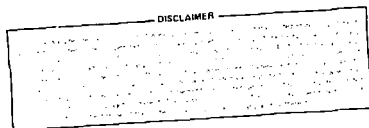PROCEEDINGS

of the workshop

ALGORITHMS AND COMPUTER CODES FOR
ATOMIC AND MOLECULAR QUANTUM SCATTERING THEORY

Volume I

Sponsored by the

NATIONAL RESOURCE FOR COMPUTATION IN CHEMISTRY

Lawrence Berkeley Laboratory
Berkeley, California 94720

Held at
Argonne National Laboratory
June 25-27, 1979

NRCC Proceedings No. 5

Edited by: Lowell Thomas

CONTENTS

INVITED PRESENTATIONS

FOREWORD

The National Resource for Computation in Chemistry (NRCC) was
established as a division of the Lawrence Berkeley Laboratory (LBL) in
October 1977. The functions of the NRCC may be broadly categorized as
follows: (1) to make information on existing and developing computa-
tional methodologies available to all segments of the chemistry community,
(2) to make state-of-the-art computational facilities (hardware and soft-
ware) accessible to the chemistry community, and (3) to foster research
and development of new computational methods for application to chemical
problems.

Workshops are planned as an integral part of the NRCC's program.
A workshop in the titled area was judged timely by key members in the
field and led to a planning meeting held February 23-24, 1979 at the
University of Utah at Salt Lake City. In addition to the co-chairmen,
Professor John Light, University of Chicago, and Dr. Lowell Thomas, NRCC,
we were pleased to have the participation of Dr. B. Robert Johnson,
Aerospace Corporation, and Dr. G. Parker, University of Chicago.

As the site for this workshop, we sought Argonne National Laboratory
which has maintained active interest in the development of the NRCC. We
are indebted to Dr. Michael V. Nevitt, Deputy Director, Argonne National
Laboratory for making the ANL available for this purpose.

Finally, I wish to express my thanks to the co-chairmen for their
considerable efforts in organizing the workshop and in editing this
Proceedings.

The National Resource for Computation in Chemistry is funded jointly by the Basic Energy Sciences Division of the U.S. Department of Energy and the National Science Foundation.

William A. Lester, Jr.
Director, NRCC

Workshop participants. *Left to right:* L. Thomas, N. Harvey, J. Light, D. Secrest, G. Parker, A. Askar, D. Malik, R. Gordon, M. Alexander, R. Walker, M. Redmon, A. Allison, F. Krogh, H. Rabitz, B. Johnson, H. Shin, T. Schmalz, Y. Yung, D. Truhlar, W. Lester, Jr.

List of Participants

Professor Millard Alexander
Department of Chemistry
University of Maryland
College Park, MD 20742

Dr. Arthur Allison
Department of Computing Science
University of Glasgow
Glasgow, United Kingdom

Professor Attila Askar
Mathematics Department
Bogazici University
Bebek, Istanbul
Turkey

Professor Roy G. Gordon
Department of Chemistry
Harvard University
Cambridge, MA 02138

Dr. Nancy Mullaney Harvey
Division of Chemistry &
  Chemical Engineering
California Institute of Technology
Pasadena, CA 91125

Dr. B. Robert Johnson
The Aerospace Corporation
Chemistry and Physics Laboratory
El Segundo, CA 90245

Dr. F. T. Krogh
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91130

Dr. William A. Lester, Jr.
National Resource for Computation
  in Chemistry
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Professor John C. Light
The James Franck Institute and
Department of Chemistry
University of Chicago
Chicago, IL 60637

Dr. David J. Malik
University of Illinois
School of Chemical Sciences
Urbana, IL 61801

Dr. Gregory A. Parker
The James Franck Institute and
Department of Chemistry
University of Chicago
Chicago, IL 60637

Professor Herschel Rabitz
Department of Chemistry
Princeton University
Princeton, NJ 08540

Dr. Michael J. Redmon
Battelle Columbus Laboratory
Chemical Physics Group
505 King Avenue
Columbus, OH 43016

Dr. Tom G. Schmalz
The James Franck Institute and
Department of Chemistry
University of Chicago
Chicago, IL 60637

Professor Don Secrest
School of Chemical Science
University of Illinois
Urbana, IL 61801

Professor Hyung K. Shin
University of Nevada, Reno
Department of Chemistry
Reno, NV 89557

Dr. Lowell D. Thomas
National Resource for Computation
  in Chemistry
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Professor Donald G. Truhlar
Department of Chemistry
University of Minnesota
Minneapolis, MN 55455

Dr. Robert B. Walker
T-12
Los Alamos Scientific Laboratory
Los Alamos, NM 87545

Dr. Yat Yan Yung
Physics Department
Pacific Lutheran University
Tacoma, WA 98444

PROGRAM

ALGORITHMS AND COMPUTER CODES FOR
ATOMIC AND MOLECULAR QUANTUM SCATTERING THEORY

Co-Chairmen: John Light and Lowell Thomas

Schedule

June 25, 1979

Morning	Chairman, John Light
8:45	Introduction Michael V. Nevitt, Deputy Director ANL William A. Lester, Jr., Director NRCC
9:00	Overview Don Secrest, University of Illinois
10:00	Coffee Break
10:15	Gordon's Method Roy Gordon, Harvard University
11:05	An Accelerated Gordon Algorithm for Inelastic Collisions Millard H. Alexander
11:15	The Log Derivative and Renormalized Numerov Algorithms B. Robert Johnson, The Aerospace Corporation
12:15 pm	Lunch
Afternoon	Chairman, Lowell Thomas
1:30	DeVogelaere's Method William A. Lester, Jr., NRCC
2:15	New Developments in Methods for the Numerical Solution of the Radial Schrödinger Equation Arthur Allison, University of Glasgow
3:15	Coffee Break
3:30	R-Matrix Recursion Methods: Continuous and L^2 Corrections John Light, University of Chicago
4:30	Finite Element Methods in Quantum Dynamics Herschel Rabitz, Princeton University
5:30	Adjourn

June 26, 1979

Morning	Chairman, John Light
9:00	Recent Developments at the NRCC William A. Lester, Jr., NRCC
9:30	Presentation of Test Problems Lowell Thomas, NRCC
10:30	Coffee Break
10:45	Accessing the NRCC Library and Test Problems Lowell Thomas, NRCC
12:00 noon	Lunch
Afternoon	Chairman, Lowell Thomas
1:30	A Variable Interval Variable Step Method for Linear Second Order Coupled Differential Equations Gregory Parker, University of Chicago
2:30	Integral Equation Methods for Inelastic Scattering Don Secrest, University of Illinois
3:30	Coffee Break
3:45	Applications of Close Coupling Algorithms to Electron-Atom, Electron-Molecule, and Atom-Molecule Scattering Don Truhlar, University of Minnesota and Nancy Mullaney Harvey, California Institute of Technology
4:45	Reference Potentials with Integral Equations Michael Redmon, Battelle Columbus Laboratories
5:30	Adjourn
6:00	Dinner

June 27, 1979

Morning	Chairman, William A. Lester, Jr.
9:00	Variable-Order, Predictor-Corrector Methods F. T. Krogh, Jet Propulsion Laboratory
10:00	An L^2 Approach to R-Matrix Propagation Robert Walker, Los Alamos Scientific Laboratory
10:45	Coffee Break
11:00	A Variation-Iteration Method for a Single Column of the S-Matrix Lowell Thomas, NRCC
11:45	Final Remarks
12:00 noon	Adjourn

Talks have been scheduled to allow 15 minutes for discussion. Persons desiring the give short presentations relating to one of the talks should make arrangements with the floor chairman of that session. Manuscripts of the scheduled talks may be mailed to the NRCC, c/o Close Coupling Workshop, or presented to one of the co-chairmen at the workshop. The time allotted for each talk includes 15 minutes for discussion. Manuscripts for unscheduled presentations will also be accepted for inclusion in the workshop report, and should be presented to one of the co-chairmen at the workshop.

PREFACE TO VOLUME I

The goals of this workshop are to identify which of the existing computer codes for solving the coupled equations of quantum molecular scattering theory perform most efficiently on a variety of test problems, and to make tested versions of those codes available to the chemistry community through the NRCC software library. To this end, many of the most active developers and users of these codes have been invited to discuss the methods and to solve a set of test problems using the LBL computers.

The first volume of this workshop report is a collection of the manuscripts of the talks that were presented at the first meeting held at the Argonne National Laboratory, Argonne, Ill. June 25-27, 1979. It is hoped that this will serve as an up-to-date reference to the most popular methods with their latest refinements and implementations.

Many of the codes will be used to solve the test problems on the CDC 7600 computer at LBL. A second meeting will be held in late October or early November of 1979 at Berkeley to discuss and compare the performance of the different codes with respect to the tests. A second report will then be issued containing the results and conclusions drawn about them. The two reports together should then serve as a useful guide to both the inexperienced person wishing to do calculations of this type and the forefront researcher wishing to advance the state of the art.

August 1979

John C. Light
Lowell D. Thomas

OVERVIEW

Don Secrest
School of Chemical Sciences
University of Illinois
Urbana, Illinois 61801

Since the early days of exact quantum scattering calculations a great variety of methods for solving scattering problems have been developed. One of the aims of this workshop is to identify viable methods and discuss the advantages of each.

We have all had the experience of attempting a new problem with one of our favorite techniques and for some reason seen it fail, only to find that a method we were sure was inferior gave beautiful results. Thus it is clear that there is no one best method. Each method is suited to some class of scattering problems. As I see it, our mission is to identify a small collection of programs suited to each class of scattering problems and make comparative studies of these methods to determine the most appropriate role of each method in inelastic scattering calculations.

The programs we will be discussing are quite general for the solution of the coupled differential equations,

$$(- \frac{d^2}{dR^2} + \frac{\ell_i(\ell_i+1)}{R^2} - k_i^2) \, F(R)_{iI} = - \sum_{i'} V(R)_{ii'} \, F(R)_{i'I} \tag{1}$$

with appropriate boundary bonditions. This 's the form of the close coupling equations for rotation, vibration or rotation-vibration and a number of other problems. It should be emphasized that many approximate formulations of scattering such as the centrifugal sudden, the energy sudden and the infinite order sudden approximation lead to equations of the identical form of Eq. (1), and the methods we are discussing are applicable to a large variety of systems and approximations.

I shall discuss some of the methods I am familiar with and try to give an impression as to the circumstances under which each should be used. I shall not attempt to discuss, or even mention all of the methods described in the literature, but shall merely classify the various methods into broad categories and discuss at least one method in each category. The various methods in each category differ in detail, and one or another of them may be better for particular situations, but in general one method is about as good as another within each category.

There are basically two different numerical approaches in common use for solving the coupled equations. One approach is to solve the equations numerically either in their differential equation form, or the equivalent integral equation form. I shall call this approach the approximate solution approach. The other approach is to approximate the potential matrix $\underline{V}$ in some acceptable manner and solve the coupled equations exactly. I shall refer to this as the approximate potential approach.

For each of these two approaches there are two techniques for developing the solution. This leads to four categories of method which are discussed below.

The first of these techniques I shall call solution-following. This consists of starting the solution well into the non-classical region of the problem where the potential energy is greater than the total energy and proceeding to follow the solution step by step into the asymptotic region. This is probably the most common technique for the approximate solution approach, and is exemplified by the DeVogelaere method developed by Lester,[1-3] the Sams and Kouri[4,5] method, the method of Choi and Tang,[6] and many others. This technique as applied to the approximate potential approach is exemplified by the methods of Gordon,[7,8] Light,[9-11] and Wilson.[12]

The second technique I shall refer to as <u>invariant imbedding</u>, though all the methods I shall discuss under this heading were derived without the use of the concept of invariant imbedding. These techniques consist of solving the problem scattering from a piece of the potential. Then a connection technique is used for combining the R matrices for parts of the potential into R matrices for larger portions of the potential, until the scattering matrix for the entire potential is developed.

The approximate solution approach with this technique was first implemented by the amplitude density[13,14] method. This method is no longer used as a numerical method except in special circumstances. The amplitude density connection formulas are still of great use however in connecting solutions obtained by different methods in different regions of the interaction potential. The log-derivative method of Johnson[15] is the principal method in this category still in general use.

The category defined by the approximate potential approach and the invariant imbedding technique has a single member, the R-matrix method of Light and Walker.[16] This method has been used mostly to date for reactive scattering, but is also a valuable method for inelastic scattering in certain circumstances.

The four categories are summarized in Table 1. The methods of each of these categories have properties which commend them to use in particular circumstances. The <u>approach</u> one uses is decided by the accuracy required and the number of solutions required at different energies. In general the approximate solution approach is capable of higher accuracy at reasonable machine cost. The approximate potential approach often gives acceptable accuracy, in many cases two or sometimes 3 significant figures in practical

problems with very few integration steps. The chief advantage of the approximate potential approach is that it allows large integration step sizes. Of course high accuracy can also be obtained by taking smaller steps, but all advantage over the approximate solution approach is lost when this is done, as the work required to improve accuracy grows much faster in the approximate potential approach than in the approximate solution approach. The approximate potential approach, though it allows larger steps than the approximate solution approach, also requires much more work per step than the approximate solution approach for the first calculation. Thus the two approaches are of the same order of difficulty for one solution of low accuracy. The approximate potential approach has the advantage that much of the work done for the first solution may be saved and the problem may be solved at another energy with very little effort. If a large number of solutions to the scattering problem are to be found at different energies, the approximate potential approach is ideal. If only a few solutions are required or high accuracy is required, the approximate solution approach can be an order of magnitude less time consuming.

The solution-following techniques are plagued by instabilities requiring measures to be taken from time to time in the progress of the solution to ensure stability. The invariant imbedding techniques are inherently stable, though often time consuming.

Recently two very different techniques have appeared in the literature which do not fit this scheme very well at all. One is the iteration technique of Lowell Thomas[17] which allows the solution of huge coupled systems which would be very difficult indeed by direct methods, and the other is the finite element technique which has been pursued in its application to the scattering problem by Rabitz and Askar.[18,19] These have been added as appendages to Table 1.

There is no need to go through the details of any particular method as that will be covered by the speakers during the conference. Let me sketch out in very general terms how the techniques are approached and point out problems which arise.

Let me start with the solution following technique. How we implement it depends on the prob.m we are solving. For an inelastic scattering problem the potential is usually large near the origin. The wavefunction is zero at the origin and grows exponentially in the nonclassical region. So we can start here with a small wavefunction and integrate toward large R. When we get to the asymptotic region presumably we can compute the R matrix or the S matrix from the asymptotic form of the solution.

Several problems arise. First though the wavefunction is small in the nonclassical region unless we are solving a potential scattering problem there is more than one radial function. There are N solutions to the problem which decay near the origin. The technique used then is to start a complete independent set of solutions. When we get to the asymptotic region we can form linear combinations of these solutions. One of these combinations will correspond to the solution we are interested in. (Thus we get a complete set of solutions all at the same total energy whether we want them or not. This problem is not present in Lowell Thomas's[17] method which I will say more about later.) Now in order to do this it is necessary that the set of solutions we end up with be linearly independent. This is difficult to achieve sometimes. When we start we don't know how to pick the starting values and it probably wouldn't help much if we did. It is very easy to pick an independent set of starting conditions. But any set we pick will be a random mix in general of all solutions. All solutions are growing exponentials as we integrate toward large R, and each set will

contain some contribution from the most rapidly growing solution. This solution will rapidly dominate and all columns of the solution matrix will start looking like the fastest growing solution. This would be alright if we used infinite precision in solving the equations, but we never do. Thus after we start integrating it is necessary to alter this process before we lose precision. This is extremely easy to do. One merely forms new combinations of the solutions early in the integration which are orthogonal or in some similar way strongly independent. In the early days one performed these stabilizations every so often. It was found one needed to stabilize every 10 steps or so. This got time consuming. Roy Gordon developed some efficient stabilization codes. It was later found that one need not stabilize often--soon after starting, then much later, and 3 to 5 stabilizations were found to be enough. Thus the need for efficient stabilization techniques vanished.

This is a good time to say a word about iterative methods. One would hope by interation to obtain the one or two solutions he is interested in instead of the complete set. If he were to start integrating at small R and integrate toward large R, small numerical errors would introduce unwanted solutions which would grow and defeat the scheme. Thus an iterative technique must start at large R and integrate toward small R. This of course requires a r asonably good guess.

Now for comparison purposes I would like to give my impression of the invariant embedding technique.

In this technique one breaks the R space into segments and takes the potential to be zero outside of these regions. Then for this small region one solves the scattering problem in one step. Usually one nowadays

obtains the R matrix for the problem. One then solves the next adjacent
region and finds an R matrix for it. Then the two R matrices are combined
exactly to obtain an R matrix for the combined region. One proceeds in
this way into the asymptotic region and the problem is solved. The com-
bining of R matrices is an exact procedure but requires a full matrix
inversion at each step. But the technique is completely stable. This is
very important for some problems. In a typical rotation vibration problem one
starts integration in the nonclassical region. If he is using a solution
following technique and starts too deep into the nonclassical region he
finds it necessary to stabilize often. If he is very deep into the non-
classical region he must stabilize every step. In fact it is possible to
start so deep that one cannot stabilize the solution at all. For most
inelastic problems this is not serious and can be overcome by starting the
solution nearer the classical region. But there are problems such as
curve crossing problems for which it is not possible to start near the
classical region. For some curve crossing problems the classical turning
point of one of the curves is deep into the nonclassical region of others.
In cases of this sort the Log-Derivative method or one of the other in-
variant imbedding techniques is ideal. The invariant imbedding techniques
require a lot of computational effort at each step but they are entirely
stable.

Now let us look at the approximate potential approaches. The potential
is broken up into piecewise continuous polynomials or constant steps depending
on the method used. This simple form of the potential allows an exact solu-
tion of the equations in each region. There is still coupling between
the equations and it is necessary to diagonalize the potential matrix at each

step. Actually it is only possible to diagonalize the potential at a single point in the region, but the potential remains nearly diagonal for a reasonably large distance on either side of the diagonalization point. For the next region the potential must again be diagonalized. In general a different transformation diagonalizes the potential in each region. Thus as the wavefunction is progated, or as the R matrix is advanced if one is using an invariant embedding technique, it is necessary to transform the result at each region boundary. In the solution following technique, even though the wavefunction for each region is computed analytically it is necessary to stabilize every so often.

The great advantage of the approximate potential approach is that large steps may be taken in regions where the potential is varying slowly. The steps may be larger than the wavelength of the solution. If calculation at high energies is necessary the approximate solution approaches require more steps than at low energy as the wavelength is shorter at high energy. The approximate potential approach however may use the same step size at high energy as at low. In fact it may even be possible to use a larger step size at higher energy due to the fact that the solution is not as sensitive to small changes in the potential at high collision energies as it is at lower energies.

If a series of calculations are to be performed at different energies, the work in the approximate potential approach to determine the step size, diagonalize the potential matrix and the transformation from one region to the next need be done only for the first calculation. It may be saved in the computer and used for all future calculations.

In the nonclassical region the potential is rapidly varying and the approximate potential approach must use small steps comparable to those required for the approximate solution approach.

In comparing the approximate solution and approximate potential approaches we can say that the approximate potential approaches allow very large steps at least where the potential is slowly varying. A great deal more work is required per step in approximate potential approaches than in approximate solution approaches however. Thus for a calculation at only a few energies the approximate solution approaches seem appropriace. Since the work of the first calculation in the approximate potential approach may be saved and used at other energies, for large numbers of calculations the approximate potential methods are the best. For higher accuracy the approximate solution approaches seem to improve as stepsize is reduced faster than approximate potential approaches.

The approximate potential approach requires small steps where the potential is varying rapidly. Thus a method which is superior to both the approximate potential and approximate solution approaches is to use the approximate potential approach where the potential is slowly varying and the approximate solution approach in a rapidly varying regime.

The use of iteration to solve large systems is a new technique and may prove to be a major breakthrough for accurate scattering calculations. As this method is applied to the many important problems which cannot be approached by the direct methods we will find either that it is the direction exact calculations will take in the future or it is too unwieldy to be practical.

The finite element method is well suited to some problems which are refractory to the methods we have been discussing. The dissociation problem which leads naturally to a two independent variable problem comes immediately to mind. As we develop more familiarity with this method its role in scattering theory will become clearer.

References

1. W. A. Lester, Jr., DeVogelaere's method for the numerical integration
    of second order differential equations without explicit first
    derivatives: Applications to coupled equations arising from the
    Schrödinger equation, J. Comp. Phys. $\underline{3}$, 322-325 (1968).

2. W. A. Lester, Jr., Calculation of cross sections for rotational exci-
    tation of diatomic molecules by heavy particle impact: Solution of
    the close-coupled equations, in "Methods of Computational Physics,"
    (B. Alder, S. Fernbach and M. Rotenberg, eds.) Vol. 10, pp. 211-243
    (1971).

3. W. A. Lester, Jr., The N coupled-channel problem, in "Modern Theoretical
    Chemistry," (W. H. Miller, ed.) Vol. 1, pp. 1-32, Plenum Press, New
    York (1976).

4. W. N. Sams and D. J. Kouri, Noniterative solution of integral equations
    for scattering. I. Single channels, J. Chem. Phys. $\underline{51}$, 4809-4814 (1969).

5. W. N. Sams and D. J. Kouri, Noniterative solution of integral equations
    for scattering. II. Coupled channels, J. Chem. Phys. $\underline{51}$, 4815-4819 (1969).

6. B. H. Choi and K. T. Tang, Inelastic collisions between an atom and
    a diatomic molecule. I. Theoretical and numerical considerations for
    the close coupling approximation, J. Chem. Phys. $\underline{63}$, 1775-1782 (1975).

7. R. G. Gordon, A new method for constructing wavefunctions for bound
    states and scattering, J. Chem. Phys. $\underline{51}$, 14-25 (1969).

8. R. G. Gordon, Quantum scattering using piecewise and analytical solutions, in "Methods of Computational Physics," (B. Alder, S. Fernbach, and M. Rotenberg, eds.) Vol. 10, pp. 81-110 (1971).

9. C. C. Rankin and J. C. Light, Quantum solution of collinear reactive systems: $H + Cl_2 \rightarrow HCl + Cl$, J. Chem. Phys. 51, 1701-1719 (1969).

10. G. Miller and J. C. Light, Quantum calculations of collinear reactive triatomic systems. II. Theory, J. Chem. Phys. 54, 1635-1642 (1971).

11. J. C. Light, Quantum calculations in chemically reactive systems, in "Methods of Computational Physics," (B. Alder, S. Fernbach and M. Rotenberg, eds.) Vol. 10, pp. 111-142 (1971).

12. A. S. Cheung and D. J. Wilson, Quantum vibrational transition probabilities in atom-diatomic molecule collisions, J. Chem. Phys. 51, 3448-3457 (1969).

13. B. R. Johnson and D. Secrest, The solution of the nonrelativistic quantum scattering problem without exchange, J. Math. Phys. 7, 2187-2195 (1966).

14. D. Secrest and B. R. Johnson, Exact quantum-mechanical calculation of a collinear collision of a particle with a harmonic oscillator, J. Chem. Phys. 45, 4556-4570 (1966).

15. B. R. Johnson, The multichannel log-derivative method for scattering calculations, J. Comp. Phys. 13, 445-449 (1973).

16. J. C. Light and R. B. Walker, An R matrix approach to the solution of coupled equations for atom-molecule reactive scattering, J. Chem. Phys. 65, 4272-4282 (1976).

17. L. D. Thomas, J. Chem. Phys. <u>70</u>, 2979 (1979).

18. H. Rabitz, A. Askar and A. S. Cakmak, Chem. Phys. <u>26</u> 61 (1978).

19. A. Askar, A. S. Cakmak and H. Rabitz, Chem. Phys. <u>33</u>, 267 (1978).

Table I

	Approximate Solution Approach	Approximate potential approach	
	Very little work per step. Accuracy increases rapidly as step size is reduced. Extreme accuracy practical	Lost of work per step for first solution. Subsequent solutions little work. Large steps when potential is varying slowly.	Iteration Thomas[17]
			Finite Elements Rabitz and Askar
Solution Following Technique Requires stabilization	DeVogalaere[1] Sams and Kouri[4] Choi and Tang[6]	Gordon[7,8] Light[9-11]	
Invariant Imbedding Technique completely stable Requires a matrix inversion at each step	Log-Derivative[15]	R-Matrix[16]	

PIECEWISE ANALYTIC SOLUTIONS TO QUANTUM CLOSE-COUPLING EQUATIONS:
A REVIEW OF RECENT DEVELOPMENTS

Roy G. Gordon
Department of Chemistry
Harvard University
Cambridge, Massachusetts 02138

"Divide and Conquer" is the approach used for the piecewise analytic solution of the quantum close-coupling equations. The problem is divided into a set of simpler model problems which approximate the original problem piecewise in a set of intervals. The model problems are solved analytically _ᴀ each interval. These approximate solutions are joined together continuously to form the complete solutions which satisfy the proper boundary conditions.

The main advantages of the method are greater computational speed and accuracy, compared to purely numerical methods. Also, the wave function is available in a relatively convenient and compact form for use in calculating matrix elements, when needed.

Since the original method[1] has been reviewed[2] in some detail, and a computer program based on it has been distributed widely,[3] the present paper will focus on more recent developments. These include the development of piecewise quadratic approximations to a useful form. The following sections discuss briefly these new results, for constructing the analytic zero-th order and first-order solutions using a piecewise quadratic potential. Then finally we sketch the piecewise analytic solution to equations with both 1st and 2nd derivatives, such as those that often arise in reactive scattering theory.

II. Zero-th order solutions for piecewise quadratic potentials

The originally developed method[1] uses a set of linear approximations
to the potential function, to generate the zeroth-order approximate
solutions. Deviations of the actual potential from these linear approx-
imations then generate the first-order perturbation corrections,[1] which
will also be discussed in section III below.

The analytic solutions to this linear model problem are the Airy
functions, which may also be considered as special cases of Bessel
functions. Accurate and efficient algorithms for evaluating the two
Airy functions and their first derivatives, were developed,[1] using
generalized Gaussian quadrature techniques.

An obvious improvement in accuracy would be gained by using piece-
wise quadratic approximations to the potential. Indeed, it might seem
at first sight that such an approach might rely on the vast amount of
study already given to the harmonic oscillator equation. Unfortunately,
one needs the less-studied solutions at non-eigenvalue energies, and
two linearly independent solutions. Mathematically, these functions were
first defined and studied by Weber; however, accurate and efficient
algorithms for their numerical evaluation were not available for all
ranges of the two arguments "a" and "x". ("a" is a parameter related to
the energy, and "x" is related to the distance variable.) Power series
and asymptotic series are valid and useful in certain ranges of a and x,
but large areas of the a,x plane could not be treated with previous
methods. Extensions of the Guassian quadrature method,[1] which was so

successful for the Airy functions, did not succeed in covering all the parameter values, either.[4] A successful algorithm for the complete useful ranges of a and x has been achieved by a combination of uniform asymptotic series for large $|a|$, and recursion relations for smaller values of $|a|$, when needed.[5] Care is needed in choosing the suitable directions for the recursion relations. For some parameter values, recursions in the complex energy plane are required! At present, the program produces function and derivative values accurate to about seven digits. Higher accuracy could be obtained, if necessary, at the cost of longer running times.

III. First-order solutions and accuracy

The zeroth-order solutions discussed above solve exactly model problems involving piecewise polynomial potentials. First-order perturbation theory is then used to evaluate the (small) effect of the deviation between the actual potential and the model potential. There are two reasons for evaluating this first-order perturbation correction: (1) to improve the accuracy of the zeroth-order results, and (2) to decide how large the intervals may be chosen, in the definition of the model potential.

The first-order perturbation corrections can be evaluated analytically, after the perturbing potential (the difference between the true potential and the polynomial potential) has been approximated as a higher-order polynomial. The formulas for the perturbation integrals involve only the zero-th order functions and derivatives at the end points of the intervals. These integrals were originally derived for the constant and linear cases,[1] and the formulas were later simplified to about half as many terms.[6] We have recently found that the perturbation integrals can also be evaluated analytically for the piecewise quadratic potentials.[7]

IV. Choice of Method

    With the development of constant, linear and quadratic potential
models, and zeroth or first order solutions based on these models, one
has a wide range of possible methods. The higher order methods offer
higher accuracy and/or smaller numbers of intervals, but with a higher
calculation cost per interval. The choice of an optimum method
depends both on the problem being solved and on the accuracy required
in the solution.

    For a problem involving a single equation to solve (single
scattering channel, elastic scattering) or a small number of channels
(e.g., $N \lesssim 10$), there is a clear trend. The higher the accuracy
required, the higher the order of the optimum method. This is indicated
qualitatively in the figures. At equal numbers of steps, the higher
order methods are generally more accurate. When the comparison is made
at equal computation time, however, lower order methods are more accurate
when larger errors are satisfactory, but the higher order methods become
more efficient when higher accuracy is demanded.

    For problems involving many channels, (e.g., $N \gtrsim 20$), the compu-
tation time becomes dominated by matrix operations, such as multiplications
or inversions. These matrix operations require on the order of N^3
arithmetic operations, while the special function evaluations only are
called 2N times per interval. Thus for these large problems with many
channels, it is always advantageous to use a higher-order method to
propagate the solutions within an interval.

V. Equations with both first and second derivatives

The piecewise analytic methods available in the past have been limited to equations, such as those for non-reactive scattering, in which first derivatives are absent. Many formulations of reactive scattering theory,[8] however, introduce first-derivative terms as well. We have recently shown how the piecewise analytic methods can be adapted to these equations, as well.[9] The coefficient of the first derivative term may be a linear function of the independent variable, and the potential may be a linear or quadratic function in each interval. The basic solutions for this case turn out to be Weber functions, which can now be evaluated accurately and efficiently.[5]

References

1. R.G. Gordon, J. Chem. Phys. __51__, 14 (1969).

2. R.G. Gordon, in Methods in Computational Physics, edited by
    Manuel Rotenberg (Academic Press, New York, 1971) Vol. 10,
    pp. 81-109.

3. "Coupled Channel Scattering Matrices," Program #187, Quantum
    Chemistry Program Exchange, Indiana University, Bloomington, Indiana.

4. R.G. Gordon, unpublished results, 1970; R.W. Redding, J. Chem.
    Phys. __60__, 1392 (1974).

5. Z. Schulten, R.G. Gordon and D.G.M. Anderson, A Numerical Algorithm
    for the Weber Parabolic Cylinder Functions $U(a,x)$, $V(a,x)$ and $W(a,\pm x)$,
    preprint attached.

6. A. Rosenthal and R.G. Gordon, J. Chem. Phys. __64__, 1621 (1976).

7. R. Shirts, Thesis, Harvard University, 1979, Ch. 3, Appendix A,
    attached; Paul Hunt and R.G. Gordon, to be published.

8. G. Miller and J.C. Light, J. Chem. Phys. __54__, 1635 (1971).

9. R.B. Shirts, R.G. Gordon and S. Bienstock, "Piecewise Analytic
    Solutions of Second-Order Linear Boundary Value Problems Using
    Weber Parabolic Cylinder Functions," preprint attached.
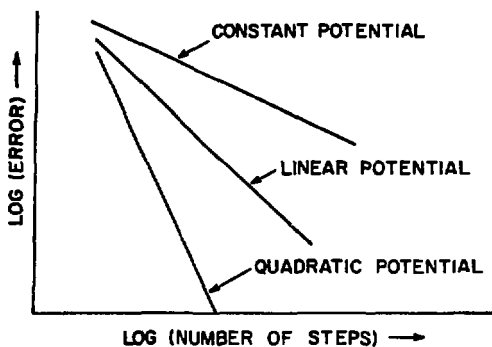
Fig. 1. Integration error vs. number of intervals used (qualitative), for various zeroth-order potentials.
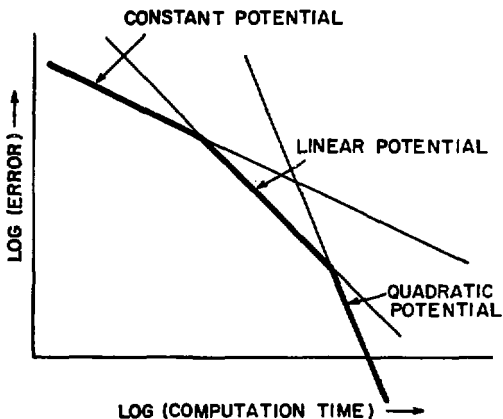
Fig. 2. Integration error vs. computation time (qualitative), for various zeroth-order potentials.

A NUMERICAL ALGORITHM FOR THE EVALUATION OF WEBER PARABOLIC
CYLINDER FUNCTIONS U(a,x), V(a,x), AND W(a,±x)

Z. Schulten*
Max-Planck-Institut für biophysikalische Chemie
D-3400 Göttingen, Fed. Rep. Germany

R. G. Gordon
Department of Chemistry, Harvard University,
Cambridge, Massachusetts 02138

D. G. M. Anderson
Committee Applied Mathematics, Harvard University,
Cambridge, Massachusetts 02138

Abstract

Weber's Parabolic Cylinder Functions $U(a,x)$, $V(a,x)$, and
$W(a,\pm x)$ have recently found wide application as approximations
to quantum mechanical wavefunction   propagating through
potential wells or barriers. Available algorithms for
their numerical evaluation are inapplicable in some ranges
of the two arguments. In this paper we present a new
algorithm, based on the combined use of Olver's [13]
uniform asymptotic expansions and Whittaker's [23] complex
recursion relations, to extend their range of usefulness.
The algorithm generates greater than single precision
values of the functions and their derivatives over the
whole range of arguments.

1. Introduction

As approximations to wavefunctions in quantum mechanical calculations, Weber's parabolic cylinder function have received considerable attention. They are used in WKB-type problems involving two or more transition or turning points [2,3,7-9,13,21]. Our interest in the evaluation of these functions grew out of our work on piecewise analytical solutions for the Schroedinger equation [5], whereby we approximate the potential locally by a polynomial function. Then we use the continuity conditions to form the complete wavefunction as a composite of the local wavefunctions. Previously, we had been able to use only piecewise linear polynomial approximations, giving rise to a basis of Airy functions. A more accurate approximation to the potential can be formed by a quadratic polynomial. The resulting wavefunction is a solution of Weber's complex linear second order differential equation [22],

$$\frac{d^2}{dz^2} D_\nu(z) + (\nu + \frac{1}{2} - \frac{1}{4}z^2)D_\nu(z) = 0. \qquad (1.1)$$

$D_\nu(z)$ is Whittaker's notation for parabolic cylinder functions, and its value is determined upon specifying a point in the two dimensional complex space (z,ν). $D_\nu(z)$ is an entire function of both variables. Throughout the text ν and z will denote complex variables while their real counterparts will be denoted   a , a real parameter, and x a real independent variable. It is clear that for special

values of the variables (x,a), (1.1) can be transformed into either the equation for the generalized harmonic oscillator functions, which we write as $U(a,x)$ and $V(a,x)$, or the equation for propagation through a potential barrier, with a set of solutions $W(a,\pm x)$.

To allow the greatest flexibility on using (1.1) or specifically its two distinct real forms as approximations to more complicated differential equations, we must be able to evaluate these functions for arbitrary values of ν and z (a and x). The number of numerical studies on the Weber functions is voluminous [4,10,11,14-17]. While there exist asymptotic formulas for large magnitudes of the parameter ν and/or the spatial variable z, and power series for small magnitudes of ν and z, there are still ranges for which, heretofore, no accurate or convenient means of evaluation existed. Extrapolation from a table of values [4,11] is both inefficient and inconvenient. Employing an algorithm developed by Gordon, integral representations for the $U(a,x)$ and $V(a,x)$ have been evaluated by Gaussian quadrature for small values of the parameter a [18-20]. Attempts to extend the variable range by recurrence relations, however, leads to instabilities in the recessive functions.

The most recent thorough analysis of the asymptotic behaviour ($|\nu|$ large) of the parabolic cylinder functions, and in particular $U(a,x)$, $V(a,x)$, $W(a,\pm x)$, is in a series of papers by Olver [13-15]. While his asymptotic representations are

valid only for large $|\overset{.}{\nu}|$, they have the advantage of being uniform in the spatial variable z. Consequently we were left only with the problem of devising a convenient method to evaluate these functions and their derivatives when the parameters are in the moderate range.

In this paper we present an algorithm for the computation of the parabolic cylinder functions $U(a,x)$, $V(a,x)$, $W(a,\pm x)$ and their derivatives for arbitrary values of the variables. Those regions of the (x,a) plane, previously inaccessible by accurate and efficient computational techniques, are covered by a set of complex recurrence relations first derived by Whittaker. For large $|a|$, the uniform asymptotic formulas of Olver are employed directly, and for moderate $|a|$ they are employed to generate starting values for the recurrence relations at some large initial index. In section 2 we define our choice of standard functions for the algorithms. In section 3 we examine the recurrence relations and specify the special discrete paths of the complex recurrence index in the ν-plane. In section 4 we outline the numerical evaluation of $U(a,x)$ and $V(a,x)$, and in section 5 $W(a,\pm x)$.

2. Choice of Standard Functions

For most physical problems only the two real standard forms
of (1.1) are of importance. The first, obtained by setting
$v + \frac{1}{2} = -a$ and $z = x$, is the generalized harmonic oscillator
equation

$$\frac{d^2 D_{-a-\frac{1}{2}}(x)}{dx^2} - (\frac{1}{4} x^2 + a)D_{-a-\frac{1}{2}}(x) = 0. \qquad (2.1)$$

Following Miller's criteria [10] for standard solutions we take
as the two linearly independent solutions $U(a,x)$ and $V(a,x)$

$$U(a,x) = D_{-a-\frac{1}{2}}(x) \qquad (2.2)$$

$$V(a,x) = \frac{1}{\pi} \Gamma(\frac{1}{2} + a)\{D_{-a-\frac{1}{2}}(x) \sin a\pi + D_{-a-\frac{1}{2}}(-x)\}. \qquad (2.3)$$

When a is not an half-integer, $U(a,\pm x)$ are an alternate pair
of solutions.

The second standard form, obtained by setting $v + \frac{1}{2} = -ia$ and
$z = xe^{\frac{-\pi i}{4}}$, describes the propagation through or over a
parabolic potential barrier

$$\frac{d^2 D_{-ia-\frac{1}{2}}(xe^{\frac{-\pi i}{4}})}{dx^2} + (\frac{1}{4} x^2 - a)D_{-ia-\frac{1}{2}}(xe^{\frac{-\pi i}{4}}) = 0. \qquad (2.4)$$

Two real solutions to (2.4), $W(a, \pm x)$, can be defined in terms of a complex function $E(a,x)$, which with its complex conjugate forms another pair of linearly independent solutions:

$$E(a,x) = \sqrt{2} \; e^{\frac{\pi a}{4} + \frac{i}{2} (\frac{\pi}{4} + \phi)} D_{-ia - \frac{1}{2}} (xe^{\frac{-\pi i}{4}}) \qquad (2.5)$$

$$x \geq 0$$

$$= k^{-\frac{1}{2}} W(a,x) + ik^{\frac{1}{2}} W(a,-x) \qquad (2.6)$$

$$E(a,-x) = k^{-\frac{1}{2}} W(a,-x) + ik^{\frac{1}{2}} W(a,+x) \qquad x \leq 0 \qquad (2.7)$$

where $\phi = \arg\Gamma(\frac{1}{2} + ia)$ and $k = \sqrt{1 + e^{2\pi a}} - e^{\pi a}$.

For a few convenient values of the parameters a, we have plotted in Fig. 1 the functions $U(a, \pm x)$, $V(a,x)$, and $W(a, \pm x)$ with respect to the x variable. We make reference to these graphs in order to emphasize the characteristic behavior and disparate nature of the standard functions. To describe the propagation through a potential barrier, the set of functions most closely satisfying Miller's criteria are $k^{-\frac{1}{2}} W(a,x)$ and $k^{\frac{1}{2}} W(a,-x)$.

3. Recurrence Relations

The following recurrence formulas for the parabolic cylinder
function $D_\nu(z)$,

$$D_{\nu+1}(z) - zD_\nu(z) + \nu D_{\nu-1}(z) = 0, \tag{3.1}$$

$$D_\nu'(z) + \frac{z}{2} D_\nu(z) - \nu D_{\nu-1}(z) = 0, \tag{3.2}$$

$$D_\nu'(z) - \frac{z}{2} D_\nu(z) + D_{\nu+1}(z) = 0, \tag{3.3}$$

were derived by Whittaker [23] from the contour integral
representation

$$D_\nu(z) = -\frac{\Gamma(\nu+1)}{2\pi i} e^{-\frac{1}{4} z^2} \int_\infty^{0+} e^{-zt - \frac{1}{2}t^2} (-t)^{-\nu-1} dt$$

$$-\pi < \arg(-t) < \pi, \; \mathrm{Re}(\nu) < 0. \tag{3.4}$$

Integration by parts of equation (3.4) yields (3.1).
Differentiating formally under the integral we obtain (3.2).
We solve for $D_{\nu-1}(z)$ from equation (3.1) and substitute the
expression into (3.2) to obtain the last relation (3.3).
Unless otherwise stated, the prime notation will denote
differentiation with respect to the independent variable
x or z, which is indicated in the argument of the function.
These relations are valid for all complex values of ν and z.

When both variables are real, two sets of relations can be
derived from equations (3.1) - (3.3). In terms of the real
valued functions $U(a,x)$ and $V(a,x)$ the recursion formulas are
as follows:

$$U(a-1,x) - xU(a,x) - (a + \frac{1}{2})U(a+1,x) = 0, \qquad (3.5)$$

$$U'(a,x) + \frac{1}{2} xU(a,x) + (a + \frac{1}{2})U(a+1,x) = 0, \qquad (3.6)$$

$$U'(a,x) - \frac{1}{2} xU(a,x) + U(a-1,x) = 0, \qquad (3.7)$$

and

$$V(a+1,x) - xV(a,x) - (a - \frac{1}{2})V(a-1,x) = 0, \qquad (3.8)$$

$$V'(a,x) - \frac{1}{2} xV(a,x) - (a - \frac{1}{2})V(a-1,x) = 0, \qquad (3.9)$$

$$V'(a,x) + \frac{1}{2} xV(a,x) - V(a+1,x) = 0. \qquad (3.10)$$

The stable direction to use the recurrence relations can be
determined from a graph of $U(a,x)$, $V(a,x)$ vs. a for constant
values x, as in Fig. 2, or by an analysis of the asymptotic
form of (3.5) and (3.8). In general, to provide a balance
among the terms in the relations, the following inequalities
must hold:

$$|U(a-1,x)| > |U(a+1,x)|$$
$$\text{as } a \to +\infty, \; x > 0.$$
$$|V(a+1,x)| > |V(a-1,x)|$$

For a positive, $U(a,x)$ decays exponentially, and $V(a,x)$ grows exponentially. For a negative, they are oscillatory functions whose moduli are either strictly increasing or decreasing functions of a. Since the stable recurrence process is in the direction of increasing function values, the relations should be used in a backward (decreasing a) direction to evaluate $U(a,x)$ and a forward (increasing a) direction to evaluate $V(a,x)$.

