Practical Points Concerning the Solution of the Schrödinger Equation

JOHN M. BLATT[1]

Department of Theoretical Physics,
Hebrew University, Jerusalem, Israel

and

Department of Nuclear Physics,
Weizmann Institute, Rehovoth, Israel

ABSTRACT

We consider the numerical solution of the one-dimensional Schrödinger equation in a potential of the type occurring in molecular spectroscopy, i.e., with both an inner and an outer classical turning point. Some practical tricks are described which, it is hoped, may prove useful to others. These involve choice of a step size, changing step size, iteration on the eigenvalue, setting upper and lower bounds on the eigenvalue, determining a useful range of x for the integration, etc. Formulas are derived for the value of the function, and of the first derivative, to be used in conjunction with the Noumerov method.

I. INTRODUCTION

In connection with a calculation of molecular energy levels of the quarkium molecule ion, it proved necessary to solve the one-dimensional (radial) Schrödinger equation many times over, for different values of the vibrational quantum number N (= number of nodes in the radial wavefunction) and of the angular momentum number L. The calculation being of an exploratory nature, we were not interested in extreme accuracy for the wavefunction, and were willing to trade accuracy against machine time. The methods described in this paper are, however, useful for any desired accuracy, the modification in the choice of parameters being obvious in each case.

We shall discuss the various problems encountered, and methods used to solve

---

[1] On study leave from the University of New South Wales, Kensington, NSW, Australia.

these problems, in an order designed to facilitate the reading of the paper, rather than in completely logical order. These problems are as follows.

(1) How to choose a step size, how to decide when the step size needs changing, and how to carry out this change. In this connection, a midpoint formula is developed for use with the Noumerov method.

(2) How to iterate on the eigenvalue when already close to it. The usual variational method is put into a convenient form, and a formula is developed for the derivative of the wavefunction, to be used with the Noumerov method (it should be noted that the Nuomerov method itself does not give the derivative at all).

(3) How to narrow down the search for the desired eigenvalue (given number of nodes N) in the initial stages.

(4) How to decide on a practical range of the independent variable x (or r) for the numerical integration.

All these processes, many of them involving choices, are to be carried out by the machine, without operator intervention during the course of the computation.

The method described here are doubtlessly not all new, though some of them probably are. In a field as well worked as this one, it would be a major job searching through the literature so as to give proper credit at all stages. We thought it best to describe methods which we have actually used, and which have worked effectively, without trying to determine who originated each of the methods described.

II. STATEMENT OF THE PROBLEM, AND A QUICK REVIEW OF THE NOUMEROV METHOD

We write the one-dimensional Schrödinger equation in the form

$$d^2u/dx^2 = f(x)u(x), \qquad (2.1)$$

where

$$f(x) = (2M/\hbar^2)[V(x) - E]. \qquad (2.2)$$

Here $V(x)$ is the potential energy function, M is the reduced mass of the problem, and $\hbar$ is Planck's constant divided by 2π. The potential $V(x)$ approaches zero in the limit of large positive x, it is negative for intermediate values of x, and becomes positive and large for small positive x. Formally speaking, we desire a solution $u(x)$ which is bounded and square-integrable on the positive x axis, with $u(0) = 0$. In practice, we need not consider values near $x = 0$ at all, since $u(x)$ becomes ex-

ponentially small in that region, due to the "repulsive core" of the potential $V(x)$. The energy E in (2.2) is an eigenvalue, to be determined so that the solution $u_N(x)$ of (2.1) is not only square-integrable (and therefore approaches zero as x approaches infinity), but has exactly N nodes (zeros) on the positive x axis.

The eigenvalues E_N in question are negative. For any E, eigenvalue or not, there are two values of x, called the classical turning points, at which

$$f(x) = 0 \quad \text{classical turning points.} \tag{2.3}$$

The desired solution $u(x)$ has increasing exponential behavior for $0 < x < x_1$, oscillatory behavior between the two turning points [where $f(x)$ is negative], and decreasing exponential behavior for $x > x_2$.

In the remainder of this Section, as well as in Section III, we shall concentrate on the problem of numerical integration of the differential equation (2.1), without worrying about the eigenvalue problem aspects; that is, we assume that $f(x)$ is a given function, which is large and positive for small x, becomes negative in the range $x_1 < x < x_2$ [where x_1 and x_2 are the solutions of (2.3)], and then becomes positive again, approaching a constant positive value as x approaches infinity.

