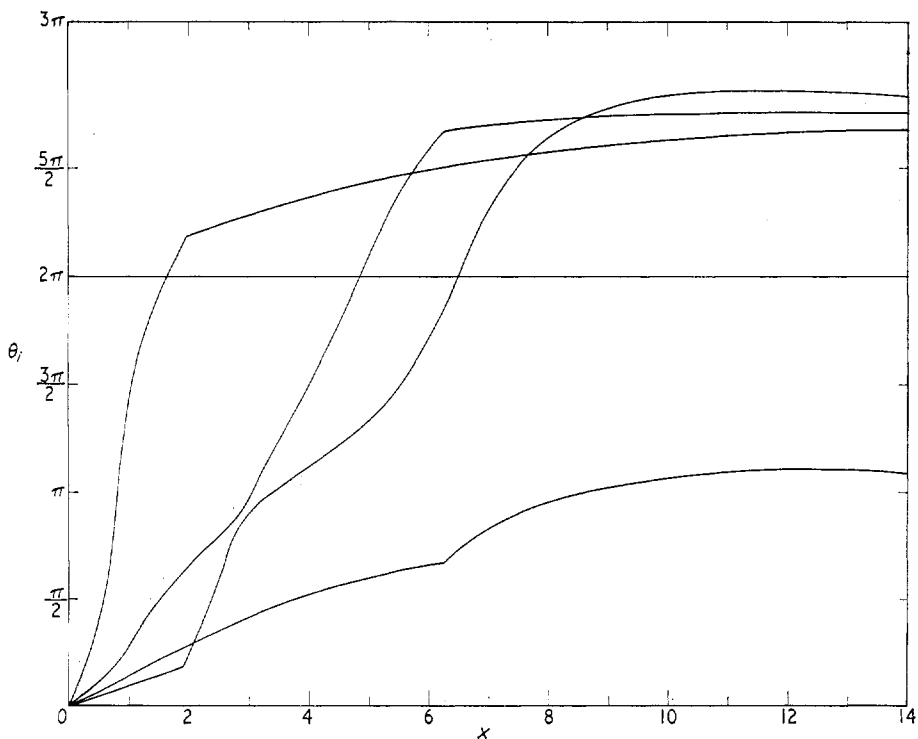


Figure 5. Angles θ_i in the case $n = 4$ at the 3σ eigenvalue ($\lambda = -0.17757$).

x ; on the other hand, if $n \geq 3$ there is no value of x for which \mathbf{F} is positive definite, and this is reflected in figures 4 and 5, in which the angles $\bar{\theta}_i$ and θ_i respectively have been plotted as functions of x at the 3σ eigenvalue in the case $n = 4$. It will be seen that in both figures some 'crossings' occur; this does not, as might perhaps be thought, imply any violation of theorem II in the appendix, which simply provides a way of determining the angles $\bar{\theta}_i$ and θ_i unambiguously and of uniquely labelling the various branches of the resulting curves. In figure 5 we also observe an 'avoided crossing'. It will clearly be of interest to investigate further the conditions under which such features occur.

Let us now consider the accurate determination of the 3σ eigenvalue. In the scalar case $n = 1$, $x = 0$ is not a singular point, and the eigenvalue may be determined by locating the zero of

$$\bar{D}(0, \lambda) = \bar{y}(0, \lambda)$$

as discussed in § 3. The results are given in table 1, and the rapid convergence of the

Table 1. Determination of the 3σ eigenvalue (λ_2) for $n = 1$

Iteration No.	λ_2
1	-0.172886
2	-0.174737
3	-0.174459
4	-0.174448
5	-0.174448

These values were obtained with a fixed integration interval $h = 0.025$; the values obtained with $h = 0.05$ were found to be identical to the number of decimal places quoted.

procedure is at once apparent. When $n \geq 2$, there is a singularity at $x = 0$ arising from the first (centrifugal) term in (41), so it is necessary to adopt one of the matching procedures discussed in § 3. The results for $n = 4$, with matching point $x_m = 7.0$, are given in