Unfortunately no analogous set of real recurrence relations is known to exist for either set of independent solutions $k^{\mp\frac{1}{2}} W(a,\pm x)$ or $E(a,x)$, $E^{*}(a,x)$. However, we observed that since the recurrence relations are valid for complex values of v and z, Equations (3.1) - (3.3) could be used to recur on $D_{-iA-\frac{1}{2}}$ $(xe^{-\frac{\pi i}{4}})$ where A is a complex parameter, $A = a \pm iN$, $N = 0, 1, \ldots N_{max}$. The raising and lowering of the index is now in unit intervals along a line parallel to the imaginary axis in the A-plane. Upon recurring to the real axis, i.e. $A = a$, Eqs. (2.5) and (2.6) are used to recover $E(a,x)$ and consequently $k^{\mp 1/2} W(a,\pm x)$.

In practice we find it more convenient to express the recurrence relations in terms of the complex function

$$
\begin{aligned}
\tilde{E}(A,x) &= \sqrt{2}\, e^{\frac{\pi A}{4}}\, e^{\frac{i}{2}\, [\phi(R_e A) + \frac{\pi}{4}]}\, D_{-iA-1/2}(xe^{-\frac{\pi i}{4}}) \\
&= e^{\frac{i}{2}[\phi(R_e A) - \phi(A)]}\, E(A,x)
\end{aligned}
\tag{3.11}
$$

where $\phi(A) = \arg\Gamma(\frac{1}{2} + iA)$, and $E(A,x)$ is the analytical continuation of the function $E(a,x)$ defined in Eq. (2.5). With this substitution, Eqs. (3.1) - (3.3) now have the simple form

$$\tilde{E}(A+i,x) - x\tilde{E}(A,x) - i(\frac{1}{2} + iA)\tilde{E}(A-i,x) = 0, \qquad (3.12)$$

$$\tilde{E}'(A,x) - \frac{ix}{2}\tilde{E}(A,x) + (\frac{1}{2} + iA)\tilde{E}(A-i,x) = 0, \qquad (3.13)$$

$$\tilde{E}'(A,x) + \frac{ix}{2}\tilde{E}(A,x) - i\tilde{E}(A+i,x) = 0. \qquad (3.14)$$

When A is real ($A = a$), $\tilde{E}(a,x)$ reduces to the function $E(a,x)$.

For any value of x, the modulus of $\tilde{E}(A,x)$ decreases as the imaginary component of A is made more negative. One should then evaluate $\tilde{E}(A,x)$ and $\tilde{E}'(A,x)$ at $A = a - iN_{max}(a,x)$, for some large integer N_{max}, which is in general dependent on a and x, and use the recurrence relations in the forward direction to raise the index N_{max} times to determine $E(a,x)$ and $E'(a,x)$.

Returning to the original recurrence formulas in terms of the general parabolic cylinder function $D_\nu(z)$, the recurrence indices in the above relations for $U(a,x)$, $V(a,x)$ and $E(A,x)$ can be represented naturally as particular paths in the ν-plane. Defining the initial conditions for $U(a,x)$ and $V(a,x)$ as $\nu_\pm = -a_\pm - \frac{1}{2}$, the various numerically stable paths for the recurrence process are indicated in Fig. 3.

4. Evaluation of $U(a,x)$ and $V(a,x)$

The combined usage of Olver's asymptotic formulas and the
recurrence relations constitutes a convenient and accurate
algorithm for the evaluation of the generalized Harmonic
oscillator functions $U(a,x)$, $V(a,x)$ and their derivatives
for arbitrary real values of a and x. Performing the
calculations in double precision arithmetic on an IBM (360)
and an Univac (1800), values of the functions and their
derivatives are obtained to fourteen significant digits.
The choice of the method depends solely on the value of a,
which we have divided into two complementary domains:

Asymptotic Region	Non-asymptotic Region
$a > 11$ or $a < -41$	$-41 \leq a \leq 11$

When a falls within the asymptotic region, Olver's formulas
are employed directly. When the functions are needed for
non-asymptotic values of a, the asymptotic formulas
provide the starting values for the recurrence relations
at the initial indices $a_{\pm}$. These have been experimentally
determined:

$$a_{+} = a + N \geq 11 \quad U(a,x)$$

$$a_{-} = a - M \leq -41 \quad V(a,x).$$

In the case of $U(a,x)$, the function and its derivative,
both evaluated at a_+, are first used to obtain $U(a_+-1,x)$
from Eq. (3.7). The index is then lowered N-1 times to a
employing Eq. (3.5) in accordance with the stable direction
for recurring discussed earlier. The symmetry of the parabolic
cylinder functions with respect to x (see Fig. 1) makes
it necessary to develop an algorithm only for x in the
right-half plane $(x \geq 0)$. The computational algorithm is
summarized in Fig. 4.

For completeness we will now specify which of Olver's
asymptotic series are employed in the algorithm and comment
on any difficulties that arose in their evaluation. Since
we used more terms than Olver originally presented, we have
recorded the necessary expansion coefficients in the
Appendix. The asymptotic representations are derived from
an analysis of the normal equation,

$$\frac{d^2w}{dt^2}(\mu,t) = \mu^4(t^2 - 1)\ w(\mu,t) \qquad (4.1)$$

where μ and t are complex, and $|\mu| \gg 0$.

For a positive, (4.1) is brought into the real standard
differential equation for $U(a,x)$ by the following trans-
formations of the dependent and independent variables:
set $t = -iz$, $\mu = i\eta$ and $y(\eta,z) = w(i\eta,-iz)$ to obtain

$$\frac{d^2y}{dz^2}(\eta,z) - \eta^4(z^2 + 1) \ y(\eta,z) = 0, \qquad (4.2)$$

and $x = z\eta\sqrt{2}$, $a = \frac{1}{2}\eta^2$, $y(a,x) = y(\sqrt{2a}, \frac{x}{2\sqrt{a}})$ to obtain

$$\frac{d^2y}{dx^2}(a,x) - (\frac{1}{4}x^2 + a) \ y(a,x) = 0. \qquad (4.3)$$

A satisfactory pair of solutions for (4.3) when a is positive is $U(a,x)$ and $U(a,-x)$. To preserve Olver's notation, the asymptotic formulas are more conveniently written as functions of η and z, and in the following section, we will denote $U(a,x)$ equivalently by $U(\frac{1}{2}\eta^2, z\eta\sqrt{2})$.

Equation (4.3) exhibits no transition point characteristics for real, positive a, and the functions can be expressed in terms of elementary functions [13]:

$$U(a,x) = U(\frac{1}{2}\eta^2, \eta z\sqrt{2}) \sim \frac{\bar{g}(\eta) e^{-\eta^2 \bar{\xi}(z)}}{(z^2+1)^{\frac{1}{4}}} \sum_{s=0}^{\infty} \frac{\bar{u}_s(z)}{(z^2+1)^{\frac{3s}{2}}} \cdot \frac{1}{\eta^{2s}} \qquad (4.4)$$
$$a > 0$$

$$U'(a,x) = U'(\frac{1}{2}\eta^2, \eta z\sqrt{2}) \sim \frac{-\eta\bar{g}(\eta)}{\sqrt{2}} (z^2+1)^{\frac{1}{4}} e^{-\eta^2 \bar{\xi}(z)} \sum_{s=0}^{\infty} \frac{\bar{v}_s(z)}{(z^2+1)^{\frac{3s}{2}}} \cdot \frac{1}{\eta^{2s}}$$
$$(4.5)$$

The auxiliary function $\bar{g}(\eta)$ is calculated from the asymptotic expansion

$$\frac{1}{g(\eta)} \sim 2^{+\frac{1}{4}\eta^2} + \frac{1}{4} e^{+\frac{1}{4}\eta^2} \eta^{-\frac{1}{2}\eta^2} + \frac{1}{2}\left(1 + \sum_{j=0}^{\infty} \frac{g_{2j+1}}{\eta^{4j+2}}\right), \qquad (4.6)$$

where $\bar{g}(\eta) = e^{\pi i\left(\frac{1}{4} + \frac{1}{4}\eta^2\right)} g(i\eta)$, and $\bar{\xi}(z)$ is given by

$$\bar{\xi}(z) = \frac{1}{2} z(z^2 + 1)^{\frac{1}{2}} + \frac{1}{2} \ln[z + (z^2 + 1)^{\frac{1}{2}}]. \qquad (4.7)$$

The functions $\bar{u}_s(z)$ and $\bar{v}_s(z)$ are defined in terms of the polynomial functions $u_s(z)$ and $v_s(z)$

$$\bar{u}_s(z) = i^s u_s(-iz)$$

$$\bar{v}_s(z) = i^s v_s(-iz) \qquad (4.8)$$

It is a tedious but straightforward exercise to determine the coefficients $u_s(z)$ and $v_s(z)$ from a set of recurrence relations. We record the first seven terms along with the constants g_s in Table II in the Appendix. The branches of the multi-valued functions are well defined upon specifiying $\arg\mu = \frac{\pi}{2}$ ($\arg\eta = 0$) and $z\varepsilon\ e^{\frac{\pi i}{2}}\ S(\frac{\pi}{2})$, the domain shown in Fig. 5a.

In our algorithm the factor $\bar{g}(\eta)$ is omitted thereby altering the normalization, but eliminating an unnecessary calculation if only relative values of the parabolic cylinder functions are needed at different values of x for fixed a. When x < 0, equation (4.4) is a valid asymptotic representation

for the linearly independent solution $U(a,-x)$. The corres-
ponding asymptotic formula for $V(a,x)$ can be straight-
forwardly derived by substitution of the expressions for
$U(a,x)$ and $U(a,-x)$ into (2.3).

When a is negative, the change of variables $t = z$, and
$\mu = \eta$, followed by $\eta^2 = -2a$ and $x = \eta z \sqrt{2}$, again transforms
(4.1) into (4.3). However, equation (4.3) now possesses
two real transition points at $x_{T.P.} = \pm 2\sqrt{|a|}$. Olver's asymptotic
expansions for $U(a,x)$ and $U'(a,x)$, which are uniform for all x
to the right of the left hand transition point, are in
terms of Airy functions, $Ai(\eta^{\frac{4}{3}}\zeta)$ and $Bi(\eta^{\frac{4}{3}}\zeta)$ [13]:

$$U(a,x) = U(\frac{-1}{2}\eta^2, \eta z\sqrt{2}) \sim 2\pi^{\frac{1}{2}} \eta^{\frac{1}{3}} g(\eta)\Phi(\zeta) \left\{ Ai(\eta^{\frac{4}{3}}\zeta) \sum_{s=0}^{\infty} \frac{A_s(\zeta)}{\eta^{4s}} \right.$$

$$+ \frac{Ai'(\eta^{\frac{4}{3}}\zeta)}{\eta^{\frac{8}{3}}} \sum_{s=0}^{\infty} \frac{B_s(\zeta)}{\eta^{4s}} \left. \right\} , \qquad a < 0 \qquad (4.9)$$

$$U'(a,x) = U'(\frac{-1}{2}\eta^2, \eta z\sqrt{2}) \sim \frac{(2\pi)^{\frac{1}{2}} \eta^{\frac{2}{3}} g(\eta)}{\Phi(\zeta)} \left\{ \frac{Ai(\eta^{\frac{4}{3}}\zeta)}{\eta^{\frac{4}{3}}} \sum_{s=0}^{\infty} \frac{C_s(\zeta)}{\eta^{4s}} \right.$$

$$+ Ai'(\eta^{\frac{4}{3}}\zeta) \sum_{s=0}^{\infty} \frac{D_s(\zeta)}{\eta^{4s}} \left. \right\} , \qquad (4.10)$$

where $\Phi(\zeta) = (\frac{\zeta}{z^2-1})^{\frac{1}{4}}$. The variable ζ is given by

$$\frac{2}{3}\,\zeta^{\frac{3}{2}} = \frac{1}{2}\,z\,(z^2 - 1)^{\frac{1}{2}} - \frac{1}{2}\,\ln(z + (z^2 - 1)^{\frac{1}{2}}) \text{ for } z > 1$$

$$\text{(4.11)}$$

$$\text{and } \frac{2}{3}(-\zeta)^{\frac{3}{2}} = \frac{1}{2}\,\arctan(z(z^2-1)^{\frac{1}{2}}) - \frac{1}{2}\,z(z^2-1)^{\frac{1}{2}} \text{ for } 0 \leq z < 1$$

with $\zeta = 0$ at the turning point $z = 1$. The coefficient functions $A_s(\zeta)$, $B_s(\zeta)$, $C_s(\zeta)$, and $D_s(\zeta)$ are determined from the following series:

$$A_s(\zeta) = \sum_{m=0}^{2s} b_m \zeta^{-3m/2} \mathscr{A}_{2s-m}(z) \qquad \zeta^{1/2} B_s(\zeta) = -\sum_{m=0}^{2s+1} a_m \zeta^{-3m/2} \mathscr{A}_{2s-m+1}(z)$$

$$\text{(4.12)}$$

$$\zeta^{-1/2} C_s(\zeta) = -\sum b_m \zeta^{-3m/2} \mathscr{B}_{2s-m+1}(z) \qquad D_s(\zeta) = \sum_{m=0}^{2s} a_m \zeta^{-3m/2} \mathscr{B}_{2s-m}(z)$$

where $\mathscr{A}_0(z) = 1$, $\mathscr{A}_s(z) = \dfrac{u_s(z)}{(z^2-1)^{3s/2}}$, $\mathscr{B}_s(z) = \dfrac{v_s(z)}{(z^2-1)^{3s/2}}$

and $a_m = \dfrac{(2m+1)(2m+3)\ldots(6m-1)}{m!\,(144)^m}$, $b_m = -\dfrac{6m+1}{6m-1}\,a_m$, with $a_0 = 1$.

The analogous expansion for $V(a,x)$ can be obtained, after some manipulation, from the connection formulas for parabolic cylinder functions. We will record here only the results.

$$V(a,x) = V(-\frac{1}{2}\,\eta^2, \eta z\sqrt{2}) \sim \frac{2\pi^{\frac{1}{2}}\,\eta^{\frac{1}{3}}\,g(\eta)}{\Gamma(\frac{1}{2} - a)}\,\phi(\zeta)\left\{ \text{Bi}(\eta^{\frac{4}{3}}\zeta) \sum_{s=0}^{\infty} \frac{A_s(\zeta)}{\eta^{4s}} \right.$$

$$\left. + \frac{\text{Bi}'(\eta^{\frac{4}{3}}\zeta)}{\eta^{\frac{8}{3}}} \sum_{s=0}^{\infty} \frac{B_s(\zeta)}{\eta^{4s}} \right\}. \qquad \text{(4.13)}$$

$$V'(a,x) = V'(\frac{-1}{2}\eta^2, \eta z\sqrt{2}) \sim \frac{(2\pi)^{\frac{1}{2}}\eta^{\frac{2}{3}}g(\eta)}{\Gamma(\frac{1}{2}-a)\Phi(\zeta)} \left\{ \frac{Bi(\eta^{\frac{4}{3}}\zeta)}{\eta^{\frac{4}{3}}} \sum_{s=0}^{\infty} \frac{C_s(\zeta)}{\eta^{4s}} \right.$$

$$\left. + Bi'(\eta^{\frac{4}{3}}\zeta) \sum_{s=0} \frac{D_s(\zeta)}{\eta^{4s}} \right\}. \qquad (4.14)$$

The multi-valued functions are well c ined and the above expressions valid for $z \in T(0)$. $T(0)$ is the unshaded region in Fig. 5b.

As x approaches the right hand transition point $2\sqrt{|a|}$ or as $z \to 1$ the factor $(\frac{\zeta}{z^2-1})^{1/4} \equiv \Phi(\zeta)$ remains well defined and has a finite value at $z = 1$. However, in a computational sense, the limit process is ill-conditioned since it is dependent upon including more and more terms in the series and the cancellation which must exist between them. Numerical experiments revealed a preferred direction in passing through the transition region. The asymptotic formulas maintain their accuracy longer when the transition point is approached from the left. Consequently for z within a neighborhood of 1,

$$z^+ = 1 + \delta_1 > 1 > 1 - \delta_2 = z^- \quad \text{where } \delta_1, \delta_2 > 0, \delta_1 > \delta_2, \quad (4.15)$$

the asymptotic series is evaluated at z^- and then contunued through the transition point region from z^- by a Taylor series expansion, including seventh-order derivatives.

5. Evaluation of W(a,x) and E(a,x)

As with the U(a,x) and V(a,x) functions, one would like
to use Olver's uniform asymptotic series to evaluate
$k^{\mp 1/2}$ W(a,±x) or E(a,x) for large magnitudes of a and devise
a recursion scheme, which is dependent on a only, x being
treated as a parameter, to cover the complementary region
of non-asymptotic values of a:

$$a < -40 \text{ or } a > 20 \qquad \text{Asymptotic Region}$$
$$\tag{5.1}$$
$$-40 \leq a \leq 20 \qquad \text{Complementary Region}$$

We have already outlined such a scheme in section 3 involving
the complex recurrence relations for E(A,x) or $D_{-iA-1/2}(xe^{-\pi i/4})$.
For asymptotic values of A, Olver developed series for
E(A,x) or $D_{-iA-1/2}(xe^{-\pi i/4})$ which are uniformly valid with
respect to the x variable. For a in the complementary region,
these complex asymptotic series generate values of $\tilde{E}(A,x)$
and $\tilde{E}'(A,x)$ at $A = a - iN_{max}$ which are subsequently employed
in the recurrence relation (3.14) to obtain $\tilde{E}(A+1,x)$. E(a,x)
is evaluated by raising the index N-1 times using Eq. (3.12)
and E'(a,x) is determined at the end from Eq. (3.13).

The Olver expansions for E(A,x) and W(a,±x) are also derived
from the normal equation

$$\frac{d^2 w}{dt^2} (\mu,t) = \mu^4 (t^2 - 1) \ w(\mu,t) \tag{5.2}$$

in which μ and t are complex variables. The following change of variables transforms (5.2) into the desired form:

set $t = -iz$, $\eta = e^{\frac{-i\pi}{4}} \mu$, and $y(\eta,z) = w(\eta e^{\frac{-\pi i}{4}},-iz)$ to obtain

$$\frac{d^2y}{dz^2} (\eta,z) = - \eta^4(z^2 + 1) \, y(\eta,z) \tag{5.3}$$

and $x = \eta z\sqrt{2}$, $A = -\frac{1}{2} \eta^2$, $y(A,x) = y(\sqrt{-2A}, \frac{x}{2\sqrt{-A}})$ to obtain

$$\frac{d^2y}{dz^2} (A,x) = - (\frac{1}{4} x^2 - A) \, y(A,x). \tag{5.4}$$

In analogy to Eq. (2.5) we can express $E(A,x)$ in terms of the principal solution of Eq. (5.4)

$$E(A,x) = \sqrt{2} \, e^{\frac{\pi A}{4} + \frac{i}{2} (\phi + \frac{\pi}{4})} U(iA,xe^{\frac{-\pi i}{4}}), \tag{5.5}$$

where we have made use of Whittaker's notation

$$U(A,z) = D_{-A-1/2}(z) . \tag{5.6}$$

When A takes on complex values, (5.4) exhibits no real transition point characteristics, and the asymptotic series for $E(A,x)$ can be expressed in terms of elementary functions. The expansions when $Re(A)$ is negative are [13]

$$U\left(\frac{-i\eta^2}{2}, e^{\frac{-\pi i}{4}} z\eta\sqrt{2}\right) \sim \frac{g(\eta e^{\frac{\pi i}{4}})}{(z^2+1)^{\frac{1}{4}} e^{\frac{-\pi i}{4}}} e^{\eta^2(i\bar{\xi}(z)+\pi/4)} \sum_{s=0}^{\infty} (-1)^s \frac{u_s(-iz)}{(z^2+1)^{\frac{3s}{2}}} \cdot \frac{1}{\eta^{2s}}$$

and

$$\text{(5.7)}$$

$$U'\left(\frac{-i\eta^2}{2}, e^{\frac{-\pi i}{4}} \eta z\sqrt{2}\right) \sim -\frac{\eta}{2} g(\eta e^{\frac{\pi i}{4}}) (z^2+1)^{\frac{1}{4}} e^{\eta^2(i\bar{\xi}(z)+\frac{\pi}{4})} e^{\frac{-\pi i}{4}}$$

$$\cdot \sum_{s=0}^{\infty} \frac{(-1)^s v_s(-iz)}{(z^2+1)^{\frac{3s}{2}}} \cdot \frac{1}{\eta^{2s}} \qquad (5.8)$$

The functions $g(\eta)$, $\bar{\xi}(z)$ are as defined in the previous section, and the polynomial coefficients $u_s(t)$, $v_s(t)$ for $s \leq 7$ are given in the Appendix. The region of validity of the above expressions is $z \in e^{\frac{\pi i}{2}} S(\arg\mu)$, where $\arg\mu$ varies with N as follows:

$$0 \leq N < \infty$$

$$0 \leq \arg\eta < \frac{\pi}{4} \quad \text{where} \quad -\frac{1}{2} \eta^2 = a - iN, \ a < 0. \quad (5.9)$$

$$\frac{\pi}{4} \leq \arg\mu < \frac{\pi}{2}$$

The domain $S(\arg\mu)$, with the appropriate branch cuts, is shown in Fig. 6a.

When $\text{Re}(A) \geq 0$, $\text{Im}(A) \neq 0$, the corresponding transformation of variables which lead to asymptotic series for $U(iA, e^{-i\pi/4} x)$ or $E(A,x)$ is

$$\eta = e^{\frac{\pi i}{4}} \mu, \quad z = t$$

$$x = \eta z\sqrt{2}, \quad A = \frac{1}{2} \eta^2 \tag{5.10}$$

The series expansions in terms of these redefined variables are [13]

$$U(\frac{i\eta^2}{2}, e^{\frac{-\pi i}{4}} \eta z\sqrt{2}) \sim \frac{g(\eta e^{\frac{-\pi i}{4}})}{(z^2-1)^{\frac{1}{4}}} e^{i\eta^2\xi} \sum_{s=0}^{\infty} i^s \frac{u_s(t)}{(z^2-1)^{\frac{3s}{2}}} \cdot \frac{1}{\eta^{2s}} \tag{5.11}$$

$$U'(\frac{i\eta^2}{2}, e^{\frac{-\pi i}{4}} \eta z\sqrt{2}) \sim \frac{-\eta}{\sqrt{2}} g(\eta e^{\frac{-\pi i}{4}}) (z^2-1)^{\frac{1}{4}} e^{i\eta^2\xi} \sum_{s=0}^{\infty} \frac{i^s v_s(t)}{(z^2-1)^{\frac{3s}{2}}} \cdot \frac{1}{\eta^{2s}} \tag{5.12}$$

where $\xi(z) = \frac{1}{2} z(z^2-1)^{1/2} - \frac{1}{2} \ln[z + (z^2-1)^{1/2}]$.

The range of the arguments of η and μ as a function of N is

$$0 \leq N < \infty$$

$$0 \geq \arg\eta > \frac{-\pi}{4} \quad \text{where } \frac{1}{2} \eta^2 = a - iN, \ a \geq 0 \tag{5.13}$$

$$\frac{-\pi}{4} \geq \arg\mu > \frac{-\pi}{2}$$

The branch cuts that define the multi-valued function $\xi(z)$ and form the boundaries of the domain $S(\arg\mu)$ are labeled in Fig. 6b.

Evaluation of the complex function $E(A,x)$ for subsequent use in the recurrence relations with a complex index is only

necessary when the energy parameter a lies within the complementary region defined in (5.1). When a < -40, the functions $k^{\mp\frac{1}{2}} W(a,\pm x)$ can be evaluated directly from the components of $E(a,x)$ (A real) using (2.6), (5.5) and (5.7). When a lies in the positive asymptotic range, a ≥ 20, $k^{\mp\frac{1}{2}} W(a,\pm x)$ and their derivatives are computed by Olver's real uniform asymptotic series in terms of Airy functions:

$$W(a,x) = W(\tfrac{1}{2}\eta^2, \eta z\sqrt{2}) \sim \frac{\pi^{\frac{1}{2}} \eta^{\frac{1}{3}} \ell(\eta)}{2^{\frac{1}{2}} \cdot \frac{\pi\eta}{2^{\frac{1}{4}}}} \left(\frac{\zeta}{z^2-1}\right)^{\frac{1}{4}} \left\{ Bi(-\eta^{\frac{4}{3}}\zeta) \sum_{s=0}^{\infty} (-)^s \frac{A_s(\zeta)}{\eta^{4s}} \right.$$

$$\left. + \frac{Bi'(-\eta^{\frac{4}{3}}\zeta)}{\eta^{\frac{8}{3}}} \sum_{s=0}^{\infty} (-)^s \frac{B_s(\zeta)}{\eta^{4s}} \right\} , \qquad (5.14)$$

$$W'(a,x) = W'(\tfrac{\eta^2}{2}, \eta z\sqrt{2}) \sim \frac{\pi^{\frac{1}{2}} \eta^{\frac{2}{3}} \ell(\eta)}{2 e^{\frac{\pi\eta}{4}}} \left(\frac{z^2-1}{\zeta}\right)^{\frac{1}{4}} \left\{ \frac{Bi(-\eta^{\frac{4}{3}}\zeta)}{\eta^{\frac{4}{3}}} \sum_{s=0}^{\infty} (-)^s \frac{C_s(\zeta)}{\eta^{4s}} \right.$$

$$\left. - Bi'(-\eta^{\frac{4}{3}}\zeta) \sum_{s=0}^{\infty} (-)^s \frac{D_s(\zeta)}{\eta^{4s}} \right\} , \qquad (5.15)$$

$$W(a,-x) = W(\tfrac{\eta^2}{2}, -\eta z\sqrt{2}) \sim \frac{\pi^{\frac{1}{2}} \eta^{\frac{1}{3}} \ell(\eta)}{2^{-\frac{1}{2}} e^{-\frac{\pi\eta}{4}}} \left(\frac{\zeta}{z^2-1}\right)^{\frac{1}{4}} \left\{ Ai(-\eta^{\frac{4}{3}}\zeta) \sum_{s=0}^{\infty} (-)^s \frac{A_s(\zeta)}{\eta^{4s}} \right.$$

$$\left. + \frac{Ai'(-\eta^{\frac{4}{3}}\zeta)}{\eta^{\frac{8}{3}}} \sum_{s=0}^{\infty} (-)^s \frac{B_s(\zeta)}{\eta^{4s}} \right\} , \qquad (5.16)$$

$$W'(a,-x)=W'(\frac{\eta^2}{2},-\eta z\sqrt{2})\omega\frac{\pi^{\frac{1}{2}}\,n^{\frac{2}{3}}\,\ell(\eta)}{e^{\frac{-\pi\eta^2}{4}}}(\frac{z^2-1}{\zeta})^{\frac{1}{4}}\left\{\frac{\text{Ai}(-n^{\frac{4}{3}}\,\zeta)}{n^{\frac{4}{3}}}\sum_{s=0}^{\infty}(-)^s\frac{C_s(\zeta)}{n^{4s}}\right.$$

$$\left.-\text{Ai}'(-n^{\frac{4}{3}}\,\zeta)\sum_{s=0}^{\infty}(-)^s\frac{D_s(\zeta)}{n^{4s}}\right\}\tag{5.17}$$

where $W'(a,-x) = \dfrac{dW(a,-x)}{dx}$ and $\ell(\eta) \neq 2^{\frac{1}{2}}\,e^{\frac{\pi\eta^2}{8}}\,e^{i(\frac{\phi(a)}{2}-\frac{\pi}{8})}\,g(\eta e^{\frac{-\pi i}{4}})$

with $\phi(a) = \arg\Gamma(\frac{1}{2} + ia)$.

The expressions are valid for z in the region $T(\frac{-\pi}{4})$, where $\arg\mu = \frac{-\pi}{4}$ and $\arg\eta = 0$ (Fig. 6c). The expansion coefficients $A_s(\zeta)$, $B_s(\zeta)$, $C_s(\zeta)$, $D_s(\zeta)$ are evaluated by recurrence relations recorded in the previous section.

By truncating the series (5.14) - (5.17) at a finite value of s, the partial cancellation between consecutive terms which allows the asymptotic representations to be evaluated even at the right transition point, z = 1, cannot occur as $x \rightarrow x_{TP} = 2\sqrt{a}$. Furthermore, as the turning point is approached, evaluation of the series for A_s, B_s, C_s, and D_s (Eq. (4.12)) becomes ill-conditioned. Near x_{TP} the terms in the series e.g. $b_m\,\zeta^{-3m/2}\mathscr{B}(z)_{2s-m+1}$, are large with alternating signs, while the series sums are much smaller than the individual terms. The round-off error is circumvented by employing a Taylor series approximation within the turning region for a > 0 as was done in the computation of U(a,x) and V(a,x) in section 4. Use, however, of the

Taylor series produces function values at x_{TP} which are the least accurate in the entire algorithm (8 vs. 11 significant digits), and clearly a table of values for $A_s(0)$, $B_s(0)$, $C_s(0)$, and $D_s(0)$ would improve the evaluation.

The asymptotic series for $W(a,\pm x)$ and $W'(a,\pm x)$ (Eqs. (5.14) - (5.17)) have variable accuracy over the asymptotic region, as was determined from evaluating the Wronskian and spot-checking against expansions for large and small x found in the NBS handbook [1]. The accuracy of calculations performed with double-precision arithmetic on an IBM 360 and an UNIVAC 1108, expressed as significant digits, is indicated in Fig. 7 for the 3 term ($s \leqslant 2$) and 4 term ($s \leqslant 3$) series. The accuracy increases with increasing x since $\lim_{x \to \infty} A_s \to 0$ for $s > 1$ and $\lim_{x \to \infty} B_s \to 0$. At the origin, the coefficients A_s, B_s, C_s, D_s are of order 1, e.g. $A_s(t=0) = \sum_{m=0}^{s} b_{2m} \zeta_{(0)}^{-3m} \mathscr{A}(0)_{2s-2m}$, and the asymptotic series, truncated at the $s = j$th term, are correct to $O[\frac{1}{(2a)^{2j+2}}]$.

In the complementary region (5.1), the algorithm using $s = 6$ in Eqs. (5.7), (5.8), (5.11) and (5.12), e.g. series include terms up to u_6 and v_6, and $j = 2$ in the series for $g(\eta)^{-1}$, Eq. (4.6), generates greater than single precision values of $E(a,x)$ or equivalently $k^{\mp 1/2} W(a,\pm x)$, i.e. 11 - 14 significant digits. For $a > 0$ and $x < x_{TP}$, numerical difficulties exist in determining the imaginary part of $E(a,x)$. The difficulties arise from the disparity in size

between the real and imaginary components of $E(a,x)$ (see Fig. 1e),

$$\frac{\text{Re } E(a,x)}{\text{Im } E(a,x)} = \frac{k^{-\frac{1}{2}} W(a,x)}{k^{\frac{1}{2}} W(a,-x)} \sim 2e^{\pi a} \quad a > 0 \quad 0 \leq x \leq x_{TP}.$$

In the last step of the recurrence process, two numbers with the same order of magnitude as Re $E(a,x)$ are substracted to determine Im $E(a,x)$. If the calculations are performed in double precision, the use of the complex recurrence relations must be restricted to a < 3.5 in order to guarantee single precision values of Im $E(a,x)$.

There are a number of ways to circumvent this problem. One is to recur on $E(A,-x)$ instead of $E(A,x)$ for small positive x. A more efficient solution is to employ the complex recurrence relations in the exponential region, $0 \leq x \leq 2\sqrt{a}$, to obtain just the dominant real component $k^{-1/2} W(a,x)$ and $k^{-1/2} W'(a,x)$. The ratio $\frac{W'(a,-x)}{W(a,-x)} = y$ can be evaluated from Miller's [11] non-linear differentiation equation for the derivative log function,

$$\frac{dy}{dx} + y^2 + 1/4 \ x^2 - a = 0$$

where $\quad y_o = \frac{W'(a,o)}{W(a,o)} = -2^{1/2} \left| \frac{\Gamma(\frac{3}{4} + \frac{1}{2} ia)}{\Gamma(\frac{1}{4} + \frac{1}{2} ia)} \right| ,$

using fourth-order Runge-Kutta integration with stepsizes $\leq$ 0.005. From the Wronskian relation

$$W(a,x) \frac{W'(a,-x)}{W(a,-x)} - W'(a,x) = \frac{1}{W(a,-x)}$$

one can solve for $W(a,-x)$ and subsequently $W'(a,-x)$ in a numerically stable fashion.

The computational methods to evaluate at least single-precision values (11 - 14 significant digits) of the parabolic cylinder functions $k^{\mp 1/2} W(a,\pm x)$ and their derivatives are summarized in Fig. 8. Within the exponential region $0 \le x \le 2\sqrt{a}$, for $1 < a < 20$ only the functions $k^{-1/2} W(a,x)$ and $k^{-1/2} W'(a,x)$ are obtained from the complex recurrence relations for $E(A,x)$. The imaginary components $k^{1/2} W(a,-x)$ and $k^{+1/2} W'(a,-x)$ are disgarded and determined instead from Miller's derivative-log method discussed above. The number of steps needed for the recursion relations as a function of a and x is recorded in Table I. For small $|a|$ and large x, i.e. $x^2 \gg 4a$, a region in which $k^{\mp 1/2} W(a,\pm x)$ are oscillating functions, the values have been checked against the Miller's modulus-phase expressions [11,1]:

$$k^{-1/2} W(a,x) + ik^{1/2} W(a,-x) = Fe^{ix}$$

$$k^{-1/2} W'(a,x) + ik^{1/2} W'(a,-x) = -Ge^{i\psi}.$$

For $0 < |a| \le 4$, $x > 200$, Olver's exponential expansions, Eqs. (5.7 - 8) and Eqs. (5.11 - 12) with $N \le 2$ agree with Miller's expressions to at least 11 significant digits.

Such an agreement is surprising considering that Olver's representations are supposing valid for asymptotic values of $|a|$.

REFERENCES

1. M. Abramowitz, and I.A. Stegun (editors), <u>National Bureau of Standards Handbook of Mathematical Functions</u>, U.S. Government Printing Office, Washington, D.C., 1970.

2. R.J. Bienieck, J. Phys. B: Atom. Molec. Phys. <u>17</u>, L266-L271 (1974).

3. M.S. Child, Molec. Phys. <u>28</u>, 495-501 (1974).

4. L. Fox, "Table of Weber Functions for Large Arguments", National Physical Laboratory Math. Tables No. 4, Her Majesty's Stationery Office, London, England, 1960.

5. R.G. Gordon "Quantum Scattering Using Piecewise Analytic Solutions", <u>Methods in Computational Physics</u>, Vol. 10, pp. 81-109, Academic Press Inc., New York, :971.

6. R.G. Gordon, J. Chem. Phys. <u>51</u>, 14-25 (1969).

7. J. Heading, J. Phys. A: Math., Nucl. Gen. <u>6</u>, 958-973 (1973).

8. W. Hecht, J. Math. Phys. <u>14</u>, 1519-1521 (1973).

9. W.P. Latham and R.W. Redding, J. Comp. Phys. <u>16</u>, 66 (1974).

10. J.C.P. Miller, Proc. Camb. Philos. Soc. <u>48</u>, 428-435 (1952).

11. J.C.P. Miller, "Tables of Weber Parabolic Cylinder Functions", National Physical Laboratory, Her Majesty's Stationery Office, London, England, 1955.

12. S.C. Miller and R.H. Good, Jr., Phys. Rev. <u>91</u>, 174-179 (1973).

13. F.W.J. Olver, J. Res. NBS (U.S.A.) <u>63B</u>, 131-169 (1959).

14. F.W.J. Olver, Phil. Trans. Roy. Soc. London Ser. <u>A 278</u>, 137-185 (1975).

15. F.W.J. Olver, "Asymptotics and Special Functions", Academic Press, New York, 1974.

16. M.K. Kerimov, <u>Certain New Results of the Theory of Weber Functions</u>, Translation 71-86, Bell Laboratories, 1968 and references therein.

17. Antra Priede and Gabriel Allen, "Fortran Program for Computation of Weber Functions and First Derivatives", NASA TN D-3833 (Clearinghouse for Scientific and Technical Information, Springfield, Virginia, 1968).

18. R.W. Redding, J. Chem. Phys. <u>60</u>, 649-655 (1974).

19. R.W. Redding, J. Chem. Phys. <u>60</u>, 1392 (1974).

20. R.W. Redding and W.P. Latham, J. Comp. Phys. <u>20</u>, 256 (1976).

21. M.J. Richardson, Phys. Rev. A. <u>8</u>, 781-787 (1973).

22. F. Weber, Math. Annal. <u>1</u>, 1-36 (1969).

23. E.T. Whittaker, Proc. London Math. Soc. <u>35</u>, 417-427 (1903).
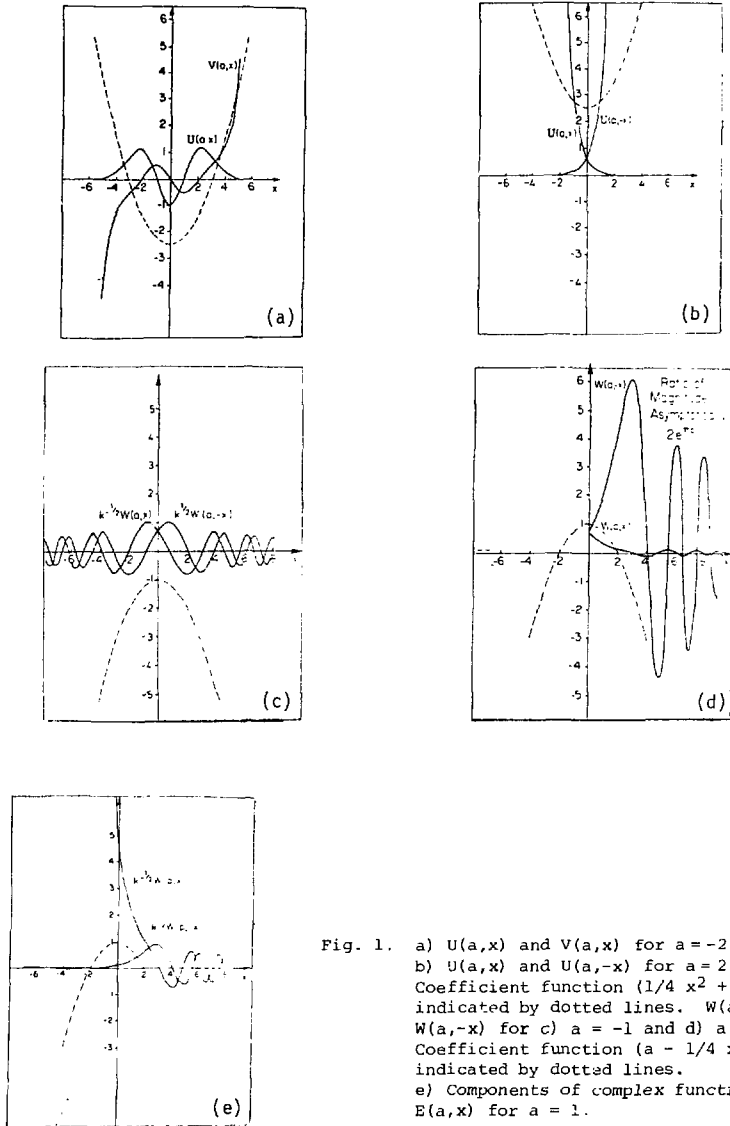
Fig. 1. a) $U(a,x)$ and $V(a,x)$ for $a = -2.5$ and
b) $U(a,x)$ and $U(a,-x)$ for $a = 2.5$.
Coefficient function $(1/4\ x^2 + a)$ is
indicated by dotted lines. $W(a,x)$ and
$W(a,-x)$ for c) $a = -1$ and d) $a = 1$.
Coefficient function $(a - 1/4\ x^2)$ is
indicated by dotted lines.
e) Components of complex function
$E(a,x)$ for $a = 1$.
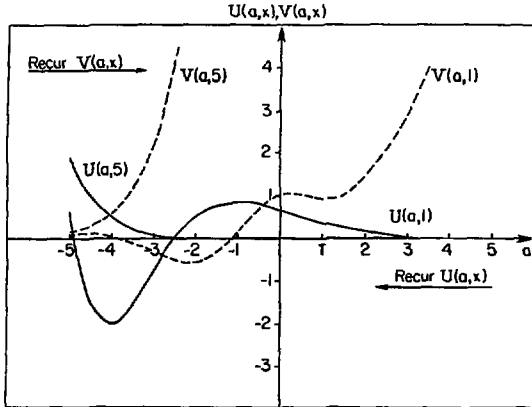
Fig. 2. U(a,x) and V(a,x) as function of a.
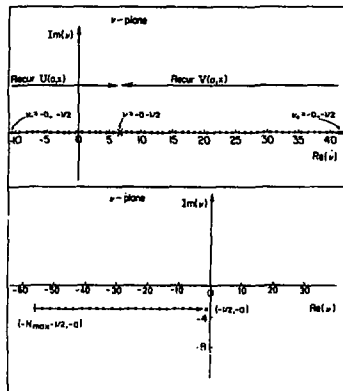
Fig. 3. Typical paths for recurrence index in evaluation of a) U(a,x),
V(a,x) where $\nu = -a - \frac{1}{2}$ for $-41 < a < 11$ and b) E(a,x) where
$\nu = -iA - \frac{1}{2} = -i(a-iN) - \frac{1}{2}$ for $-40 < a < 20$.
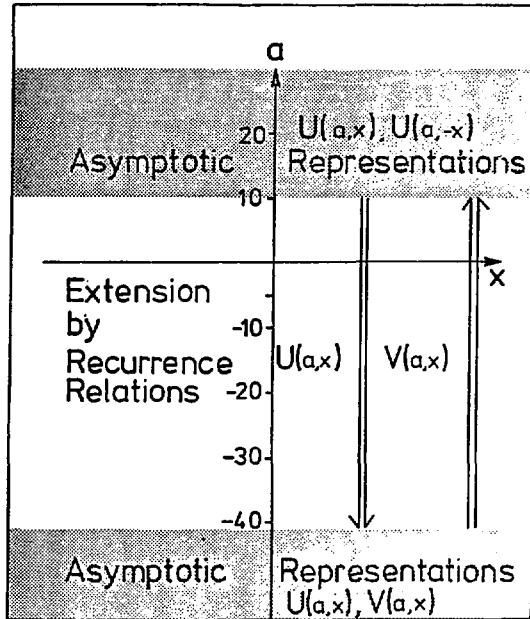
Fig. 4.   Computational methods to evaluate U(a,x) and V(a,x) and their derivatives.
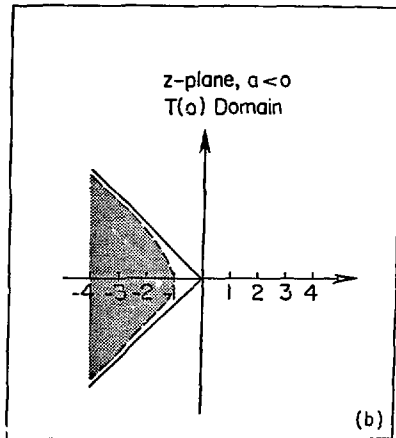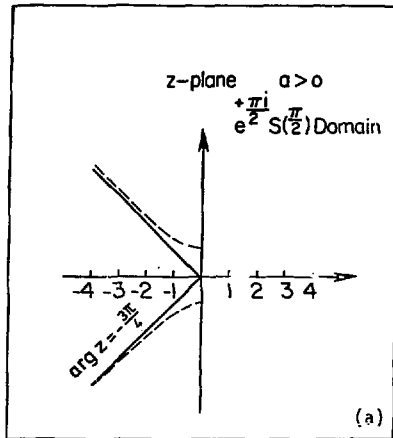
Fig. 5. Domain of asymptotic expansions for a) $U(\frac{1}{2}\eta^2, \eta z\sqrt{2})$ where $\frac{1}{2}\eta^2 = a$ and b) $U(-\frac{1}{2}\eta^2, \eta z\sqrt{2})$ and $V(-\frac{1}{2}\eta^2, \eta z\sqrt{2})$ where $\frac{1}{2}\eta^2 = -a$.
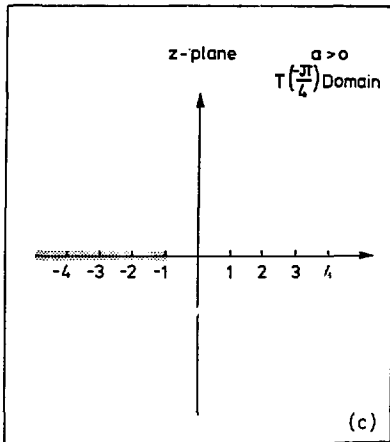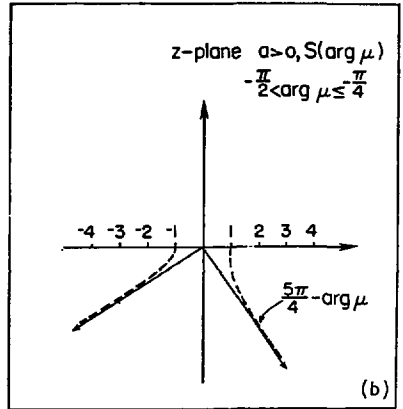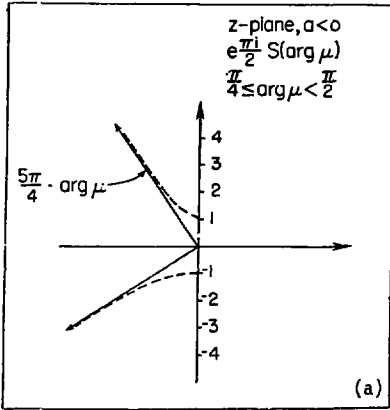
Fig. 6. Domain of asymptotic expansions for a) $U(-i\eta^2,$ $\exp(-i\pi/4)\ z\eta\sqrt{2})$ and b) $U(i\eta^2,\ \exp(-i\pi/4)\ z\eta\sqrt{2})$. c) Domain of asymptotic expansions for $W(1/2\ \eta^2,$ $\pm\eta z\sqrt{2})$ where $1/2\ \eta^2 = a$.
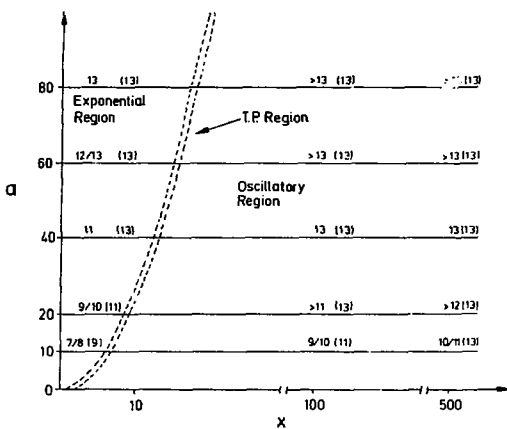
Fig. 7. Accuracy (number of significant digits) of 3 (4) term asymptotic
series for $W(a,\pm x)$, $a > 0$, Eqs. (5.14) and (5.15).
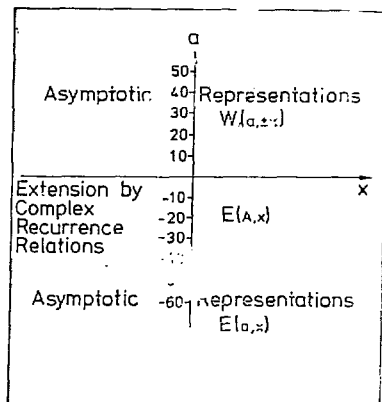
Fig. 8. Computational methods to evaluate $k^{\pm 1/2} W(a,\pm x)$ and their deriva-
tives.