Considered purely as a differential equation, Eq. (2.1) is linear, second-order, and self-adjoint (does not involve the first derivative explicitly). There is one "canonical" method for its numerical solution, the Noumerov method, which is so clearly superior to all other methods that no other method should be seriously considered in practice. To establish notation for later use, we review this method briefly here, and follow this with a brief reminder why this is the method of choice.

We start from the Taylor expansion of $u(x + h)$ around the point x,

$$u(x + h) = \sum_{n=0}^{\infty} \frac{h^n}{n!} u^{(n)} \tag{2.4}$$

where $u^{(n)}$ is the n'th derivative of $u(x)$ evaluated at the point x. We obtain

$$\frac{1}{2} [u(x + h) + u(x - h)] = u + \frac{1}{2} h^2 u^{(2)} + \frac{h^4}{4!} u^{(4)} + \frac{h^6}{6!} u^{(6)} + \dots \tag{2.5}$$

and, differentiating twice,

$$\frac{1}{2} [u^{(2)} (x + h) + u^{(2)} (x - h)] = u^{(2)} + \frac{1}{2} h^2 u^{(4)} + \frac{h^4}{4!} u^{(6)} + \dots . \tag{2.6}$$

We now multiply (2.6) by the factor $\frac{1}{12} h^2$, and subtract the result from (2.5). This eliminates the term proportional to $u^{(4)}$. We replace the second derivative

$u^{(2)}$, wherever it occurs, by $f(x)u(x)$ according to (2.1). Introducing the notation

$$T(x) = \frac{h^2}{12} f(x) = \frac{h^2}{12} \frac{2M}{\hbar^2} [V(x) - E],\tag{2.7}$$

we thus arrive at the basic formula of the Noumerov method:

$$[1 - T(x + h)]u(x + h) + [1 - T(x - h)]u(x - h)$$

$$= [2 + 10T(x)]u(x) - \frac{h^6}{240} u^{(6)} + \dots .\tag{2.8}$$

If $u(x)$ and $u(x - h)$ are known, $u(x + h)$ can be found directly from this equation if the error term, proportional to $u^{(6)}$, is ignored; the values of $T(x)$ are known, of course.

This is the method of choice for the integration of (2.1) because it is the highest-order method which is at the same time a three-point method. Lower-order methods such as the Runge–Kutta method (error of order $h^4 u^{(4)}$) lead to smaller net intervals h, and hence longer integration times and more roundoff errors. Methods involving more than three adjacent function values should be avoided like the plague, since they are frequently unstable. They amount to replacing the original second-order differential equation by a difference equation of a higher order. The difference equation then has spurious roots which bear no relation to properties of the original differential equation; if one of the spurious roots gives rise to an exponentially increasing spurious contribution (and this is a frequent circumstance), the solution is completely submerged in numerical errors after comparatively few steps. If the Noumerov method with a given step size h has insufficient accuracy, the remedy lies in decreasing the step size, not in going to some other method.

III. Changing Net Size

In view of the large positive values of the potential $V(x)$ for small x, we must start the integration with a rather small net size h. Unless we are prepared to change net size as we go out, we will then waste a lot of machine time in the region where $f(x)$ is small in absolute value. It is therefore necessary to decide, during the course of the integration, whether the net size can be increased with safety, whether it must perhaps be decreased, and to program in the necessary steps for carrying out these operations.

In the neighborhood of a point x at which $f(x)$ is positive and varying slowly, the solution $u(x)$ has roughly exponential behavior $\exp(\pm ax)$ with $a \simeq [f(x)]^{1/2}$.

In a region of negative $f(x)$, the behavior of $u(x)$ is oscillatory, of type $\sin(kx - b)$ with $k \simeq [-f(x)]^{1/2}$. In either case, we obtain the estimate

$$u^{(6)} \sim [f(x)]^3 u(x). \tag{3.1}$$

We substitute this estimate into the error term of (2.8), and use the notation (2.7) to obtain

$$\text{Relative error per step} = \frac{\text{Error in } u(x+h)}{u(x)} \sim -\frac{72}{10}[T(x)]^3. \tag{3.2}$$

The relative error per step that we are prepared to tolerate depends on the total number of steps we anticipate having to take, and on the accuracy with which we wish to know the final wavefunction. Crudely speaking, the number of steps we shall have to take is proportional to the number of nodes N in the wavefunction, since a given accuracy in $u(x)$ decides in the main the number of steps h needed per half-wavelength of $u(x)$, and the number of half-wavelengths can be estimated as $N + \frac{1}{2}$ for our present purpose.