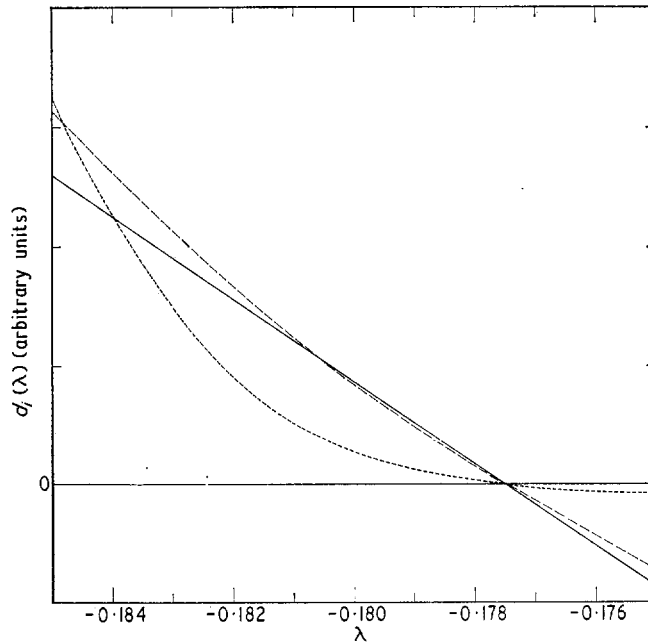


Figure 6. Functions $d_i(\lambda)$ in the case $n = 4$ near the 3σ eigenvalue. ---, $i = 1$; full curve, $i = 2$; - · - ·, $i = 3$.

table 2. Method II exhibits the most rapid convergence, that of method III is almost as good, while that of method I is decidedly inferior. These differences reflect the different forms of the functions $d_i(\lambda)$ in the vicinity of the $3s\sigma$ eigenvalue, as shown in figure 6.

Table 2. Determination of the $3s\sigma$ eigenvalue (λ_3) for $n = 4$

Method		I	II	III
Initial bounds		-0.180 -0.175	-0.185 -0.175	-0.185 -0.175
Iteration No.	h			
1	0.100	-0.176040	-0.177478	-0.176876
	0.050	-0.176075	-0.177531	-0.176920
	0.025		-0.177537	
2	0.100	-0.179053	-0.177507	-0.177521
	0.050	-0.179178	-0.177561	-0.177575
	0.025		-0.177567	
3	0.100	-0.178380	-0.177507	-0.177507
	0.050	-0.178502	-0.177560	-0.177560
	0.025		-0.177567	
4	0.100	-0.177850		-0.177507
	0.050	-0.177952		-0.177560
5	0.100	-0.177571		
6	0.100	-0.177510		
7	0.100	-0.177507		

For the spherically symmetric approximation $n = 1$ we obtain the eigenvalue $-0.174\ 448$, which we believe to be more accurate than the value $-0.174\ 476$ derived by Chen (1958). After extrapolation to zero integration interval we obtain the value $-0.177\ 569$ for $n = 4^\dagger$, which may be compared with the exact value ($n \rightarrow \infty$) of $-0.177\ 68$ (Bates *et al.* 1953). The calculations illustrate not only the accuracy of the single-centre expansion method for excited states, but also the effectiveness and relative simplicity of the computational procedures we have proposed. At the same time it would be surprising if the theory of § 2 did not admit of several refinements and extensions, and it will be interesting to see how far various features appearing in the present calculations are reproduced in other applications.

Acknowledgments

The research reported has been sponsored by the Air Force Cambridge Research Laboratories, under Contract AF 61(052)-780 through the European Office of Aerospace Research, OAR, United States Air Force. One of us (J. McI. C.) wishes to thank the Ministry of Education, Northern Ireland, for a research grant.