Table I

Recurrence Index N for $E(A, x)^a$, $A = a - i \, N$

a	x	N
$-40 < a < 20$	$0 \leq x \leq 20$	60
	$20 \leq x \leq 100$	10
$-40 < a < 4$	$100 < x \leq 200$	10
$-2 \leq a \leq 2$	$200 < x \leq 500$	4

[a] $E(A,x)$ is evaluated from Eq. (5.7) for $a < o$ and Eq. (5.11) for $a > o$, using s=6. $g(\eta)$ is determined from Eq. (4.6) with $j = 2$.

TABLE II

EXPANSION COEFFICIENTS FOR THE POLYNOMIAL FUNCTIONS US AND VS

S	K	US(K)	VS(K)
1	1	-.25000000000000000000+00	.25000000000000000000+00
	3	.41666666666666666666-01	.41666666666666666666-01
2	0	.12586805555555555556+00	-.12413194444444444400
	2	.21614583333333333333+00	-.28385416666666666667+00
	4	-.78124999999999999999-02	.13020833333333333333-01
3	1	-.62521701388888888890+00	.62521701388888888890+00
	3	-.36650028935185185853+00	.57490596064814814814+00
	5	-.68207465277777777761-01	-.87738715277777777759-01
	7	.43858506944444444435-01	.43858506944444444435-01
	9	-.97463348765432096644-02	-.97463348765432096644-02
4	0	.32082672747074331200+00	-.30439028641814557840+00
	2	.30794612931616512400+01	-.33343841929445988040+01
	4	.12794328544077932400+01	-.14436563217351466100+01
	6	-.38927368666640946820-02	-.38167820055258634950-02
	8	-.80711835696373445520-02	.13858069902584876200-01
	10	.18274377809351851810-01	-.30453290648919753020-02
5	1	-.50846283398537969440+01	.50846283739853796540+01
	3	-.17876599593337921400+02	.19573475616022527720+02
	5	-.55871591324668853000+01	.52646639725759532240+01
	7	.66285492562322604100+00	.57298206743296112400+00
	9	-.54273463000088799600+00	-.56678057817073755340+00
	11	.29826426071534784000+00	.30283285551872746940+00
	13	-.92386621783833120150-01	-.92386621763303120950-01
	15	.12315482904510827800-01	.12315482904510827800-01
6	0	.25819438650657535540+01	-.25026844747804341000+01
	2	.62685920892021549970+02	-.64673700517678340200+02
	4	.12171794602206850000+03	-.12970634337197194100+03
	6	.25849627983363680900+02	-.26175555642156022540+02
	8	-.31866840674067167400+00	.43098635711732221300+00
	10	.86014735294561076500-01	-.16025349184236465800+00
	12	-.50765538086973749610-01	.90371705133164107300-01
	14	.17130390289065369900-01	-.29063920602130204300-01
	16	-.23097155445957802100-02	.36495259076593360900-02
7	1	-.86572189672073719800+02	.86572189672073720000+02
	3	-.74619316747727207800+03	.77505056403462999000+03
	5	-.93585512027292787700+02	.96412286561702598200+03
	7	-.14684264579958295540+03	.17372863592713045140+03
	9	-.10763440603293235010+02	-.13249296507746644400+02
	11	.12951388793046040465+02	.13102236447268131610+02
	13	-.99843945489400516940+01	-.10074544460722579770+02
	15	.53274823635874539040+01	.53676136711746055700+01
	17	-.18638925777686603540+01	-.18696666606201520540+01
	19	.39714080613590644570+00	.39714080613590644570+00
	21	-.37621933917705567300-01	-.37622933917705567300-01

Table III

Expansion Coefficients for the Function $g(\eta)^{-1}$

s	g_{2s+1}
0	0.416666666666666666-01
1	- 0.974633487654320964-02
2	0.123184829045108278-01
3	- 0.378229339177055673-01

APPENDIX

EVALUATION OF COEFFICIENTS FOR ASYMPTOTIC REPRESENTATIONS

OF WEBER'S PARABOLIC CYLINDER FUNCTIONS

The coefficient functions u_s and v_s in the asymptotic series

representation of $U(a,x)$ for $a > 0$ and $E(A,x)$ for complex A

are either totally even or odd polynomials of maximum degree 3s.

We shall denote the argument of u_s and v_s generally by t

and evaluate them by a set of recurrence formulas [13].

The recurrence formula for $u_s(t)$ given in (4.8) is

$$(t^2-1)u_s'(t) = 3stu_s(t) = r_{s-1}(t) \qquad (A.1)$$

where

$$8r_s(t) = (3t^2+2)u_s(t) - 12(s+1)tr_{s-1}(t) + 4(t^2-1)r_{s-1}'(t)$$

and $u_o(t) = 1$.

The functions $r_s(t)$ are determined first from equation (A.2)

and then substituted into (A.1). Since the coefficient

functions $u_s(t)$ are polynomials of degree 3s, they can be

derived from equation (A.1) by matching powers of t. For

s even, the coefficients of the leading term t^{3s} in $u_s(t)$

are zero.

Once $u_s(t)$ and $r_s(t)$ are known, the coefficient function

$v_s(t)$ is evaluated from the relation

$$v_s(t) = u_s(t) + \frac{1}{2} tu_{s-1}(t) - r_{s-2}(t) \qquad (A.3)$$

where $v_o(t) = 1$.

Olver provided the functions $u_s(t)$ and $v_s(t)$ for $s = 0,1,2,3$.

In order to increase the accuracy of the asymptotic series for moderate values of a, at least two more terms should be included. In Table II we have recorded the coefficients for the polynomials for $s = 1,2 \ldots 7$, obtained from our program solving (A.1) – (A.3).

The constants g_s appearing in (4.6) are defined to be

$$g_s = \lim_{|t| \to \infty} \frac{u_s(t)}{(t^2-1)^{3/2 \ s}} \ .$$

Hence $g_{2s} = 0$ and g_{2s+1} is the coefficient of the leading power t^{6s+3} in $U_{2s+1}(t)$ and can be obtained from Table II.

PIECEWISE ANALYTIC SOLUTION OF SECOND ORDER
LINEAR BOUNDARY VALUE PROBLEMS
USING WEBER PARABOLIC CYLINDER FUNCTIONS[*]

Randall B. Shirts[+]
Roy G. Gordon
and
Sergio Bienstock

Department of Chemistry
Harvard University
Cambridge, Mass. 02138

Abstract

We present a method for obtaining accurate piecewise analytic
solutions of a general second order linear differential equation in-
cluding both first and second derivatives and a small inhomogeneous
term. The method approximates the coefficient functions by piecewise
linear functions and transforms the equation into Weber's equation in
each sub-interval. Algorithms for evaluation of Weber parabolic
cylinder functions are then employed to give analytic expressions for
the solution in each sub-interval. Error analysis and application to
a test problem with known solution show the error to be of fourth order
in the sub-interval size and of at least sixth order when a first order
perturbation term is included.

I.   Introduction

Many problems in applied mathematics can be reduced to solving the second-order linear boundary value problem

(1) $$y'' + f(x)y' + g(x)y = h(x) \quad \text{in} \quad (a,b),$$

with general mixed boundary conditions

(2) $$\sum_{j=1}^{2} (\alpha_{ij} y^{(j-1)}(a) + \beta_{ij} y^{(j-1)}(b)) = \gamma_i \quad i = 1,2.$$

The problem (1) is most commonly solved by the method of finite differences, the solution being obtained at a relatively large number of mesh points in (a,b).

Another method is to approximate the coefficient functions f, g, and h by suitable approximants in a series of N sub-intervals given by the partition $\pi = \{a = x_1 < x_2 < \ldots < x_N = b\}$ where the sub-intervals are (x_n, x_{n+1}), $n = 1,2, \ldots .N$. Pruess[1] has shown that if $|\pi|$ is the maximum sub-interval size, and f, g, and h are approximated by a m-th order polynomial by interpolating at the roots of the (m + 1)st degree Legendre polynomial transformed to (x_n, x_{n+1}), then the error in y is of order $|\pi|^{2m+2}$ for $|\pi|$ sufficiently small. For m = 0, the basis solutions are exponential and trigonometric functions. For m > 0, the basis solutions are special functions which are less readily evaluated. In this work, we show that piecewise analytical solution of (1) for the case of m = 1 can be readily accomplished and does, in fact, give errors of fourth order. We also demonstrate that corrections can easily be added to the zeroth order solution which yield errors of at least sixth order.

Several authors have treated problems less general than (1).
The most commonly studied case is the radial Schroedinger equation in
which $f(x) = h(x) = 0$. Gordon[2,3] first investigated this case and
subsequently obtained accurate solutions for $m = 1$ in which Airy
functions are the basis solutions. Other authors have developed the
theory for perturbation corrections to the zeroth order solution for
$m = 0$.[4-8] Luthey[9] has developed the computational methods for the
$m = 2$ case, including algorithms for evaluation of the Weber para-
bolic cylinder functions. In this work, we will transform (1) into a
problem closely related to the $m = 2$ case of the radial Schroedinger
equation and then use Weber parabolic cylinder functions as the basis
solutions.

We emphasize the piecewise analytic character of the solutions
we obtain. After solution, one has an analytic representation of y
for the entire interval which is continuous and has a continuous
first derivative. This analytic character has important advantages
over a finite difference solution in which one has only numerical
values of y at a series of mesh points. The solution has a closed
form expression in each sub-interval which may be differentiated, inte-
grated, or evaluated at arbitrary points within the sub-interval. In
addition, because of the high order of the errors, one can use con-
siderably fewer sub-intervals than in other numerical methods. In
such a case, the approximate solution is completely specified by a
small number of parameters.

In Section II we will formally present the method of solution.
Section III will consist of a brief error analysis. Section IV will
apply the method to a problem with a known solution. Section V will
discuss the advantages of the method and summarize our conclusions.

II.   Formal Presentation of the Method

Given the partition π of (a,b) into sub-intervals of arbitrary, not necessarily uniform length, we interpolate the functions f and g (assumed to be real valued) in each sub-interval by linear interpolants $\hat{f}$ and $\hat{g}$ at the zeros of the second degree Legendre polynomial transformed to that sub-interval. This procedure is equivalent to linear least squares approximation up to fourth order in the interval size. The difference functions ε_1 and ε_2 are defined in each sub-interval by:

(3) $$f \equiv \hat{f} + \varepsilon_1, \text{ and } g \equiv \hat{g} + \varepsilon_2, \quad i = 1, 2, \ldots N$$

The canonical method for analytic solution of equations with $f \neq 0$ is to transform the equation into normal form in which the first derivative term is absent. Since we approximate the coefficient function f by a linear function, we transform the equation by the change of dependent variable as if $\hat{f}$ were exactly equal to f, i.e,

(4) $$y = u \exp(- \frac{1}{2} \int \hat{f} dx)$$

after which (1) can be rewritten in the form

(5) $$u'' - u(\frac{\hat{f}^2}{4} + \frac{\hat{f}'}{2} - \hat{g}) = h \exp(\frac{1}{2} \int \hat{f} \ dx) - \lambda u' \varepsilon_1 + \lambda u(\frac{\hat{f}}{2} \varepsilon_1 - \varepsilon_2)$$

where we have added the factor λ on the right hand side as a perturbation parameter which will be taken to be unity. Since $\hat{f}$ is linear, the coefficient of u on the left hand side of (5) is quadratic. A linear change of independent variable can be used to transform (5) to the first standard form of Weber's equation:

(6) $$\frac{d^2 u}{dz^2} - u(\frac{z^2}{4} + a) = \tilde{h}(z) - \lambda u' \varepsilon_1 r^{-1} + \lambda u r^{-2}(\frac{\hat{f}}{2} \varepsilon_1 - \varepsilon_2)$$

where $z = r(x - \frac{1}{2}(x_i + x_{i+1})) + s$, $\tilde{h}(z(x)) = hr^{-2}\exp(\frac{1}{2}\int \hat{f}\,dx')$, and r, s, and a are determined by the coefficients of x^2, x^1, and x^0 on the left hand side of (5).

We can now expand u in perturbation series in λ:

$$(7) \qquad u = {}^{(0)}u + \lambda{}^{(1)}u + \lambda^2\,{}^{(2)}u + \ldots$$

For f and g sufficiently smooth, such series will always converge in some radius of convergence in λ which we assume to be greater than unity. Substituting (7) into (6) and separating in equations in each power of λ gives the following hierarchy of equations:

$$(8) \qquad {}^{(0)}u'' - {}^{(0)}u(\frac{z^2}{4} + a) = \tilde{h}(z),$$

$$ {}^{(i)}u'' - {}^{(i)}u(\frac{z^2}{4} + a) = F_i(z), \quad i = 1, 2 \ldots$$

where $F_i(z) = {}^{(i-1)}ur^{-2}(\frac{1}{2}\hat{f}\epsilon_1 - \epsilon_2) - {}^{(i-1)}u'r^{-1}\epsilon_1$. Standard variation of constants or Green's function techniques yield the following expressions for the solutions:

$$(9) \qquad {}^{(0)}u(z) = U(a,z)[A_0 - \frac{1}{w}\int^z \tilde{h}(z')V(a,z')dz'] +$$

$$ V(a,z)[B_0 + \frac{1}{w}\int^z \tilde{h}(z')U(a,z')dz']$$

$$ {}^{(i)}u(z) = U(a,z)[A_i - \frac{1}{w}\int^z dz'V(a,z')F_i(z')] +$$

$$ V(a,z)[B_i + \frac{1}{w}\int^z dz'U(a,z')F_i(z')],$$

$$ i = 1, 2 \ldots$$

where U and V are Weber parabolic cylinder functions and w is their Wronskian. The integrals in which $\tilde{h}$ appears are evaluated by two point Gauss-Legendre quadrature, which is equivalent to approximating $\tilde{h}$ by a

linear function at the transformed zeros of the second degree Legendre polynomial. The integrals in the higher order solutions can be evaluated by higher order quadratures, or it is possible to evaluate them analytically in terms of the integrals $\int ABz^n dz$ (A,B = U, V, U', or V') for which indefinite integrals can be evaluated.[10]

It is necessary to link up the solutions in each sub-interval in such a way that the global solution and its first derivative are continuous and the boundary condition (2) is satisfied. This is done by adjusting the constants A_i and B_i in each sub-interval. We set A_i and B_i equal to zero for $i \geqslant 1$ by fixing the lower limit of integration in (9) at x_n. The most efficient method to determine the 2N remaining constants A_0 and B_0 is to set up a 2-vector of independent solutions at a and use a shooting method to propagate each component of the vector from sub-interval to sub-interval, making each component and its derivative continuous at the mesh points. This linking requires the solution of a 2 x 2 system of linear equations at each of the N - 1 interior mesh points. When this is complete, a linear combination of the component solutions can be found to satisfy the boundary condition (2).

In some instances, it is convenient or necessary to propagate the vector of solutions from b to a rather than the reverse. This is done simply by resetting the lower limit of integrations in (9) at x_{n+1} rather than x_n.

We will investigate the accuracy of the solutions obtained by truncating the perturbation series (7) after the first and second term. The solution with one term of Eq. (7) we call the zeroth order solution. We call the second term the first order correction.

III. Error Analysis

In this discussion, we define the error to be the maximum abso-
lute deviation from the correct result among a set of points, i.e. the
supremum or L_∞ norm of these deviations. The order of error for an
approximate solution we define to be the first power in the sub-interval
size which is neglected. For the zeroth order solution, the error can
be most easily estimated from the order of the first order correction.
In each sub-interval, this correction consists of a sum of integrals
of the form:

(10)
$$\int_{x_n}^{x_{n+1}} AB\varepsilon(x)dx,$$

where A and B are either $U(a,s + r(x - \frac{1}{2}(x_n + x_{n+1})))$ or $V(a,s + r \times$
$(x - \frac{1}{2}(x_n + x_{n+1})))$ or their derivatives, and $\varepsilon(x)$ is a function which
has zeros at the zeros of the transformed second degree Legendre poly-
nomial. It is well known[11] that the error in k-point Gauss-Legendre
integration is proportional to $(x_{n+1} - x_n)^{2k+1}$, and for k = 2, the
quadrature sum for (10) vanishes exactly since the Gaussian pivots are
precisely the zeros of the integrand. Thus the integrals must be of
order at least $(x_{n+1} - x_n)^5$. Errors of the fifth order in each sub-
integral give rise to global errors of fourth order. We conclude that
the zeroth order solution in (9) has errors of fourth order. Expanding
the functions in Taylor series about the midpoint of the interval
yields precisely the same result. Pruess has shown that this fourth
order accuracy is the maximum attainable order of accuracy for linear
approximations of the coefficient functions.[1] We have demonstrated a
method to construct a piecewise analytic solution with this maximum

attainable accuracy. This analysis is only applicable to the error at the mesh points x_n. In practice we observe errors of fourth order at non-mesh points as well, although we can theoretically justify errors of only third order by our simple analysis. We observe errors of fourth order in the derivative of the solution at mesh points, but errors of second order in the derivative at non-mesh points.

The order of approximation of the first order corrected solution is estimated by examination of the second perturbation corrections. These consist of integrals of the form

(11)
$$\int_{x_n}^{x_{n+1}} A\epsilon(x)\ ^{(1)}u(x)dx.$$

The function $^{(1)}u(x)$ is of order $(x_{n+1} - x_n)^3$ for linearly inter-polated coefficient functions since it consists of integrals over an interval range of order $(x_{n+1} - x_n)$ each with an integrand of order $(x_{n+1} - x_n)^2$. Integrals of the form (11) were shown above to be of order $(x_{n+1} - x_n)^5$, but since the $^{(1)}u(x)$ multiplier is of order $(x_{n+1} - x_n)^3$, the net order of the integrals for sufficiently small sub-intervals is $(x_{n+1} - x_n)^8$. This gives rise to global errors of seventh order. In practice, we observe errors of at least sixth order in function value and function derivative at both grid points and non-grid points. We are unable to distinguish between sixth and higher order errors numerically due to lack of precision in evaluating Weber parabolic cylinder functions.

IV.   Application

We take as an example the following problem treated by Pruess[1]:

(12)
$$y'' + \frac{4x}{1+x^2} y' + \frac{2}{1+x^2} y = 0 \quad \text{in } (0,1/2),$$

$$y'(0) = 0; \; y(1/2) = 8000,$$

which has the known solution $y = 10^4/(1 + x^2)$.

We used sub-intervals of equal length for computational convenience and for comparison with the results of Ref. 1. The calculation was performed in double precision on the departmental PDP 11-45. Table 1 lists the relative error $|(y - y_\pi)/y|$ and $|(y' - y'_\pi)/y'|$ where y_π is an approximate solution and primes denote derivatives. In this calculation, first order corrections were evaluated by four point Gauss-Legendre quadrature. We note that the relative error in our calculation is limited by the precision to which our present routines evaluate the Weber parabolic cylinder functions. This precision is approximately 8.E-7 for U and V functions and 4.E-6 for their derivatives. This precision can easily be improved, but has not yet been implemented in our present program versions. To within this tolerance, our zeroth order function errors agree with those of Pruess. The important feature of Table 1 is that the error in the first order corrected solution is three orders of magnitude less than the zero order solution without subdivision of the interval (0,1/2). Even the derivative is accurate to five figures. It was necessary to extend the interval to (0,1) in order to estimate the order of errors in this case.

V.   Discussion and Conclusion

The use of piecewise analytic methods for $m \geqslant 1$ requires the evaluation of special functions which are slower than simple exponential or trigonometric functions. The zeroth order solution can be constructed in N sub-intervals with only 2N function evaluations of the Weber parabolic cylinder functions and their derivatives. The difficulty of computing special functions is offset in many cases by the higher order of error of this method, which requires significantly fewer sub-intervals.

In practice, the most efficient choice of partition is not equal length sub-intervals. The most useful criterion that we have devised is to require that the first order perturbation correction in each sub-interval be of equal magnitude relative to the zeroth order solution. This criterion allows the mesh to be adjusted for equal relative error per sub-interval at the time of propagation of the solution vector as discussed in Sec. II. This criterion has been successfully used by the authors in applications to the radial Schroedinger equation, chemical kinetics, and charge transport in semiconduct~·

It should be understood that there is nothing which prevents the approximation of g by a piecewise quadratic function. In general, this might require the use of the second standard form of Weber's equation with basis solutions $W(a, \pm x)$. Although we have algorithms[9] for evaluation of these functions, they are considerably slower than those for the functions U and V. Furthermore, such a quadratic approximation to g would not increase the order of accuracy. Such a method might be successful only when g dominates f sufficiently.

Some uneasiness sometimes arises over the fact that the piecewise linear approximants $\hat{f}$ and $\hat{g}$ are discontinuous at the mesh points x_n.

These discontinuities result only in a discontinuous second derivative
of y. Such a discontinuity is seldom important in second order dif-
ferential equations.

In some instances when the slope of f vanishes in a sub-interval,
Eq. (6) reduces to the Airy equation. The use of Airy basis solutions
does not effect the order of accuracy of the solution. If, in addition,
the slope of g vanishes in the same sub-interval, then Eq. (6) has
exponential and trigonometric solutions. These occurrences are usually
only special cases which happen rarely and do not affect the global
absolute error appreciably.

Although the analysis we have presented has been concerned only
with the univariate case, e.g. one equation (or set of uncoupled equa-
tions), the generalization to M coupled equations is straightforward.
Much of the theoretical and error analysis has been completed by Luthey[9]
and Smooke.[7] In the multivariate case, the amount of computation goes
up as M^3. Since the number of Weber or Airy function evaluations neces-
sary is proportional to NM, the fraction of computer time spent doing
function evaluations becomes negligible for large M. In such a case,
a higher order solution such as we have presented becomes much more
desirable.

We have demonstrated a method of solving approximately the second
order linear boundary value problem using piecewise analytic solutions
in terms of Weber parabolic cylinder functions. We have demonstrated
by brief error analysis and numerical example that the zeroth order
solution is of fourth order in the step size, and the first order cor-
rected solution is of at least sixth order. We believe that in many
applications the present method is more efficient than other methods
for comparable accuracy. We also feel that the analytic character of

our solution gives it advantages not possessed by fully numerical
solution methods, in that the solution can be evaluated at arbitrary
points by closed form expression and is completely specified by a
small number of parameters.

References

1. S.A. Pruess, "Solving Linear Boundary Value Problems by Approximating the Coefficients," Math. Comp., V. 27, 1973, pp. 551-561.

2. R.G. Gordon, "New Method for Constructing Wavefunctions for Bound States and Scattering," J. Chem. Phys., V. 55, 1969, pp. 14-25.

3. R.G. Gordon, "Quantum Scattering using Piecewise Analytic Solutions," in Methods of Computational Physics, Vol. 10 (Edited by B.J. Alder, S. Fernbach, and M. Rotenberg), Academic Press, New York, 1971, pp. 81-109.

4. J. Canosa and R. Gomes de Oliveira, "A New Method of the Solution of the Schrödinger Equation," J. Comput. Phys., V. 5, 1970, pp. 188-207.

5. J.P. Riehl, D.J. Diestler, and A.F. Wagner, "Comparison of Perturbation and Direct-Numerical-Integration Techniques for th Calculation of Phase Shifts for Elastic Scattering," J. Comput. Phys., V. 15, 1974, pp. 212-225.

6. G. Adam, L.G. Ixaru, and A. Corcionei, "A First-Order Perturbative Numerical Method for the Solution of the Radial Schrödinger Equation," J. Comput. Phys., V. 22, 1976, pp. 1-33.

7. M.D. Smooke, unpublished Ph.D. thesis, Harvard University, 1978.

8. S.A. Pruess, "Higher Order Approximations to Sturm-Liouville Eigenvalues," Numer. Math., V. 24, 1975, pp. 241-247.

9. Z.A. Luthey, unpublished Ph.D. thesis, Harvard University, 1974.

10. P.M. Hunt, private communication.

11. M. Abrahamowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, New York, 1970, formula 25.4.30.

Table 1

Relative error in function and derivative for piecewise analytic approximation to the known solution of Eq. (12). Numbers in parentheses are limited by the precision of Weber function evaluation.

$|\pi|$ =	1/2	1/4	1/8	1/16
Pruess (Ref. 1)	1.01E-3	4.42E-5	2.59E-6	1.60E-7
zero order function	1.01E-3	4.46E-5	(3.14E-6)	(7.,2E-7)
zero order derivative	2.02E-2	5.98E-3	1.57E-3	4.01E-4
zero order derivative				
(at mesh points only)	6.19'-5	4.76E-5	(2.43E-6)	(2.37E-6)
first order function	2.23E-6	(3.41E-7)	(5.50E-7)	(6.12E-7)
first order derivative	3.24E-5	(2.16E-6)	(2.88E-6)	(3.78E-6)

QCOL/MK2: AN ACCELERATED GORDON
ALGORITHM FOR INELASTIC COLLISIONS*

Millard H. Alexander
Department of Chemistry
University of Maryland
College Park, MD 20742

The usual quantum formulation of inelastic molecular
collisions at low energy leads to a set of coupled second- ·
order ordinary differential equations, commonly called the
close-coupled (CC) equations. In matrix notation these are

$$[\frac{d^2}{dR^2} \underset{=}{1} + \underset{=}{k}^2 - \underset{=}{V}(R)]\underset{=}{u}(R) = 0 \ , \tag{1}$$

where $\underset{=}{1}$ is the unit matrix, $\underset{=}{k}^2$ is the diagonal wavevector matrix
and $\underset{=}{V}(R)$ is the Hermitian matrix of the coupling potential plus
the centrifugal barrier. Gordon has developed[1-3] a widely-used
program for the efficient numerical solution of these equations,
which is based on propagating the solution matrix $\underset{=}{u}(R)$ outward
through a series of intervals. Within each interval this matrix
is subjected to an orthogonal rotation, $\underset{\equiv}{C}_m$, which is chosen to
diagonalize the sum of the wavevector and potential matrices,
$\underset{=}{k}^2 - \underset{=}{V}(R)$, at the midpoint of the interval, R_m.

Expanding the transformed potential matrix in a power series
about R_m, one can write the transformed CC equations as

$$[\frac{d^2}{dR^2} \underline{\underline{1}} + \underline{\underline{\lambda}} - \underline{\underline{G}}(R-R_m) - \frac{1}{2}\underline{\underline{H}}(R-R_m)^2 + \ldots] \underline{\underline{v}}_m(R) = 0 \; , \tag{2}$$

where $\underline{\underline{\lambda}}$ is the (diagonal) transform of $\underline{\underline{k}}^2 - \underline{\underline{V}}(R_m)$

$$\underline{\underline{\lambda}} = \underline{\underline{C}}_m^T [\underline{\underline{k}}^2 - \underline{\underline{V}}(R_m)] \underline{\underline{C}}_m \tag{3}$$

and $\underline{\underline{G}}$ and $\underline{\underline{H}}$ are, respectively the transforms of the first and second derivatives of $\underline{\underline{V}}(R)$, evaluated at R_m, namely

$$\underline{\underline{G}} = \underline{\underline{C}}_m^T [\frac{d}{dR} \underline{\underline{V}}(R)]_{R_m} \underline{\underline{C}}_m \tag{4}$$

and

$$\underline{\underline{H}} = \underline{\underline{C}}_m^T [\frac{d^2}{dR^2} \underline{\underline{V}}(R)]_{R_m} \underline{\underline{C}}_m \; . \tag{5}$$

By neglecting the off-diagonal elements of $\underline{\underline{G}}$, the entire $\underline{\underline{H}}$ matrix, and higher derivatives of $\underline{\underline{V}}(R)$, these equations can be uncoupled and the (diagonal) solution matrix expressed in terms of Airy functions. This is equivalent to replacing the transformed potential matrix by a diagonal, piecewise linear matrix. For the nth channel one has

$$(\underline{\underline{v}}_m)_{nn'} = \delta_{nn'} v_n^m(R) = A_n Ai[\alpha_n(R+\beta_n)]+B_n Bi[\alpha_n(R+\beta_n)] \tag{6}$$

where $\alpha_n = (G_{nn})^{1/3}$ and

$$\beta_n = -[(\lambda_n - \frac{1}{24} H_{nn}\Delta_m^2)/G_{nn} + R_m] \tag{7}$$

where Δ_m is the width of the mth interval. The term containing H_{nn} is added to provide an optimal piecewise linear potential.[2,4]

The coefficients A_n and B_n are determined by solution-matching at the boundary of the previous interval. Prior to this matching, the solutions in the previous interval must be backtransformed into the original basis and subsequently transformed into the basis which diagonalizes $\underline{\underline{k}}^2 - \underline{\underline{V}}(R_m)$. This step involves multiplication by the interval-to-interval transformation matrix, $\underline{\underline{T}}_m$, where

$$\underline{\underline{T}}_m = \underline{\underline{C}}_m \, \underline{\underline{C}}_{m-1}^T \, . \tag{8}$$

For N coupled equations propagation across a given interval will require, in the limit of large N, approximately $(6\frac{1}{6})N^3$ multiplications, reflecting the following matrix operations

i) $\frac{2}{3} N^3$ reduction of $\underline{\underline{k}}^2 - \underline{\underline{V}}(R_m)$ to tridiagonal form

ii) N^3 backtransformation of eigenvectors of tridiagonal matrix to obtain $\underline{\underline{C}}_m$

iii) N^3 formation of transformation matrix $\underline{\underline{T}}_m$

iv) $\frac{3}{2} N^3$ unitary transformation of derivative matrix [Eq. (4)]

v) $2N^3$ multiplication of solution matrix in previous interval, $\underline{\underline{y}}_{m-1}$, and its derivative by transformation matrix $\underline{\underline{T}}_m$ [Eq.(8)].

At nearby collision energies,[2] or, in the case of any "average-ℓ" decoupling approximation, at other values of the orbital angular momentum $\underline{\ell}$,[5,6] the same transformation matrices can be used. Propagation across the mth interval then necessitates only the $2N^3$ multiplications corresponding to step (v).

In established versions of the Gordon program,[3] the initial propagation across each interval is accompanied by the determina-

tion of the first-order perturbation corrections to the solution
matrix, $\underline{v}_m(R)$ arising from the neglected off-diagonal terms in
the transformed first-derivative matrix, $\underline{G}$ [Eq.(4)], and diagonal
terms in the second-derivative matrix, $\underline{H}$ [Eq.(5)]. The largest
of the diagonal and offdiagonal corrections, CDIAG and COFF,
are then used to choose the size of the next interval.[2]

The actual determination of these perturbation corrections
suffers from the following drawbacks:

i) Considerable computational effort and a significant
fraction of the total object code is devoted solely to this
somewhat minor aspect of the overall calculation. Additionally,
mathematical instabilities can arise in certain applications.[7]

ii) Although the addition of these perturbation corrections
should in principle render the "first-order" solutions more ac-
curate than the "zeroth-order" solutions, in our experience the
improvement is small, perhaps because of the approximations
which are made in the evaluation of the crucial off-diagonal
corrections.[1,2]

iii) For most problems it is desirable to carry out sub-
sequent solutions of the CC equations using the already deter-
mined transformation matrices. The interval width must then
by necessity be small enough to ensure sufficient accuracy not
only in the first-order but also in the zeroth-order solutions.

For these reasons we have modified the original Gordon
program to eliminate the determination of the first-order cor-
rections. The relative magnitudes of the neglected off-diagonal
and diagonal corrections, which are used to determine the next

step size, can still be adequately estimated by integrating the neglected terms in the $\underline{G}$ and $\underline{H}$ matrices over the interval in question, in a manner similar in spirit to Gordon's work.[1,2] Specifically, since

$$\int_{R_m-\Delta_m/2}^{R_m+\Delta_m/2} (R-R_m)^2 dR = \Delta_m^3/12 \quad , \tag{9}$$

the average integrated magnitude of the diagonal elements of the second-derivative terms in Eq.(2) is

$$H_{av} = \frac{\Delta_m^3}{24N} \sum_{n=1}^{N} |H_{nn}| \quad . \tag{10}$$

A comparison of this quantity with the average magnitude of the eigenvalues λ_n, will now provide a value for CDIAG, the estimate of the effect of the neglected diagonal second derivative terms. We thus have

$$CDIAG = \frac{\Delta_m^3}{24 \, \|\lambda\|_1} \sum_{n} |H_{nn}| \tag{11}$$

where the matrix norm in the denominator is defined by[8]

$$\|\lambda\|_1 = \sum_{n=1}^{N} |\lambda_n| \quad , \tag{12}$$

and is equal to N times the average eigenvalue magnitude.

Similarly, a value for the parameter COFF, which is a measure of the effect of the neglected off-diagonal terms in

$\underline{\underline{G}}$ can be obtained by comparing to $\|\lambda\|_1$ the integrated magnitude of the largest (in magnitude) off-diagonal element of $\underline{\underline{G}}$. Since

$$\int_{R_m-\Delta_m/2}^{R_m+\Delta_m/2} |(R-R_m)| \, dR = \Delta_m^2/4, \tag{13}$$

we find

$$COFF = \frac{N\Delta_m^2}{4 \, \|\lambda\|_1} \, \max \, (|G_{ij}|)_{i\neq j} \quad . \tag{14}$$

To incorporate the above changes we have made the following modifications in the original Gordon program:

i) Propagation across a given interval, even for the initial energy, is achieved by the subroutine SPROP. The lengthy subroutines STEP, DPROP, and FLAT are eliminated.

ii) The algorithm for determination of the width of the next interval, Δ_{m+1}, is unchanged with the exception that the key input parameters CDIAG and COFF are now determined from Eqs. (11) and (14).

In addition we have made one other modification:

iii) In the subroutine DTRANS, which performs the transformation of Eq. (4), the option which allows additional transformations in the case of nearly degenerate eigenvalues was eliminated. In practice we have found this option to provide little gain in accuracy or speed and, in some cases, to contribute possibly to instabilities in the propagated solution matrix.

On the UNIVAC 1100 at the University of Maryland these modifications yielded savings of ∿4100 words of object code as well as N^2+20N words of required data storage. The modified program, designated QCOL/MK2, was tested both on the Lester-Bernstein model atom-rigid rotor collision system,[9] and on the collision of two HF molecules at $E_{tot}=8000$ cm^{-1} and J=400.[10] The former calculations were performed on the UNIVAC 1108 at Maryland and the latter on the CDC 7600 at LBL. In both cases the use of Eqs. (11) and (14) resulted in a distribution of interval widths very similar to that predicted by the error criteria in the original Gordon program. Table I displays some representative times for determination of an S-matrix at one value of the total angular momentum.

Figure 1 displays the convergence of the root-mean-square error of the inelastic and elastic transition probabilities as a function of CPU time for the Lester-Bernstein model problem at a 9-channel level. The error quantities are

$$\Delta_{elas} = [\sum_{j,\ell} \Delta^2_{j\ell,j\ell} /9]^{\frac{1}{2}} \quad , \tag{15}$$

for the elastic transitions and

$$\Delta_{inel} = [\sum_{\substack{j\neq j' \\ \ell\neq\ell'}} \Delta^2_{j\ell,j'\ell'} /72]^{\frac{1}{2}} \quad , \tag{16}$$

for the inelastic transitions. Here $\Delta_{j\ell,j'\ell'}$ is the absolute deviation of the calculated transition probability $|S^J_{j\ell,j'\ell'}|^2$

from the exact value, taken from Stechel, Walker, and Light.[11]
Each plotted point corresponds to a different choice of input
parameters.

If one desires only moderate accuracy, it is obvious that
the present modifications offer a substantial decrease in CPU
time, especially for the crucial initial energy calculation. The
degree of convergence of both the original and modified Gordon al-
gorithms appears to be substantially similar. The larger scatter
in the QCOL/MK2 generated points in Figs. 1 and 2 at longer CPU
times (high accuracy) probably reflects the fact that Eqs. (11)
and (14) are only *estimates* of the error and thus are ultimately
less accurate for the prediction of step sizes than the actual
corrections to the propagated solution, which are determined in
the original Gordon code. We are presently working on the develop-
ment of alternate procedures for the even more accurate prediction
of step sizes.

In summary, we have developed a simple modification of the
original Gordon program which can result in a substantial saving in
computer time. The new QCOL/MK2 code should be best suited to the
rapid generation of S-matrices of moderate accuracy.

1. R. G. Gordon, J. Chem. Phys. $\underline{51}$, 14 (1969).

2. R. G. Gordon, Meth. Comput. Phys. $\underline{10}$, 81 (1971).

3. R. G. Gordon, QCOL, Program 187, Quantum Chemistry Program Exchange (QCPE), Department of Chemistry, Indiana University, Bloomington, IN.

4. M. H. Alexander and R. G. Gordon, J. Chem. Phys. $\underline{55}$, 4890 (1971).

5. M. H. Alexander, J. Chem. Phys. $\underline{61}$, 5167 (1974).

6. R. B. Walker and J. C. Light, Chem. Phys. $\underline{7}$, 84 (1975).

7. M. H. Alexander, J. Comput. Phys. $\underline{20}$, 248 (1976).

8. E. Isaacson and H. B. Keller, "Analysis of Numerical Methods," (John Wiley, N.Y., 1966).

9. W. A. Lester and R. B. Bernstein, J. Chem. Phys. $\underline{48}$, 4896 (1968).

10. A. E. DePristo and M. H. Alexander, J. Chem. Phys. $\underline{66}$, 1334 (1977).

11. E. B. Stechel, R. B. Walker, and J. C. Light, J. Chem. Phys. $\underline{69}$, 3518 (1978).
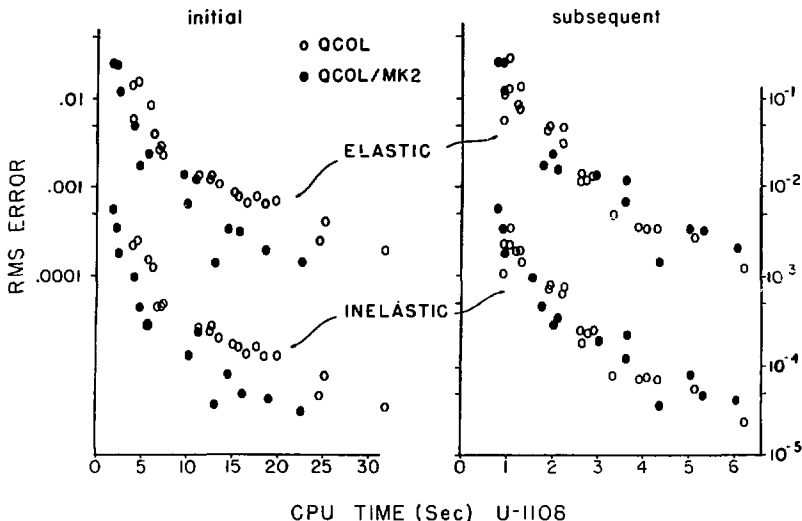
Fig. 1. Plot of the root-mean-square error [Eqs. (15) and (16)] for the Lester-Bernstein model problem (9-channels) as a function of CPU time (UNIVAC 1108). The left and right panels refer, respectively, to initial and subsequent energy calculations. For clarity the inelastic points have been displaced downward one decade. Consequently, the magnitude of the error in the *elastic* probabilities should be read from the *left* ordinate; the magnitude of the error in the *inelastic* probabilities, from the *right* ordinate.