To give an example, suppose we anticipate having to take some 500 steps altogether, and we wish to know the function $u(x)$ to 1% accuracy. We can then tolerate relative errors of up to 2×10^{-5} per step, which by (3.2) means $|T(x)| \leq 0.01$ is a safe upper limit. Since the actual error committed is proportional to the cube of $T(x)$, we gain accuracy rapidly as $T(x)$ decreases in absolute value. Thus, we also want to avoid $T(x)$ becoming too small; otherwise, we are wasting steps. Suppose we decide to double the step size h when $|T(x)|$ falls below one fifth the value of its upper limit, i.e., below 0.002 in our example. Doubling the step size multiplies $T(x)$ [Eq. (2.7)] by 4. Thus the new value of T after the net-size doubling is 0.008, still below the upper limit 0.01. If $T(x)$ decreases steadily in absolute value as x increases (as it does in our problem until we reach the inner turning point), the sooner the net-size doubling is done, the better.

These estimates depend, of course, on integrating in such a way that the error made at a given step does not tend to perpetuate itself, with compound interest, during subsequent steps. In practice, this means we must not integrate "against the grain" of the differential equation: e.g., for x larger than the outer turning point, the solution we want has decreasing exponential behavior; if we integrate in that region, every little error we make admixes to this desired solution an undesired portion proportional to an increasing exponential. After comparatively few steps, the desired solution is utterly submerged by the exponentially increasing error terms. The remedy is obvious, however. In this region of x we must integrate inwards rather than outwards. With this understood, our crude method of error estimation is adequate.

It should be noted that the condition deduced from the estimate (3.2) is a condition on $T(x)$, i.e., on the coefficient in the differential equation, not on the solution $u(x)$ directly. This highly desirable behavior is associated with the linearity of the differential equation (2.1), and is sometimes not brought out clearly in books on numerical analysis where the emphasis is frequently on the solution of nonlinear differential equations. Since the condition which determines the choice of step size is on $T(x)$, the regions of x in which different step sizes are required can be determined as soon as $T(x)$ is known for all x to sufficient accuracy—i.e., as soon as we have a trial value of the energy E [which appears in (2.7)] sufficiently close to the true value of E. Thus, in principle, $T(x)$ need not be tested at every point during the integration; but in practice, such a test takes little time.

For the practical point of view, the only step-size changes of interest are halving and doubling of the step size. It is imperative that $T(x)$ should be calculated in minimum time. The effective way of doing so is to store, in core memory, the values of

$$U(x) = h^2 2MV(x)/12\hbar^2 \tag{3.3}$$

for all netpoints $x_n = nh$ on the finest net size h. The computation of $T(x)$ for this net size is then a straight table-lookup followed by subtraction of a constant:

$$T(x) = U(x) - h^2 2ME/12\hbar^2 \tag{3.4}$$

If we restrict net-size changes to doublings and halvings, and if h in (3.3) is the finest net size ever used [i.e., the net size appropriate for the point at which we start the outwards integration, where $f(x)$ and $T(x)$ are largest], then the effect of net-size changes amounts to multiplication of the right side of (3.4) by an appropriate power of 4, a fast operation in a binary machine. If h is chosen as indicated, the possible net-size halvings during the course of the integration will never lead us to a net size h' smaller than h, and there will never be a need for interpolation in the table of $U(x)$.

Doubling the net size is trivial: all we need to do is to carry along, during the integration, $u(x - 2h)$ as well as $u(x - h)$. When h is doubled, as a result of a test on $T(x + h)$, the value of $u(x - 2h)$ is stored into the position reserved for $u(x - h)$, $T(x)$ is multiplied by 4, and $T(x - h') = T(x - 2h)$ as well as $T(x + h') = T(x + 2h)$ are obtained by table-lookup followed by multiplication by a new power of 4.