Appendix

We gather together here certain results concerning the eigenvalues of unitary matrices such as (7) and (26), obtained by adapting and extending various theorems proved by Atkinson (1964).

[†] Cohen and Coulson (1961) quote an even lower value of $-0.177\ 65$ for $n = 3$; however, this was obtained with a relatively large integration interval, and extrapolation to zero interval would undoubtedly yield a value higher than that which we quote for $n = 4$.

Let $\mathbf{Y}(x, \lambda)$ be a solution of (4) defined by certain conditions at one of the end points $x = c$ ($c \equiv a$ or $c \equiv b$), such that at that point the matrix $\mathbf{Y}^*\mathbf{V}$ is Hermitian and $(\mathbf{V} - i\mathbf{Y})$ has an inverse, where \mathbf{V} is defined by (6). We then have:

Theorem I. The matrix

$$\mathbf{W} = (\mathbf{V} + i\mathbf{Y})(\mathbf{V} - i\mathbf{Y})^{-1} \tag{A1}$$

exists in $[a, b]$, is unitary, and satisfies the equation

$$\frac{\partial \mathbf{W}(x, \lambda)}{\partial x} = i\mathbf{W}(x, \lambda)\mathbf{\Omega}(x, \lambda) \tag{A2}$$

where $\mathbf{\Omega}$ is the Hermitian matrix

$$\mathbf{\Omega} = 2(\mathbf{V}^* + i\mathbf{Y}^*)^{-1}\{\mathbf{V}^*\mathbf{R}^{-1}\mathbf{V} + \mathbf{Y}^*(\lambda\mathbf{P} + \mathbf{Q})\mathbf{Y}\}(\mathbf{V} - i\mathbf{Y})^{-1}. \tag{A3}$$

Theorem II. If the arguments $\theta_i(x, \lambda)$ of the eigenvalues of \mathbf{W} are fixed so that they satisfy (17) at $x = c$, then they may be continued uniquely and continuously so that they satisfy (17) for all x in $[a, b]$.

Theorem III. For a given value of λ , $\mathbf{\Omega}$ is positive definite for values of x (if they exist) such that $[\lambda\mathbf{P}(x) + \mathbf{Q}(x)]$ is positive definite, and for such x it follows that the angles $\theta_i(x, \lambda)$ are monotonic increasing functions of x . But in any case, $\mathbf{\Omega}$ can be shown to be positive definite when applied to eigenvectors of \mathbf{W} associated with an eigenvalue $+1$, and so the $\theta_i(x, \lambda)$ are at least monotonic increasing functions of x when passing through multiples of 2π .

Theorem IV. For all real λ

$$\frac{\partial \mathbf{W}(x, \lambda)}{\partial \lambda} = i\mathbf{W}(x, \lambda)\hat{\mathbf{\Omega}}(x, \lambda) \tag{A4}$$

where $\hat{\mathbf{\Omega}}$ is the Hermitian matrix

$$\hat{\mathbf{\Omega}} = 2(\mathbf{V}^* + i\mathbf{Y}^*)^{-1} \left[\int_c^x \mathbf{Y}^*(t, \lambda)\mathbf{P}(t)\mathbf{Y}(t, \lambda) dt + \mathbf{X}(\lambda) \right] (\mathbf{V} - i\mathbf{Y})^{-1} \tag{A5}$$

with

$$\mathbf{X}(\lambda) = \mathbf{V}^*(c, \lambda) \frac{\partial \mathbf{Y}(c, \lambda)}{\partial \lambda} - \mathbf{Y}^*(c, \lambda) \frac{\partial \mathbf{V}(c, \lambda)}{\partial \lambda}. \tag{A6}$$

Then for all x for which the matrix in square brackets in (A5) is positive definite (or negative definite), the $\theta_i(x, \lambda)$ are monotonic increasing (or decreasing) functions of λ . Since \mathbf{P} has been assumed to be positive definite, the integral in (A5) is either positive definite or negative definite according as to whether $c = a$ or $c = b$. Hence a *sufficient* (but not necessary) condition for the $\theta_i(x, \lambda)$ to be monotonic increasing (or decreasing) functions of λ is that the matrix $\mathbf{X}(\lambda)$ should be positive (or negative) semi-definite.