Table I.   Representative CPU times (seconds) for determination
of an S-matrix.

N$_{channels}$	initial energy		subsequent energy
	QCOL	QCOL/MK2	
	UNIVAC 1108[a]		
9	7.0	3.6	1.5
16	24.6	11.6	4.2
20	39.8	19.5	6.6
25	74.6	32.1	10.4
	CDC 7600[b]		
10	0.51	0.36	0.11
17	1.5	1.0	0.25
28	4.8	3.0	0.71
44	13.2	8.9	2.0
72	53.5	34.5	8.⁻

a) Lester-Bernstein model, Ref. 9.

b) Rotationally inelastic collision between two HF molecules, Ref. 1⁰.
    For most meaningful comparison with atom-diatom calculations the
    values shown are exclusive of times required for computation of
    the more complex diatom-diatom coupling potential.

THE LOG DERIVATIVE AND RENORMALIZED
NUMEROV ALGORITHMS

B. R. Johnson
Chemistry and Physics Laboratory
The Ivan A. Getting Laboratories
THE AEROSPACE CORPORATION
El Segundo, Calif. 90245

The Log-Derivative and Renormalized Numerov Algorithms

I. Introduction

Two algorithms for solving the coupled channel differential equations which arise in atomic and molecular scattering theory will be presented. They are the log-derivative method[1,2] and the renormalized Numerov method.[2,3] Both these algorithms share the following desirable properties: They are simple and easy to implement, no special difficulties are encountered with closed channels, the step size can be easily changed and no linear dependence or overflow difficulties arise when propagating the solution though classically forbidden regions.

The log-derivative method will be discussed first, then the renormalized Numerov method and finally a model calculation using both these algorithms will be discussed and features of the two methods compared.

II. Log-Derivative Method

The "coupled-channel Schroedinger equation" is most conveniently written in the following matrix differential equation form:

$$\left[I \frac{d^2}{dx^2} + Q(x) \right] \Psi(x) = 0 \tag{1}$$

where

$$Q(x) = (2\mu/\hbar^2)\left[EI - V(x) \right]. \tag{2}$$

Here, I is the unit matrix, μ is the reduced mass, $V(x)$ is the symmetric potential matrix which has the centrifugal potential and the diagonal threshold energy terms included in it and E is the total energy. The wave function $\Psi(x)$ is a square-matrix function of x. The log-derivative matrix is defined to be

$$y(x) = \Psi'(x)\,\Psi^{-1}(x), \tag{3}$$

where the prime means differentiation with respect to x. Differentiating Eq. (3) and using Eq. (1) to eliminate the second derivative term, we obtain the matrix Ricatti equation

$$y'(x) + Q(x) + y^2(x) = 0, \tag{4}$$

This equation cannot be integrated by the usual numerical techniques for solving first order differential equations because $y(x)$ diverges for certain values of x. This is illustrated by the solution to the simple one channel problem in which Q is a constant:

$$y(x) = Q^{1/2}\text{ctn}\,(Q^{1/2}x). \tag{5}$$

This function is infinite at the points $x = n\pi Q^{-1/2}$ and the usual numerical algorithms for solving a first order differential equation cannot propagate the solution across these points. Our algorithm has no difficulty propagating the solution across these singular points.

The algorithm is as follows:

$$y_n = (I + h\, y_{n-1})^{-1}\, y_{n-1} - h^{-1}\, U_n \tag{6}$$

where h is the spacing between the N + 1 grid points $x_o, x_1,...x_N$ and where

$$U_n = \begin{cases} (h^2/3)\, Q\,(x_n), & n = 0, N \\ 2\,(h^2/3)\, Q\,(x_n), & n = 2, 4,...N-2 \\ 8I + 8\left[I - (h^2/6)\, Q\,(x_n)\right]^{-1}, & n = 1, 3, ...N-1 \end{cases} \tag{7}$$

Eq. (6) is a two term recurrence relation that can be iterativly solved once the term y_o is specified. The initial term is related to the initial value of the log-derivative function by the relation

$$y_o = y(x_o) - h^{-1}\, U_o. \tag{8}$$

The calculated value of y_n is equal to $y(x_n)$ only at the final integration point n =N. This is no problem however, since only this value is needed to calculate the S-matrix.

In actual practice it is somewhat more convenient and efficient to solve for the quantity

$$z_n = I + hv_n \tag{9}$$

Subsitituting this into Eq. (6) it is easy to show that

$$z_n = (2I - U_n) - z_{n-1}^{-1} \tag{10}$$

The initial term is calculated from the relation

$$z_0 = (I - U_0) + hv(x_0) \tag{11}$$

For most scattering problems, this leads to $z_0^{-1} = 0$. The matrix $y(x_N)$ is obtained from z_N in a final calculation

$$y(x_N) = h^{-1}(z_N - I) \tag{12}$$

The reaction matrix K is defined by the asymptotic behavior of the wave function. In the region $x \geq x_N$, in which all but the centrifugal part of the potential has become negligible, the wave function is

$$\Psi(x) \underset{x \geq x_N}{=} J(x) + N(x)K. \tag{13}$$

The matrices $J(x)$ and $N(x)$ are diagonal. The matrix elements of the open channels are made up of Riccati-Bessel functions

$$\left[J(x)\right]_{ij} \quad = \quad \delta_{ij} \; k_j^{-\frac{1}{2}} \; \widehat{j}_{\ell_j}(k_j x) \tag{14}$$

$$\left[N(x)\right]_{ij} \quad = \quad \delta_{ij} \; k_j^{-\frac{1}{2}} \; \widehat{n}_{\ell_j}(k_j x) \tag{15}$$

and the matrix elements for the closed channels are made up of modified spherical Bessel functions of the first and third kinds

$$\left[J(x)\right]_{ij} = \delta_{ij} \; (k_i x)^{\frac{1}{2}} \; I_{\ell_j + \frac{1}{2}}(k_j x) \tag{16}$$

$$\left[N(x)\right]_{ij} = \delta_{ij} \; (k_i x)^{\frac{1}{2}} \; K_{\ell_j + \frac{1}{2}}(k_j x) \tag{17}$$

where k_j is the channel wave number. Differentiate Eq. (13) with respect to x, then multiply from the right by the inverse of this equation, set $x = x_N$, and solve the resulting equation for K in terms of $y(x_N)$.

$$K = -\left[y(x_N) \; N(x_N) \; - \; N'(x_N)\right]^{-1} \times \left[y(x_N) \; J(x_N) \; - \; J'(x_N)\right]. \tag{18}$$

The matrix K is an augmented reaction matrix containing elements connecting closed as well as open channels, e.g., K can be written in the form

$$K \quad = \quad \begin{pmatrix} K_{oo} & K_{oc} \\ K_{co} & K_{cc} \end{pmatrix} \tag{19}$$

where K_{oo}, K_{oc}, K_{co}, and K_{cc} are open-open, open-closed, closed-open, and closed-closed submatrices of K. The S-matrix is given in terms of the open-open submatrix, K_{oo}, by the familiar formula

$$S = (I + iK_{oo})^{-1} (I - iK_{oo}). \tag{20}$$

Each of the matrices in Eq. (18) can be partitioned into open-open, open-closed, closed-open and closed-closed submatrices similar to the partitioning of K in Eq. (19). Written in partitioned form, Eq. (18) is

$$\begin{pmatrix} K_{oo} & K_{oc} \\ K_{co} & K_{cc} \end{pmatrix} = -\begin{pmatrix} y_{oo}N_o - N'_o & y_{oc}N_c \\ y_{co}N_o & y_{cc}N_c - N'_c \end{pmatrix}^{-1}$$

$$\times \begin{pmatrix} y_{oo}J_o - J'_o & y_{oc}J_c \\ y_{co}J_o & y_{cc}J_c - J'_c \end{pmatrix}$$

Since the S-matrix depends only K_{oo}, it is clear that the calculation can be simplified somewhat by only computing the left hand column of partitions of the K-matrix. The equation then becomes

$$\begin{pmatrix} K_{oo} \\ K_{co} \end{pmatrix} = -\begin{pmatrix} y_{oo}N_o - N'_o & y_{oc}N_c \\ y_{co}N_o & y_{cc}N_c - N'_c \end{pmatrix}^{-1} \begin{pmatrix} y_{oo}J_o - J'_o \\ y_{co}J_o \end{pmatrix} \tag{21}$$

The matrix J_c, which is the closed channel part of $J(x_N)$ is not used in this equation. Thus, the closed channel functions defined by Eq. (16) are not needed to calculate K_{oo}. However, the closed channel elements of $N(x_N)$ defined by Eq. (17) are still required.

The closed channel functions defined by Eq. (17) decrease exponentially with increasing x. This is a possible source of numerical difficulty. The problem is easily eliminated by redefining the closed channel elements of both $N(x)$ and $N'(x)$ by multiplying these functions by increasing exponential functions which just cancel the exponential decrease. That is, we make the following simple replacement

$$\left[N(x)\right]_{ii} \longrightarrow \left[N(x)\right]_{ii} \exp\left(k_i x\right) \tag{22a}$$

$$\left[N'(x)\right]_{ii} \longrightarrow \left[N'(x)\right]_{ii} \exp\left(k_i x\right) \tag{22b}$$

It should be noted that after this replacement is made, $N'(x)$ is no longer the first derivative of $N(x)$. It is easily verified that replacing $N(x)$ and $N'(x)$ by the expressions given in Eq's. (22a) and (22b) will leave K_{oo} unchanged. These modified closed channel functions can be easily calculated from recurrence relations.

We have seen from Eq. (21) that the closed channel elements of $J(x)$ are not required to calculate K_{oo} but the closed channel elements of $N(x)$ are, in general, required. The need for these functions can also be eliminated, but only if the value of x_N (see Eq. (18)) is sufficiently large. It can be shown that the elements of $y_{co}(x)$ and $y_{oc}(x)$ must eventually approach zero exponentially as x increases. Thus, the open and closed channel parts of Eq. (21) decouple and only the open channel elements of $N(x)$ are required to calculate K_{oo}.

III. Renormalized Numerov Method

The matrix Numerov algorithm is an efficient method that can be used to obtain numerical solutions of Eq. (1). The basic formula is the three term recurrence relation

$$[I - T_{n+1}]\Psi_{n+1} - [2I + 10T_n]\Psi_n + [I - T_{n-1}]\Psi_{n-1} = 0 \tag{23}$$

where

$$\Psi_n \equiv \Psi(x_n) \tag{24}$$

and

$$T_n = -(h^2/12)Q(x_n) \tag{25}$$

Here h is the spacing between the $N + 1$ equally spaced grid points x_0, $x_1, \ldots x_N$ and the square matrix $Q(x)$ is defined by Eq. (2). Equation (23) is derived by an obvious generalization of the derivation of the ordinary Numerov algorithm to matrix quantities.

The renormalized Numerov algorithm is derived from Eq. (23) by making two transformations. First define the matrix

$$F_n = [I - T_n]\Psi_n \tag{26}$$

and substitute into Eq. (23). This gives

$$F_{n+1} - U_n F_n + F_{n-1} = 0. \tag{27}$$

where

$$U_n = (I - T_n)^{-1} (2I + 10T_n). \tag{28}$$

Next, define the ratio matrix

$$R_n = F_{n+1} F_n^{-1}. \tag{29}$$

Substitute this into Eq. (27) to obtain the two term recurrence relation

$$R_n = U_n - R_{n-1}^{-1}. \tag{30}$$

This is the basic equation of the renormalized Numerov method. It can be solved once the value of the initial term R_0 is specified. In scattering problems, the usual case is to assume the initial values of the wavefunction are $\Psi(x_0) = 0$ and $\Psi(x_1) \neq 0$. The corresponding value of the initial inverse ratio matrix is $R_0^{-1} = 0$. (For exceptions to this rule see Appendix D in Ref. 3 .)

The matrix U_n, defined by Eq. (28), is symmetric. It follows from this and the symmetry of R_0^{-1} and also from Eq. (30) that the matrix R_n is also symmetric. For computational convenience, Eq. (28) can be reformulated as a symmetric matrix inversion problem. Define

$$W_n = I - T_n, \tag{31}$$

then

$$U_n = 12W_n^{-1} - 10I. \tag{32}$$

Thus, at each grid point we must invert two symmetric matrices.

Equation (30) can be solved iterativly to obtain R_N. The value of R_{N-1}^{-1} is also readily available and can be saved at the last integration point. Using these two quantities the log-derivative matrix can easilv be calculated by means of the formula[3]

$$y(x_n) = h^{-1} (A_{n+1} R_n - A_{n-1} R_{n-1}^{-1})(I - T_n) \tag{33}$$

where T_n is defined by Eq. (25) and

$$A_n = (0.5 I - T_n)(I - T_n)^{-1} \tag{34}$$

The S-matrix can then be calculated from this log-derivative matrix by the techniques outlined in section II.

Another method for calculating the S-matrix, which avoids calculating the log-derivative matrix, is also possible. Multiply Eq. (13) by $(I - T_n)$ to obtain

$$F_n = j(x_n) + n(x_n) K \tag{35}$$

where we have defined

$$j(x_n) = (I - T_n) J(x_n) \tag{36}$$

and

$$n(x_n) = (I - T_n) N(x_n) \tag{37}$$

Evaluate Eq. (35) at x_N and x_{N+1}, calculate the ratio matrix $R_N = F_{N+1} F_N^{-1}$, then solve the resulting equation for K in terms of R_N

$$K = -[R_N n(x_N) - n(x_{N+1})]^{-1} [R_N j(x_N) - j(x_{N+1})] \tag{38}$$

This equation is similar to Eq. (18) and can be partitioned and solved in exactly the same way. If x_N is large enough, the open and closed channel parts of this equation decouple; if not, they are coupled and the closed channel elements of $n(x_N)$ will be required. In order to avoid any possible numerical difficulty, the closed channel elements of $N(x)$ in Eq. (37) should be the modified functions defined in Eq. (22a).

The K-matrix and S-matrix computed by the renormalized Numerov algorithm will by symmetric only to within the truncation error of the calculation. In fact, one can obtain an estimate of the magnitude of the truncation error from the error in symmetry. This is in contrast to the log derivative method where there is no relationship between truncation error and the symmetry of the S-matrix.

IV. Example Calculations

In this section several of the characteristics of the log-derivative and renormalized Numerov algorithms will be elucidated and compared by applying them to a model problem.

The model is the atom-collinear harmonic oscillator system described by Secrest and Johnson.[4] The Schroedinger equation for this problem is

$$\left[-\frac{1}{2m}\left(\frac{\partial^2}{\partial x^2}\right) - \frac{1}{2}\left(\frac{\partial^2}{\partial y^2}\right) + \frac{1}{2}y^2 + V(x-y) - \frac{1}{2}E\right]\Psi = 0$$

where the interaction potential is

$$V(x-y) = A \exp\left[-\alpha(x-y)\right]$$

This problem was recently solved very accurately by Stechel, Walker and Light[5] for the particular set of parameters: $A = 41000$, $\alpha = 0.3$, $m = 2/3$, and $E = 6.9$, 8.0. They used a six channel expansion of the wave function and an integration range from $x = 0-100$. We have chosen this same set of parameters for our model problem. However, we only solved the $E = 6.0$ case.

Accurate, converged values of the transition probabilities are given in Table I. These values were calculated using the renormalized Numerov algorithm with 2000 points and a grid spacing h = 0.05. Since the transition probability matrix, computed by the renormalized Numerov method, is symmetric only to within the truncation error, we have symmetrized the results by averaging P_{mn} and P_{nm}. It is these averaged probabilities that are given in Table I. The inaccuracy of any of these numbers is no greater than two digits in the last place shown.

In Figs. 1 and 2 we show the relative truncation error of the calculated transition probabilities as a function of the grid spacing. There are several features to observe: The error curves are almost linear (on a log-log scale) with a slope very close to 4. This is consistent with the fact that both algorithms are fourth-order methods. Next we note that for a given grid spacing, the renormalized Numerov method is more accurate than the log-derivative method. Alternatively, in order to obtain the same relative error, we must use a smaller grid spacing with the log-derivative method. The worst case is the 0-1 transition, where the ratio of log derivative to renormalized Numerov grid spacings must be about 0.63 in order to obtain equal relative errors. The best case is the 1-2 transition where this ratio is about 0.88.

On the other hand, the average CPU time per grid point is less for the log derivative method than for the renormalized Numerov method by an approximately constant ratio of about 0.76. This ratio is easy to understand. It is approximately the ratio of the number of matrix inversions. Two inversions per grid point are required for the renormalized Numerov method whereas, on the average, only 1.5 inversions per grid point are required for the log derivative method. The only other procedure that might

require much computer time at each grid point is the calculation of the potential matrix. In the present model problem this is almost negligible. However, if it were not, it would tend to make the time ratio per grid point less favorable to the log derivative method.

The fact that less time per grid point is required, approximately compensates for the increased number of points required by the log derivative method. Based on the figures given, the 1-2 transition could be calculated more efficiently using the log derivative method whereas the 0-1 transition could be calculated faster with the renormalized Numerov method.

The average CPU time per grid point as a function of the number of channels is plotted in Fig. 3. These curves can be approximately extrapolated for N larger than 20 channels by the formulas

$$T_{RN} = 0.208*10^{-5}N^{2.8}$$

and

$$T_{LD} = 0.158*10^{-5}N^{2.8}$$

By using the information in Figs. 1, 2 and 3 and given the integration range, the time required to calculate solutions of various accuracies can be determined.

REFERENCES

1.   B. R. Johnson, J. Comp. Phys. 13, 445 (1973).

2.   B. R. Johnson, J. Chem. Phys. 67, 4086 (1977).

3.   B. R. Johnson, J. Chem. Phys. 69, 4678 (1978).

4.   D. Secrest and B. R. Johnson, J. Chem. Phys. 45, 4556 (1966).

5.   E. B. Stechel, R. B. Walker, and J. C. Light J. Chem. Phys. 69, 3518 (1978).

ATOM - COLLINEAR HARMONIC OSCILATOR

RENORMALIZED NUMEROV
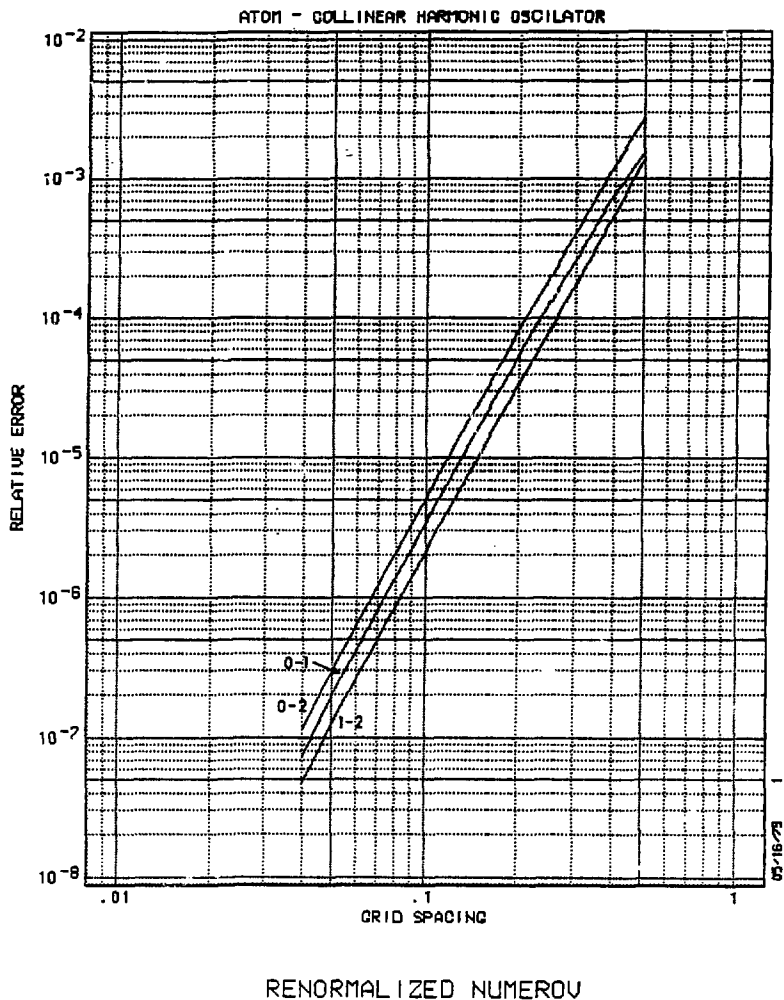
Figure 1.

LOG. DERIVATIVE

Figure 2.

CPU TIME VS. NUMBER OF CHANNELS (CDC 7600)
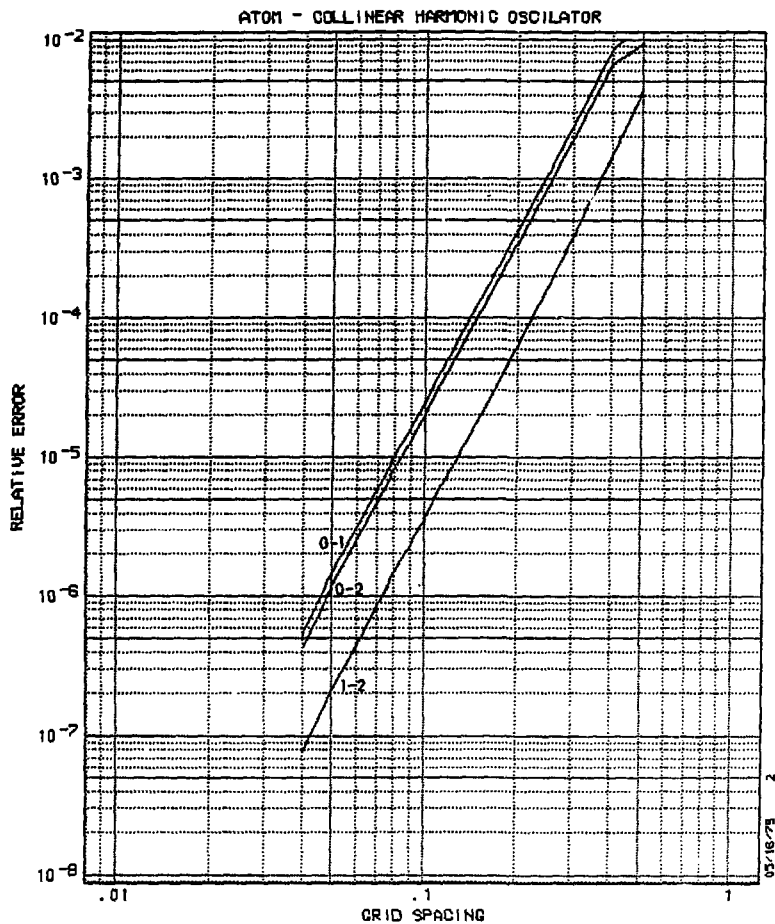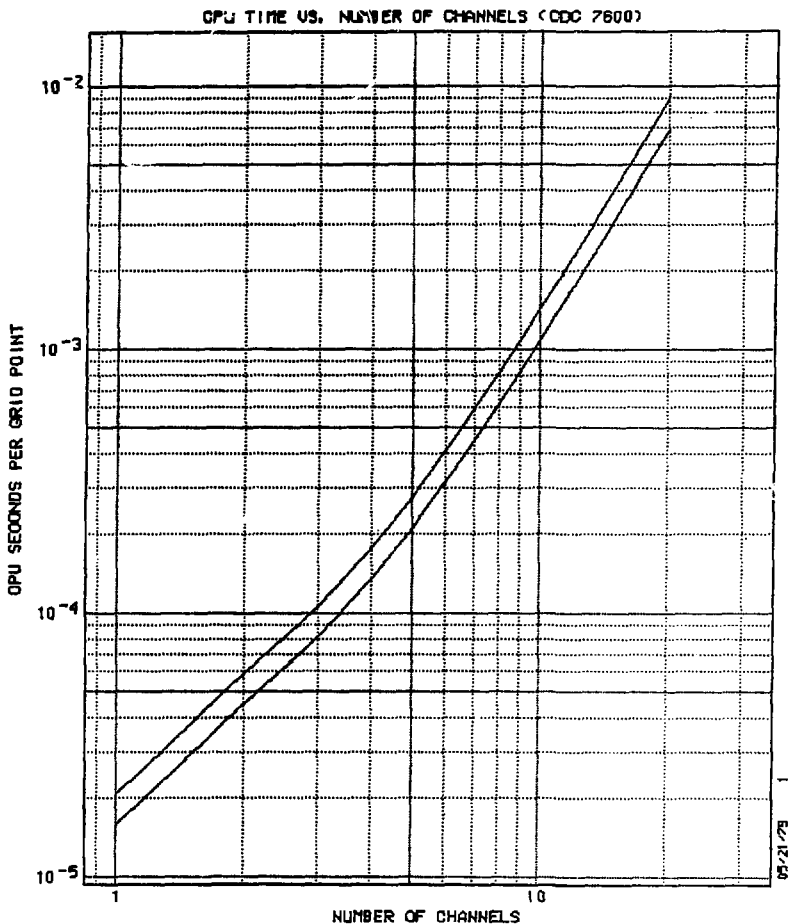
UPPER CURVE = REN NUM. LOWER CURVE = LOG DERIV.

Figure 3.

Table I		Transition Probabilities

$\underline{n}$	$\underline{m}$	P_{nm}
0	0	0.97788564
1	1	0.97699265
2	2	0.999096929
0	1	$0.2210932*10^{-1}$
1	2	$0.898031*10^{-3}$
0	2	$0.503948*10^{-5}$

DEVOGELAERE'S METHOD

William A. Lester, Jr.
National Resource for Computation in Chemistry
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

For a differential equation of the form

$$y'' = f(x,y)$$

DeVogelaere's method[1] consists of cyclic use of the equations

$$y_{1/2} = y_0 + \frac{1}{2} y_0' + \frac{1}{6} (F_0 - \frac{1}{4} F_{-1/2}) \quad \{4\} \tag{1}$$

$$y_1 = y_0 + hy_0' + \frac{1}{6} (F_0 + 2F_{1/2}) \quad \{5\} \tag{2}$$

$$hy_1' = hy_0' + \frac{1}{6} (F_0 + 4F_{1/2} + F_1) \quad \{6\} \tag{3}$$

where

$$F_p = h^2 f(x_p, y_p) \ . \tag{4}$$

The method requires $F_{-1/2}$ from the previous step and therefore is not self-starting. For the initial step $F_{-1/2}$ may be obtained from

$$y_{-1/2} = y_0 - \frac{1}{2} hy_0' + \frac{1}{8} F_0 \quad \{3\} \tag{5}$$

In Eqs. (1) - (5) h is the interval. We further note that the method requires only two evaluations per interval. Following Ref. 1, we use the symbol $\{n\}$ which is equivalent to $O(h^n)$.

Very recently Coleman and Mohamed have determined truncation error estimates which permit automatic error control.[2] A FORTRAN program incorporating these features is available from Computer Physics Communication.[3]

The generalization to a system of differential equations is straight-forward: For

$$y_i'' = f_i(x, y_k) \qquad i, k = 1, 2, \ldots N \tag{6}$$

one obtains, in place of Eqs. (1) - (3)

$$y_{i,1/2} = y_{i,0} + \frac{h}{2} y_{i,0}' + \frac{1}{6} (F_{i,0} - \frac{1}{4} F_{i,-1/2}) \tag{7}$$

$$y_{i,1} = y_{i,0} + h y_{i,0}' + \frac{1}{6} (F_{i,0} + 2F_{i,1/2}) \tag{8}$$

$$h y_{i,1}' = h y_{i,0}' + \frac{1}{6} (F_{i,0} + 4F_{i,1/2} + F_{i,1}) \tag{9}$$

where

$$F_{i,p} = h^2 f_i(x_p, y_{k,p}) \qquad i, k = 1, 2, \ldots N \tag{10}$$

The initial step requires

$$y_{i,-1/2} = y_{i,0} - \frac{1}{2} h y_{i,0}' + \frac{1}{8} F_{i,0} \tag{11}$$

Scraton[4] has indicated that by use of Radau's closed quadrature formula for any value of n, i.e.,

$$\int_{x_0}^{x_1} g(x) dx = h[W_0 g(x_0) + \sum_{r=1}^{n-1} W_r g(x_0 + \alpha_r h) + W_n g(x_1)] , \quad \{2n+1\} \tag{12}$$

it can be deduced that

$$y_1 = y_0 + h y_0' + W_0 F_0 + \sum_{r=1}^{n-1} W_r (1 - a_r) F_{a_r} \qquad \{2n+1\} \tag{13}$$

$$h y_1' = h y_0' + W_0 F_0 + \sum_{r=1}^{n-1} W_r F_{a_r} + W_n F_1 \qquad \{2n+2\} \tag{14}$$

It is possible to write down a set of equations for the unknowns W_r depending on the order of accuracy required.[5]

For n=1 one obtains the simple trapezium rule, which leads to

$$y_1 = y_0 + hy_0' + \frac{1}{2} F_0 \qquad \{3\} \qquad (15)$$

$$hy_1' = hy_0' + \frac{1}{2} (F_0 + F_1) \qquad \{4\} \qquad (16)$$

This case requires no starting procedure.

For n=2, the Radau formula is Simpson's rule and the corresponding equations are DeVogelaere's method Eqs. (1) - (3). Thus, as Scraton[4] states, higher order cases of this type can be regarded as generalizations of DeVogelaere's method.

It is worthwhile examining the n=3 case. which to the author's knowledge has not been used in collision studies. Following Scraton,[4] Radau's four point formula (n=3) is

$$\int_{x_0}^{x_1} g(x)dx = \frac{h}{12} (g_0 + 5g_a + 5g_{1-a} + g_1) \qquad \{7\} \qquad (17)$$

where

$$a = \frac{5 - \sqrt{5}}{10} = 0.2763,9320$$

This leads to

$$y_a = y_0 + 0.2763,9320 \ hy_0' + 0.0645,7768 \ F_0$$

$$- 0.0387,4353 \ F_{-a} + 0.0187,1643 \ F_{a-1}$$

$$- 0.0063,5398 \ F_{-1} \qquad \{6\} \qquad (18)$$

$$y_{1-a} = y_0 + 0.7236,0680 \ hy_0' + 0.2971,1983 \ F_a$$
$$-0.1294,4272 \ F_0 + 0.1098,7164 \ F_{-a} - 0.0157,4536 \ F_{a-1} \quad \{6\} \quad (19)$$

$$y_1 = y_0 + hy_0' + \frac{1}{12} F_0 + 0.3015,0283 \ F_a + 0.1151,6383 \ F_{1-a} \quad \{7\} \quad (20)$$

$$hy_1' = hy_0' + \frac{1}{12} (F_0 + 5F_a + 5F_{1-a} + F_1) \quad \{8\} \quad (21)$$

For the initial step, one can obtain F_{-a}, F_{a-1}, and F_{-1} to an adequate degree of accuracy from the following values of y:

$$y_{-1/2} = y_0 - \frac{1}{2} hy_0' + \frac{1}{8} F_0 \quad \{3\} \quad (22)$$

$$y_{-1} = y_0 - hy_0' + \frac{1}{6} (F_0 + 2F_{-1/2}) \quad \{5\} \quad (23)$$

$$y_{-a} = y_0 - 0.2763,9320 \ hy_0' + 0.0286,1197 \ F_0$$
$$+ 0.0121,3107 \ F_{-1/2} - 0.0025,4644 \ F_{-1} \quad \{5\} \quad (24)$$

$$y_{a-1} = y_0 - 0.7236,0680 \ hy_0' + 0.1180,5469 \ F_0$$
$$+ 0.1612,0227 \ F_{-1/2} - 0.0174,5356 \ F_{-1} \quad \{5\} \quad (25)$$

To my knowledge the first application of the DeVogelaere algorithm, Eqs. (1-3), to single channel scattering was by Bernstein et al.[6] in their studies of barrier penetration and resonance effects. A program for the multi-channel case was written by me[7] and formed the step-wise propagation part of the code used in some of the earliest convergence tests of coupled-channel solutions for the atom-rigid rotor problem.[8]

The multi-channel version has since been used by McGuire[9] for coupled-channel studies of atom-rotor systems and coupled-states studies of atom-vibrator systems, and by Launay[10] who noted for atom-rotor studies in a body-fixed formulation that computational time with the DeVogelaere algorithm should increase as $N^{2.5}$ instead of as N^3 in the SF representation. This savings arises from the reduced number of non-zero matrix elements in the body-fixed coupling matrix and the facile elimination of matrix multiplications involving null factors that is possible because of the cyclic structure of the algorithm. In addition, unlike most other methods, there are no matrix inversions in the Devogelaere method.

An excellent comparison of the DeVogelaere, matrix Numerov, and iterative Numerov has been given by Allison,[11] and demonstrates the advantage of the latter two methods over the former. Comparisons are also made with Gordon's linear reference potential method.[12] The interested reader is referred to Allison's paper for details.

Finally, it is noted that the multi-channel version of the $n=3$ single-channel equations derived from Radau's formula, Eqs. (17)-(25), does not appear to have been applied to scattering problems. Such application bears investigation.

References

1. R. DeVogelaere, J. Research Natl. Bur. Std. $\underline{54}$, 119 (1955).

2. J. P. Coleman and J. Mohamed, Math. Comput. $\underline{32}$, 751 (1978).

3. J. P. Coleman and J. Mohamed, Comp. Phys. Comm. $\underline{17}$, 283 (1979).

4. R. E. Scraton, Comp. J. $\underline{6}$, 368 (1964).

5. L. Collatz, The Numerical Treatment of Differential Equations, Springer Verlag, Berlin, Germany (1960), p. 61.

6. R. B. Bernstein, C. F. Curtiss, S. Iman-Rahajoe, and H. T. Wood, J. Chem. Phys. $\underline{44}$, 4072 (1966).

7. W. A. Lester, Jr., J. Computational Phys. $\underline{3}$, 322 (1968); Theoretical Chemistry Institute, University of Wisconsin, Report No. WIS-TCI-285R (1968). The Appendix contains a description and listing of a FORTRAN subprogram.

8. W. A. Lester, Jr. and R. B. Bernstein, Chem. Phys. Letters $\underline{1}$, 207, 347 (1967); J. Chem. Phys. $\underline{48}$, 4896 (1968).

9. See, for example, P. McGuire, Chem. Phys. $\underline{13}$, 81 (1976) and references to McGuire contained therein.

10. See, for example, J.-M. Launay, J. Phys. B: Atom, Molec. Phys. $\underline{9}$, 1823 (1976).

11. A. M. Allison, J. Comput. Phys. $\underline{6}$, 376 (1970).

12. R. G. Gordon, Meth. Comput. Phys. $\underline{10}$, 81 (1971).

NEW DEVELOPMENTS IN METHODS FOR THE NUMERICAL SOLUTION
OF THE RADIAL SCHRÖDINGER EQUATION

Arthur Allison
Department of Computer Science
University of Glasgow

Introduction

For large values of the independent variable r, the solution of the radial Schrödinger equation usually behaves in some predictable manner-- either exponentially decaying appropriate to a closed channel or oscillatory in the case of an open channel.

One of the major disadvantages of the well known Numerov method, or indeed any of the usual linear multistep integration formulae, is that it cannot exploit this known behavior. This is because the formulae are based on polynomial approximation and polynomials do not easily approximate exponential or trigonometric functions. In this context the Numerov method integrates polynomials of degree up to five exactly and thus a basis set for this method is 1, r, r^2, r^3, r^4, r^5.

The known asymptotic form of the solution is, of course, assumed when the effect of the potential is dying away, usually slowly, and a linear fit to the potential is valid over some reasonably large radial distance. These are the conditions under which the stepwise perturbative methods, such as that of Gordon [2], work extremely well. Thus recent comparisons between Numerov and the stepwise perturbative methods have tended to favor the latter.

It is possible to modify the Numerov formula, while retaining its computational simplicity, and circumvent the problems mentioned above using exponential fitting.

Exponentially fitted methods

The concept of exponential fitting arose from the study of the problem of solving sets of stiff differential equations where one of the characteristic values had modulus much greater than the others and hence the step size of integration was constrained to unacceptably small values. Liniger and Willoughby [4] developed a method which fitted this single large eigenvalue by adjusting a parameter in their numerical formula, thus allowing use of an interval size determined by the smaller eigenvalues.

A modification of a definition by Lambert [3] would read "A numerical method is said to be exponentially fitted at a value λ_0 if when the method is applied to the scalar test problem $y" = \lambda^2 y$, $y(r_0) = y_0$, $y'(r_0) = y'_0$ with exact initial conditions, it yields the exact theoretical solution in the case when $\lambda = \lambda_0"$.

This approach, meshing with earlier work by Gautschi [1] on fitting with trigonometric polynomials, has been developed by Lyche [5], and recently Raptis and Allison [6] have applied these ideas to the solution of the radial Schrödinger equation. We looked for an exponentially fitted analogue of the Numerov formula and we found the formula

$$y_{r+1} - 2y_r + y_{r+1} = h^2 \{\beta_2(h)y"_{r+1} + \beta_1(h)y"_r + \beta_0(h)y"_r\}$$

where

$$\beta_0(h) = \beta_2(h) = \frac{(1 - \exp(\lambda h))^2 - \lambda^2 h^2 \exp(\lambda h)}{\lambda^2 h^2 (1 - \exp(\lambda h))^2}$$

$$\beta_1(h) = 1 - 2\beta_0(h) \ .$$

The functions 1, r, r^2, r^3, $\exp(\lambda r)$, $\exp(-\lambda r)$ are integrated exactly and the numerical method would be exact for the differential equation $y'' = \lambda^2 y$. To gain this advantage we have had to let our coefficients β depend on the interval size h and so they have to be recalculated any time the interval size is changed.

There will be exponentially fitted analogues of most of the linear multistep methods; for example, the Hartree method

$$y_{r+1} - 2y_r + y_{r-1} = h^2 y_r'',$$

with its basis set of 1, r, r^2, r^3, gives rise to the formula

$$y_{r+1} - 2y_r + y_{r-1} = h^2 \beta(h) y_r''$$

$$\beta(h) = \frac{1}{\lambda^2 h^2} (\exp(\lambda h) + \exp(-\lambda h) - 2)$$

with basis set 1, r, $\exp(\lambda h)$, $\exp(-\lambda h)$.

Furthermore, for any particular method, there may be several different choices of the coefficients β corresponding to different basis sets.

Implementation

The computational simplicity of the Numerov formulae has been retained and the main new problem is to decide on the optimal value of the parameter λ. In the asymptotic region λ will clearly be set to the wave number, but in regions where the potential is rapidly varying the correct decision is not clear. However any reasonable choice of λ will still be as good as Numerov. These points will be discussed further.

Extrapolation methods

Under the heading of new developments I will raise the topic of extrapolation methods which have been used in calculations of atomic polarizabilities by Stewart [7]. This scheme has used the Hartree method with extrapolation on the $O(h^2)$ global truncation error. My colleague, Dr. M. J. Jamieson, and I have been thinking recently along these lines and some discussion would be valuable.

REFERENCES

[1] W. Gautschi   Numer. Math. 3(1961) 381.

[2] R. G. Gordon,   J. Chem. Phys. 51(1969) 14.

[3] J. D. Lambert,   Computational Methods in Ordinary Differential
                              Equations (Wiley, New York, 1973).

[4] W. Liniger and   IBM Research Report RC-1970 (1967).
     R. A. Willoughby

[5] T. Lyche,   Numer. Math. 19(1972) 65.

[6] A. Raptis and   Computer Phys. Commun. 14(1978) 1.
     A. C. Allison,

[7] R. F. Stewart   J. Phys. B 8(1975) 1.

R-MATRIX RECURSION METHODS: CONTINUOUS
AND L² CORRECTIONS*

J. C. Light, T. G. Schmalz, and J. V. Lill
The James Franck Institute and The Department of Chemistry
The University of Chicago, Chicago, Illinois 60637

I. Introduction

Most problems in atomic and molecular scattering require the solution of the Schrödinger equation for accurate results. Although there are numerous specific approaches and many approximation schemes to reduce the complexity of the equations, the most commonly used techniques result in a set of coupled ordinary second order differential equations to be solved over a specified range of the independent (scattering) variable for the coefficient functions of the known (perhaps varying) basis set expansion. For inelastic scattering only the solutions regular at the origin are required and, for N_0 open channels, an $N_0 \times N_0$ matrix containing information equivalent to the R-matrix, log derivative matrix, or the K-matrix is required for the complete physically meaningful (open channel) S matrix to be evaluated. Thus here we are concerned with the solution of a set of matrix equations of the form

$$\left\{ -\underline{\underline{I}} \frac{d^2}{dR^2} + \underline{\underline{V}}(R) \right\} \underline{f}(R) = E \underline{f}(R) \qquad \lim_{R \to \infty} \underline{f}(R) \text{ finite} \qquad (1)$$

where $\underline{\underline{V}}$ is real symmetric and where we assume that the basis functions corresponding to a given element, $|i\rangle$, are fixed and orthonormal over some range of R. As shown recently[1] these

are not unduly restrictive conditions as even reactive and charge transfer problems can be handled in this fashion locally, with known transformation procedures between local sectors.

Although a number of methods are available for solving problems of this type numerically, the R-matrix recursion approach[2-4] developed over the last few years has proven useful for a variety of problems because of its relative simplicity, large step size for slowly varying potentials, and inherent stability in non-classical regions. In this paper we shall briefly re-derive the basic R-matrix equations in terms of matrix Green's functions (see Schneider and Walker[5]) and show how analytic perturbative corrections can increase the step size and accuracy. Results from a model rotational problem will be presented. Finally we will discuss both remaining problems and possible future improvements via L^2 corrections.