Halving the net size, at first sight, is more troublesome. If we decide that the absolute value of $T(x + h)$ is too large for comfort, and to introduce the halved net size $h' = \frac{1}{2}h$, we require $u(x - h') = u(x - \frac{1}{2}h)$ to continue the integration.

The values of $u(x)$ known to us are $u(x - h)$ and $u(x)$. We thus require an accurate formula for midpoint interpolation—accurate to the same order as the Noumerov method itself—for otherwise we lose the advantage of the Noumerov method.

Letting $x_0 = x - \frac{1}{2}h$ be the point at which we desire to know u, our problem can be restated as follows: find $u(x_0)$, given values of $u(x_0 + h')$ and $u(x_0 - h')$, and given that $u(x)$ satisfies the differential equation (2.1).

The solution, though exceedingly simple, does not, to our knowledge, appear in the literature: it consists in using the basic formula of the Noumerov method, (2.8), to solve for $u(x)$! The accuracy is then obviously the same as the accuracy of the Noumerov method. No additional function values need be stored and net-size halving is now as simple as net-size doubling. For the sake of the record, we write down the midpoint interpolation formula explicitly[2]:

$$u(x) = \frac{[1 - T(x + h)]u(x + h) + [1 - T(x - h)]u(x - h)}{2 + 10T(x)} + \frac{h^6 u^{(6)}}{480} + \dots. \quad (3.5$$

IV. Iteration on the Eigenvalue when We Are Close; a Derivative Formula

Let $v(x)$ be a "trial function" which is sufficiently close to the true function $u(x)$. We leave the definition of "sufficiently close" to Section V. Then the usual "variation method" of quantum mechanics asserts that an improved value of the energy eigenvalue E can be found by quadrature over the known function $v(x)$, namely,

$$\frac{2ME}{\hbar^2} \simeq \frac{\int v(x)\{-(d^2v/dx^2) + [2MV(x)/\hbar^2]v(x)\}\, dx}{\int [v(x)]^2\, dx}. \quad (4.1)$$

Written in this form, the expression is rather awkward. However, we have seen already that, for purely numerical reasons, we must integrate outwards for small x, and inwards for large x. We now make a virtue out of this necessity: Let Q be a trial value for the energy E, hopefully close to the true value of E. Then integrate the differential equation

$$\frac{d^2v}{dx^2} = \frac{2M}{\hbar^2}\,[V(x) - Q]v(x) \equiv f_Q(x)v(x), \quad (4.2)$$

first outwards from some sufficiently small value of x until we reach a joining

---

[2] In the error term, we may replace $2 + 10T(x)$ by 2.

point $x = x_0$, then inwards from some sufficiently largue value of x until we reach the same joining point. The outwards integration is started in such a way that we obtain the exponentially increasing solution (see Section VII for the method), the inwards integration is started in such a way that we obtain the exponentially decreasing solution (which of course increases in the direction of integration). In either solution, there is one free multiplicative constant. We can, and do, readjust this constant at the end so that $v(x)$ turns out to be continuous at $x = x_0$, the joining point. In practice, x_0 is chosen to be the minimum of the potential $V(x)$.

The fact that this $v(x)$ is not yet the true solution to the eigenvalue problem manifests itself as a discontinuity in the first derivative $v'(x)$ at $x = x_0$. The left-hand value v_L' obtained from the outwards integration fails to equal the right-hand value v_R' obtained from the inwards integration. The second derivative d^2v/dx^2 therefore has a delta-function singularity at $x = x_0$, which makes a finite contribution to the integral in the numerator of (4.1). Except for this delta-function contribution, the result would be just Q, the trial value for the energy, as can be seen by substituting (4.2) into (4.1). Putting things together, we obtain the simple iteration formula

$$\frac{2ME}{\hbar^2} \simeq \frac{2MQ}{\hbar^2} - \frac{v(x_0)(v_R' - v_L')}{\int [v(x)]^2 dx}. \tag{4.3}$$

The Simpson-rule sums necessary for the evaluation of the integral in the denominator can be accumulated during the process of solving the differential equation (4.2), and can be multiplied by the appropriate factors to make $v(x)$ continuous at $x = x_0$, without any problem.