References

ATKINSON, F. V., 1964, *Discrete and Continuous Boundary Problems* (New York: Academic Press).
 BAILEY, P. B., 1966, *SIAM J. Appl. Math.*, **14**, 242–9.
 BANKS, D. O., and KUROWSKI, G. J., 1968, *Math. Comput.*, **22**, 304–10.
 BATES, D. R., LEDSHAM, K., and STEWART, A. L., 1953, *Phil. Trans. R. Soc. A*, **246**, 215–40.
 BISHOP, D. M., 1967, *Advances in Quantum Chemistry*, **3**, 25–59, Ed. P.-O. Löwdin (New York: Academic Press).
 BLATT, J. M., 1967, *J. Comput. Phys.*, **1**, 382–96.
 BLISS, G. A., and SCHOENBERG, I. J., 1931, *Am. J. Math.*, **53**, 781–800.
 BRAND, L., 1966, *Differential and Difference Equations* (New York: John Wiley), pp. 582–90.
 BURKE, P. G., and SCHEY, H. M., 1962, *Phys. Rev.*, **126**, 147–62.
 BURKE, P. G., and SMITH, K., 1962, *Rev. Mod. Phys.*, **34**, 458–502.
 CHEN, T. C., 1958, *J. Chem. Phys.*, **29**, 347–55.
 COHEN, M., and COULSON, C. A., 1961, *Proc. Camb. Phil. Soc.*, **57**, 96–103.

- COOLEY, J. W., 1961, *Math. Comput.*, **15**, 363-74.
- FOX, L., 1960, *Boundary-value Problems in Differential Equations*, Ed. R. E. Langer (Madison, Wisconsin: University of Wisconsin Press), pp. 243-56.
- 1961, *Proc. Camb. Phil. Soc.*, **57**, 103-6.
- GODART, M., 1966, *Math. Comput.*, **20**, 399-406.
- GREENSTADT, J., 1955, *Math. Tabl. Natn. Res. Coun., Wash.*, **9**, 47-52.
- HAUK, P., and PARR, R. G., 1965, *J. Chem. Phys.*, **43**, 548-52.
- JOY, H. W., and HANDLER, G. S., 1965, *J. Chem. Phys.*, **42**, 3047-51.
- KOPAL, Z., 1961, *Numerical Analysis*, 2nd edn (London: Chapman and Hall).
- LESTER, W. A., 1968, *J. Comput. Phys.*, **3**, 322-6.
- LEVINE, R. D., JOHNSON, B. R., MUCKERMAN, J. T., and BERNSTEIN, R. B., 1968, *J. Chem. Phys.*, **49**, 56-64.
- LOTKIN, M., 1956, *Q. Appl. Math.*, **14**, 267-75.
- MAYERS, D. F., 1962, *Numerical Solution of Ordinary and Partial Differential Equations*, Ed. L. Fox (Oxford: Pergamon Press), pp. 87-94.
- MORSE, M., 1930, *Math. Ann.*, **103**, 52-69.
- MULLIKEN, R. S., 1964, *J. Am. Chem. Soc.*, **86**, 3183-97.
- NACHAMKIN, J., 1968, *Can. J. Phys.*, **46**, 445-6.
- PRÜFER, H., 1926, *Math. Ann.*, **95**, 499-518.
- VON SEGGERN, M., and TOENNIES, J. P., 1969, *Z. Phys.*, in the press.
- TEMKIN, A., 1963, *J. Chem. Phys.*, **39**, 161-9.
- DE VOGELAERE, R., 1955, *J. Res. Natn. Bur. Stand.*, **54**, 119-25.