II.   R-Matrix (Green's Function) Recursion Method

If we consider the general solution of Eq. 1 over an interval of length h, $R_i - \frac{h}{2} \leqslant R \leqslant R_i + \frac{h}{2}$, a simple formal method is to construct the Green's function matrix for the interval. Let us add the Bloch operator[6] to both sides of Eq. 1:

$$\left\{ -\underline{\underline{I}} \frac{d^2}{dR^2} + \underline{\underline{V}}(R) - E\underline{\underline{I}} + \underline{\underline{I}} \, \mathcal{L} \right\} \underline{f} = \underline{\underline{I}} \, \mathcal{L} \, \underline{f} \tag{2}$$

where $L(R) = - \delta(R-R_i + h/2)\frac{d}{dR} + \delta(R-R_i - h/2)\frac{d}{dR}$. The operator on the left is now Hermitian since

$$\int_{R_i - \frac{h}{2}}^{R_i + \frac{h}{2}} g(R) \left[-\frac{d^2}{dR^2} + \mathcal{L} \right] h(R) \, dR = \int_{R_i - \frac{h}{2}}^{R_i + \frac{h}{2}} \left(\frac{d}{dR} g \right) \left(\frac{d}{dR} h \right) dR \tag{3}$$

The Green's function matrix for Eq. (2) satisfies the equation

$$\left\{ -\underset{=}{I} \frac{d^2}{dR^2} + \underset{\sim}{V}(R) - E \underset{=}{I} + \underset{=}{I} L \right\} \underset{=}{g}(R,R') = \underset{=}{I} \delta(R-R') \tag{4}$$

$$R_i - \frac{h}{2} \le R, R' \le R_i + \frac{h}{2}$$

Thus the solution of Eq. (2) is

$$\underset{\sim}{f}(R) = \int_{R_i - \frac{h}{2}}^{R_i + \frac{h}{2}} \underset{=}{g}(R,R') \, L(R') \underset{\sim}{f}(R') \, dR'$$

$$= - \underset{=}{g}\left(R, R_i - \tfrac{h}{2}\right) \frac{d}{dR'} \underset{\sim}{f}(R') \Big|_{R' = R_i - \frac{h}{2}} + \underset{=}{g}\left(R, R_i + \tfrac{h}{2}\right) \frac{d}{dR'} \underset{\sim}{f}(R') \Big|_{R' = R_i + \frac{h}{2}} \tag{5}$$

Evaluating this for $R = R_i - h/2 \equiv R_i^-$ and $R = R_i + h/2 \equiv R_i^+$, and writing in matrix form, we have

$$\begin{pmatrix} \underset{\sim}{f}(R_i^-) \\ \underset{\sim}{f}(R_i^+) \end{pmatrix} = \begin{pmatrix} \underset{=}{g}(R_i^-, R_i^-) & \underset{=}{g}(R_i^-, R_i^+) \\ \underset{=}{g}(R_i^+, R_i^-) & \underset{=}{g}(R_i^+, R_i^+) \end{pmatrix} \begin{pmatrix} -\underset{\sim}{f}'(R_i^-) \\ \underset{\sim}{f}'(R_i^+) \end{pmatrix}$$

$$= \begin{pmatrix} \underset{=}{r_1} & \underset{=}{r_2} \\ \underset{=}{r_3} & \underset{=}{r_4} \end{pmatrix} \begin{pmatrix} -\underset{\sim}{f}'(R_i^-) \\ \underset{\sim}{f}'(R_i^+) \end{pmatrix} \tag{6}$$

Equation (6) is easily recognized as the defining relation for a generalized (non-diagonal) sector r-matrix[2-5] giving the values of the functions on the boundaries in terms of the derivatives there (e.g. Eq. (9) of Ref. 4 or Ref. 5). Thus the exact sector r-matrix is just made up of the exact Green's function matrix (satisfying zero derivative b.c.'s) evaluated at the appropriate sector boundaries.

For inelastic scattering problems only the Green's function (R_4 matrix[3-5]) evaluated at the final (asymptotic) boundary is required, and this is given by recursion of the sector Green's functions:[2]

$$\underline{G}(R_i^\dagger, R_i^\dagger) = \underline{g}(R_i^\dagger, R_i^\dagger) - \underline{g}(R_i^\dagger, R_i^-) \underset{=}{Z} \underline{g}(R_i^-, R_i^\dagger)$$

$$\underset{=}{Z} \equiv \left[\underline{g}(R_i^-, R_i^-) + \underline{T}_{i-1,i}^T \underline{G}(R_{i-1}^\dagger, R_{i-1}^\dagger) \underline{T}_{i-1,i} \right]^{-1} \quad (7a)$$

where $\underline{T}_{i-1,i}$ is the matrix transformation between the basis used in sector i-1 and that in sector i. The initial $\underline{G}$, $\underline{G}(0,0)$ satisfies the regular b.c.'s at the origin although normally one starts near, but inside, all the classical turning points with a $\underline{G}$ which satisfies the appropriate exponentially decreasing b.c.'s there. For inelastic scattering the relationship between $\underline{G}$ and the wavefunction coefficients is

$$\underline{f}(R_i^\dagger) = \underline{G}(R_i^\dagger, R_i^\dagger) \underline{f}'(R_i^\dagger) \quad (7b)$$

As pointed out earlier,[4,5] the sector Green's functions (or r-matrices) can be evaluated by any means which yields accurate results over the sector - L^2 expansion,[2,5,7] approximation by an exact Green's function for a model diagonal potential,[2-4] etc. In what follows we give an elementary analytic perturbative method which is quite effective in increasing the step size and the accuracy, particularly for accurate calculations.

The problem is to evaluate the solution of Eq. (4) accurately over a step of length h. We use the straightforward approach of diagonalizing the potential matrix at the center of the step, evaluating the first or first and second derivative matrices in the diagonal representation, and computing the zero order Green's function corresponding to the diagonal constant potential. We then evaluate the first order perturbation corrections due to the linear and quadratic (if necessary) terms in the Taylor series expansion of the potential matrices. Various refinements of this **will be discussed in the last section.**

We take, for convenience,

$$X = R - (R_i - \frac{h}{2})$$

$$\underline{T}_i^T \underline{V}(R_i)\underline{T}_i = \lambda^2 \ (\text{diagonal})$$

$$\underline{T}_i^T \underline{V}^{(n)}(R_i)\underline{T}_i = \underline{\mu}^{(n)}$$

$$\underline{k}^2 = -\underline{\lambda}^2 + E \underline{I}$$

where $\underline{V}^{(n)}$ is the nth derivative of the potential matrix, and $\underline{T}_i$ is the orthogonal transformation which diagonalizes $\underline{V}$ at the center of the step. In the diagonal basis we write the equation for the exact Green's function matrix as

$$\left\{ -\underline{I}\frac{d^2}{dx^2} - \underline{k}^2 \right\} \underline{g}(x,x') = -\left\{ \sum_{n=1}^{\infty} \frac{\underline{\mu}^{(n)}}{n!} (x-\frac{h}{2})^n \right\} \underline{g}(x,x') + \underline{I}\,\delta(x-x')$$

$$= -\left(\underline{\mu} - \underline{\mu}^{(o)} \right) \underline{g}(x,x') + \underline{I}\,\delta(x-x') \qquad (8)$$

where we require $\underline{g}$ to satisfy zero derivative b.c.'s at $x = 0,h$. The zero order Green's function we take to satisfy the homogeneous part of Eq. (8)

$$\left\{ -\underline{\underline{I}} \frac{d^2}{dx^2} - \underline{\underline{k}}^2 \right\} \underline{\underline{q}}^0(x,x') = \underline{\underline{I}} \, \delta(x-x') \tag{9a}$$

$$\left[\underline{\underline{q}}^\bullet(x,x') \right]_{ij} = \delta_{ij} \, \frac{\cos(k_i \, x_<) \cos[k_i(x_> - h)]}{k_i \sin(k_i h)} \tag{9b}$$

where $x_<$ ($x_>$) have their normal meaning of the lesser (greater) of x,x'. It should be noted that if channel i is closed at this point, k_i is merely replaced by (ik_i) in Eq. (9b) and the resulting Green's function is still real and symmetric.

The exact Green's function satisfying these b.c.'s in the interval $0 \leqslant x \leqslant h$ is given by the integral equation

$$\underline{\underline{q}}(x,y) = \underline{\underline{q}}^\bullet(x,y) - \int_0^h \underline{\underline{q}}^0(x,z) \left(\underline{\underline{U}}(z) - \underline{\underline{U}}^{(o)} \right) \underline{\underline{q}}(z,y) dz \tag{10}$$

The first order perturbed Green's function is then

$$\underline{\underline{q}}^{(l)}(x,y) = \underline{\underline{q}}^\bullet(x,y) - \int_0^h \underline{\underline{q}}^0(x,y) \left(\underline{\underline{U}}(z) - \underline{\underline{U}}^{(o)} \right) \underline{\underline{q}}^\bullet(z,y) \, dz \tag{11}$$

$$= \underline{\underline{q}}^\bullet(x,y) + \underline{\Delta}\underline{\underline{q}}(x,y)$$

The specific integrals required are then

$$\left[\underline{\Delta}\underline{\underline{q}}(0,0) \right]_{ij} = -C_{ij} \int_0^h \cos k_i(z-h) \cos k_j(z-h) \left(\underline{\underline{U}}(z) - \underline{\underline{U}}^{(o)} \right)_{ij} dz$$

$$\left[\underline{\Delta}\underline{\underline{q}}(0,h) \right]_{ij} = -C_{ij} \int_0^h \cos k_i(z-h) \cos k_j z \left(\underline{\underline{U}}(z) - \underline{\underline{U}}^{(o)} \right)_{ij} dz$$

$$= \left[\underline{\Delta}\underline{\underline{q}}(h,0) \right]_{ji}$$

$$\left[\underline{\Delta}\underline{\underline{q}}(h,h) \right]_{ij} = -C_{ij} \int_0^h \cos k_i z \cos k_j z \left(\underline{\underline{U}}(z) - \underline{\underline{U}}^{(o)} \right)_{ij} dz$$

where $\qquad C_{ij} = \left[k_i \, k_j \, \sin k_i h \, \sin k_j h \right]^{-1}$

The explicit formulae for constant, linear, and quadratic terms in the residual potential matrix, $\underline{U}(z) - \underline{U}^{(o)}$, are given in the Appendix. The use of the first order corrected Green's function matrix in the recursion equation (Eq. 7) then yields a better approximation to the global Green's function matrix, $\underline{G}$.

III. Numerical Considerations

In implementing the above scheme there are a number of considerations, minor in theory, but important in practice, in determining the accuracy and speed of the calculation. We first look at the order of the error for fixed k's and small h. Using the expansion (Eq. 8) of the potential about the value at the center of the step we see that, for example, the integrals for $\Delta \underline{g}(h,h)$ become, to quadratic terms in the expansion of $\underline{U}$,

$$\lim_{h \to 0} \left[\Delta \underline{g}(h,h) \right]_{ij} = \frac{-1}{k_i^2 k_j^2 h^2} \int_0^h \left(1 - \frac{k_i^2 z^2}{2}\right)\left(1 - \frac{k_j^2 z^2}{2}\right)\left\{ U_{ij}^{(1)}\left(z - \frac{h}{2}\right) \right.$$
$$\left. + \left(U_{ij}^{(2)} / 2 \right)\left(z - \frac{h}{2}\right)^2 \right\} dz$$

$$= \frac{1}{k_i^2 k_j^2 h^2} \left\{ U_{ij}^{(2)} \left[\frac{h^3}{24} - \frac{(k_i^2 + k_j^2)h^5}{60} \right] + \frac{U_{ij}^{(1)}(k_i^2 + k_j^2)h^4}{24} \right\} \tag{12}$$

Since the zero order diagonal terms, $\underline{g}^0$, are of order h^{-1}, the leading $U^{(2)}$ term is of order h with respect to this. This term, which represents an average error of the potential due to the quadratic terms, can be taken into account exactly (to all

orders) by adding $\underline{V}^{(2)}(R_i)\frac{h^2}{24}$ to $V(R_i)$ before diagonalization, and subsequently substracting the perturbation integrals due to the constant perturbation, $\underline{U}^{(2)}\frac{h^2}{24}$. This shifts the eigenvalues to include the "average" $V^{(2)}$ contribution, and the 1st order perturbation correction no longer has terms of $O(h^2)$ with respect to the zero order. The lowest order (in h) terms neglected now are order h^4 or higher w.r.t. the h^{-1} of the zero order. They result from the 1st order p.t. contributions due to $\underline{U}^{(4)}$ and from mixed first and second derivative contributions in second order perturbation theory. Thus, for small step sizes the convergence is very rapid with step size.

As is usual with large matrix problems, however, the goal is to reduce the overall number of matrix operations required for a given level of accuracy. There are a rather large number of matrix operations per step in this method, particularly at the first energy where the potential evaluations, diagonalization, and transformations must be done. Specifically, the method requires the matrix operations shown in Table I per step. Also shown are the operations required in the unperturbed method.

Table I

Matrix Operation	1st E		Other E's	
	Unperturbed	Perturbed	Unperturbed	Perturbed
Diagonalization	1	1	0	0
Inversion and Multiplication $(A^{-1}\ B)$	1	1	1	1
Multiplications	3 (2.5)	8 (6)	2 (1.5)	3 (2)

( ) Using symmetry of R-matrix and potential

As can be seen, the first energy is significantly more work, but at subsequent energies the perturbed calculation requires only about 30% more work per step. As will be shown in the next section, the perturbed method is clearly superior for relatively accurate calculations.

A final consideration which may be of considerable importance in practice is the choice of step size to be used in building up the complete solution. Since the approach outlined here is a perturbative one, the size of the perturbation corrections must in some way be controlled to insure the accuracy of the solution. The leading order term in Eq. (12) after the eigenvalues have been shifted is proportional to $\underline{U}^{(1)}h^3$ relative to the zero order solution, so that the simplest step size algorithm is just to choose h such that h^3 times the largest element of $\underline{U}^{(1)}$ is kept less than some constant. While this algorithm will eventually insure convergence, the approach to the converged limit may not be at all smooth. Such an algorithm may also not be optimum since it is not sensitive to the partial cancellation of the perturbation integrals when hk is large. In addition, systems for which the potential goes through a minimum present special problems since the small value of $U^{(1)}$ near the minimum may predict too large a step beyond the minimum. Careful design of a step size algorithm to deal with these problems may ultimately lead to important improvements in the efficiency of the method.

IV. Results

To test the speed and efficiency of the R-matrix propagation method with perturbation corrections, calculations have been carried out on a well studied model rotational problem described by Lester and Bernstein.[8] The problem consists of an atom and a rigid rotor interacting through the potential

$$V(R, \theta) = \left(1 + a\, P_2(\cos\theta)\right) V(R) \tag{13}$$

with $V(R)$ a Lennard-Jones potential, and reduced parameters $a = 0.4$, $E/\epsilon = 1.5$, $r_m^2 \mu/I = 2.0$, $2\mu\epsilon\, r_m^2/\hbar^2 = 500$. Here E is the total energy, r_m and ϵ are the minimum and well depth of the Lennard-Jones, μ is the reduced mass of the atom relative to the rotor, and I is the moment of inertia of the rotor.

Although these parameters do not model any real system, even within the restriction of the potential to the form (13), they were chosen because they result in many small but not negligible transition probabilities together with a few large ones and are thus a sensitive test of the method, and because accurate results for this system are available in the literature.[4,8] The calculations reported here were performed for total angular momentum equal to 8 and including all channels with rotational quantum number upto and including 4, resulting in 9 coupled channels. With this many channels the time required to perform the calculations is essentially proportional to the number of matrix operations required. Integrations were carried out from 4.6 to 80 bohr.

In Fig. 1 the root mean square error in the 9x9 transition probability matrix is shown versus the number of matrix operations for the perturbed and unperturbed R-matrix propagation methods. The step size algorithm used is one which controls the size of certain perturbation matrix elements appearing in an integral equation approach to this problem, and may or may not be optimal for R-matrix calculations. However, it seems to be somewhat more efficient than the simple algorithm discussed in the last section. The calculations shown are then essentially second energy calculations: the number of equivalent matrix multiplications (not using symmetry), is $3^{1/3}$ for the unperturbed calculations and $4^{1/3}$ for the perturbed calculations. Also shown is the result using perturbation theory to correct only the diagonal elements of the sector R-matrix, which results in no additional matrix operations.

The "exact" answers were obtained by a 1400 step $(\sim 4200$ matrix operations) calculation by an integral equation method described later in this conference by Parker et al. It is estimated that they are accurate to at least r.m.s. error $\lesssim 10^{-10}$.

One disturbing feature of both the perturbed and unperturbed R-matrix calculations is the non-monotonic convergence toward the correct answer, particularly for large step sizes. (It should be noted that the smallest number of steps shown in Fig. 1 is 14 and that the perturbed method reaches the 10^{-4} level (r.m.s. error) at about 50 steps). For many scattering calculations an assured accuracy of three significant figures for the larger probability matrix elements and of two significant figures for the smaller

ones is sufficient. Unfortunately it is seen from Fig. 1 that in this regime (r.m.s. error 10^{-3} to 10^{-4}) the convergence is still quite oscillatory. This is probably due to the appearance in the recursion formula of the "squares" of the first order perturbation matrices uncompensated by higher order perturbation terms. Possible solutions to this problem are discussed in the next section.

IV. L^2 Corrections

As is well known there is an alternative representation of a sector Green's function in terms of the discrete complete set of eigenfunctions of the Hamiltonian plus Bloch operator. This is, of course, the basis of the standard L^2 R-matrix approach of Wigner and Eisenbud[7] and many others. The eigenfunctions are approximated by diagonalization of the Hamiltonian matrix in a finite basis (c.f. R. B. Walker's paper). The number of "translational" basis functions required per internal state depends on the region (step) size, and accuracy is often improved by adding the Buttle correction[9] which is just the $\underline{\underline{g}}^o$ for the set of discrete states not included in the L^2 basis used. We propose to "invert" the process and to use simple and small L^2 corrections to the perturbed $\underline{g}$ we have calculated in the last section in order to take the largest perturbation terms into account to all orders by diagonalization.

We assume, as is usual in L^2 methods, that we start with a complete direct product orthonormal basis in the internal and translational functions labeled by the two indices α,n respec-

tively. The exact Green's function can be represented in terms of the exact eigenfunctions of the Hamiltonian plus Bloch operator in this basis, and the Green's function matrix we desire is the projection onto the appropriate internal states evaluated at the boundaries as before. We take the internal basis to be that which diagonalizes the Hamiltonian at the center of the step, and the translational basis to be $(X = R - R_i + h/2$ as before)

$$|n\rangle = \begin{cases} h^{-1/2} & n = 0 \\ \sqrt{2/h} \, \cos(n\pi x/h) & n = 1, 2, \cdots \end{cases} \tag{14}$$

The matrix elements of the Hamiltonian in the direct product basis will be

$$\langle \alpha, n | H_\circ | \beta, m \rangle = \delta_{\alpha\beta} \, \delta_{nm} \left(\epsilon_\alpha + \frac{n^2 \pi^2}{h^2} \right)$$

$$\langle \alpha, n | \mathcal{L} | \beta, m \rangle = 0$$

$$\langle \alpha, n | V | \beta, m \rangle = V_{\alpha n}^{\beta m} \tag{15}$$

In a finite (truncated) internal basis $(0 \leq \alpha \leq M)$, the zero order Green's function is

$$g^\circ(\eta, x) = \sum_{\alpha=0}^{M} \sum_{n=0}^{\infty} \frac{|\alpha n\rangle \langle \alpha n|}{\epsilon_\alpha + \frac{n^2 \pi^2}{h^2} - E} = \sum_{\alpha=0}^{M} \frac{|\alpha\rangle \langle \alpha| \cos k_\alpha x \, \cos k_\alpha (x-h)}{k_\alpha \sin k_\alpha h} \tag{16}$$

where η is the set of arguments of the internal states $\{|\alpha\rangle\}$ This projects to the zero order diagonal Green's function matrix used before.

We now want to correct the Green's function matrix for only
some of the large perturbations, i.e. we do not want to diag-
onalize a very large Hamiltonian matrix in this basis as is
done in standard L^2 methods. The equation for the Green's
function is equivalent to that used before:

$$\left(H_o - E\right) g\left(\eta, \bar{x}\right) = \int(\eta - \eta')\delta(x-x') - V \, g\left(\eta, \eta'; x, x'\right) \tag{17}$$

We now add to both sides of the equation a non-local projection
operator containing the offending matrix elements which we want
to treat more accurately than in perturbation theory:

$$V^P = \sum_{\substack{\{\alpha n\} \\ \{\beta, w\}}} |\alpha n\rangle \langle \alpha n | V | \beta w \rangle \langle \beta w | \tag{18}$$

where we are free to choose which elements are included subject
only to the constraint that V^P must be Hermitian, i.e. for each
$\langle\alpha n|V|\beta m\rangle$ element included, its transpose must also be included.
Thus we have

$$\left(H_o - E + V^r\right) g\left(\eta, \eta'; x, x'\right) = \int(\eta - \eta')\delta(x - x') - \left(V - V^r\right) g\left(\eta, \eta'; x, x'\right) \tag{19}$$

We can now find the e-values and e-functions of the operator on
the l.h.s. by diagonalizing the matrix of $H_o - V^P$ in the basis
included in V^P only. Thus for this portion of the L^2 space, we
determine the eigenfunctions, $\{|\gamma\rangle\}$:

$$\left(H_o + V^r\right)|\gamma\rangle = \epsilon_\gamma|\gamma\rangle \tag{20}$$

The e-functions, $|\gamma\rangle$, are obtained by the diagonalizing transformation from the $|\alpha,n\rangle$ included in our basis:

$$|\gamma\rangle = \mathcal{J}|\alpha,n\rangle \tag{21a}$$

$$\mathcal{J}^T(H_o + V^P)\mathcal{J} = \epsilon_\gamma \tag{21b}$$

For example we might restrict V^P to one diagonal term in the internal states and two translational states, $|0\rangle$, $|1\rangle$; e.g.

$$V^P = |\alpha,0\rangle V_{\alpha 0}^{\alpha 0}\langle\alpha,0| + |\alpha,1\rangle V_{\alpha 1}^{\alpha 0}\langle\alpha,0|$$

$$+ |\alpha,1\rangle V_{\alpha 1}^{\alpha 1}\langle\alpha,1| + |\alpha,0\rangle V_{\alpha 0}^{\alpha 1}\langle\alpha,1|$$

In this case the matrix $H_0 + V^P$ is diagonal except for the $|\alpha,0\rangle$, $|\alpha,1\rangle$ block:

$$\underline{\underline{H}}_o + \underline{\underline{V}}^P = \begin{pmatrix} \ddots & & & \\ & \epsilon_\alpha + V_{\alpha 0}^{\alpha 0} & V_{\alpha 0}^{\alpha 1} & & 0 \\ & V_{\alpha 1}^{\alpha 0} & \epsilon_\alpha + \frac{\pi^2}{h^2} + V_{\alpha 1}^{\alpha 1} & \\ & 0 & & \ddots \end{pmatrix}$$

$$\underline{\underline{J}} = \begin{pmatrix} \underline{\underline{I}} & & 0 \\ & \underline{\underline{T}}_{0,1}^{\alpha} & \\ 0 & & \underline{\underline{I}} \end{pmatrix}$$

and the Green's function for $H_0 + V^P$ would be the diagonal Green's function matrix for all states but α , with the $\underline{\underline{g}}_{\alpha\alpha}^0$ element replaced by

$$g^{\flat}_{\alpha\alpha} = g^{o}_{\alpha\alpha} + \sum_{\gamma=1}^{2} \frac{|\gamma\rangle\langle\gamma|}{\epsilon_{\gamma} - E} - \sum_{n=0}^{1} \frac{|\alpha n\rangle\langle\alpha n|}{\epsilon_{\alpha} + \frac{n^{2}\pi^{2}}{h^{2}} - E}$$

where $|\gamma\rangle$ is the e-fn of the $|\alpha\phi\rangle, |\alpha 1\rangle$ block of $H_0 + V^P$.

The "zero order" Green's function matrix is now taken to be that determined by the homogeneous soln of Eq. (19):

$$\left(\underset{\approx}{H}_o + \underset{\approx}{V}^P - E \underset{\approx}{I} \right) \underset{\approx}{g}^{\flat} = \underset{\approx}{I} \, \delta(x - x') \tag{22}$$

This Green's function matrix, $\underset{\approx}{g}^D$, is given by

$$\underset{\approx}{g}^{P} = \underset{\approx}{g}^{o} + \underset{\approx}{g}_{\gamma} - \underset{\approx}{g}_{n} \tag{23a}$$

where $\underset{\approx}{g}^o$ is given by Eq. (9),

$$\left[\underset{\approx}{g}_{\gamma} \right]_{\alpha\beta} = \langle\alpha| \sum_{\gamma} \frac{|\gamma\rangle\langle\gamma|}{\epsilon_{\gamma} - E} |\beta\rangle \tag{23b}$$

$$\left[\underset{\approx}{g}_{n} \right]_{\alpha\beta} = \langle\alpha| \sum_{\{\alpha,n\}} \frac{|\alpha n\rangle\langle\alpha n|}{\epsilon_{\alpha} + \frac{n^{2}\pi^{2}}{h^{2}} - E} |\beta\rangle$$

$$= \delta_{\alpha\beta} \sum_{\{n\}} \frac{|n\rangle\langle n|}{\epsilon_{\alpha} + \frac{n^{2}\pi^{2}}{h^{2}} - E} \tag{23c}$$

That this is, in fact, the Green's function matrix for Eq. (23) can be seen by recognizing that $H_0 + V^P - E$ operating on a translational basis function not in $\{\alpha,n\}$, ($\{\tilde{\alpha},n\}$), yields the same as $H_0 - E$. Thus we have, in the (n,X) representation,

$$\left(H_{\circ} + V^{P} - E \right) | \gamma \rangle = \left(\epsilon_{\gamma} - E \right) | \gamma \rangle$$

$$\left(H_{\circ} + V^{P} - E \right) | \widetilde{\alpha n} \rangle = \left(\frac{n^2 \pi^2}{h^2} + \epsilon_{\alpha} - E \right) | \widetilde{\alpha n} \rangle \qquad (24)$$

$$= \left(\epsilon^{\circ}_{\widetilde{\alpha n}} - E \right) | \widetilde{\alpha n} \rangle$$

and thus

$$\left(H_{\circ} + V^{P} - E \right) \left\{ g^{\circ} + \sum \frac{|\gamma \rangle \langle \gamma|}{\epsilon_{\gamma} - E} - \sum_{\{\alpha n\}} \frac{|\alpha n \rangle \langle \alpha n|}{\epsilon^{\circ}_{\alpha n} - E} \right\}$$

$$= \sum_{\{\widetilde{\alpha n}\}} \left(\epsilon^{\circ}_{\widetilde{\alpha n}} - E \right) \frac{|\widetilde{\alpha n} \rangle \langle \widetilde{\alpha n}|}{\epsilon^{\circ}_{\widetilde{\alpha n}} - E} + \sum \left(\epsilon_{\gamma} - E \right) \frac{|\gamma \rangle \langle \gamma|}{\epsilon_{\gamma} - E}$$

$$= \sum_{\{\widetilde{\alpha n}\}} |\widetilde{\alpha n} \rangle \langle \widetilde{\alpha n}| + \sum |\gamma \rangle \langle \gamma| = \delta(\gamma - \gamma') \delta(x - x')$$

We note the last equality holds if the basis $|\alpha n \rangle + |\widetilde{\alpha n} \rangle =$ $|\gamma \rangle + |\widetilde{\alpha n} \rangle$ is complete. Since the $|\alpha \rangle$ basis is, in practice, truncated, it holds in matrix form in the truncated projection of the internal states, $|\alpha \rangle$.

We now return to the exact equation (19) for g, and solve it formally using g^{D} as our "zero order" Green's function. In what follows we omit the matrix notation and write only $\int g^{D} V g$, for example, to mean

$$\int_{R_i - \frac{h}{2}}^{R_i + h/2} g^{D}(x, z) \, \underline{V}(z) \, g(z, x') \, dz$$

The integral equation for g is

$$g = g^p - \int g^p (V - V^p) g \tag{25}$$

Iterating once we have

$$g^{1D} = g^0 + g_\gamma - g_n - \int (g^0 + g_\gamma - g_n)(V - V^p)(g^0 + g_\gamma - g_n) \tag{26}$$

Since V^p has no projection on g^0-g_n, this simplifies somewhat to

$$g^{1D} = g^0 + g_\gamma - g_n - \int (g^0 + g_\gamma - g_n) V (g^0 + g_\gamma - g_n) + \int g_\gamma V^p g_\gamma \tag{27}$$

We note that if V^p contained __all__ the matrix elements of V in the L^2 basis, $\{|\alpha\, n\rangle\}$, then $V^p = V^{cp}$ (the complete projection of V), and V-V^{cp} would have __no__ projection on g_γ or g_n, leading to

$$g^{1D} = g^0 + g_\gamma - g_n - \int g^0 V g^0 + \int g_n V^{cp} g_n$$
$$- \int (g^0 - g_n) V (g_\gamma - g_n) - \int (g_\gamma - g_n) V (g^0 - g_n) \tag{28}$$

In this equation the original perturbation is corrected by eliminating the perturbation matrix elements used in the diagonalization ($\int g_n V^{cp} g_n$), and by eliminating the untransformed off-diagonal perturbations $\left[\ \int (g^0 - g_n) V\ g_n\ \right]$ and replacing them with the transformed off-diagonal perturbations $\left[\int (g^0 - g_n) V\ g_\gamma\right]$. When $V^p \neq V^{cp}$, additional terms arise:

$$g^{ip} = g^o + g_\gamma - g_n - \int g^o V g^o + \int g_n V^p g_n$$
$$- \int (g^o - g_n) V (g_\gamma - g_n) - \int (g_\gamma - g_n) V (g^o - g_n)$$
$$+ \int g_n (V^{cp} - V^p) g_n - \int g_\gamma (V^{cp} - V^p) g_\gamma \qquad (29)$$

Equation (29) is a generalization of the Buttle correction in two ways. First, it corrects for the fact that V^p, not V^{cp}, is used to determine g_γ, and second, it includes the perturbation corrections for the infinite set of $|\tilde{\alpha} n\rangle$ of translational functions not included in the L^2 basis.

In practice the use of Eq. (29) would not be feasible because it introduces corrections to off-diagonal elements as a result of the L^2 correction to a single diagonal element. Since it is usually only diagonal elements that become large enough that the continuous perturbation theory breaks down, we would like to have an approximate method of correcting a diagonal element alone. Such a method may be derived as follows. Equation (11) gives the continuous perturbation correction to g^o from all of V through 1st order only as

$$\Delta g = - \int g^o V g^o$$

while Eq. (23a) gives the L^2 correction to g^o from V^p only to all orders of perturbation theory as

$$\Delta g' = g^p - g^o = g_\gamma - g_n$$

If we simply add these two corrections to g^o we include both
effects at once, but we see that the 1st order correction due
to V^p has been counted twice, once in each correction. We
must therefore subtract it off, yielding the following approxi-
mate equation for g:

$$g \approx g^o - \int g^o V g^o + g_\gamma - g_n + \int g_n V' g_n \quad (30)$$

We note that this has the first five terms of Eq. (29) omitting
only the (small) cross terms. Preliminary tests using only
diagonal (in α) corrections and only the (0,1) translational
states indicate that it sometimes helps a lot (a factor of 10
in the r.m.s. error), but investigations are still in progress.

Appendix:

Let $V_{ij}^{(n)}$ be the i-j^{th} element of the nth derivative of $2/\mu$
times the potential (i.e. the potential in K^2 units). Define
$c_{ij}^{(n)}$ to be

$$C_{ij}^{(n)} = \frac{-V_{ij} \, h^{n+1}}{n! \, k_i \, k_j \, \sin z_i \, \sin z_j}$$

where h is the step length, K_i and K_j are the wave vectors for
the channels, and $z_i = hK_i$. The perturbations to the r-matrices
are then given as

$$\left(\Delta r_4^{(n)} \right)_{ij} = C_{ij}^{(n)} \; \frac{1}{h^{n+1}} \int_0^h \cos(k_i x) \cos(k_j x) \left(x - \frac{h}{2}\right)^n dx$$

$$\left(\Delta r_1^{(n)} \right)_{ij} = (-1)^n \left(\Delta r_4^{(n)} \right)_{ij}$$

$$\left(\Delta r_2^{(n)} \right)_{ij} = C_{ij}^{(n)} \; \frac{1}{h^{n+1}} \int_0^h \cos k_i (x-h) \cos k_j x \; \left(x - \frac{h}{2}\right)^n dx$$

$$\left(\Delta r_3^{(n)} \right)_{ij} = \left(\Delta r_2^{(n)} \right)_{ji}$$

Thus we need evaluate only two integrals for each n, which can
be written in dimensionless form as (let $y = x/h$)

$$\left(I_1^{(n)} \right)_{ij} = \int_0^1 \left(y - \frac{1}{2}\right)^n \cos z_i y \; \cos z_j y \; dy$$

$$\left(I_2^{(n)} \right)_{ij} = \int_0^1 \left(y - \frac{1}{2}\right)^n \cos z_i (y-1) \cos z_j y \; dy$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} (y)^n \cos z_i \left(y - \frac{1}{2}\right) \cos z_j \left(y + \frac{1}{2}\right) dy$$

We also define the following quantities:

$$X = h^2 \left(k_i^2 + k_j^2 \right)$$

$$Y = h^2 \left(k_i^2 - k_j^2 \right)$$

$$S = h \left(k_i + k_j \right)$$

$$D = h \left(k_i - k_j \right)$$

The explicit expressions for the integrals $I_1^{(n)}$ and $I_2^{(n)}$ are given for n = 0,1,2 for i=j, for i≠j, for $K_i \gtrsim K_j$, and for K_i and K_j small on the next pages.

0.1 $\left(I_1^{(0)}\right)_{ii} = \frac{1}{2}\left[1 + (1/z_i)\sin z_i \cos z_i\right]$

$\left(I_1^{(0)}\right)_{ij} = \frac{1}{2}\left[(1/S)\sin S + (1/D)\sin D\right]$

$\qquad = \frac{1}{y}\left[z_i \sin z_i \cos z_j - z_j \sin z_j \cos z_i\right]$

0.1 a) $D \to 0$

$\left(I_1^{(0)}\right)_{ij} = (1/S)\sin(S/2)\cos(S/2) + \frac{1}{2} - \frac{D^2}{12}$
$\qquad\qquad\qquad + D^4/240 - D^6/10080$

b) $z_i, z_j \to 0$

$\left(I_1^0\right)_{ij} = 1 - (z_i^2 + z_j^2)/6 + (z_i^4 + z_j^4 + 6z_i^2 z_j^2)/120$
$\qquad\qquad - (z_i^6 + z_j^6 + 15 z_i^4 z_j^2 + 15 z_i^2 z_j^4)/5040$

0.2 $\left(I_2^{(0)}\right)_{ii} = \frac{1}{2}\left[\cos z_i + (1/z_i)\sin z_i\right]$

$\left(I_2^{(0)}\right)_{ij} = (1/S)\sin(S/2)\cos(D/2) + (1/D)\sin(D/2)\cos(S/2)$

$\qquad = \frac{1}{y}\left[z_i \sin z_i - z_j \sin z_j\right]$

0.2 a) $D \to 0$

$$(I_2^{(0)})_{ij} = \cos(S/2)\left[\frac{1}{2} - \frac{D^2}{48} + \frac{D^4}{3840} - \frac{D^6}{645\,120}\right]$$

$$+ \frac{\sin(S/2)}{S}\left[1 - \frac{D^2}{8} + \frac{D^4}{384} - \frac{D^6}{46\,080}\right]$$

b) $z_i, z_j \to 0$

$$(I_2^{(0)})_{ij} = 1 - (z_i^2 + z_j^2)/6 + (z_i^4 + z_j^4 + z_i^2 z_j^2)/120$$

$$- (z_i^6 + z_j^6 + z_i^4 z_j^2 + z_i^2 z_j^4)/5\,040$$

<u>1.1</u> $(I_1^{(1)})_{ii} = (1/4z_i)\sin z_i \cos z_i - (1/4z_i^2)\sin^2 z_i$

$$(I_1^{(1)})_{ij} = (1/2)\left[\frac{\cos S}{S^2} + \frac{\sin S}{2S} - \frac{1}{S^2} + \frac{\cos D}{D^2}\right.$$

$$\left. + \frac{\sin D}{2D} - \frac{1}{D^2}\right]$$

$$= \frac{1}{y^2}\left[X(\cos z_i \cos z_j - 1) + 2z_i z_j \sin z_i \sin z_j\right.$$

$$\left. + \frac{y}{2}\left(z_i \sin z_i \cos z_j - z_j \sin z_j \cos z_i\right)\right]$$

1.1 a) $D \to 0$

$$(I_1^{(1)})_{ij} = -(1/S^2)\sin^2(S/2) + (1/2S)\sin(S/2)\cos(S/2)$$

$$- D^2/48 + D^4/720 - D^6/26\,880$$

b) $z_i, z_j \to 0$

$$(I_1^{(1)})_{ij} = -(z_i^2 + z_j^2)/24 + (z_i^4 + z_j^4)/360 + z_i^2 z_j^2/60$$

$$- (z_i^6 + z_j^6)/13\,440 - (z_i^2 + z_j^2)z_i^2 z_j^2/896$$

__1.2__ $(I_2^{(1)})_{ii} = 0$

$$(I_2^{(1)})_{ij} = -\frac{\sin(D/2)}{2}\left(\frac{\cos(S/2)}{S} - \frac{2\sin(S/2)}{S^2}\right)$$

$$- \frac{\sin(S/2)}{2}\left(\frac{\cos(D/2)}{D} - \frac{2\sin(D/2)}{D^2}\right)$$

$$= -(X/y^2)(\cos z_i - \cos z_j) - (1/2y)(z_i \sin z_i + z_j \sin z_j)$$

1.2 a) $D \to 0$

$$(I_2^{(1)})_{ij} = -\frac{D}{2}\left\{ \sin(S/2)\left[-\frac{1}{12} + D^2/480 - D^4/53760\right]\right.$$

$$+ \frac{\cos(S/2)}{S}\left[1/2 - D^2/48 + D^4/3840\right]$$

$$\left. + \frac{\sin(S/2)}{S^2}\left[-1 + D^2/24 - D^4/1920\right]\right\}$$

1.2 b) $z_i, z_j \to 0$

$$(I_2^{(1)})_{ij} = -y\left[-1/24 + (z_i^2 + z_j^2)/360 - \frac{(3z_i^4 + 3z_j^4 + 4z_i^2 z_j^2)}{40\,320}\right]$$

$$\underline{2.1} \quad (I_1^{(2)})_{ii} = \frac{1}{24} + \frac{\sin z_i \cos z_i}{8 z_i} + \frac{\cos^2 z_i}{4 z_i^2} - \frac{\sin z_i \cos z_i}{4 z_i^3}$$

$$(I_1^{(2)})_{ij} = \frac{\sin S}{8 S} + \frac{(\cos S + 1)}{2 S^2} - \frac{\sin S}{S^3}$$

$$+ \frac{\sin D}{8 D} + \frac{(\cos D + 1)}{2 D^2} - \frac{\sin D}{D^3}$$

$$= \frac{1}{\gamma^3} \left\{ -4X \left(z_i \sin z_i \cos z_j - z_j \sin z_j \cos z_i \right) \right.$$

$$+ 2\gamma \left(z_i \sin z_i \cos z_j + z_j \sin z_j \cos z_i \right) + X\gamma \left(1 + \cos z_i \cos z_j \right)$$

$$\left. + 2\gamma \left(z_i \sin z_i z_j \sin z_j \right) + \frac{\gamma^2}{4} \left(z_i \sin z_i \cos z_j - z_j \sin z_j \cos z_i \right) \right\}$$

2.1 a) $D \to 0$

$$(I_1^{(2)})_{ij} = \frac{\sin(S/2)\cos(S/2)}{4 S} + \frac{\cos^2(S/2)}{S^2} - \frac{2\sin(S/2)\cos(S/2)}{S^3}$$

$$+ \frac{1}{24} - \frac{D^2}{120} + \frac{11 D^4}{20\,160} - \frac{11 D^6}{725\,760}$$

2.1 b) $z_i, z_j \to 0$

$$(I_1^{(2)})_{ij} = \frac{1}{12} - \frac{z_i^2 + z_j^2}{60} + \frac{11(z_i^4 + z_j^4 + 6 z_i^2 z_j^2)}{10\,080}$$

$$- 11(z_i^6 + z_j^6 + 15 z_i^4 z_j^2 + 15 z_i^2 z_j^4)/362\,880$$

$\underline{2.2}$ $\quad (I_2^{(2)})_{ii} = \left(\frac{1}{24} + \frac{1}{4 z_i^2} \right) \cos z_i + \left(\frac{1}{8 z_i} - \frac{1}{4 z_i^3} \right) \sin z_i$

$(I_2^{(2)})_{ij} = \cos(D/2) \left[\frac{\cos(S/2)}{S^2} + \frac{\sin(S/2)}{4S} - \frac{2 \sin(S/2)}{S^3} \right]$

$\qquad + \cos(S/2) \left[\frac{\cos(D/2)}{D^2} + \frac{\sin(D/2)}{4D} - \frac{2 \sin(D/2)}{D^3} \right]$

$\qquad = \frac{1}{\gamma^3} \left\{ \left(-4X + \frac{\gamma^2}{4} \right) \left(z_i \sin z_i - z_j \sin z_j \right) \right.$

$\qquad \left. + X\gamma \left(\cos z_i + \cos z_j \right) + 2\gamma \left(z_i \sin z_i + z_j \sin z_j \right) \right\}$

$2.2 \, a) \quad D \to 0$

$(I_2^{(2)})_{ij} = \left[\sin(S/2) \left(\frac{1}{4S} - \frac{2}{S^3} \right) + \frac{\cos(S/2)}{S^2} \right] \left[1 - \frac{D^2}{8} + \frac{D^4}{384} - \frac{D^6}{46\,080} \right]$

$\qquad + \cos(S/2) \left[\frac{1}{24} - \frac{D^2}{320} + \frac{D^4}{21\,504} - \frac{D^6}{3\,317\,760} \right]$

$2.2 \, b) \quad z_i, z_j \to 0$

$(I_2^{(2)})_{ij} = \frac{1}{12} - \frac{z_i^2 + z_j^2}{60} + \frac{11 z_i^4 + 11 z_j^4 + 3 z_i^2 z_j^2}{10\,080}$

$\qquad - \frac{11 z_i^6 + 11 z_j^6 + 3 z_i^4 z_j^2 + 3 z_i^2 z_j^4}{362\,880}$

References

1. E. B. Stechel, T. G. Schmalz, and J. C. Light, "Quantum Theory of Exchange Reactions: Use of Nonorthogonal Bases and Coordinates," J. Chem. Phys. (to appear June 15, 1979).

2. D. J. Zvijac and J. C. Light, Chem. Phys. $\underline{12}$, 237 (1976).

3. J. C. Light and R. B. Walker, J. Chem. Phys. $\underline{65}$, 4272 (1976).

4. E. B. Stechel, R. B. Walker, and J. C. Light, J. Chem. Phys. $\underline{69}$, 3518 (1978).

5. B. I. Schneider and R. B. Walker, J. Chem. Phys. $\underline{70}$, 2466 (1979).

6. C. Bloch, Nucl. Phys. $\underline{4}$, 503 (1957).

7. The L^2 method was originally proposed by E. P. Wigner, Phys. Rev. 70, 15, 606 (1946); E. P. Wigner and L. Eisenbud, ibid. $\underline{72}$, $2\overline{9}$ (1947).

8. W. A. Lester, Jr., and R. B. Bernstein, J. Chem. Phys. $\underline{48}$, 4896 (1968).

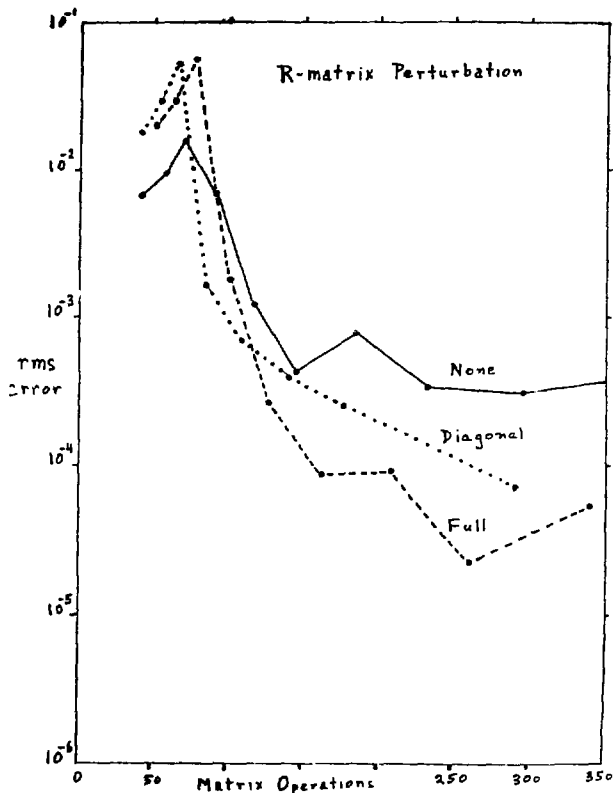9. P. J. A. Buttle, Phys. Rev. $\underline{160}$, 719 (1967).