However, Eq. (4.3) is useless unless we have an accurate value for the derivative $dv/dx = v'(x)$. The Noumerov method by itself fails to give us such a value; in fact, the Noumerov method is built on the fact that the first derivative does not appear explicitly in the differential equation.

One method is to integrate the second derivative d^2v/dx^2 numerically; however, this is both awkward and productive of numerical inaccuracies.

A better method, which is new to our knowledge, can be developed by using reasoning similar to that of the Noumerov method itself. We start by developing a low-accuracy formula, so as to show the basic idea; we then improve the method so as to get a derivative formula with an error term of order $h^9 v^{(9)}$.

Returning to the Taylor-series expansion (2.4), we compute

$$A_1 \equiv \frac{1}{2} [u(x + h) - u(x - h)] = hu' + \frac{h^3}{3!} u^{(3)} + \frac{h^5}{5!} u^{(5)} + \dots . \tag{4.4}$$

Taking the second derivative on both sides, and multiplying by $\frac{1}{6}h^2$, we obtain

$$B_1 \equiv \left(\frac{h^2}{12}\right) [u''(x+h) - u''(x-h)] = \frac{h^3}{6} u^{(3)} + \frac{h^5}{36} u^{(5)} + \dots \quad (4.5)$$

We subtract (4.5) from (4.4) and use the differential equation to replace u'' by $f(x)u(x)$, to obtain the first derivative formula:

$$hu' = [1 - T(x+h)]u(x+h) - [1 - T(x-h)]u(x-h) + (7/360)h^5u^{(5)} + \dots$$
$$(4.6)$$

Here $T(x)$ is defined by (2.7) with $f(x) = f_Q(x)$ defined by (4.2). The error term in (4.6) may be sufficiently small in some cases. It is, however, of order $h^5u^{(5)}$, poorer than the basic error of the Noumerov method.

At substantially no expense in machine time, the accuracy of the first derivative, and hence of (4.3), can be improved significantly, simply by using function values at $x + 2h$ and $x - 2h$. We define

$$A_2 = \tfrac{1}{2}[u(x+2h) - u(x-2h)] \quad (4.7)$$

and

$$B_2 = \tfrac{1}{12}h^2[u''(x+2h) - u''(x-2h)]$$
$$= T(x+2h)u(x+2h) - T(x-2h)u(x-2h) \quad (4.8)$$

We then write down the Taylor expansions of A_1, A_2, B_1, and B_2, carrying terms up to order $h^9u^{(9)}$ inculsive. We eliminate the terms proportional to $h^ku^{(k)}$ with $k = 3, 5, 7$ and solve for hu'. The procedure is tedious but strightforward; we quote only the result which is

$$hu' = \frac{16}{21}\left(-A_1 + \frac{37}{32}A_2 - \frac{37}{5}B_1 - \frac{17}{40}B_2\right) - \frac{4016}{35}\frac{h^9u^{(9)}}{9!} + \dots \quad (4.9)$$

Thus, by integrating a mere two steps beyond the joining point $x = x_0$, we can determine the value of the first derivative to an accuracy substantially better than the basic accuracy of the Noumerov method. (We note that the instability problems involved in the use of higher-order schemes do *not* arise here: we wish to find u' at one point $x = x_0$, whereas instability arises only if we wish to replace a differential equation by a difference equation of higher order, and to integrate over many steps.)

We now have an iteration scheme of second order for the eigenvalue E: starting from a trial eigenvalue Q, near to E, we find an improved approximation to E

from (4.3). The improvement is second order, i.e., the error of E is proportional to the square of the error of the trial value Q, once Q is close enough.

The iteration is terminated conveniently when we begin to hunt, i.e., when the new correction $E\text{-}Q$ is both sufficiently small for safety (say, less than 10^{-5} of E in absolute value) and no smaller than the previous correction in absolute value. The one unnecessary iteration can be saved after a bit of experience, by setting a straight upper limit on $|E\text{-}Q|$, and terminating as soon as $|E\text{-}Q|$ falls below this upper limit.

> [Although much of what we have done can be used also for coupled systems of equations, such as arise for example in the deuteron problem with tensor forces, the termination condition in that case can *not* be based on the size of the correction to the eigenvalue E. For coupled equations, involving unknown functions $u_1(x)$ and $u_2(x)$, say, the ratio u_1/u_2 at the starting point of the integration must be determined as well as the eigenvalue E. Convergence of the ratio u_1/u_2 is then a more stringent test than convergence of E, and the more stringent test must be applied if an accurate solution is desired. This problem does not arise in the situation discussed by us here, where the trial function $v(x)$ and the trial eigenvalue Q are directly linked through the differential equation (4.2), with nothing but simple multiplicative factors at our disposal in the determination of $v(x)$.]

V. HOMING IN ON THE EIGENVALUE FROM FAR

The iteration procedure described in the preceding section works only if the trial energy Q is already quite close to the true eigenvalue E_N, associated with the eigenfunction $u_N(x)$ which has the desired number N of nodes. We require a procedure to get to this stage, as well as a method of recognizing when we have gotten there.

By far the easiest, and usually not far from the quickest, method for this purpose is the elementary device of halving the interval. Let us suppose that we know two energies Q_1 and Q_2 such that the desired eigenvalue E_N lies certainly between them:

$$Q_1 < E_N < Q_2. \tag{5.1}$$

We then try the value

$$Q = \tfrac{1}{2}(Q_1 + Q_2) \tag{5.2}$$

and ascertain whether Q lies above or below the desired E_N. If Q lies above E_N, we replace Q_2 by Q and repeat the process; if Q lies below E_N, we replace Q_1 by Q and repeat the process. At each stage, we gain exactly one binary digit of accuracy in the energy. Unless Q_1 and Q_2 are very bad limits indeed, a few stages

of halving suffice to get us close enough to E_N so that we can use the faster, second-order iteration scheme of Section IV.

The first step to ascertain where Q lies in relation to E_N is to count the nodes of the trial function $v(x)$ generated by (4.2). As we generate $v(x)$, we count each node and accumulate. If the node count, at any stage, exceeds N, then the trial value Q was too high. Conversely, if at the end of generating $v(x)$, the node count is below N, then Q was too low. Since we generate $v(x)$ in two stages, so to speak, first by integrating out, then by integrating in, a bit of care is required to avoid double counting of nodes occurring right at the joining point x_0.

While node counting is enough to reject completely unsuitable values of Q and to decide whether such values are high or low, a correct node count by itself is not sufficient to allow us to use the procedure of Section IV. We have found the following simple scheme quite adequate to ensure convergence:

(1) Ascertain whether the number of nodes in $v(x)$ equals N; if not, proceed with halving.

(2) If the node count agrees, proceed to evaluate the second-order correction E-Q from (4.3) and (4.9); the *sign* of the correction is right, even if the magnitude is far off. Thus, if E-Q turns out to be positive, Q was too low, and we replace Q_1 by Q; if E-Q turns out to be negative, Q was too high, and we replace Q_2 by Q.

(3) Now compute the new $E = Q + (E - Q)$ predicted by the second-order iteration scheme. If this new value of E lies between Q_1 and Q_2, it is safe to use. If not, the next trial value is determined by (5.2).

In this way, we combine the safety of the halving scheme with the speed of the second-order iteration. Two cautions should be observed, however. (1) Once the correction $E - Q$ gets very small, it is not necessary to readjust Q_1 and Q_2, and it is actually safer not to, since round-off errors can then lead to trouble. (2) To avoid trouble of an unforeseen sort, it is desirable to keep a count of the number of iterations, and to go out with a failure indication once this count exceeds 50, say. This should not happen, but things which should not happen sometimes do; a simple iteration count will prevent looping of the program.

VI. SETTING UPPER AND LOWER BOUNDS ON THE EIGENVALUE

The halving method of Section V requires bounds Q_1 and Q_2 on the true eigenvalue E_N. In this section, we discuss methods of setting such bounds.

The simplest case, and the one occurring most of the time, is that we already

possess a list of true eigenvalues E_M with $M = N - 1, N - 2, N - 3, ..., N - k$, say. Clearly $Q_1 = E_{N-1}$ is a lower bound for E_N. Ordinary polynomical extrapolation of the list E_M to some depth $j \leq k$ [in practice, $j = \text{Min}(k, 4)$ is adequate] yields a prediction for E_N, which we denote by W_N. We then put

$$Q_1 = E_{N-1}, \quad Q_2 = E_{N-1} + 2(W_N - E_{N-1}). \tag{6.1}$$

Unless the polynomical extrapolation is utterly unjustified, the factor 2 in (6.1) ensures that Q_2 lies above the true E_N. Furthermore, when we start the halving procedure with *this* Q_1 and Q_2, the first value tried, by (5.2), is the predicted value, $Q = W_N$. If the prediction is accurate, this first trial value is close enough to the truth to permit use of the second-order iteration scheme, which then yields full convergence in 3 or 4 steps.