Fig. 1: Root mean square error of the probability matrix vs.
the number of equivalent matrix multiplications for the
unperturbed, diagonally perturbed, and full perturbed
R-matrix propagation. The problem is the model Lester-
Bernstein[8] 9-state rotational problem described in
the text.

FINITE ELEMENT METHODS IN QUANTUM DYNAMICS

Herschel Rabitz
Department of Chemistry
Princeton University
Princeton, New Jersey 08544

I. Introduction

Traditional quantum scattering theory has mainly been based on solving the close coupling equations derived from the original Schrodinger equation. Although the basic coupled equations are quite standard, it is useful for our purposes here to briefly review the matter. The Schrodinger equation can be expressed as

$$H(R,r_1,r_2,...) \ \psi(R,r_1,r_2,...) \ = \ E\psi(R,r_1,r_2,...) \tag{1}$$

where R, r_1, r_2,... are the coordinates for the system. A particular coordinate R (usually referred to as the radial translational coordinate) is singled out. Eq.(1) is a many variable second order partial differential equation. In order to reduce it further the wavefunction is expanded in a complete set of functions, in the coordinates r_1, r_2,...

$$\psi \ = \ \sum_n a_n(R) \ \phi_{n_1}(r_1) \ \phi_{n_2}(r_2) \ ... \tag{2}$$

where the unknown coefficients $a_n(R)$ are obtained by solving the following set of coupled <u>ordinary</u> differential equations.

$$\underline{\underline{H}}(R) \ \underline{a}(R) \ = \ E \ \underline{a}(R) \tag{3}$$

The chief difficulty in solving these equations is associated with the matrix algebra during their numerical solution. This expense grows as $\sim N^3$ where N is the total number of terms in the expansion of Eq.(2).

The philosophy behind the finite element method is based on solving fewer coupled equations than above, and correspondingly replacing the matrix algebra by labor involved in solving partial differential equations. The basic idea is apparent by considering the alternate expansion of the wavefunction in Eq.(4)

$$\psi = \sum_n b_n(R,r_1) \phi_{n_2}(r_2) \ldots \tag{4}$$

where an additional coordinate r_1 has been singled out to make the expansion coefficients a function of two variables. These coefficients are obtained by solving the following set of coupled _partial_ differential equations.

$$\underline{\underline{H}}(R,r_1) \underline{b}(R,r_1) = E \underline{b}(R,R_1) \tag{5}$$

The potential for savings through Eq.(5) over Eq.(3) is apparent if the r_1 coordinate is assumed to require ten basis functions in the expansion of Eq.(2). In this case its contribution to the computational expense is approximately a factor of 10^3. Provided the numerical methods for treating Eq.(5) can stay within this factor of 10^3, a computational advantage can be achieved.

The availability of new finite element methods borrowed from the engineering disciplines[1] shows considerable promise for the partial differential equation approach. This technique is presently under development for applications to quantum mechanics, and a full assessment of its usefulness cannot be made at this time. However, applications to certain quantum mechanical problems have already been made as will be discussed

below. The aim of this presentation is to give a status report on the
present development of finite element methods in quantum mechanics and also
indicate possible further directions the field is expected to go in.

II. Essentials of the Finite Element Method

Various methods for treating partial differential equations have been
suggested in the literature, and perhaps most notable amongst these is the
finite difference method. There are certain suggestive similarities between
the finite difference and finite element methods. However, the two methods
differ significantly, and this point is clearly made in a paper by Morton.[2]
In essence, the finite difference method attempts to approximate the
differential operators while the finite element method directly approximates
the sought after function. The difference between these two approaches will
become particularly clear in a specialized application arising in three
dimensional scattering discussed below. It suffices to say at this point
that the same questions on the choice of partial differential equation method
have been considered in related engineering problems (e.g., water wave
scattering), and ultimately the finite element method was the technique of
choice.

The basic theory of the finite element method is apparent from
considering the Schrodinger equation in two variables, x, y. Indeed, such
a partial differential equation in two variables will exactly describe a
collinear reactive system. Higher variable equations will be discussed
later. The Schrodinger equation in these two variables is

$$\frac{-\hbar^2}{2m} \nabla^2 \phi + V\phi = E\phi \tag{6}$$

In order to conveniently derive the finite element equations, consider the standard variational functional

$$I = \int_A \phi(H - E) \phi dA \tag{7}$$

where A is the domain of interest in the x-y plane. Setting the first variation of I equal to zero we obtain

$$\delta I = \int_A \delta\phi \left[\frac{-\hbar^2}{2m} \nabla^2 + V - E\right]\phi \, dA = 0 \tag{8}$$

Gauss's theorem applied to Eq.(8) yields

$$0 = \int_A \left[\frac{\hbar^2}{2m} \vec{\nabla}\delta\phi \cdot \vec{\nabla}\phi + (V - E)\phi\right] dA - \frac{\hbar^2}{2m} \oint_L \delta\phi \frac{\partial\phi}{\partial n} \, dS \tag{9}$$

where the last integral is integrated over the boundary L of the x-y domain. The derivative in the integrand in the latter term is normal to the boundary. It is convenient to treat the scattering event as a boundary value problem whereby $\delta\phi = 0$ on the boundary since ϕ is prescribed along L. Therefore the boundary integral vanishes. The overall area A is now broken into <u>finite element</u> areas A_k

$$A = \sum_A A_k \tag{10}$$

In this fashion the total variation in Eq.(8) can now be expressed as

$$\delta I = 0 = \sum_k \delta I_k \tag{11}$$

where

$$\delta \bar{I}_k = \int_{A_k} \left[\frac{\hbar^2}{2m} \vec{\nabla} \delta \phi_k \cdot \vec{\nabla} \phi_k + \delta \phi_k (V - E) \phi_k \right] dA \tag{12}$$

The shape of the areas A_k is arbitrary, but generally triangles or quadralaterals are utilized, as shown in Figure 1. A topological covering based on the use of triangles seems to be the most convenient for general application. An overall meshing for a reactive problem is illustrated in Figure 2, where the element numbering scheme is also shown.[3] The actual generation of a geometrical covering and numbering scheme can be done manually or in an automated fashion.

The unknowns in the finite element method are the wavefunction ϕ at points along the boundaries between the finite elements. Since ϕ exists throughout the region of each element A_k, an interpolating function is needed in each element. Polynomials or various transcendental functions could be used, but the former is usually the choice for convenience and simplicity. Consideration of a particular element in further detail is shown in Figure 3. A general theory for handling the algebra of any arbitrary element is desired and this naturally leads to the consideration of oblique local coordinates ξ_1, ξ_2 which can be related to the cartesian coordinates x,y. Details of these transformations are given in the literature,[3] and it suffices here to indicate that the arbitrary point P

in Figure 3 can be located in reference to the oblique coordinates by appropriate projections (the dashed lines in the figure). In addition, for reasons of symmetry it is conventional to introduce the third redundant coordinate ξ_3. As an illustration of these ideas

$$\phi_k(\xi_1,\xi_2,\xi_3) = \sum_{\ell=1}^{6} U_\ell(\xi_1,\xi_2,\xi_3)\phi_{k\ell} = \underline{U}\cdot\underline{\phi}_k \equiv <U|\phi_k> \tag{13}$$

gives a quadratic polynomial representation of the wavefunction in the k-th element. It is convenient to use a scaler product notation as in Eq.(13) where the components of the vector $\underline{U}$ are the local interpolating polynomial given in

$$U = \left[\xi_1(2\xi_1-1),\ \xi_2(2\xi_2-1),\ \xi_3(2\xi_3-1),\ 4\xi_1\xi_2,\ 4\xi_2\xi_3,\ 4\xi_3\xi_1 \right] \tag{14}$$

In a similar fashion the potential in the k-th element can also be expanded in the local interpolating polynomial, although this is strictly not an essential part of the finite element treatment. If this expansion is performed, the resultant potential is expressed as

$$V_k(\xi_1,\xi_2,\xi_3) = \sum_{\ell=1}^{6} U_\ell(\xi_1,\xi_2,\xi_3)V_{k\ell} = <U|V_k> \tag{15}$$

The first variation in the k-th element can be expanded out in terms of the x,y coordinates

$$\delta I_k = \int_{A_k} \left[\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial x}\ \delta\phi_k\ \frac{\partial}{\partial x}\ \phi_k + \frac{\partial}{\partial y}\ \delta\phi_k\ \frac{\partial}{\partial y}\ \phi_k \right) + \delta\phi_k(V - E)\phi_k \right]dA \tag{16}$$

When each function in the integrand of Eq.(16) is represented in its polynomic form we obtain

$$\phi_k(x,y) = <U|\phi_k>$$

$$\frac{\partial \phi_k}{\partial x} = <\frac{\partial U}{\partial x}|\phi_k> \qquad \text{etc.,} \tag{17}$$

all the integrals can be expressed in terms of the following analytic integral expression

$$\int_{A_k} \epsilon_1^p \epsilon_2^q \epsilon_3^r \, dA = \frac{p! \; q! \; r!}{(p + q + r + 2)!} \, 2A_k \tag{18}$$

Using the compact bra-ket notation the following expression is obtained

$$\delta I_k = <\delta\phi_k \int_{A_k} \left[\frac{\hbar^2}{2m} |\vec{\nabla}U> \cdot <\vec{\nabla}U| + |U>(<V_k|U> - E)<U| \right] \phi_k> \, dA \tag{19}$$

which can be evaluated once and for all regardless of the element index k. In matrix notation this result can be further expressed as

$$\delta I_k = <\delta\phi_k|H_k - EU_k|\phi_k> \tag{20a}$$

$$H_k = \int_{A_k} \left[\frac{\hbar^2}{2m} |\vec{\nabla}U> \cdot <\vec{\nabla}U| + |U><V_k|U><U| \right] \, dA \tag{20b}$$

$$U_k = \int_{A_k} |U><U| \, dA \tag{20c}$$

The unknown nodal values of the wavefunction $\phi_{k\ell}$ can now be obtained by demanding continuity across element boundaries. In practice this is achieved by the generation of a continuity matrix which shows the connection between the global and local numbering schemes for the nodal wavefunctions. As a result of this treatment an overall Euler equation is obtained

$$\sum_k \delta I_k = 0$$

(21)

$$(\underline{\underline{H}} - E\underline{\underline{U}}) \cdot \underline{\phi} = 0$$

where $\underline{\underline{H}}$ and $\underline{\underline{U}}$ are the finite element representations of the Hamiltonian and unit operator, respectively. These latter matrices are strictly independent of the energy E which enters simply as shown in Eq.(21). Therefore changing the total energy does not require a regeneration of the finite element matrices.

A pictorial representation of Eq.(21) is shown in Figure 4a where the matrix is non-squar due to the fact that the N_R nodal values on the boundary are actually specified in the problem as known. Utilizing the latter fact, however, the equations in Figure 4a can be rewritten as an inhomogeneous set of linear equations in terms of the $N - N_B$ unknown nodal points for the interior of the scattering region. The choice of a linearly independent set of boundary conditions generates a series of equations like Figure (4b) with different inhomogeneous terms. This is illustrated in Figure (4c) which is ultimately solved to yield scattering information. The $\underline{\underline{H}} - E\underline{\underline{U}}$ matrix in Figure (4c) is banded with a width related to the connectivity of the nodal elements and the order of the interpolating polynomial.

III. Application to H + H₂ Collinear Reactive Scattering

As a test of the above formalism, reactive calculations were performed
for H + H₂ on the Porter-Karplus surface. Details of this calculation are
in the literature[3] and the results are shown as the points labelled + in
Figure 5. The other curves and points are various close coupling calculations.
The finite element results are quite good, but more important was the
behavior observed concerning the inclusion of closed channels in the boundary
region. It was found that the finite element calculations did not change
regardless of whether closed channels were included in the outer region. In
contrast to this result, conventional close coupling on this problem has
required the use of closed channels to achieve conversion. This behavior
points out that the finite element wavefunctions inherently include an
arbitrary number of open and closed channels in the interior region of
importance. This is achieved by not explicitly attempting an eigenfunction
expansion as discussed in Section I.

IV. Higher Dimensional Problems

Problems other than collinear scattering will inherently involve more
than two coordinates. There are basically two ways to proceed.

      1.   A single partial differential equation in all
           coordinates.

      2.   Coupled partial differential equations in fewer
           coordinates.

The same trade-offs discussed in Section I are involved in deciding which
of these two is optimal for a particular class of problems. At this stage

the outcome of this competition is not entirely clear, but the nature of the trade-off is best illustrated by a few examples. Considering co-planar atom-diatom scattering there is a total of six coordinates from which two are subtracted for conservation of center of mass motion and one for conservation of angular momentum. This leaves a total of three independent coordinates such that the Schrodinger equation could be treated as a single three coordinate partial differential equation. An alternative would be to expand the angular dependence in a basis and treat coupled two coordinate partial differential equations.

The critical need to reduce the Schrodinger equation to its minimum number of independent coordinates is most clearly illustrated in atom-diatom three dimensional scattering. In this case there should be four independent coordinates taking into account the conservation relations. It is a trivial matter to remove the conservation of center of mass motion. However, the reduction associated with the conservation of angular momentum is not a trivial exercise. In normal close coupling this question is a moot point since the conservation of angular momentum is explicitly built into the choice of basis functions. The problem now concerns finding how conservation of angular momentum manifests itself in the partial differential operator. This problem is very similar to that arising in classical mechanics where conservation of angular momentum is often most conveniently treated in the action-angle formulation. We have found it expedient to proceed in a similar fashion in the quantum case, and this results in the following Hamiltonian operator

$$H = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial R^2}\right) + V(r,R,\gamma) + \hbar^2 \, J(J+1)/2\mu R^2$$

$$+ \left[-\hbar^2 \frac{\partial^2}{\partial\gamma^2} - \hbar^2 \cot\gamma \frac{\partial}{\partial\gamma} + \csc^2\gamma \; \hat{\omega}^2\right]\left(\frac{1}{2m\gamma^2} + \frac{1}{2\mu R^2}\right)$$

$$- \left\{\hat{\omega}^2 - \tfrac{1}{2}e^{iq_\omega} \, \lambda_+^{J\hat{\omega}}\left[\frac{\partial}{\partial\gamma} + \cot\gamma(\hat{\omega}+1)\right] + \tfrac{1}{2}e^{-iq_\omega} \, \lambda_-^{J\hat{\omega}}\left[\frac{\partial}{\partial\gamma} - \cot\gamma(\hat{\omega}-1)\right]\right\}\bigg/\mu R^2$$

(22)

where R is the distance between the atom–diatomic, r is the diatomic vibrational coordinate and γ is the angle between r and R. The particular operators of interest in this equation are

$$\lambda_{\pm}^{J\hat{\omega}} = \left[\hbar^2 \, J(J + 1) - \hat{\omega}(\hat{\omega} \pm 1)\right]^{\frac{1}{2}}$$ (23)

where

$$\hat{\omega} = i\hbar \frac{\partial}{\partial q_\omega}$$ (24)

and q_ω is conjugate to the projection of $\vec{J}$ on the body fixed axis $\vec{R}$. The operator in Eq.(23) is a square root differential operator which is rather unusual in dealing with the Schrodinger equation. Nevertheless it is well behaved and it can be shown that this Hamiltonian will yield the usual coupled scattering equations upon expanding the wavefunction in an appropriate basis. The numerical treatment of the square root operators raises some interesting problems which must be tackled. This type of operator most clearly illustrates the basic differences between the finite

difference and finite element methods. The finite difference approach would inherently require some sort of discretization of the operator while the finite element method can leave the operator in tact.[4]

Finally, perhaps the most powerful use of finite element theory may come from its combination with conventional close coupling methods. As illustrated above by the $H + H_2$ problem, the finite element approach inherently includes closed channels which can be significant in the region where the particles are at short range. In the asymptotic and near asymptotic regions on the other hand, the close coupling method is very efficient since the equations are nearly diagonal. Therefore, it is suggestive that an optimal procedure could best utilize the strong points of both close coupling and the finite element method. This approach is a topic of future research, but its schematic illustration is shown in Figure 6.

V. Conclusion

In summary, the finite element method appears to have a viable role in tackling certain classes of molecular dynamics problems. In particular, reactive scattering, including dissociation, seems to be the most attractive field for application. In this case, since no global basis function expansions are made, the complexities associated with choosing a coordinate system (e.g., natural collision coordinates, etc.) are no longer a problem. Although finite element theory is not new to the engineering sciences, it is a new approach in molecular dynamics. Much work needs to be done to develop and optimally utilize its potential in this new area.

References

1. K. Bathe and E. Wilson, **Numerical Methods in Finite Element Analysis** (Prentice Hall, 1976).

2. K. W. Morton, Comp. Phys. Comm. $\underline{12}$, 99 (1976).

3. A. Askar, A. Cakmak and H. Rabitz, Chem. Phys. $\underline{33}$, 267 (1978).

4. S. Augustin, M. Demiralp, H. Rabitz, A. Askar and A. Cakmak, to be published.
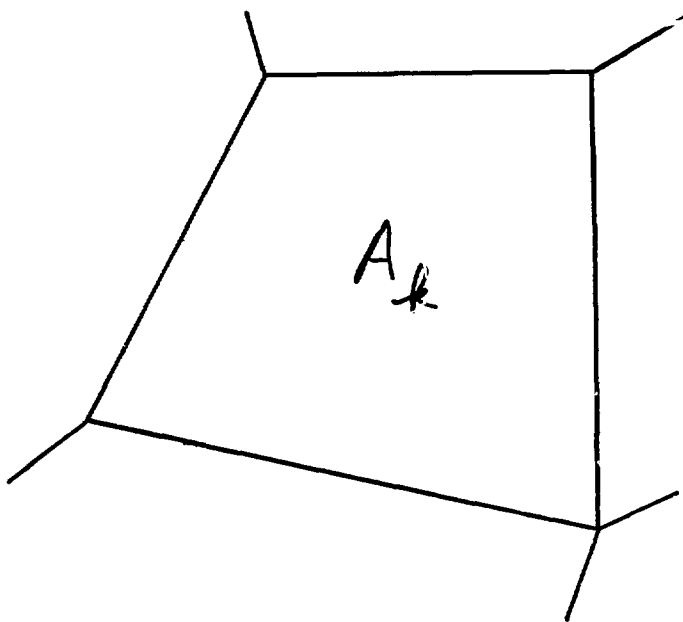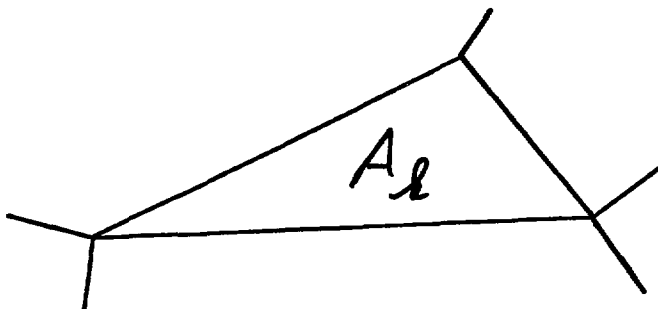
Fig. 1. Two possible local finite element geometries.
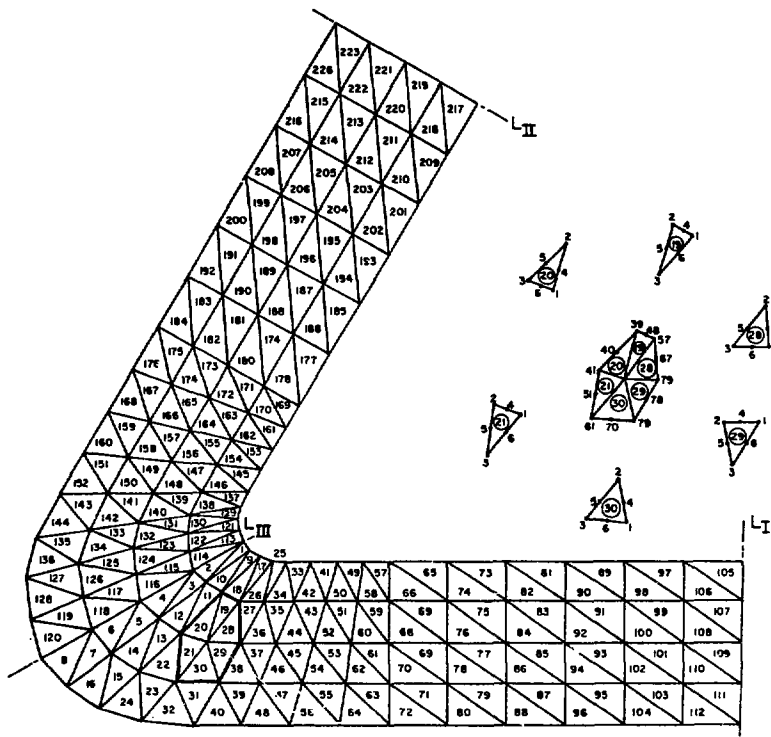
Fig. 2. A typical discretization of the two dimensional space for a collinear reactive scattering problem. The numbering is associated with the nodes and elements.
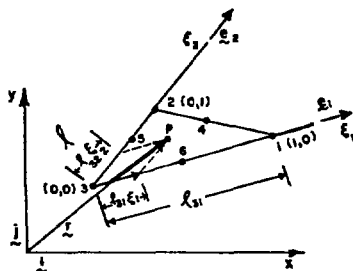
Fig. 3. The cartesian and local oblique coordinate system for an arbitrary triangular element.

$$(4a)$$

$$(4b)$$

$$(4c)$$

Fig. 4. Schematic of the finite element representation of the Schrödinger equation. N is the total number of nodal points, and N_B is the number of boundary nodes.
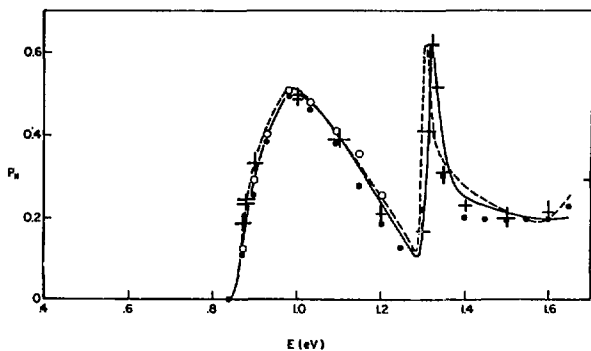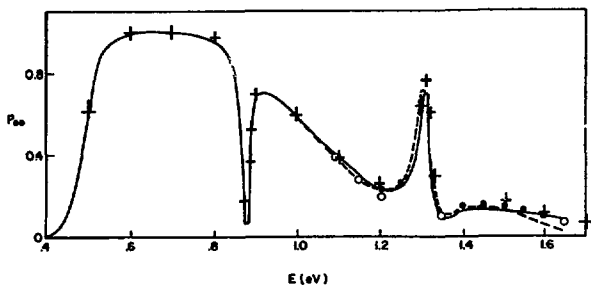
Fig. 5. Comparison of finite element (+) transition probabilities with close coupling (——, ----, o, •) for the $H + H_2$ reactive system.
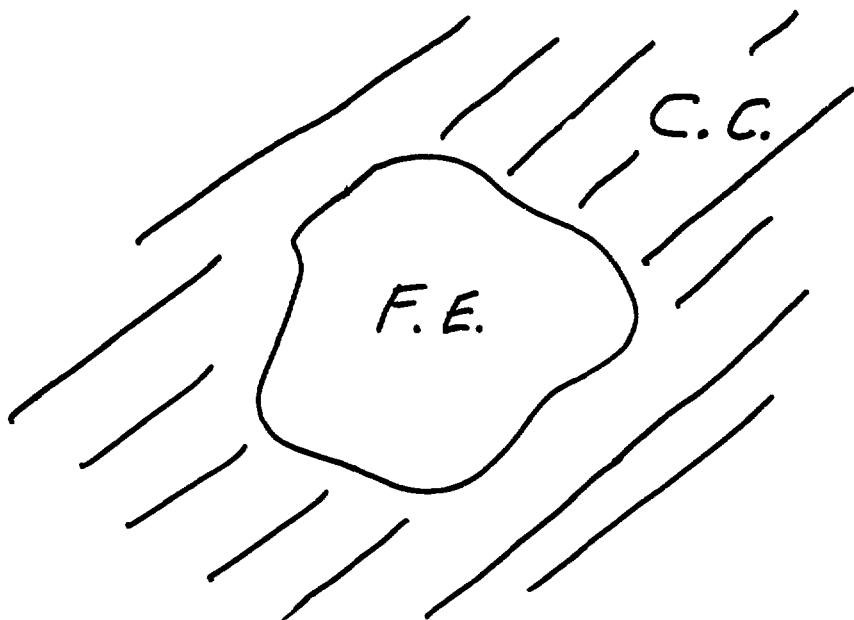
Fig. 6. A schematic representation of how the finite element and close coupling methods could be combined into an overall hybrid scheme.

TEST PROBLEMS
FOR
THE NRCC WORKSHOP ON
ALGORITHMS AND COMPUTER CODES FOR
ATOMIC AND MOLECULAR QUANTUM SCATTERING THEORY

Lowell D. Thomas
NRCC, Lawrence Berkeley Laboratory
Berkeley, Ca. 94720

I. Introduction

One of the principal objectives of this workshop is to identify which of the existing computer codes for solving sets of coupled Schroedinger equations perform most efficiently and on which types of scattering problems. To do this in an unbiased fashion, it will be necessary for the different codes to solve the same problems, using the same potential energy programs, on the same computer, and with roughly the same accuracy. To this end, several test problems have been prepared and these will be discussed below.

These test problems have been chosen based on discussions which were held during a planning meeting at Salt Lake City, Utah, February 23-24, 1979, and further discussions at the workshop itself at Argonne, June 25-27, 1979. It was agreed that the problems should be realistic, three-dimensional problems, employing ab initio potentials if possible. It was also decided to use only non-reactive collision problems. Since the chief purpose here is to study the performance of the coupled equations codes, there is no necessity for adding the other complications of reactive collisions. The parameters to be varied are:

1) the number of channels,

2) the number of closed channels,

3) the collision energy, and

4) the radial range of the potential.

With these considerations in mind, three chemical systems, He-H_2, Li(+)-CO and electron-N_2, have been selected and a variety of tests for each devised to vary the above mentioned parameters over representative ranges. For each system, four basis sets were chosen to give N, the number of channels, in the range 2 to 32. J, the total angular momentum quantum number, is fixed for all four basis sets, and chosen so that the impact parameter is small. The number of closed channels is varied at the same time as the energy. That is, when the energy is low, many channels are closed. When it is high most or all are open. Low energy can be defined by requiring the wave length to be roughly the same as the range of the strong interaction region. High energy can be defined by requiring the wave length to be one-fifth of the low energy wave length.

The calculation of the potential energy matrices for these tests has been programed in a fashion which is hopefully flexible enough to be used by all of the codes without undue modification. The most serious omission is, I believe, the exclusion of a curve crossing problem. This is largely due to the finite computing resources available to the workshop.

II. Description of the Tests

Test 1

Test 1 is the rotational and vibrational excitation of H_2 by He impact. An accurate CI potential is available which includes all of the nuclear degrees of freedom, and converged quantum calculations of the inelastic transition probabilities have been done [1]. This is currently the only chemical system which has been this thoroughly studied, and hence is a good representation of the state of the art.

The potential energy function is first expanded in Legerdre polynomials,

$$V(x,r,\theta) = \Sigma \; V_{\lambda}(x,r) \; P_{\lambda}(\cos\theta) \;, \tag{1}$$

where x, r, and θ are the H - H vibrational, He-H_2 radial, and He-H_2 angular coordinates, respectively. The Legendre coefficients are then written in a separable expansion,

$$V_{\lambda}(x,r) = \underset{\mu}{\Sigma} \; Q_{\lambda\mu}(x) \; B_{\lambda\mu}(r) \;. \tag{2}$$

Matrix elements of the Q-terms between H_2 vibrational wave functions are

$$A_{\lambda\mu}(nj, \; n'j') = \langle nj|Q_{\lambda\mu}(x)|n'j'\rangle \;. \tag{3}$$

Matrix elements of the Legendre polynomials between coupled spherical hamonics are

$$F_{\lambda}(j\ell, \; j'\ell';J) = \langle j\ell J|P_{\lambda}(\cos\theta)|j'\ell'J\rangle \;. \tag{4}$$

Therefore, defining the G-coefficients,

$$G_{\lambda\mu}(nj\ell, \; n'j'\ell') = A_{\lambda\mu}(nj,n'j') \times F_{\lambda}(j\ell, \; j'\ell'; J) \;, \tag{5}$$

we have the desired matrix elements,

$$V_{ij}(r) = \langle n_i j_i \ell_i |V(x,r,\theta)|r_j j_j \ell_j \rangle$$

$$= \sum_{\lambda\mu} \; G_{\lambda\mu}(n_i j_i \ell_i, \; n_j j_j \ell_j) \times B_{\lambda\mu}(r) \;. \tag{6}$$

The A-coefficients were computed with the formula

$$A_{\lambda\mu}(nj,n'j') = \sum_{\nu} C_{\lambda\mu}^{\nu} D^{\nu}(nj,n'j') \; , \tag{7}$$

where the Q-functions have been expanded in powers of the H-H displacement from the equilibrium position, x_e

$$Q_{\lambda\mu}(x) = \sum_{\nu} C_{\lambda\mu}^{\nu} (x-x_e)^{\nu} \; . \tag{8}$$

The constants, $C_{\lambda\mu}^{\nu}$, were taken from Ref. 1. The D-coefficients are then the matrix elements of $(x-x_e)^{\nu}$ between the H_2 vibrational wave functions

$$D^{\nu}(nj,n'j') = \langle nj|(x-x_e)^{\nu}|n'j' \rangle \; . \tag{9}$$

These matrix elements were evaluated numerically using exact numerical wave functions for the Kolos-Wolniewicz H_2 potential [2].

The four basis sets chosen are shown in Table 1. These give 2, 8, 18, and 28 channel problems, respectively. The total angular momentum quantum number, J, is equal to 4 for all the basis sets. When H_2 is in its ground state and the relative kinetic energy is E = .0224 a.u., this corresponds to an impact parameter of b = .38 a.u. At this energy the wave length is .6 a.u. which corresponds roughly to 5 wave lengths in the strong interaction region. For this system only the high energy E = .0224 is used. For the lower energy corresponding to one wave length in the strong interaction region, there are no open vibrational states of H_2.

Test 2

Test 2 is designed to test the codes against a long-range potential. The problem is the rotational excitation of a rigid rotor CO molecule by

Li(+) impact. This system was chosen because the charge-dipole and -quadrapole interactions lead to r^{-2} and r^{-3} dependence in the off diagonal potential matrix elements and because an analytic fit is available to a configuration-interaction potential energy surface [3].

Using the above criteria for low and high energy leads to 0.000079 a.u. and 0.00195 a.u., respectively. Fixing the impact parameter at b=4 leads to J=5 for the lower energy and J=25 for the higher. Four basis sets were chosen for each energy. An attempt was made to keep the basis sets realistic, much as one might do in a convergence test. However, because of the close spacing of the CO rotational energy levels, it has not been possible to retain all of the open channels. Therefore, a type of decoupled l-dominant decoupling scheme [4] was employed. The basis sets are shown in Tables 2 and 3. For E = 0.0000779, all channels with j > 2 are closed. For E = 0.00195, only the channel with j = 15 is closed.

The interaction potential is expanded in Legendre polynomials

$$V(r,\theta) = \sum_{\lambda} V_{\lambda}(r) \, P_{\lambda}(\cos\theta) \, , \tag{10}$$

where r and θ are the radial and angular Li(+)-CO coordinants, respectively. Matrix elements of the Legendre polynomials, $P_{\lambda}(\cos\theta)$, between coupled spherical harmonics are given by Eq. (4). The potential energy matrix elements are then given by

$$V_{ij}(r) = \sum_{\lambda} F_{\lambda}(j_i \ell_i, \, j_j \ell_j) \, V_{\lambda}(r) \, . \tag{11}$$

Test 3

This is a test of the codes against a short-range potential. The
problem is identical to that of test 2, except that the integration range
is $0 < r < 7$. That is, the integration is to be stopped and an S-matrix
computed by matching to spherical Bessel functions, at $r = 7.0$.

This artificial truncation of the potential is a slight contradiction
of our intention to use only realistic test problems. However, the numerical
difficulties of integrating the coupled equations are so drastically dif-
ferent for the interior and exterior regions of the potential that this
truncation is justified in order to identify those codes which perform best
on the interior region.

Test 4

Electron-molecule scattering is physically quite different from atom-
molecule scattering because of the extreme differences in mass as well as
because of the exchange problem arising from having identical particles in
the system. It is therefore worthwhile to exercise the codes on an electron
scattering problem, since it may well be that the same codes are not the
best for electron- or for atom-molecule scattering.

Onda and Truhlar [5] have constructed a potential for electron-N_2
scattering which features a local approximation to the exchange potential and
an asymptotically correct long-range polarization potential. The potential
has been expanded in Legendre polynomials as for test 2, and spline fits
made to the radial coefficients. Since the local approximation to the
exchange potential is energy dependent, this particular potential energy
matrix is only meaningful when used at 30 eV (relative kinetic energy
when N_2 is in its ground rotational state). The four basis sets are shown
in Table 4.

References

[1] A. W. Raczkowski and ▓▓ ▓▓▓▓▓ ▓▓▓▓ Jr., Chem. Phys. Lett., <u>47</u>, 45 (1977)
    Note: In eq(2c) 8.146 ▓▓▓▓▓ ▓▓▓▓▓ be .008146.

[2] W. Kolos and L. Wolniewicz, J. Chem. Phys., <u>43</u>, 2429 (1965).

[3] L. D. Thomas, W. P. Kraemer and G. H. F. Diercksen, Chem. Phys., <u>30</u>, 33 (1978).

[4] A. E. DePristo and M. H. Alexander, J. Chem. Phys., <u>64</u>, 3009 (1976).

[5] K. Onda and D. G. Truhlar, Chem. Phys. Lett., <u>58</u>, 512 (1978).

Table 1. Basis sets for test one. n, j, and ℓ refer to the vibrational, rotational angular momentum, and orbital angular momentum quantum numbers, respectively. # is a serial count of the channels.

n j ℓ	#	n j ℓ	#	n j ℓ	#	n j ℓ	#	Target energy levels (a.u.)
0 0 4	1	0 0 4	1	0 0 4	1	0 0 4	1	0.0
		2 2	2	2 2	2	2 2	2	.001616
		4	3	4	3	4	3	
		6	4	6	4	6	4	
				4 0	5	4 0	5	.005329
				2	6	2	6	
				4	7	4	7	
				6	8	6	8	
				8	9	8	9	
				6 2	10	6 2	10	.011009
				4	11	4	11	
				6	12	6	12	
				8	13	8	13	
				10	14	10	14	
						8 4	15	.018473
						6	16	
						8	17	
						10	18	
						12	19	
1 0 4	2	1 0 4	5	1 0 4	15	1 0 4	20	.018971
		2 2	6	2 2	16	2 2	21	.020506
		4	7	4	17	4	22	
		6	8	6	18	6	23	
						4 0	24	.024034
						2	25	
						4	26	
						6	27	
						8	28	
Total #	2		8		18		28	

Table 2. Basis sets for tests 2 and 3 at $E = .000079$ a.u. and J=5. j and ℓ are the rotational and orbital angular momentum quantum numbers, respectively. The arrows indicate that all channels down to that point are included in the basis set.

j	ℓ	Basis sets 1	2	3	4	Target energy levels (a.u.)
0	5					0.0
1	4 6	↓				1.76-5
2	3 5 7		↓			5.28-5
3	2 4 6 8					1.06-4
4	1 3 5 7 9			↓		1.76-4
5	0 2 4 6 8 10					2.64-4
6	1 3 5 7 9 11				↓	3.70-4
Total #		3	6	15	27	

Table 3. Basis sets for tests 2 and 3 at $E = .00195$ a.u. and J=25. j and ℓ are the rotational and orbital angular momentum quantum numbers, respectively. The arrows indicate that all channels down to that point are included in the basis set.

j	ℓ	Basis sets 1	2	3	4	Target energy levels (a.u.)
0	25					0.0
1	24 26	↓				1.76-5
2	23 25 27		↓			5.28-5
3	22 24 26 28					1.06-4
4	21 23 25 27 29			↓		1.76-4
5	20 22 24 26					2.64-4
6	19 21 23			↓		3.70-4
7	18 20					4.33-4
8	17					6.33-4
9	16					7.92-4
10	15					9.68-4
11	14					1.16-3
12	13					1.37-3
13	12					1.60-3
14	11					1.84-3
15	10				↓	2.11-3
Total #		3	10	22	32	

Table 4. Basis sets for test 4 at $E = 1.1025$ a.u. and J=5. j and ℓ are the rotational and orbital angular momentum quantum numbers, respectively. The arrows indicate that all channels down to that point are included in the basis set.

j	ℓ	Basis sets 1	2	3	4	Target energy levels (a.u.)
0	5					0.0
2	3 5 7					5.46-5
4	1 3 5 7 9					1.82-4
6	1 3 5 7 9 11					3.82-4
8	3 5 7 9 11 13					6.56-4
10	5 7 9 11 13 15					1.00-3
Total	#	4	15	21	27	

A VARIABLE INTERVAL VARIABLE STEP METHOD FOR THE SOLUTION OF LINEAR SECOND ORDER COUPLED DIFFERENTIAL EQUATIONS*

Gregory A. Parker, Thomas G. Schmalz,
and John C. Light
The James Franck Institute and The Department of Chemistry
The University of Chicago,
Chicago, Illinois 60637

The usual quantum mechanical formulations of scattering as well as many bound state problems lead to a set of coupled linear second order differential equations. This set of differential equations can conveniently be written in matrix notation as

$$\left[\underset{\sim}{1} \frac{d^2}{dR^2} + \underset{\sim}{k}^2 - \underset{\sim}{V}(R) \right] \underset{\sim}{U}(R) = \underset{\sim}{Q} \tag{1}$$

where $\underset{\sim}{1}$ is the identity matrix, k^2 is a matrix of wavevectors, and $\underset{\sim}{V}(R)$ is the potential energy. Numerous methods have been developed to obtain the solution matrix, $\underset{\sim}{U}(R)$, to these equations which approach the problem from a variety of directions. Most current methods can be categorized in several ways. For example, we can divide them according to whether the basis is fixed (diabatic) or is transformed during the calculation (quasi-adiabatic), whether the stepsize algorithm is dependent on the variation of the solution matrix or the potential matrix, or whether the solution is propagated across an elementary interval as a diagonal or full matrix. Each approach, whether De Vogeleare,[1]

Numerov[2], Airy function propagation[3], integral equations[4], or R-matrix propagation[5] has certain advantages for particular types of problems, depending on the potential, the number of channels, the number of energies to be run, etc.

Although it is obvious that one algorithm will not be optimum for all problems, we have tried, in this paper, to combine some of the most successful features of several of these methods into a single new method. This new method can be considered either as a generalization of such single interval methods as distorted wave Born approximation (DWBA)[6] or infinite order sudden plus perturbation corrections (IOS)[7,8] to cases in which more than one interval (over the range of R) is required for numerical accuracy; or a generalization of the Gordon,[3] Magnus,[9] or R-matrix[5] diagonalization methods to much higher accuracy (via perturbation theory); or finally, a generalization of the Sams and Kouri[4] integral equation method to include both a quasi-adiabatic (re-diagonalized) basis and analytic perturbations to the local Green's function solution.

The success of this method, particularly for large numbers of channels and energies, is based on the reduction of the number of matrix operations required for accurate solutions. Since a change of basis requires two matrix multiplications, a stabilization of the solution matrix requires about one, and a diagonaliztion about 5/3, one wants to use a fixed basis over as large an interval as possible. On the other hand a basis for which the residual (after transformation) interaction matrix (RIM) contains

large diagonal and off-diagonal elements (after a short distance) will require frequent perturbation evaluations (an N^2 process) <u>and</u> frequent recoupling of the solution matrix (equivalent to at least two matrix multiplications).

Thus the method presented in this paper attempts (and we believe succeeds) in balancing these factors to obtain an algorithm which reduces markedly the number of matrix (multiplication) operations while not increasing the number of steps required. This is accomplished by using standard techniques in ways that are innovative in several respects. First, the <u>intervals</u> over which a fixed basis (without stabilization) is used are large. Within each interval first order perturbative corrections to the zero order solutions are evaluated. In the limit of one interval only, the solution reduces to IOS plus perturbative corrections or DWBA (or something else) depending on the basis chosen. Second, since for many problems the diagonal elements of the residual interaction matrices (RIM's) change faster (and become larger) across an interval than the off-diagonal elements, analytic approximations to the solutions of the diagonal elements of the RIM's are generated for a number of <u>steps</u> (one or more) within the interval. Finally, the remaining perturbations (both diagonal away from the center of the steps and off-diagonal) are evaluated and <u>summed</u> over steps to give the overall perturbation over the interval. These are then used in an integral equation formalism to evaluate the solution and derivative matrices at the end of the interval. These are stabilized and propagated by the R-matrix (inverse

log derivative) evaluation at the interval boundary and then joined across interval boundaries by the transformation matrices (since different bases are used in each interval).

Accuracy is maintained by controlling both the step size and interval size dynamically as the calculation is done at the first (lowest) energy. Since the potential diagonalizations need not be repeated at other energies, the transformation matrices, RIM's, etc., are saved for subsequent calculations. In the following, we discuss first the formal solution in an interval and its propagation across the interval boundaries within the framework of integral equations and R-matrix evaluation. We then discuss the evaluation of the perturbatively corrected solution over the steps within the interval, and the step size and interval size algorithms used. In the final section we give the results of an application to the standard nine channel rigid rotor problem of Lester and Bernstein.[10] A few speculative comments follow.