Use of (6.1) requires at least two known eigenvalues, E_{N-1} and E_{N-2}. Thus, an alternative procedure is required at the beginning of the run. The simplest choice at the very start is

$$Q_1 = \text{Min}[V(x)], \quad Q_2 = 0. \tag{6.2}$$

These are perfectly safe upper and lower bounds on all bound-state energies. There are two troubles, however: (1) Quite a few halving steps may be required if such generous bounds are used [this is not a very serious trouble, since (6.1) can be used after the second eigenvalue has been determined]: and (2) The choice of the range of integration (minimum and maximum values of x) for finding $v(x)$, which is discussed in Section VII, depends on the trial energy. A range of integration suitable for $Q = \frac{1}{2}\text{Min}[V(x)]$, the first trial value generated from the choice (6.2) by means of (5.2), is generally a most unsuitable range of integration for the desired eigenvalue E_N, particularly so if $N = 0$, i.e., if we wish to start by generating the ground state. Thus, if (6.2) is used, the range of integration must be readjusted during the halving process until we are down to the right number of nodes.

Better limits than (6.2) are available if we want to generate all eigenvalues E_N from $N = 0$ onwards. $Q_1 = \text{Min}[V(x)]$ is then a fairly close lower bound for the true E_0, but $Q_2 = 0$ would be a very bad upper bound. A simple scheme consistes in approximating $V(x)$ near its minimum by a quadratic polynominal (oscillator potential) and determining the ground state energy of this oscillator, $\frac{1}{2}(h\nu)$. We then set $Q_2 = Q_1 + h\nu$, the extra factor 2 serving as a safety measure. An alternative, and even safer, scheme, is to use (4.1) with a simple trial function $v(x)$, e.g.,

$$v(x) = \exp[-a(x - x_0)^2]. \tag{6.3}$$

No matter what value of a we choose in (6.3), the right side of (4.1) is an upper limit Q_2 to the true ground-state energy E_0. The two methods can be combined by determining the parameter a in (6.3) from the oscillator-potential fitting, the function (6.3) being just of the right form for the ground-state wavefunction of an oscillator potential.

Although this method is safe, it is probably overly elaborate for the time gained. With a bit of experience, it is possible to make a reasonable guess at the zero-point energy $E_0 - \mathrm{Min}[V(x)]$, and to set a generous upper limit Q_2 which is nonetheless far nearer to the true E_0 than the trivial choice $Q_2 = 0$, without going to all the trouble of evaluating (4.1) numerically for the function (6.3).

Once the lowest eigenvalue E_0 is known, a safe upper limit for the eigenvalue E_1 is

$$E_1 \le Q_2 = E_0 + 3\{E_0 - \mathrm{Min}[V(x)]\}. \tag{6.4}$$

The factor 3 is exactly right for a square-well potential, and is an overestimate (hence giving an upper bound) for all other potentials; for an oscillator potential, the correct factor would be 2, to that 3 is a perfectly safe choice for an upper limit.

Once E_0 and E_1 are known, extrapolation becomes possible with more and more accuracy as further eigenvalues are accumulated. We thus fall back on the choice (6.1) which ensures very rapid convergence.

VII. STARTING THE INTEGRATION: DECIDING ON THE RANGE OF INTEGRATION

The outwards integration must start at a value of x less than the inner turning point x_1 [the lower root of (2.3)]; the inwards integration must start at a value of x larger than the outer turning point x_2 [the higher root of (2.3)]. In this section, we discuss the choice of these starting points, and hence the choice of the total range of integration for the wavefunction. We also discuss how the integrations are started so as to get the desired solution, i.e., the exponentially increasing solution for the outwards integration, the esponentially decreasing solution for the inwards integration.

We discuss the second point first; that is, let us suppose we have chosen a starting value of x, call it $x = a$, for the outward integration. Clearly $v(a)$ can be set arbitrarily, since one multiplicative factor is free. To get going with the Noumerov method, we require an approximation to $v(a + h)$ for the exponentially increasing solution.