Intervals

We begin by dividing the range of interest into a series of intervals. For each interval i the solution matrix is transformed to a new basis by an orthogonal transformation T_i with the solution $G_i(R)$ in the new basis given by

$$G_i(R) = T_i^t \, U(R). \tag{2}$$

The transformation matrix T_i is determined by diagonalizing the interaction

$$\underset{\sim}{k}^2 - \underset{\sim}{V}(R_i) \tag{3}$$

where $\underset{\sim}{R_i}$ is an estimate of the midpoint of the interval (the actual interval that is taken may be several times larger than the predicted interval length). The transformed set of equations for interval i are then

$$\left[\underset{\sim}{1} \frac{d^2}{dR^2} + \underset{\sim}{W}_i(R) \right] \underset{\sim}{G}_i(R) = \underset{\sim}{O} \, , \tag{4}$$

where

$$\underset{\sim}{W}_i(R) = \underset{\sim}{T}_i^{-1} \left[\underset{\sim}{k}^2 - \underset{\sim}{V}(R) \right] \underset{\sim}{T}_i \quad . \tag{5}$$

At the end of each interval an $\underset{\sim}{R}$-matrix $\underset{\sim}{R}_i$ in the local basis is determined from the transformed wavefunction and its derivative evaluated at the end of the interval, $\mathbf{R}_E$, by solving the following set of linear equations:

$$\underset{\sim}{G}_i(R_E) = \underset{\sim}{R}_i \underset{\sim}{G}_i'(R_E). \tag{6}$$

Since we may multiply the wavefunction by any arbitrary constant, we solve (6) by right multiplying $\underset{\sim}{G}_i(R_E)$ by $[\underset{\sim}{G}_i'(R_E)]^{-1}$ thus making the wavefunction numerically equal to $\underset{\sim}{R}$ and its derivative the unit matrix. As shown later this saves $2N^3/3$ operations/interval, where N is the number of coupled differential equations, while automatically stabilizing the solution for closed channels.

Expressing the interaction matrix $\underset{\sim}{W}_i$ in terms of a diagonal reference potential $\underset{\sim}{W}_{ref}^i$ (to be defined later) we have

$$\left[\underline{1} \frac{d^2}{dR^2} + \underset{\sim}{W}^i_{ref}(R) \right] \underset{\sim}{G}_i(R) = \underset{\sim}{P}_i(R) \underset{\sim}{G}_i(R) , \qquad (7)$$

where $\underset{\sim}{P}_i(R)$ is the difference between the reference potential and the interaction matrix

$$\underset{\sim}{P}_i(R) = \underset{\sim}{W}^i_{ref}(R) - \underset{\sim}{W}_i(R) . \qquad (8)$$

With the use of Green's functions[11] or equivalently variation of parameters[12] the set of differential equations (7) can be reexpressed in terms of integral equations as

$$\begin{aligned}
\underset{\sim}{G}_i(R) = &\ \underset{\sim}{g}^i_1(R) \underline{A} + \underset{\sim}{g}^i_2(R) \underline{B} \\
&+ \underset{\sim}{g}^i_1(R) \int_{R_i - \ell_i/2}^{R} \underset{\sim}{g}^i_2(R') \underset{\sim}{P}_i(R') \underset{\sim}{G}_i(R') dR' \\
&- \underset{\sim}{g}^i_2(R) \int_{R_i - \ell_i/2}^{R} \underset{\sim}{g}^i_1(R) \underset{\sim}{P}_i(R) \underset{\sim}{G}_i(R) dR ,
\end{aligned} \qquad (9)$$

where g^i_1 and g^i_2 are any two linearly independent solutions of the homogeneous reference equation

$$\left[\underline{1} \frac{d^2}{dR^2} + \underset{\sim}{W}^i_{ref}(R) \right] \begin{pmatrix} g^i_1(R) \\ or \\ g^i_2(R) \end{pmatrix} = \underset{\sim}{0} , \qquad (10)$$

and A and B are constants to be fixed by the boundary conditions for the interval, the predicted length of the interval being ℓ_i. Since an equation like (9) can be set up in each interval,

using different reference solutions g_1^i and g_2^i, the solution to the whole problem can be built up by using the known solutions for interval i-1 to fix the constants in the solution for interval i. After having solved Eq. (7) in the last interval, continuity of the wavefunction and its derivative require that

$$\underset{\sim}{R}_{i-1}^{old} = \underset{\sim}{g}_1^i(R_i - \ell_i/2)\underset{\sim}{A} + \underset{\sim}{g}_2^i(R_i - \ell_i/2)\underset{\sim}{B} \tag{11}$$

$$\underset{\sim}{1} = \underset{\sim}{g}_1^{i'}(R_i - \ell_i/2)\underset{\sim}{A} + \underset{\sim}{g}_2^{i'}(R_i - \ell_i/2)\underset{\sim}{B} . \tag{12}$$

If we take g_1^i to be the solution in interval i which is zero at $R_i - \ell_i/2$ and normalize g_2^i to insure a unit wronskian, Eq. (9) will assume the form

$$\underset{\sim}{G}_i(R) = \underset{\sim}{g}_1^i(R)\Big[\underset{\sim}{1} + \int_{R_i - \ell_i/2}^{R} \underset{\sim}{g}_2^i(R')\, \underset{\sim}{P}_i(R')\underset{\sim}{G}_i(R')dR'\Big] ,$$

$$+ \underset{\sim}{g}_2^i(R)\Big[\underset{\sim}{R}_{i-1}^{old} - \int_{R_i - \ell_i/2}^{R} \underset{\sim}{g}_1^i(R')\, \underset{\sim}{P}_i(R')\, \underset{\sim}{G}_i(R')dR'\Big] \tag{13}$$

where

$$R_{i-1}^{old} = Q_i^t\, R_{i-1}\, Q_i \tag{14}$$

and

$$Q_i = T_{i-1}^t\, T_i \tag{15}$$

transforms the R-matrix of the last interval into the new basis. By choosing the reference potential equal to the centrifugal

term

$$W_{ref}(R) = \ell(\ell+1)/R^2 \tag{16}$$

and having T_i be the identity matrix Eq. (13) can be made
identical to that given by Sams and Kouri[4] except that they
begin with purely regular boundary conditions since they are
giving the solution for the whole problem rather than just
one interval. Sams and Kouri then use a numerical quadrature
to propagate the solution.

   If the interval is sufficiently small and a good reference
potential is chosen we can solve Eq. (13) by perturbation theory,
and use

$$\underset{\sim}{G}_i^o(R) = \underset{\sim}{g}_1^i(R) + \underset{\sim}{g}_2^i(R)\, \underset{\sim}{R}_{i-1}^{old} \tag{17}$$

for G_i on the righthand side of Eq. (13). We then obtain a
perturbatively corrected wavefunction over the interval

$$\underset{\sim}{G}_i(R) = \underset{\sim}{g}_1^i(R)\left[\underset{\sim}{1} + \underset{\sim}{I}_{21}(R)\right] - \underset{\sim}{g}_2^i(R)\underset{\sim}{I}_{11}(R)$$
$$+ \left\{\underset{\sim}{g}_1^i(R)\,\underset{\sim}{I}_{22}(R) + \underset{\sim}{g}_2^i(R)\left[\underset{\sim}{1} - \underset{\sim}{I}_{12}(R)\right]\right\} \underset{\sim}{R}_{i-1}^{old} \tag{18}$$

where

$$\underset{\sim}{I}_{11}(R) = \int_{R_i-\ell_i/2}^{R} \underset{\sim}{g}_1^i(R')\, \underset{\sim}{P}(R')\underset{\sim}{g}_1(R')\, dR' \quad, \tag{19}$$

$$\underset{\sim}{I}_{21}^t(R) = \underset{\sim}{I}_{12}(R) = \int_{R_1-\ell_i/2}^{R} \underset{\sim}{g}_1^i(R')\, \underset{\sim}{P}_i(R')\underset{\sim}{g}_2^i(R)\, dR' \quad, \tag{20}$$

and

$$\underset{\sim}{I}_{22}(R) = \int_{R_i - \ell_i/2}^{R} g_2^{i}(R) \, \underset{\sim}{P}_i(R') \, g_2^{i}(R') \, dR' \; .$$

(21)

The derivative of the wavefunction is

$$G_i'(R) = g_1^{i'}(R) \left[\underset{\sim}{1} + \underset{\sim}{I}_{21}(R) \right] - g_2^{i'}(R) \, \underset{\sim}{I}_{11}(R)$$

$$+ \left\{ g_1^{i'}(R) \, \underset{\sim}{I}_{22}(R) + g_2^{i'}(R) \left[\underset{\sim}{1} - \underset{\sim}{I}_{12}(R) \right] \right\} R_{i-1}^{old}$$

(22)

which is easily verified by direct differention of Eq. (18).
By writing Eqs. (18) and (22) in the form

$$G_i(R) = g_1^{i}(R) + \Delta g_1^{i}(R) + \left[g_2^{i}(R) + \Delta g_2^{i}(R) \right] R_{i-1}^{old}$$

(23)

$$G_i'(R) = g_1^{i'}(R) + \Delta g_1^{i'}(R) + \left[g_2^{i'}(R) + \Delta g_2^{i'}(R) \right] R_{i-1}^{old}$$

(24)

where

$$\Delta g_1^{i}(R) = g_1^{i}(R) \, \underset{\sim}{I}_{21}(R) - g_2^{i}(R) \underset{\sim}{I}_{11}(R) \; ,$$

(25)

$$\Delta g_2^{i}(R) = g_1^{i}(R) \, \underset{\sim}{I}_{22}(R) - g_2^{i}(R) \underset{\sim}{I}_{12}(R) \; ,$$

(26)

$$\Delta g_1^{i'}(R) = g_1^{i'}(R) \, \mathcal{I}_{21}(R) - g_2^{i'}(R) \mathcal{I}_{11}(R) \; , \qquad (27)$$

and

$$\Delta g_2^{i'}(R) = g_1^{i'}(R) \, \mathcal{I}_{22}(R) - g_2^{i'}(R) \mathcal{I}_{12}(R) \; , \qquad (28)$$

we can use the perturbative corrections Δg_1^i, Δg_2^i, $\Delta g_1^{i\prime}$, and $\Delta g_2^{i\prime}$ to control the size of the interval. We terminate the interval whenever the maximum of any of the perturbative corrections is greater than some tolerance TOFF

$$TOFF \geq MAX \left[\Delta g_1^i, \Delta g_2^i, \Delta g_1', \Delta g_2' \right]. \qquad (29)$$

As mentioned before, calculating an R-matrix at the end of each of the intervals saves $2N^3/3$ operations which is easily seen from Eqs. (23) and (24). Calculating an R-matrix via a linear equation solver requires $4N^3/3$ operations; however, $2N^3$ operations are saved by not having matrix multiplications with Δg_1^i and $\Delta g_1^{i\prime}$ (due to renormalization of the wavefunction to unit derivative at the end of each interval), resulting in a net savings of $2N^3/3$ operations. The total number of matrix operations/interval is 5 1/3 determined as follows:

i) $2N^3$ from transforming the R-matrix into a new basis [Eq. (14)],

ii) $2N^3$ from multiplication of Δg_2^i and $\Delta g_2^{i\prime}$ by R_i [Eqs. (23-24)], and

iii) $4N^3/3$ for determination of the R-matrix [Eq. (6)].
The number of matrix operations/interval can be reduced to 4
5/6 by symmetrizing the R-matrix in each interval since only
$3/2\ N^3$ operations are then required to transform the symmetrized
R-matrix. The lack of symmetry in the R-matrix is a direct
result of the perturbative corrections which in general are not
symmetric (and thus do not conserve flux). However, by keeping
the perturbative corrections small the R-matrix is nearly
symmetric.

The number of matrix operations given above assumed that
the diagonalizations and energy independent transformations
were already computed and stored ready for use. For the first
energy there are an additional 2 2/3 N^3 operations/interval
determined as follows:

i) $5N^3/3$ for diagonalization of Eq. (3), and

ii) N^3 for the calculation of the interval to interval
transformation matrix [Eq. (15)].

Steps

To propagate across each interval we may of course simply
choose a reference potential and evaluate Eqs. (18) and (22).
However, we can improve the accuracy of the solution by
subdividing the intervals into steps where for each step, s,
the interaction matrix $W_i(R)$ is expanded in a Taylor series
about the midpoint R_i and terms through quadratic are retained:

$$\underset{\sim}{W}_i(R) \approx \underset{\sim}{\lambda}_{i,s}^2 + \overline{\underset{\sim}{W}}_i(R_s) + \underset{\sim}{W}_i'(R_s)(R-R_s)$$

$$+ \underset{\sim}{W}''(R_s)(R-R_s)/2 \tag{30}$$

where $\lambda_{i,s}^2$ is a diagonal matrix

$$(\underset{\sim}{\lambda}_{i,s}^2)_{j,k} = \begin{cases} 0 & j \neq k \\ [\underset{\sim}{W}_i(R_s) + h_s^2 \underset{\sim}{W}_i''(R_s)/24]_{j,j} & j = k \end{cases} \tag{31}$$

$$[\overline{\underset{\sim}{W}}_i(R_s)]_{j,k} = \begin{cases} [\underset{\sim}{W}_i(R_s)]_{j,k} & j \neq k \\ [-h_s^2 \underset{\sim}{W}_i''(R_s)/24]_{j,j} & j = k \end{cases} \tag{32}$$

and h is the length of the step. For a reference potential $\underset{\sim}{W}_{ref}^i$ we use $\lambda_{i,s}^2$ which is constant over the step but not constant within an interval. The solutions of the homogeneous equations within each step are simply a linear combination of sines and cosines,

$$g_1^{i,s}(R) = \frac{a_s \sin \lambda_{i,s}\rho + b_s \cos \lambda_{i,s}\rho}{\lambda_{i,s}} \tag{33}$$

$$g_2^{i,s}(R) = \frac{c_s \sin \lambda_{i,s}\rho}{\lambda_{i,s}} + d_s \cos \lambda_{i,s}\rho \tag{34}$$

where $\rho = R - R_s - h_s/2$ for open channels and hyperbolic sines and cosines

$$g_1^{i,s}(R) = \frac{a_s \sinh \lambda_{i,s} \rho}{\lambda_{i,s}} + b_s \cosh \lambda_{i,s} \rho \tag{35}$$

$$g_2^{i,s}(R) = \frac{c_s \sinh \lambda_{i,s} \rho}{\lambda_{i,s}} + d_s \cosh \lambda_{i,s} \rho \tag{36}$$

for closed channels, with the requirement that the wronskian

$$W\left[g_1^{i,s}(R), g_2^{i,s}(R)\right] = g_2^{i,s'}(R) g_1^{i,s}(R) - g_2^{i,s}(R) g_1^{i,s'}(R) \tag{37}$$

be a unit matrix. The coefficients a_s, b_s, c_s, and d_s are obtained by requiring that the homogeneous solutions $g_1^{i,s}$ and $g_2^{i,s}$ be continuous across the steps giving $(R_s - \frac{h_s}{2} = R_{s-1} + \frac{h_{s-1}}{2})$

$$a_s = g_1^{i,s-1}(R_s - h_s/2) \tag{38}$$

$$b_s = g_1^{i,s-1}(R_s - h_s/2) \tag{39}$$

$$c_s = g_2^{i,s-1'}(R_s - h_s/2) \tag{40}$$

$$d_s = g_2^{i,s-1}(R_s - h_s/2) . \tag{41}$$

For the first step in each interval the coefficients are determined by the definition of g_1^i and g_2^i used in the interval to interval matching:

$$\underline{a}_I = \underline{d}_I = \underline{1} \tag{42}$$

$$\underline{b}_I = \underline{c}_I = \underline{0} \ . \tag{43}$$

The integrals in Eqs. (19-21) can now be expressed in terms of the homogeneous solutions over each step as

$$\underline{I}_{11} = \sum_S \int_{R_s - h_s/2}^{R_s + h_s/2} \underline{g}_1^{i,s}(R') \underline{\rho}_i(R) \underline{g}_1^{i,s}(R') dR' \ , \tag{44}$$

$$\underline{I}_{21}^t = \underline{I}_{12} = \sum \int_{R_s - h_s/2}^{R_s + h_s/2} \underline{g}_1^{i,s}(R') \underline{\rho}_i(R') \underline{g}_2^{i,s}(R) dR \ , \tag{45}$$

and

$$\underline{I}_{22} = \sum \int_{R_s - h_s/2}^{R_s + h_s/2} \underline{g}_2^{i,s}(R) \underline{\rho}_i(R) \underline{g}_2^{i,s}(R) dR' . \tag{46}$$

Since the $g_1^{i,s}$ and $g_2^{i,s}$ are simply sums of sines and cosines (or hyperbolic sines and cosines) and $P_i(R)$ is a polynomial the integrals can be evaluated analytically. The step size is controlled by keeping the perturbative corrections to the step less than TDIAG for the diagonal elements and MAX (0.5*TOFF,TDIAG) for the off-diagonal elements. With the choice of reference potential used here there are no error terms with less than h_s^4 dependence in either the diagonal or off-diagonal perturbation elements. Other reference potentials could also be used (i.e., linear[3] or quadratic) leading to more accurate unperturbed solutions $g_1^{i,s}$ and $g_2^{i,s}$. However, the analytic evaluation of all the matrix elements in Eqs. (44-46) is not known for most solutions and, since the number of operations within each step is proportional to N^2, whereas the number of operations/interval is proportional to N^3, increasing the number of steps slightly has little effect on the overall computational time. In addition the sines and cosines are rapidly evaluated on computers while other solutions (e.g., airy functions[3]) may require significantly more effort for their evaluation.

For the first energy the potential and its first and second derivatives must be transformed to the local basis giving a total of $9N^3/2$ operations/step ($3/2\ N^3$ operations/transformation since the potential energy matrices are assumed symmetric). For subsequent energies there are no N^3 type operations at each step since the number of operations for the evaluation of the perturbation integrals Eqs. (44-46) is proportional to N^2.

Shown in Table I is the number of matrix operations/interval for the first and subsequent energy runs, where one sees that the ratio of the number of operations for a subsequent to the first energy is 0.56 or less depending on the number of steps/ interval.

If there is no off-diagonal coupling, increasing the number of steps and/or intervals will converge the solution to the exact result. When there is off-diagonal coupling we have to increase the number of intervals to converge the solution to the exact result. Since within each interval we have formed distorted wave solutions[6] in the T_i basis for that interval, increasing the number of steps alone will only converge to an approximate solution. There are two interesting limits for an approximate solution with one interval over the entire range of interest. First, if T_1 is the identity matrix then the resulting solution is the distorted wave approximation which is a good approximation when the off-diagonal coupling is small. This is often the case for vibrational and electronic transi- tions. Secondly, if T_1 is the transformation which diagonalizes the potential without centrifugal and kinetic energy terms, the resulting solution is the infinite order sudden (IOS)[7,8] solution plus perturbative corrections. The IOS solution has been shown to be a good approximation when the relative kinetic energy is large compared to the relative spacings of the wave- vectors as is often the case in rotational transitions. When two intervals are taken, the program is essentially doing an IOS plus perturbative corrections at short distances and a distorted wave approximation at large distances. As the number

of intervals is increased the results converge to the exact
coupled channel solution.

Application

In order to determine the rate of convergence and accuracy
of our method we used the rigid rotor problem of Lester and
Bernstein.[10] Although the parameters of this problem are not
physically meaningful it is one for which accurate solutions
are known.[5b,13] For completeness the parameters used are as
follows: total angular momentum J=8, with rotational states
j=0,2,4, rotational constant B=.202652 x 10^{-4} a.u., relative
kinetic energy E=.759945 x 10^{-2} a.u., reduced mass μ =
.66984935 a.m.u. x 1822.828 a.u./a.m.u., anisotropy parameter
a_2=0.4, and finally the well depth and well position of the
Lennard-Jones potential are, respectively, ϵ =.50663 x 10^{-2} a.u.,
r_m=6.35716 a_o. All integrations started at 4.60 a.u. and ended
at 80.0 a.u. Our transition probabilities P(j $\ell \leftarrow$ j' ℓ') for
the 9 state rotational problem rounded to six significant figures
are given in Table II.

In Fig. 1 we have a plot of the root mean square error

$$r.m.s. = \left[\sum_{j\ell j'\ell'} \left[P\left(\stackrel{exact}{j}\ell \leftarrow j'\ell'\right) - P(j\ell \leftarrow j'\ell')\right]^2 / (N^2-1)\right] \quad (47)$$

of the transition probability vs. the average number of steps/
interval. One sees that the results converge for roughly 8 steps/
interval and that the convergence is uniform.

In Fig. 2, we have a plot of the root mean square error in

the transition probability vs. the number of matrix operations
(assumed second energy). The solid curve is the result with
no perturbative corrections and 1 step/interval which is exactly
equivalent to the R-matrix recursion method of Walker and
Light,[5] where one sees that the rate of convergence is neither
rapid nor uniform. A conversion factor of 3 1/3 matrix
operations/interval was used. The dashed curve is the result
with 1 step/interval plus perturbative corrections using a
conversion factor of 5 1/3 matrix operations/interval. The
rate of convergence has been improved as well as the uniformity
of convergence over the unperturbed R-matrix results. The
dotted curve is the result obtained from   our   method with
TOFF = 8*TDIAG. There were an average of 4.2 steps/interval
and the conversion factor for the number of matrix operations/
interval was (5 1/3 + $\frac{1}{9}$ number of steps/interval). The curve
shows a dramatic increase in the convergence rate over both
the unperturbed R-matrix results and the results using only
1 step/interval.

Speculations

Although the results shown in Fig. 2 demonstrate that this
approach succeeds admirably in increasing the accuracy while
reducing the number of matrix operations for this problem,
nothing is perfect, and we may speculate on the effort required
to produce further improvements. The present method has a
large overhead at the first energy (primarily in transforming
the potential and derivative matrices at each step). In
common with all other quasi-adiabatic basis methods it must also

take small intervals where the potential matrix is rapidly varying, even in the non-classical regions.

This suggests two changes in the algorithm, one trivial, the second less so. To reduce the potential transformation overhead at the first energy, it would probably be advantageous to evaluate the Taylor series expansion of the potential at fewer points than at every step. One requires that the potential matrix be accurately represented at each step but the current practice of evaluating and transforming $V^{(0)}$, $V^{(1)}$, $V^{(2)}$ at each step, requiring 4.5 N_s matrix multiplications per interval at the first energy may be overkill.

The second problem, small steps and intervals in (typically) the non-classical repulsive region, might be overcome by the use of a different sort of algorithm. In this region the potentials are typically varying more rapidly than the wave-functions, and a high order numerical (non-propagation) method would seem advantageous. Since the renormalized Numerov[2] is high order and only 2 matrix inversions per step and produces the log derivative matrix (R^{-1}), a hybrid using these two methods would seem to be optimal. Alternatively, a standard integral equation integrator using only trapezoidal (or Simpson's) rule could be used on a fixed basis, requiring only one matrix multiplication per step. However, since these are both true hybrid modifications of the present method with another, we present only these speculations here.

References

1) R. De Vogelaere, J. Research N.B.S. **54**, 119 (1955). See also Notes on Applied Science No. 16 Modern Computing Methods, 2nd ed. (National Physical Laboratory, Department of Scientific and Industrial Research, London, 1961), p. 88.

2) B. R. Johnson, J. Chem. Phys. **67**, 4186 (1977); **69**, 4678 (1978).

3) R. G. Gordon, J. Chem. Phys. **51**, 14 (1969); Methods in Comp. Phys. **10**, 81 (1971).

4) W. N. Sams and D. J. Kouri, J. Chem. Phys. **51**, 4809, 4815 (1969).

5) a) J. C. Light and R. B. Walker, J. Chem. Phys. **65**, 4272 (1976).

   b) E. B. Stechel, R. B. Walker, and J. C. Light, J. Chem. Phys **69**, 3518 (1978).

6) R. D. Levine, M. Shapiro, and B. R. Johnson, J. Chem. Phys. **52**, 1755 (1970).

7) T. P. Tsien and R. T Pack, Chem. Phys. Lett. **6**, 54 (1970); Chem. Phys. Lett. **8**, 579 (1971); T. P. Tsien, G. A. Parker, and R. T Pack, J. Chem. Phys. **59**, 5373 (1973); G. A. Parker and R. T Pack, J. Chem. Phys. **68**, 1585 (1978).

8) D. Secrest, J. Chem. Phys. **62**, 710 (1975); L. W. Hunter, J. Chem. Phys. **62**, 2855 (1975).

9) P. Chang and J. C. Light, J. Chem. Phys. **50**, 2517 (1969).

10) W. A. Lester, Jr., and R. B. Bernstein, J. Chem. Phys. **48**, 4896 (1968).

11) See, for example, James T. Cushing, <u>Applied Analytical</u> <u>Mathematics for Physical Scientists</u> (John Wiley and Son, Inc., New York, 1975).

12) See, for example, D. L. Kreider, R. G. Kuller, and P. R. Osterg, <u>Elementary Differential Equations</u> (Addison-Wesley, Mass., 1968).
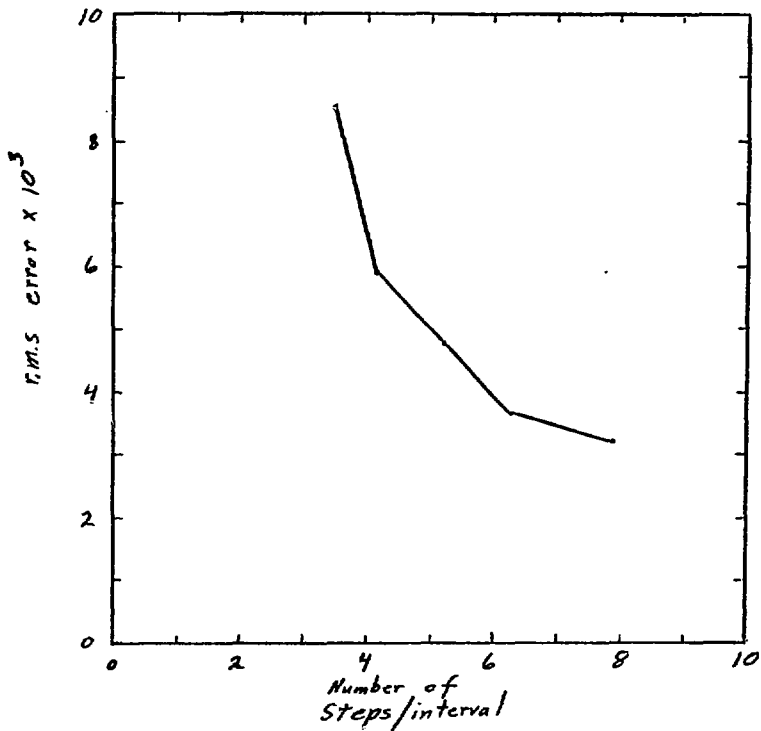
Fig. 1.   Root mean square error (eq. 47) of the transition proba-
bility versus the number of steps/interval for TOFF = .512.
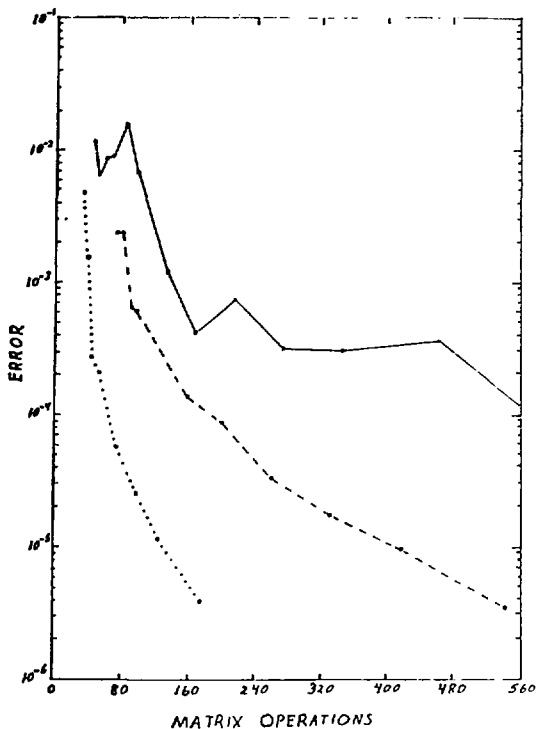
Fig. 2. Root mean square error (eq. 47) of the transition proba-
bility versus the number of matrix operations (assumed second
energy) for the Lester-Bernstein rotational problem (Ref. 10).
Solid curve (————): without perturbative corrections (3 1/3 matrix
operations/interval) equivalent to the R-matrix propagation method
(Ref. 5). Dashed curve (— — — —): current method 1 step/interval
(5 1/3 matrix operations/interval). Dotted curve (· · · ·) current
method with an average of 4 steps/interval, TOFF = 8*TDIAG (5 1/3
matrix operations/interval + $\frac{1}{9}$ number of steps/interval).

TABLE I

| Number of Steps/Interval | Matrix Operations Interval | | Ratio |
	1st Energy	Subsequent Energies	
1	9.5	5.33	0.56
2	11.0	5.33	0.48
3	12.5	5.33	0.43
4	14.0	5.33	0.38
5	15.5	5.33	0.34

TABLE II

Lester Bernstein rigid rotor rotational problem.
(See text for details.)

Transition Probabilities

j		4	2	4
ℓ		4	6	6
4	4	0.605514	0.281756	0.593056(-1)
2	6	0.281756	0.313523	0.742543(-1)
4	6	0.593056(-1)	0.742543(-1)	0.692690
0	8	0.226829(-1)	0.171931	0.159442(-1)
2	8	0.201651(-1)	0.961616(-1)	0.316285(-1)
4	8	0.729668(-2)	0.159411(-1)	0.115317
2	10	0.237556(-2)	0.371953(-1)	0.204618(-2)
4	10	0.803347(-3)	0.663197(-2)	0.862864(-2)
4	12	0.101578(-3)	0.255516(-2)	0.185634(-3)

j		0	2	4
ℓ		8	8	8
4	4	0.226829(-1)	0.201651(-1)	0.729668(-2)
2	6	0.171931	0.961616(-1)	0.159411(-1)
4	6	0.159442(-1)	0.316285(-1)	0.115317
0	8	0.384656	0.133936	0.159547(-1)
2	8	0.133936	0.433751	0.107687
4	8	0.159547(-1)	0.107687	0.656834
2	10	0.207964	0.661537(-1)	0.469303(-2)
4	10	0.180190(-1)	0.971295(-1)	0.752400(-1)
4	12	0.289113(-1)	0.153875(-1)	0.103665(-2)

j		2	4	4
ℓ		10	10	12
4	4	0.237556(-2)	0.803347(-3)	0.101578(-3)
2	6	0.371953(-1)	0.663197(-2)	0.255516(-2)
4	6	0.204618(-2)	0.862864(-2)	0.185634(-3)
0	8	0.207964	0.180190(-1)	0.289113(-1)
2	8	0.661537(-1)	0.971295(-1)	0.133875(-1)
4	8	0.469303(-2)	0.752400(-1)	0.103665(-2)
2	10	0.342219	0.314856(-1)	0.305868
4	10	0.314856(-1)	0.749201	0.128106(-1)
4	12	0.305868	0.128106(-1)	0.635143

Table II (Continued). Real Part of the S-matrix

j		4	2	4
ℓ		4	6	6
4	4	0.528309	-0.528565	0.227472
2	6	-0.528565	-0.497361	-0.252315
4	6	0.227472	-0.252315	-0.573729
0	8	0.104741	0.120203	0.119976
2	ℓ	-0.778045(-1)	0.202571(-1)	-0.636897(-1)
4	8	0.111511(-1)	-0.112659	0.116215
2	10	0.195871(-1)	0.144675	-0.415927(-1)
4	10	-0.208697(-1)	-0.811259(-1)	0.918147(-1)
4	12	-0.998008(-2)	-0.450312(-1)	0.423519(-2)

j		0	2	4
ℓ		8	8	8
4	4	0.104741	-0.778045(-1)	0.111511(-1)
2	6	0.120203	0.202571(-1)	-0.112659
4	6	0.119976	-0.636897(-1)	0.116215
0	8	-0.526697(-1)	0.296390	0.119123
2	8	0.296390	-0.452396(-1)	0.176317
4	8	0.119123	0.176317	-0.725772
2	10	0.451322	-0.238758	-0.158431(-1)
4	10	0.534475(-1)	0.305490	-0.211871
4	12	-0.924876(-1)	0.322496(-1)	0.145718(-1)

j		2	4	4
ℓ		10	10	12
4	4	0.195871(-1)	-0.208697(-1)	-0.998008(-2)
2	6	0.144675	-0.811259(-1)	-0.450312(-1)
4	6	-0.415927(-1)	0.918147(-1)	0.423519(-2)
0	8	0.451322	0.534475(-1)	-0.924876(-1)
2	8	-0.238758	0.305490	0.322436(-1)
4	8	-0.158431(-1)	-0.211871	0.145718(-1)
2	10	0.360917	0.177006	0.157255
4	10	0.177006	0.232537	-0.937555(-1)
4	12	0.157255	-0.937555(-1)	0.750511

Table II (Continued)   <u>Imaginary Part of the S-matrix</u>

j	ℓ	4 / 4	2 / 6	4 / 6
4	4	-0.571317	0.487372(-1)	0.869590(-1)
2	6	0.487372(-1)	-0.257208	0.102915
4	6	0.869590(-1)	0.102915	-0.602931
0	8	0.108223	0.396841	0.393689(-1)
2	8	-0.118792	-0.309437	0.166048
4	8	0.846896(-1)	0.570000(-1)	-0.319078
2	10	-0.446308(-1)	-0.127532	-0.177829(-1)
4	10	0.191782(-1)	0.100277(-1)	-0.140959(-1)
4	12	-0.140572(-2)	-0.229641(-1)	0.129498(-1)

j	ℓ	0 / 8	2 / 8	4 / 8
4	4	0.108223	-0.118792	0.846896(-1)
2	6	0.396841	-0.309437	0.570000(-1)
4	6	0.393689(-1)	0.166048	-0.319078
0	8	0.617966	0.214684	-0.420054(-1)
2	8	0.214684	0.657042	0.276765
4	8	-0.420054(-1)	0.276765	0.360679
2	10	-0.653609(-1)	-0.956467(-1)	0.666485(-1)
4	10	-0.123136	0.616904(-1)	-0.174215
4	12	-0.142679	0.111119	-0.287109(-_)

j	ℓ	2 / 10	4 / 10	4 / 12
4	4	-0.446308(-1)	0.191782(-1)	-0.140572(-2)
2	6	-0.127532	0.100277(-1)	-0.229641(-1)
4	6	-0.177829(-1)	-0.140959(-1)	0.129498(-1)
0	8	-0.653609(-1)	-0.123136	-0.142679
2	8	-0.956467(-1)	0.616904(-1)	0.111119
4	8	0.666485(-1)	-0.174215	-0.287109(-1)
2	10	0.460388	0.124333(-1)	-0.530225
4	10	0.124333(-1)	0.833743	0.634071(-1)
4	12	-0.530225	0.634071(-1)	-0.268099

INTEGRAL EQUATIONS METHODS
FOR INELASTIC SCATTERING

Don Secrest and Kelly McLenithan
School of Chemical Sciences
University of Illinois
Urbana Illinois 61801

Integral Equation Methods

We may for the purposes of this discussion start with the radial Schrödinger equation in some appropriate reduced units,

$$(\frac{d^2}{dR^2} - \frac{\ell(\ell+1)}{R^2} + k_i^2) \, f(R) = \sum_{i'} \sum_{i''} V(R) \, f(R)$$

$$_{iI} \qquad i'i'' \quad i'I \qquad (1)$$

Of course here the subscripts are compound subscripts which depend on the system being studied. The matrix V is the potential matrix for the problem and $f(R)$ is the channel radial wavefunction which must vanish at
$i'I$
the origin. The differential equation, Eq. (1), can be cast into the form of an inhomogeneous equation by partitioning the radial wavefunction into an incoming wave and a scattered wave

$$f(R) = u_\ell(k_i R) \, \delta_{iI} + g(R).$$

$$_{iI} \qquad \qquad \qquad _{iI} \qquad (2)$$

Here $u_\ell(k_i R)$ is the regular Riccati Bessel function which satisfies the left hand side of Eq. (1) equal to zero. Substituting Eq. (2) into Eq. (1) one obtains

$$(\frac{d^2}{dR^2} - \frac{\ell(\ell+1)}{R^2} + k_i^2) \, g_{iI}(R) = \sum_{i'} V(R) f(R)$$

$$\qquad \qquad \qquad \qquad i''i'' \quad i'I \qquad (3)$$

This inhomogeneous equation may be converted into an integral equation through the use of a Green's function to give

$$g(R) = \int_0^\infty \frac{u_\ell(k_i R_<) v_\ell(k_i R_>)}{C_i} \sum_{i'} V(R') \, F(R') dR' \quad .$$

$$_{iI} \qquad \qquad \qquad \qquad \qquad \qquad i'i'' \quad i'I \qquad (4)$$

Here $v_\ell(kR)$ is another solution of the left hand side of Eq. (1) set equal
to zero. The only requirement on this solution is that it be independent
of u_ℓ. It is chosen for computational convenience. The constant C_i
is the Wronskian of these two solutions to the homogeneous equations.
The $R_<$ and $R_>$ are the lesser and greater of R and R' respectively.

Using Eq. (2) we now may write

$$f(R) = u_\ell(k_i R)\ \delta_{iI} + \int_0^\infty \frac{u_\ell(k_i R_<)\, v_\ell(k_i R_>)}{C_i} \sum_{i'} V(R')f(R')dR' \quad (5)$$
$$_{iI} \phantom{= u_\ell(k_i R)\ \delta_{iI}} \phantom{\sum_{i'}}{}_{ii'}{}_{i'I}$$

This may now be reduced to a Voltera equation by breaking the integral into
two parts

$$f(R) = u_\ell(k_i R)\ \delta_{iI} + \int_0^R \frac{u_\ell(k_i R')\, v_\ell(k_i R)}{C_i} \sum_{i'} V(R')f(R')dR'$$
$$+ \int_R^\infty \frac{u_\ell(k_i R)\, v_\ell(k_i R')}{C_i} \sum_{i'} V_{ii'}(R')F(R')dR' \quad (6)$$

We may write the third term on the right hand size of Eq. (6) as the
integral from zero to infinity minus the integral from zero to R giving

$$f(R) = u_\ell(k_i R)\ \delta_{iI} + \int_0^R \frac{u_\ell(k_i R')\, v_\ell(k_i R) - u_\ell(k_i R)\, v_\ell(k_i R')}{C_i} \times$$
$$\sum_{i'} V(R')f(R')dR' + u_\ell(K_i R) \int_0^\infty \frac{v_\ell(k_i R')}{C_i} \sum_{i'} V(R')F(R')dR' \quad (7)$$

The last integral in Eq. (7) is a constant which may be determined at the
end of the integration.

It is of importance to note that the Kernel at R'=R is zero. Thus,
if the integral is replaced by a numerical quadrature and integration
proceeds toward large R the new point is always given in terms of previous
points only.

It is convenient to rewrite Eq. (7) in a form more amenable to numerical solution,

$$f_{iI}(R) = v_\ell(k_iR)T_{iI}(R) - u_\ell(k_iR)Q_{iI}(R) \tag{8}$$

$$G_{iI}(R) = \sum_{i'} \frac{V_{ii'}(R)f_{i'I}(R)}{C_i} \tag{9}$$

where

$$T_{iI}(R) = \int_0^R u_\ell(k_iR') \; G_{iI}(R') \; dR' \tag{10}$$

and

$$Q_{iI}(R) = \int_0^R v_\ell(k_iR') \; G_{iI}(R')dR' - \int_0^\infty v_\ell(k_iR') \; G_{iI}(R')dR' - \delta_{iI} \tag{11}$$

The quadratures in Eq. (10) and (11) may be replaced by Newton-Cotes type integration formulas and Eq. (8) then given $F_{iI}(R)$ on the equally spaced set of points R_n.

$$f_{iI^n}(R_n) = v_\ell(k_iR_n) \; T_{iI^n}(R_n) - u_\ell(k_iR_n) \; Q_{iI^n}(R_n) \tag{12}$$

where

$$T_{iI}(R_n) = T_{iI^{n-j}}(R_{n-j}) + \sum_{p=1}^{j-1} w_p^j \; G_{iI^{n-j+p}}(R_{n-j+p})u_\ell(k_iR_{n-j+p}) \tag{13}$$

and

$$Q_{iI^n}(R_n) = Q_{iI^{n-j}}(R_{n-j}) + \sum_{p=1}^{j-1} w_p^j \; G_{iI^{n-j+p}}(R_{n-j+p}) \; v_\ell(k_iR_{n-j+p}) \tag{14}$$

Here w_p^j is the p-th weight for a j point integration formula. Notice
the last point is not used. It would cancel in any case when T and Q are
substituted into Eq. (12).

From Eq. (11) one sees that as R approaches infinity $Q(R)_{II}$ must
approach $-\delta_{II}$. This would clearly happen if the proper value of $Q(R)_{II}$
were used initially. Since we do not know how to pick Q initially we start
the problem with a matrix formed from a complete set of linearly independent
Q_{II} vectors and then recombine them at the end of the integration to
obtain the desired asymptotic form. This gives us of course all initial
states at the same total energy whether we want them or not. We have usually
started by choosing $Q = -I$, the diagonal matrix with -1 on the diagonal.
This integration technique is usually used for problems with a strong repul-
sive potential near the origin. In this case it is not necessary to start
the integration at the origin. It is usually started well into the non-
classical region with the starting point determined experimentally. The
exact starting point is unimportant as the result does not depend on
where one starts as long as it is far enough into the nonclassical region.
The experimenting is necessary only to avoid wasting computer time integrat-
ing in the nonclassical region when it would serve just as well to start
farther out. As the radial wavefunction f decays rapidly in the nonclass-
ical region we take for the initial value of the T matrix the zero matrix.

In principle any order integration formula may be used. My group has
studied all of the Newton-Coates formulas up to order 9. We have found
that beyond the Simpson rule the higher order formulas are more time consuming
for the same precision in the answer. The Simpson's rule is faster than
trapezoid rule in that it allows much larger steps for the same accuracy.

It tends to be unstable for potentials with a long range interaction however. The trapezoid rule is the most versatile.

Using the trapezoid rule Eq. (13) and (14) become

$$T(R_n) = T(R_{n-1}) + h \, G(R_{n-1}) \, u_\ell(k_i R_{n-1}) \tag{15}$$

$$Q(R_n) = Q(R_{n-1}) + h \, G(R_{n-1}) \, v_\ell(k_i R_{n-1}) \tag{16}$$

For Simpson's rule we use

$$T(R_n) = \overline{T}(R_{n-1}) + 4/3h \, G(R_{n-1}) \, u_\ell(k_i R_{n-1}) \tag{17}$$

$$Q(R_n) = \overline{Q}(R_{n-1}) + 4/3h \, G(R_{n-1}) \, v_\ell(k_i R_{n-1}) \tag{18}$$

and

$$\overline{T}(R_n) = T(R_{n-2}) + 2/3h \, G(R_{n-1}) u_\ell(k_i R_{n-1}) \tag{19}$$

$$\overline{Q}(R_n) = Q(R_{n-2}) + 2/3hG(R_{n-1}) v_\ell(k_i R_{n-1}) \tag{20}$$

For the next step the roles of the barred and unbarred T and Q are interchanged.