The first thing to realize is that quite sizeable errors are permitted here. An erroneous choice of $v(a + h)$ has the effect of admixing, to the desired exponentially

increasing solution, a component proportional to the other, exponentially decreasing solution. As we integrate out from a towards the turning point x_1, this erroneous component becomes smaller and smaller, whereas the desired component increases in value.

To the crude approximation needed here, the differential equation (2.1) is satisfied by

$$v(x) \simeq v(a) \exp[W(x)], \tag{7.1}$$

where

$$W(x) = \int_a^x [f(x)]^{1/2} \, dx \tag{7.2}$$

This is one step cruder than the usual WKB approximation, but is good enough for us here. We now become even cruder, by replacing the integral in (7.2) by its trapezoidal-rule approximation. The result is the following estimate for $v(a + h)$:

$$v(a + h) \simeq v(a) \exp\{[3T(a)]^{1/2} + [3T(a + h)]^{1/2}\}, \tag{7.3}$$

where $T(x)$ is defined by (2.7), and is the quantity we require in any case for the Noumerov method. We note that $f(x)$, and hence $T(x)$, are positive outside the classical turning points, so that the square roots in (7.3) are real numbers. The positive square roots should be used for the exponentially increasing solution, i.e., to get from a to $a + h$, and the negative square roots should be used for the exponentially decreasing solution, i.e., to get from the outermost point $x = b$ to $x = b - h$, at the start of the inwards integration. We note that there is no difference, in the Noumerov method, whether one integrates inwards or outwards; the basic equation, (2.8), can be solved as easily for $u(x - h)$ as it can for $u(x + h)$.

It remains to decide on suitable values of a and b. If a is too close to the inner turning point x_1, we fail to generate enough of the desired wavefunction (we miss an appreciable part of the exponential tail); if a is too far from x_1, we not only waste machine time by generating the wavefunction in a region where its value is exceedingly small and of no conceivable physical interest, but we can also get into scaling troubles: even modern machines, with floating-point facilities, do not allow an infinite range of the floating-point exponent! And once we are well and truly into the exponential region, it becomes all too easy to get into underflow troubles even on present machines.

Suppose we wish to choose the inner starting point $x = a$ so that $v(a)$ is smaller than the value at the turning point, $v(x_1)$, by a factor $\exp(A)$, with A given *a priori*. For example, we might choose $A = 16$, corresponding to a factor of roughly 10^7. We now use the estimate (7.1), (7.2) to get the condition on a:

$$\int_a^{x_1} [f(x)]^{1/2}\, dx = A.\tag{7.4}$$

The integrand is zero at the upper limit, the classical turning point x_1. We again replace the integral by a trapezoidal-rule approximation, and keep going downwards through $x_1 - h$, $x_1 - 2h$, $x_1 - 3h$, ..., until the accumulated sum exceeds A. The terms in the sum are of form $[12T(x_i)]^{1/2}$ where $T(x)$ is defined by (2.7).

The uppermost value of x, $x = b$, is determined similarly, the condition being

$$\int_{x_2}^b [f(x)]^{1/2}\, dx = A.\tag{7.5}$$

Since $f(x) = f_Q(x)$ [Eq. (4.2)] depends on the value of the trial energy Q, the turning points x_1, x_2, as well as the cutoff points a and b, depend on the value of Q. As Q increases, the outer turning point x_2 and the outer cutoff point b move further out (increase in value), whereas the inner turning point x_1 and the inner cutoff point a move inwards (decrease in value). In principle, therefore, a and b ought to be recalculated whenever the trial energy Q changes.

This, however, is neither necessary nor desirable, except in the earliest stages of the iteration process, when the trial energy is still very poor. As soon as the upper limit Q_2 has improved to the point where it yields the right number of nodes, the values of a and b associated with this upper limit Q_2 should be retained and used throughout the remainder of the iteration. First of all, this saves time. Second, if a and b are allowed to change during the later stages of the iteration process, it is possible to get into serious trouble: the truncation errors associated with replacing the differential equation by a difference equation, as well as the errors made in the very crude starting formula (7.3), can lead to hunting of the second-order iteration procedure unless a and b are kept fixed.