As was mentioned earlier u_ℓ for open channels is a Riccati spherical Bessel function,

$$u_\ell(z) = zj_\ell(z) \tag{21}$$

we chose

$$v_\ell(z) = zn_\ell(z) \tag{22}$$

a Riccati spherical Neumann function. This gives us real equations to solve. For the closed channels we would like to choose the modified Riccati Bessel functions for complex argument, but we are in the end interested in the S matrix. We chose the Neumann function for v_ℓ to obtain real equation. A Hankel function is more appropriate for S matrix boundary conditions and inthe closed channels it may be chosen real. This is more convenient as the Hankel function for closed channels is a pure decaying exponential and thus has better numerical properties than the Neumann function which is a combination of a growing and a decaying exponential. For strongly closed channels the modified Bessel function grows and the Hankel function decays rapidly and we get into trouble with overflow and underflow quickly. Thus in order to avoid this we multiply the Bessel function by e^{-z} and the modified Hankel function by e^z which keep them a reasonable magnitude. Thus for closed channels we choose

$$u_\ell(z) = izj_\ell(iz)e^{-z} \tag{23}$$

and

$$v_\ell(z) = izh_\ell^+(iz)e^z . \tag{24}$$

Of course the closed channel Q and T will be modified by a constant. Furthermore the constant is different at each point in the integration. Thus the closed channel Q and T must be multiplied by $e^{k_i h}$ and $e^{-k_i h}$ to advance the constant to the next step. With these choices of v_ℓ the constant C_i in Eqs. (4-7) and (9) is $-|k_i|$.

Ending the integration is very simple. One sees from Eq. (8) that asymptotically

$$f(R) \underset{\text{II } R \to d}{\sim} v(k_i R) \underset{\text{iI}}{T} - u_\ell(k_i R) \underset{\text{iI}}{Q} \tag{25}$$

To obtain the proper boundary condition we multiply this equation from the left by $-\underline{Q}^{-1}$ to obtain

$$\overline{\underline{R}} = -\underline{T} \; \underline{Q}^{-1} \tag{26}$$

The R matrix is obtained from $\overline{\underline{R}}$ by

$$\underline{R} = \underline{k}^{-1/2} \; \overline{\underline{R}} \; k^{1/2} \tag{27}$$

where $\underline{k}$ is a diagonal matrix with the channel wavenumbers on the diagonal. The S-matrix is obtained from

$$\underline{S} = (\underline{I} + i\underline{R}) \; (\underline{I} - i\underline{R})^{-1} \tag{28}$$

where we use only the open channel portion of $\underline{R}$ in Eq. (28). If we had used Neumann functions in the closed channels it would have been necessary to use the entire R matrix in Eq. (28), but as the closed channels were already treated using S matrix boundary conditions only the open channels of the R matrix should be used to compute the open channel S matrix.

In the early part of the calculation, in the nonclassical repulsive region of the interaction potential the radial wavefunction for each channel is a growing exponential as integration is carried toward large R. Since the integration was started with an arbitrary Q matrix each channel is a mixture of all channels. Thus the most rapidly growing channel will tend to dominate each column of the Q matrix and the matrix will tend to become singular. To overcome this difficulty it is necessary from time to

time to form linear combinations of the columns of Q which are linearly independent. It is only necessary to do this a few times during a calculation. Since it need not be done often there is no need to do it in the most efficient way. A simple approach is to multiply Eq. (12) from the right by minus the inverse of the Q matrix thus resetting Q to minus a unit matrix and replacing $\underline{T}$ by $\underline{\bar{R}}$ of Eq. (26). This is convenient since it requires no further programming. Eq. (26) must be evaluated at the end of the program in any case. The points in the calculation at which this stabilization must be performed are determined experimentally. As a rule of thumb the first stabilization is done after a few integration steps. Each succeeding stabilization is done after approximately 5 times as many steps as the previous until the integration emerges from the nonclassical region. At this point it is usually possible to drop closed channels. After emerging from the nonclassical region, if closed channels are dropped, it is not necessary to stabilize any longer.

A listing of the program is given in the appendix.

References

1. The method discussed here is that of W.N. Sams and D.J. Kouri, J.
    Chem. Phys. $\underline{51}$, 4809 (1969); $\underline{51}$, 4815 (1969), as modified by our
    proup through the years. The derivation presented here is complete
    excect for the discussion of the mixed R-matrix and S-matrix bound-
    ary conditions discussed by D. Secrest, Methods in Computational
    Physics $\underline{10}$, 243 (1971), edited by B. Alder, S. Fernbach and M. Rot-
    enberg (Academic Press, N. Y.). Unpublished contributions to this
    work have been made through the years by various members of our group,
    including Walter Eastes, Verena Stocker, Timothy Heil, George Pfeffer,
    and Joseph Eccles.

```
     PROGRAM NRCC(INPUT,OUTPUT,TAPE50,TAPE5=INPUT,TAPE6=OUTPUT)
     REAL EK(50),WK(50),PU(50),PV(50),DR(50),DT(50),ELEV(50),V(2500),
    1 A(300),R(50,50),T(50,50),G(50,50),P(50,50),K(50,50)
     COMPLEX Q(50,50),S(50,50),RDEL
     INTEGER LQ(50),JQ(50),NSTAB(100),STB,DRP,CLC
     LOGICAL STAB,KSEE,DROP
     DATA ID,IDX,NC,NSTEP,NSTABI/50,300,0,0,-1/
C
C        THIS PROGRAM PERFORMS AN EXACT CLOSE-COUPLING MOLECULAR SCATTERING
C        CALCULATION FOR UP TO ID COUPLED CHANNELS BY THE METHOD OF SAMS AND
C        KOURI AND DEVELOPED BY OTHERS.  THE SYSTEM OF COUPLED INTEGRAL EQUATIONS
C        IS SOLVED USING THE TRAPEZOID RULE.
C
C        STABILIZATIONS
C        ONE CAN CHOOSE TO APPLY STABILIZING TRANSFORMATIONS DURING THE NUMERICAL
C        INTEGRATION BY SETTING THE LOGICAL VARIABLE STAB TO .TRUE. AT EXECUTION
C        TIME.  A LIST OF STABILIZATION POINTS MUST THEN BE INPUT BY STEP NUMBER.
C        ONE CAN CHOOSE TO LOOK AT THE REACTANCE MATRIX AT EACH OF THESE POINTS
C        BY SETTING THE LOGICAL VARIABLE KSEE TO .TRUE. AT EXECUTION TIME.
C
C        DROPPING CLOSED CHANNELS
C        CLOSED CHANNELS CAN BE DELETED AFTER A CERTAIN POINT IN THE INTEGRATION
C        BY SETTING THE LOGICAL VARIABLE DROP TO .TRUE. AT EXECUTION TIME.  THE
C        STEP NUMBER OF THE POINT AFTER WHICH CLOSED CHANNELS ARE TO BE NEGLECTED
C        MUST THEN BE READ.
C
C        WRITTEN BY KELLY MCLENITHAN (JUNE, 1979).
C
     CALL SECOND(T1)
     ASSIGN 60 TO STB
     ASSIGN 90 TO DRP
C
C        DATA INPUT
C
     READ(5,1000)STAB,KSEE,DROP,XBEG,XEND,STEP
     NTOTAL=IFIX((XEND-XBEG)/STEP+0.5)
     IF(.NOT.STAB)GO TO 10
     READ(5,1010)NSST
     DO 5 I=1,NSST
   5 READ(5,1010)NSTAB(I)
     ASSIGN 55 TO STB
     NSTABI=NSTAB(1)
  10 IF(.NOT.DROP)GO TO 15
     READ(5,1010)NSDR
     ASSIGN 89 TO DRP
C
C        COMPUTE BASIS SET (ENERGY LEVELS, QUANTUM NUMBERS, AND WAVE VECTORS)
C
  15 CALL POT(E,XBEG,W,V,CP,WK,4,1,0,ICOUNT)
     CALL LIST(NQNO,XMU,JQ,LQ,ELEV)
```

```
C
C        COUNT OPEN/CLOSED CHANNELS
C
         NOPE=NTP-NQNO
         DO 20 I=1,NQNO
         WEK=WK(I)
         IF(WEK.GE.0.0)GO TO 20
         NOPE=NOPE-1
         WEK=-WEK
   20 EK(I)=SQRT(WEK)
         ASSIGN 63 TO CLC
         IF(NOPE.EQ.NQNO)ASSIGN 75 TO CLC
C
C        DATA OUTPUT
C
         WRITE(6,2000)E,XMU,NQNO,NOPE,NQNO-NOPE,XBEG,XEND,STEP,NTOTAL
         IF(.NOT.STAB)GO TO 30
         WRITE(6,2010)NSST
         DO 25 I=1,NSST
   25 WRITE(6,2020)NSTAB(I),XBEG+NSTAB(I)*STEP
   30 IF(.NOT.DROP.OR.NOPE.EQ.NQNO)GO TO 35
         WRITE(6,2030)NSDR
   35 WRITE(6,2040)
         DO 40 I=1,NQNO
   40 WRITE(6,2050)I,JQ(I),LQ(I),ELEV(I),WK(I),EK(I)
C
C        PROGRAM INITIALIZATION
C
         CALL ZEROZ(R,ID*NQNO)
         CALL ZEROZ(T,ID*NQNO)
         LUP=0
         DO 50 I=1,NQNO
         LUP=MAX0(LUP,LQ(I))
         IF(I.LE.NOPE)GO TO 45
         DR(I)=EXP(EK(I)*STEP)
         DT(I)=1.0/DR(I)
         GO TO 50
   45 DR(I)=DT(I)=1.0
   50 R(I,I)=1.0
         X=XBEG
         XMU2=2.0*XMU
         SX2=STEP*XMU2
         ILO=NOPE+1
         IST=1
C
C        BEGIN NUMERICAL INTEGRATION (TRAPEZOID RULE)
C
C        STABILIZATION TEST
C
         GO TO STB(55,60)
```

```
    55 IF(NSTEP.EQ.NSTABI)GO TO 95
C
C         COMPUTE BESSEL FUNCTIONS
C
    60 CALL BESSEL(PU,PV,EK,LQ,X,STEP,NOPE,NC,A,1.0,IDX)
       GO TO CLC(63,75)
    63 DO 70 I=ILO,NQNO
       DO 65 J=1,NQNO
       R(I,J)=DR(I)*R(I,J)
    65 T(I,J)=DT(I)*T(I,J)
       Z=EK(I)*X
       CALL MSBESJ(Z,LUP,A,1.0,IDX)
       PU(I)=Z*A(LQ(I)+1)
       CALL MSBESH(Z,LUP,A,1.0,IDX)
    70 PV(I)=Z*A(LQ(I)+1)
C
C         CONTINUE NUMERICAL INTEGRATION
C
    75 DO 80 J=1,NTP
       DO 80 I=1,NTP
    80 G(I,J)=PU(I)*R(I,J)-PV(I)*T(I,J)
       CALL POT(E,X,W,V,CP,WK,2,2,0,ICOUNT)
       DO 82 J=1,NTP
       DO 82 I=1,NTP
       IM=I-NQNO
       SUM=0.0
       DO 81 M=1,NTP
       IM=IM+NQNO
    81 SUM=SUM+V(IM)*G(M,J)
    82 P(I,J)=SUM
       DO 85 I=1,NTP
       TT=SX2/EK(I)
       TU=PU(I)*TT
       TV=PV(I)*TT
       DO 85 J=1,NTP
       T(I,J)=T(I,J)+P(I,J)*TU
    85 R(I,J)=R(I,J)+P(I,J)*TV
C
C         DROP CLOSED CHANNELS
C
       GO TO DRP(89,90)
    89 IF(NSTEP.NE.NSDR)GO TO 90
       NTP=NOPE
       ASSIGN 90 TO DRP
       ASSIGN 75 TO CLC
    90 NSTEP=NSTEP+1
       X=XBEG+NSTEP*STEP
       IF(NSTEP.LE.NTOTAL)GO TO STB(55,60)
C
C         END NUMERICAL INTEGRATION
```

```
C
        NSTEP=NTOTAL
        X=XBEG+NTOTAL*STEP
        STAB=.FALSE.
C
C        STABILIZATION AND TERMINAL REACTANCE MATRIX COMPUTATION
C
    95 IST=IST+1
        IF(IST.GT.NSST)GO TO 97
        NSTABI=NSTAB(IST)
        GO TO 98
    97 ASSIGN 60 TO STB
    98 CALL LINSYR(R,T,NTP,ID)
        DO 100 J=1,NTP
        SQEKJ=1.0/SQRT(EK(J))
        DO 100 I=1,NTP
   100 K(I,J)=-SQRT(EK(I))*T(I,J)*SQEKJ
        IF(.NOT.STAB)GO TO 110
        CALL ZEROZ(R,ID*NQNO)
        DO 105 I=1,NTP
   105 R(I,I)=1.0
C
C        OUTPUT REACTANCE MATRIX
C
        IF(.NOT.KSEE)GO TO 115
   110 WRITE(6,2060)NSTEP,X
        CALL MATPR(K,JQ,LQ,NTP,ID)
   115 IF(STAB)GO TO 60
C
C        COMPUTATION OF SCATTERING MATRIX
C
        DO 120 J=1,NOPE
        DO 120 I=1,NOPE
        KDEL=(0.0,0.0)
        IF(I.EQ.J)KDEL=(1.0,0.0)
        S(I,J)=KDEL+(0.0,1.0)*K(I,J)
   120 Q(I,J)=KDEL-(0.0,1.0)*K(I,J)
        CALL LINSYT(Q,S,NOPE,ID)
        WRITE(6,2070)
        CALL MATPC(S,JQ,LQ,NOPE,ID)
C
C        COMPUTATION OF TRANSITION PROBABILITIES
C
        WRITE(6,2080)
        DO 125 J=1,NOPE
        SUM=0.0
        DO 125 I=1,NOPE
        TPROB=S(I,J)*CONJG(S(I,J))
        SUM=SUM+TPROB
        WRITE(6,2090)JQ(J),LQ(J),JQ(I),LQ(I),TPROB
```

```
   125 IF(I.EQ.NOPE)WRITE(6,2100)SUM
       CALL SECOND(T2)
       WRITE(6,2110)T2-T1
C
C          I/O FORMAT STATEMENTS
C
C          INPUT FORMATS
C
  1000 FORMAT(3L1,3E10.0)
  1010 FORMAT(I5)
C
C          OUTPUT FORMATS
C
  2000 FORMAT(41X,#TOTAL ENERGY#,20X,1PE14.7,# HARTREE#/41X,#SYSTEM #,
      1  #REDUCED MASS#,13X,1PE14.7,# ELECTRON MASSES#/41X,#TOTAL NUMBER#
      2  ,# OF CHANNELS#,5X,I5/54X,#OPEN#,12X,I5/54X,#CLOSED#,10X,I5/#0#,
      3  14X,#NUMERICAL INTEGRATION
      4  41X,#STOP#,24X,F11.4,8X,#BOHR#/41X,#STEP SIZE#,19X,F11.4,8X,
      5  #BOHR#/41X,#TOTAL STEPS#,16X,I7)
  2010 FORMAT(41X,#STABILIZATIONS#,16X,I4/56X,#STEP#,6X,#COORDINATE #,
      1  #(BOHR)#)
  2020 FORMAT(55X,I5,9X,F11.4)
  2030 FORMAT(41X,#DROP CLOSED CHANNELS, STEP#,3X,I5)
  2040 FORMAT(#-#,56X,#BASIS SET INFORMATION#/#0#,32X,#INDEX#,5X,#J#,6X
      1  ,#L#,8X,#E(J)#,11X,#K(J)**2#,12X,#K(J)#/# #)
  2050 FORMAT(34X,I3,4X,I3,4X,I3,3X,1PE14.7,3X,1PE14.7,3X,1PE14.7)
  2060 FORMAT(#0 REACTANCE MATRIX AT STEP #,I6,#,   X =#,F11.4)
  2070 FORMAT(#- OPEN CHANNEL SCATTERING MATRIX#)
  2080 FORMAT(#-#,29X,#INITIAL STATE#,4X,#FINAL STATE#,4X,#TRANSITION #,
      1  #PROBABILITY#,10X,#UNITARITY#/33X,#J#,5X,#L#,9X,#J#,5X,#L#)
  2090 FORMAT(32X,I2,3X,I3,8X,I2,3X,I3,10X,1PE14.7)
  2100 FORMAT(#+#,90X,F14.11)
  2110 FORMAT(#-#,40X,#CPU TIME#,22X,F8.3,# SECONDS#)
       STOP
       END
       SUBROUTINE MMULT(A,B,M,N,L,C)
       REAL A(M,M),B(N,N),C(N,N)
C
C          THIS SUBROUTINE PERFORMS THE MATRIX MULTIPLICATION C=A*B, WHERE A HAS
C          BEEN GIVEN IN GENERAL COLUMN STORAGE (AS M COLUMNS, EACH COLUMN BEING
C          M ELEMENTS LONG).  THE RESULT IS RETURNED IN C.
C
       DO 2 J=1,L
       DO 2 I=1,L
       SUM=0.0
       DO 1 K=1,L
     1 SUM=SUM+A(I,K)*B(K,J)
     2 C(I,J)=SUM
       RETURN
       END
```

```
      SUBROUTINE MATPR(A,J,L,ND,ID)
      DIMENSION A(ID,ID),J(ID),L(ID)
C
C        THIS SUBROUTINE PRINTS REAL MATRIX A WITH COUPLED CHANNEL LABELS.
C
      DO 3 IBEG=1,ND,7
      IEND=MIN0(IBEG+6,ND)
      WRITE(6,1)
    1 FORMAT(3X,#J#,5X,#L#)
      WRITE(6,2) (J(ICOL),L(ICOL),ICOL=IBEG,IEND)
    2 FORMAT(#+#,15X,6(I2,3X,I3,9X),I2,3X,I3)
      DO 3 IROW=1,ND
    3 WRITE(6,4)J(IROW),L(IROW),(A(IROW,ICOL),ICOL=IBEG,IEND)
    4 FORMAT(2X,I2,3X,I3,1X,7(2X,1PE15.8))
      RETURN
      END
      SUBROUTINE MATPC(A,J,L,ND,ID)
      COMPLEX A(ID,ID)
      INTEGER J(ID),L(ID)
C
C        THIS SUBROUTINE PRINTS COMPLEX MATRIX A WITH COUPLED CHANNEL LABELS AND
C        CLEARLY DISTINGUISHES ITS REAL AND IMAGINARY PARTS.
C
      DO 4 IBEG=1,ND,7
      IEND=MIN0(IBEG+6,ND)
      WRITE(6,1)
    1 FORMAT(#0   J#,5X,#L#)
      WRITE(6,2) (J(ICOL),L(ICOL),ICOL=IBEG,IEND)
    2 FORMAT(#+#,15X,6(I2,3X,I3,9X),I2,3X,I3)
      DO 4 IROW=1,ND
      WRITE(6,3)J(IROW),L(IROW),(REAL(A(IROW,ICOL)),ICOL=IBEG,IEND)
    3 FORMAT(#0 #,I2,3X,I3,1X,7(2X,1PE15.8))
    4 WRITE(6,5)(AIMAG(A(IROW,ICOL)),ICOL=IBEG,IEND)
    5 FORMAT(11X,7(2X,1PE15.8))
      RETURN
      END
      SUBROUTINE SBESN    (Z,LMAX,S,A,ID)
C     CALCULATES THE VARIOUS SPHERICAL BESSEL FUNCTIONS UP TO ORDER IMAX
C        WITH ARGUMENT Z
C     THE DEFINITIONS ARE THOSE OF MESSIAH, VOL. I, APPENDIX B.
      LOGICAL SBJ
      DIMENSION S(ID)
      DATA LUMAX/50/
C     ID IS USED TO CHECK THAT NO OVERRUN OCCURS IN COMPUTING S; IF
C     OVERRUN OCCURS ERROR MESSAGE NUMBER 2 IS PRINTED OUT
C
C     RETURNS THE SPHERICAL NEUMAN FUNCTION
C        S(L+1) = N(L;Z)/A**L
      S(1) = COS(Z)/Z
      S(2) = (S(1)/Z + SIN(Z)/Z)/A
```

```
      AT1 = 1.0/(A*Z)
      AT2 = -1.0/A**2
      GO TO 150
C
      ENTRY MSBESH
C     RETURNS THE MODIFIED SPHERICAL HANKEL FUNCTION
C        S(L+1) = H+(L;IZ)*I**(L+1)*EXP(Z)/A**L
      S(1) = 1.0/Z
      S(2) = (S(1)**2 + S(1))/A
      AT1 = 1.0/(A*Z)
      AT2 = 1.0/A**2
C     COMPUTES THE SET OF FUNCTIONS BY UPWARD RECURSION FROM THE
C     FIRST TWO
150   LTEMP = LMAX + 1
      LP = 1
      DO 180 L=3,LTEMP
      LP = LP + 2
180   S(L) = LP*S(L-1)*AT1 + S(L-2)*AT2
      RETURN
C
      ENTRY SBESJ
C     RETURNS THE REGULAR SPHERICAL BESSEL FUNCTION
C        S(L+1) = J(L;Z)*A**L
      S1 = SIN(Z)/Z
      AT1 = 1.0/(A*Z)
      AT2 = -1.0/A**2
      SBJ = .TRUE.
      GO TO 300
C
      ENTRY MSBESJ
C     RETURNS THE MODIFIED SPHERICAL BESSEL FUNCTION
C        S(L+1) = J(L;IZ)*I**(-L)*EXP(-Z)*A**L
      ETZ = 0.0
      IF (Z .LT. 80.0) ETZ = EXP(-2.0*Z)
      S1 = (1.0 - ETZ)/(2.0*Z)
      AT1 = 1.0/(A*Z)
      AT2 = 1.0/A**2
      SBJ = .FALSE.
C     TEST TO SEE IF DOWNWARD RECURSION IS FEASIBLE
  300 IF (Z.GT.LMAX.AND.SBJ) GO TO 500
      PMAX = 1.0E8
      LP = 2*LMAX + 1
      PRD = LP/Z
      DO 310 L=2,LOMAX
      LP = LP + 2
      PRD = LP*PRD/Z
      IF (PRD .GE. PMAX) GO TO 320
310   LOVER = L
C
C        MUST USE UPWARD RECUSION
```

```
C
      IF (SBJ) GO TO 500
  315 S(1, = S1
      S(2) = (1.0/(2.0*Z) - 1.0/(2.0*Z*Z) + ETZ*(1.0/(2.0*Z)
     2   + 1.0/(2.0*Z*Z)))*A
      AT1 = -A/Z
      AT2 = A*A
      GO TO 150
C     COMPUTE THE SET OF FUNCTIONS BY DOWNWARD RECURSION STARTING
C        WITH ARBITRARY HIGH ORDER VALUES
320   LTOP = LMAX + LOVER + 10
      IF (LTOP .GT. ID) GO TO 700
      S(LTOP) = 0.0
      S(LTOP-1) = 1.0E-20
      LP = 2*LTOP - 1
      LLTOP = LTOP - 1
      DO 350 LL=2,LLTOP
      LP = LP - 2
      L = LTOP - LL
350   S(L) = LP*S(L+1)*AT1 + S(L+2)*AT2
C     NORMALIZE THE DESIRED FUNCTIONS TO THE PROPER VALUES
      SFAC = S1/S(1)
      S(1) = S1
      LLMAX = LMAX + 1
      DO 380 L=2,LLMAX
380   S(L) = S(L)*SFAC
      RETURN
C     DOWNWARD RECURSION IS NOT FEASIBLE SO INITIALIZE AND COMPUTE
C        BY UPWARD RECURSION.
  500 S(1) = S1
      S(2) = (S1 - COS(Z))*A/Z
      AT1 = A/Z
      AT2 = -A**2
      GO TO 150
  700 WRITE 2000
      WRITE 2001, Z,LMAX,A
 2000 FORMAT(#0*** ERROR *** SBES *** ARRAY TOO SMALL#,
     2  # FOR DOWNWARD RECURSION#)
 2001 FORMAT (1H ,#ARG. =#,1PE23.15,#  ORDER =#,I10,
     $  #  FACTOR#,D23.15)
      STOP
      END
      SUBROUTINE BESSEL(PU,PV,EK,L,R,H,N,NC,SJ,A,ID)
C
C        RETURNS THE SCALED RICATTI BESSEL FUNCTIONS. PU IS THE REGULAR
C          SOLUTION U*A**L, AND PV IS THE IRREGULAR SOLUTION, V*A**(-L),
C          WHERE A IS THE SCALING FACTOR DISCRIBED BELOW.
C        EK IS THE ARRAY OF WAVE NUMBERS (ONE ELEMENT PER CHANNEL).
C        L IS THE ARRAY OF ORBITAL ANGULAR MOMENTUM QUANTUM NUMBERS.
C        R IS THE RADIAL VARIABLE.
```

```
C          H IS THE STEP-SIZE OF THE INTEGRATION (MUST REMAIN CONSTANT).
C          N IS THE NUMBER OF CHANNELS (OPEN CHANNELS ONLY).
C          NC IS THE NUMBER OF STEPS SINCE SBESJ WAS CALLED (SET TO ZERO
C             BY CALLING PROGRAM FOR FIRST STEP ONLY; UPDATED INTERNALLY)
C          SJ IS A SCRATCH ARRAY FOR CALLING SBESJ AND SBESN.
C          A IS A SCALING FACTOR
C          ID IS THE DIMENSION OF SJ (NEEDED BY SBESJ)
C
      REAL EK(1),PU(1),PV(1),SJ(1)
      REAL RU(40),RV(40),Q(40),RUN(40),RVN(40),EL(40),E(40)
      REAL PUL(40), PVL(40)
      INTEGER L(1)
      IF (N .LE. 40) GO TO 2
      WRITE 1000, N
1000  FORMAT(#0BESSEL CALLED WITH N =#,I3,#. THIS IS LARGER THAN THE ARR
     1AYS IN BESSEL#)
      STOP
2     R2 = R**2
      NC = NC + 1
      IF(NC - 2) 1,11,101
    1 CONTINUE
      H2012 = H**2/12.0
      SM = 0.0
      DO 5 I = 1,N
      E(I) = EK(I)**2
      EL(I) = L(I)*(L(I) + 1)
      Z = R*EK(I)
      S = EL(I)/R2 - E(I)
      IF(S .GT. SM) SM = S
      T = 1.0 - H2012*S
      CALL SBESJ(Z,L(I),SJ,A,ID)
      PU(I) = SJ(L(I) + 1)*Z
      RU(I) = PU(I)*T
      CALL SBESN(Z,L(I),SJ,A,ID)
      PV(I) = SJ(L(I) + 1)*Z
5     RV(I) = PV(I)*T
      M = 100000
      IF(SM .GT. 0.0) M = 6.0/SQRT(SM)/H
      RETURN
11    M = M - 1
      DO 15 I = 1,N
      Z = R*EK(I)
      T = 1.0 - H2012*(EL(I)/R2 - E(I))
      CALL SBESJ(Z,L(I),SJ,A,ID)
      PU(I) = SJ(L(I) + 1)*Z
      RUN(I) = PU(I)*T
      CALL SBESN(Z,L(I),SJ,A,ID)
      PV(I) = SJ(L(I) + 1)*Z
      RVN(I) = PV(I)*T
      PUL(I) = PU(I)
```

```
      PVL(I) = PV(I)
15    Q(I) = 12.0 - 10.0*T
      RETURN
101   M = M - 1
      DO 100 I = 1,N
      SU = PUL(I)*Q(I) - RU(I)
      SV = PVL(I)*Q(I) - RV(I)
      T = 1.0 - H2012*(EL(I)/R2 - E(I))
      PU(I) = SU/T
      PV(I) = SV/T
      RU(I) = RUN(I)
      RV (I) = RVN(I)
      RUN(I) = SU
      RVN(I) = SV
      PUL(I) = PU(I)
      PVL(I) = PV(I)
100   Q(I) = 12.0 - 10.0*T
      IF(M .LE. 0) NC = 0
      RETURN
      END
      SUBROUTINE LNSYR2(A,B,N,ID,NB1)
C
C        SOLVES MATRIX EQUATION XA = B WHERE A,B,X ARE REAL N BY N
C     MATRICES
C     RESULT APPEARS IN B;  A IS DESTROYED
C     ID IS THE DIMENSION GIVEN A,B IN THE CALLING PROGRAM
C
C        NAME     -- LINSYR/LNSYR2
C
C---------------------------------------------------------------
C
C        THIS ROUTINE FUNCTIONS JUST AS BEFORE IF ENTRY LINSYR IS
C           USED. IF THE CALL IS THROUGH ENTRY LNSYR2, THE
C           B AND X MAY BE RECTANGULAR (NB1 BY N) MATRICES.
C
C---------------------------------------------------------------
C
      REAL A(ID,ID), B(ID,ID)
      NB = NB1
      GO TO 10
      ENTRY LINSYR
      NB = N
   10 CONTINUE
      NS = N - 1
      IF (NS .LE. 0) GO TO 300
      DO 200 I=1,NS
      IG = I + 1
      AT = A(I,I)
      DO 120 J=IG,N
120   A(J,I) = A(J,I)/AT
```

```
      DO 130 J=1,NB
 130  B(J,I) = B(J,I)/AT
      DO 150 K=1,N
      IF (K .EQ. I) GO TO 150
      AT = A(I,K)
      DO 140 L=IG,N
 140  A(L,K) = A(L,K) - A(L,I)*AT
      DO 145 L=1,NB
 145  B(L,K) = B(L,K) - B(L,I) *AT
 150  CONTINUE
 200  CONTINUE
      AT = A(N,N)
      DO 230 J=1,NB
 230  B(J,N) = B(J,N)/AT
      DO 250 K=1,NS
      AT = A(N,K)
      DO 250 L=1,NB
 250  B(L,K) = B(L,K) - B(L,N)*AT
      RETURN
  300 CONTINUE
      AT = A(1,1)
      DO 301 I=1,NB
  301 B(I,1) = B(I,1)/AT
      RETURN
      END
      SUBROUTINE LINSYT (A,B,N,ID)
C     SOLVES MATRIX EQUATION XA = B, WHERE A,B,X COMPLEX N BY N
C     MATRICES
C     RESULT APPEARS IN B;  A IS DESTROYED
C     ID IS THE DIMENSION GIVEN A,B IN THE CALLING PROGRAM
      COMPLEX A(ID,ID),B(ID,ID),AT
      NS = N - 1
      IF (NS .LE. 0) GO TO 300
      DO 200 I=1,NS
      IG = I + 1
      AT = A(I,I)
      DO 120 J=IG,N
 120  A(J,I) = A(J,I)/AT
      DO 130 J=1,N
 130  B(J,I) = B(J,I)/AT
      DO 150 K=1,N
      IF (K .EQ. I) GO TO 150
      AT = A(I,K)
      DO 140 L=IG,N
 140  A(L,K) = A(L,K) - A(L,I)*AT
      DO 145 L=1,N
 145  B(L,K) = B(L,K) - B(L,I) *AT
 150  CONTINUE
 200  CONTINUE
      AT = A(N,N)
```

```
      DO 230 J=1,N
230   B(J,N) = B(J,N)/AT
      DO 250 K=1,NS
      DO 250 L=1,N
      AT = A(N,K)
250   B(L,K) = B(L,K) - B(L,N)*AT
      RETURN
300   B(1,1) = B(1,1)/A(1,1)
      RETURN
      END
      SUBROUTINE ZEROZ (Z,N)
C
C         INITIALIZES THE FIRST N STORAGE LOCATIONS OF REAL MATRIX Z TO ZERO.
C
      REAL Z(N)
      DO 1 K = 1,N
  1   Z(K) = 0.0
      RETURN
      END
```

'

APPLICATIONS OF CLOSE COUPLING ALGORITHMS
TO ELECTRON-ATOM, ELECTRON-MOLECULE,
AND ATOM-MOLECULE SCATTERING

Donald G. Truhlar, Nancy Mullaney Harvey,
Kunizo Onda, and Maynard A. Brandt
Department of Chemistry and Chemical Physics Program,
University of Minnesota, Minneapolis, Minnesota 55455

Abstract. We discuss some of the details of our implementation of the
Numerov and R matrix propagation methods for close coupling calculations.
We discuss some of the successes and problems we have had applying these
and other methods in various applications, and we present some execution
times for runs we made to compare various methods.

I. Introduction

In this contribution, we summarize the procedures we currently use to solve close coupling equations for electron-atom, electron-molecule, and atom-molecule collisions, and we discuss some of our relevant past experiences in applying various numerical approaches to close coupling calculations. To facilitate comparison of these methods we also present some timing information which we gathered for this workshop. The numerical methods we compare here are the Numerov method,[1-11] the piecewise analytic method of Gordon with linear reference potential,[12-16] the integral equations algorithm of Sams and Kouri,[17-23] and the R matrix propagation method of Light and Walker with piecewise constant reference potential.[24-29]

In section II we introduce all four methods and discuss in detail our implementation of the Numerov[8] and R matrix propagation[29] methods. In section III we discuss applications of all four methods to elastic and vibrationally and rotationally inelastic electron scattering by N_2 using anisotropic model potentials. Section IV gives applications to electronically inelastic electron scattering by H. Sections V and VI discuss vibrationally and rotationally inelastic atom-diatom scattering, respectively.

II. Discussion of methods

A. Introduction. For this workshop we have compiled some comments on our experiences with close coupling calculations of electron-molecule, electron-atom, and atom-molecule scattering. In this section we provide a detailed description of our implementation of the Numerov and R matrix propagation methods. We also briefly introduce the codes which we used to apply the piecewise analytic method and the Sams-Kouri method, and we comment on our timing studies and on our use of adiabatic basis functions.

B. Numerov method. We have found the Numerov method to be convenient and reliable for many problems. Since the method is easy to apply, the computer program is relatively transparent and easy to modify and simple checks of convergence with respect to numerical parameters can be made.

The Numerov method is a hybrid finite difference method applicable to any set of second-order linear differential equations containing no first derivatives. It is a sixth-order method, i.e., if h denotes the stepsize and $f(r)$ denotes the exact solution of

$$\frac{d^2 f(r)}{dr^2} = D(r) f(r) \tag{1}$$

then the leading term in the truncation error per step is $\dfrac{h^6}{240} \dfrac{d^6 f}{dr^6}\bigg|_{r=r^*}$ where r^* is some (generally unknown) point in the interval.[4,30] Blatt[4] has suggested that the Numerov method is "the method of choice for the integration of (1) because it is the highest-order method which is at the same time a three-point method." However, the cumulative error in the Numerov method is of order h^4, which is the same order as the Runge-Kutta method which has an h^5 truncation error per step.[31,32]

Lester[33] has pointed out a minor disadvantage of the Numerov
method, namely it requires different algorithms for doubling and halving
the stepsize, respectively. For this reason, Lester chose the deVogelaere
method,[34] which is a variable stepsize method and requires only a single
algorithm for changing the stepsize. Allison[5] compared the regular Numerov,
iterative Numerov, and deVogelaere methods for a test case[36] involving rota-
tional excitation in an atom-rigid-rotator collision, and he found the iter-
ative Numerov method to be the fastest for a given precision. His tests
are significant in that they were run on the same computer using computer
codes written by the same author thereby eliminating two of the major varia-
bles usually existing in such comparisons. He also compared his Numerov
program to Gordon's program[35] using the piecewise analytic method for the
same test case, and he found for calculations of similar precision that the
execution times were comparable for problems having as many as nine channels.
Our program is based mainly on the work of Allison.[5] Further development of
the method has been carried out by Johnson who calls the resulting algorithm
the renormalized Numerov method.[34] As compared to the original matrix Numerov
method, Johnson makes two transformations. The first eliminates one matrix
multiplication and is identical to a transformation used by Allison. The
method obtained after this transformation is called the regular Numerov
method. Johnson's second tranformation is to define a ratio matrix
($Y_{n+1} \, Y_n^{-1}$ is the notation used below). This does not change the amount of
computation per step, but it does eliminate the need for stablizing trans-
formations. As discussed below, our program uses stablizing transformations
but our experience has been that these have not required a great amount of
computational effort. In the following general review of the numerical

techniques, we use Allison's equations wherever possible, but we change the notation somewhat to conform more closely with the notation used else-where in this report.

The set of N coupled second order differential equations to be solved are

$$\frac{d^2}{dr^2} \underset{\sim}{f}(r) - \underset{\sim}{D}(r) \underset{\sim}{f}(r) = \underset{\sim}{0} \tag{2}$$

with

$$D_{ij}(r) = - \left[k_i^2 - \frac{\ell_i(\ell_i + 1)}{r^2} \right] \delta_{ij} + (2\mu/\hbar^2)V_{ij}(r) \tag{3}$$

We use the convention that $\underset{\sim}{A}$ denotes a matrix with elements A_{ij}. (The columns of $\underset{\sim}{A}$ are denoted $\underline{A}$.) Equation (2) is solved subject to the boundary conditions

$$\underset{\sim}{f}(0) = \underset{\sim}{0} \tag{4}$$

and

$$\underset{\sim}{f}(r) \underset{r \to \infty}{\sim} \underset{\sim}{S}(r)\underset{\sim}{P} + \underset{\sim}{C}(r)\underset{\sim}{Q} \tag{5}$$

where, for the case of all channels asymptotically open,

$$S_{ij}(r) \underset{r \to \infty}{\sim} k_i^{-\frac{1}{2}} \delta_{ij} \sin(k_i r - \tfrac{1}{2}\ell_i\pi) \tag{6}$$

and

$$C_{ij}(r) \underset{r \to \infty}{\sim} k_i^{-\frac{1}{2}} \delta_{ij} \cos(k_i r - \tfrac{1}{2}\ell_i\pi) \tag{7}$$

A subblock $\underset{\sim}{R}$ of the reactance matrix is then given as[37]

$$\underset{\sim}{R} = \underset{\sim}{Q}\underset{\sim}{P}^{-1} \tag{8}$$

This method of asymptotic analysis is easily generalized to the case where some channels are closed.

It is interesting to note that $\underline{f}(r)$ is not unique, i.e., for any non-singular matrix $\underline{G}$, $\underline{f}(r)\underline{G}$ satisfies the eqs. (2), (4), and (5) if $\underline{f}(r)$ does. However, $\underline{R}$ is given in terms of the $\underline{f}(r)\underline{G}$ solution as

$$\underline{R} = (\underline{Q}\underline{G})(\underline{P}\underline{G})^{-1} = \underline{Q}\underline{G}\underline{G}^{-1}\underline{P}^{-1} = \underline{Q}\underline{P}^{-1} \tag{9}$$

and is the same as for the $\underline{f}(r)$ solution. Since $\underline{P}^{-1}$ is required, the columns of $\underline{f}(r)$ must be linearly independent. Furthermore, any set of N linearly independent linear combinations of the columns of $\underline{f}(r)$ is an equally valid set of solution vectors satisfying (2), (4), and (5).

We use the notation $\underline{x}_n \equiv \underline{x}(r_n)$ where $\underline{x}$ is any matrix and r_n is a grid point, and we let the stepsize between grid points be $h \equiv r_{n+1} - r_n$. At a given point, the approximate solution $\underline{f}_{n+1}$ to (2) is calculated from $\underline{D}(r)$ evaluated at the equally spaced points r_{n+1}, r_n, and r_{n-1}, respectively, and from $\underline{f}_n$ and $\underline{f}_{n-1}$ as

$$(\underline{I} - \frac{h^2}{12}\underline{D}_{n+1})\underline{f}_{n+1} = (2\underline{I} + \frac{10}{12}h^2\underline{D}_n)\underline{f}_n - (\underline{I} - \frac{h^2}{12}\underline{D}_{n-1})\underline{f}_{n-1} \tag{10}$$

Let

$$\underline{F} = -\frac{h^2}{12}\underline{D} \tag{11}$$

and

$$\underline{Y}_{n+1} = (\underline{I} + \underline{F}_{n+1})\underline{f}_{n+1} \tag{12}$$

Then (10) may be rewritten as

$$\underline{Y}_{n+1} = 2\underline{Y}_n - 12\underline{F}_n\underline{f}_n - \underline{Y}_{n-1} \tag{13}$$

with

$$\underline{f}_{n+1} = (\underline{I} + \underline{F}_{n+1})^{-1}\underline{Y}_{n+1} \tag{14}$$

We refer to (13) and (14) together as the "regular" Numerov method.

In an attempt to speed up the calculation of $(\underset{\sim}{I} + \underset{\sim}{F}_{n+1})^{-1}$ Allison proposed the following iterative method of calculating $\underset{\sim}{f}_{n+1}$. Let $\underset{\sim}{d}(r)$ and $\underset{\sim}{\Delta}(r)$ be defined to have elements given by

$$d_{ij}(r) = F_{ii}(r) \; \delta_{ij} \tag{15}$$

and

$$\Delta_{ij}(r) = -(h^2/12) \sum_{\substack{k=1 \\ k \neq j}}^{N} V_{ik}(r) \; f_{kj}(r) \tag{16}$$

so

$$\underset{\sim}{F}(r) \; \underset{\sim}{f}(r) = \underset{\sim}{d}(r) \; \underset{\sim}{f}(r) - \underset{\sim}{\Delta}(r) \tag{17}$$

Putting (17) at r_n into (13) yields

$$\underset{\sim}{Y}_{n+1} = 2\underset{\sim}{Y}_n + 12(\underset{\sim}{\Delta}_r - \underset{\sim}{d}_n \underset{\sim}{f}_n) - \underset{\sim}{Y}_{n-1} \tag{18}$$

and putting (17) at r_{n+1} into (12) and solving for $\underset{\sim}{f}_{n+1}$ yields formally

$$\underset{\sim}{f}_{n+1} = (\underset{\sim}{I} + \underset{\sim}{d}_{n+1})^{-1} \left[\underset{\sim}{Y}_{n+1} + \underset{\sim}{\Delta}_{n+1} \right] \tag{19}$$

Since $\underset{\sim}{d}(r)$ is a diagonal matrix the evaluation of $(\underset{\sim}{I} + \underset{\sim}{d}_{n+1})^{-1}$ is trivial. But $\underset{\sim}{\Delta}_{n+1}$ on the right-side of (19) depends on $\underset{\sim}{f}_{n+1}$ so Allison proposed using Gauss-Seidel iteration[38] to converge (19) at the current grid point before using (18) for the next grid point, i.e.,

$$\underset{\sim}{T}_{n+1} = \lim_{m \to \infty} \underset{\sim}{T}_{n+1}^{(m)} \tag{20}$$

for $\underset{\sim}{T}(r) = \underset{\sim}{f}(r)$ or $\underset{\sim}{\Delta}(r)$.

The Gauss-Seidel iteration procedure is a method for finding the solution $\underset{\sim}{x}$ to the system of linear algebraic equations

$$\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b} \tag{21}$$

where $\underset{\sim}{A}$, $\underset{\sim}{x}$, and $\underset{\sim}{b}$ are $N \times N$ matrices. First one makes the replacement

$$\underset{\sim}{A} = \underset{\sim}{W} - \underset{\sim}{U} - \underset{\sim}{L} \tag{22}$$

where $\underset{\sim}{W}$ is a diagonal matrix, $\underset{\sim}{U}$ is a strictly upper triangular matrix, and $\underset{\sim}{L}$ is a strictly lower triangular matrix. Putting (22) into (21) and re-arranging yields

$$\underset{\sim}{x} = \underset{\sim}{W}^{-1}\underset{\sim}{U}\underset{\sim}{x} + \underset{\sim}{W}^{-1}\underset{\sim}{L}\underset{\sim}{x} + \underset{\sim}{W}^{-1}\underset{\sim}{b} \tag{23}$$

This motivates the Gauss-Seidel iteration procedure for solving (21);

$$\underset{\sim}{x}^{(m+1)} = \underset{\sim}{W}^{-1}\underset{\sim}{U}\underset{\sim}{x}^{(m)} + \underset{\sim}{W}^{-1}\underset{\sim}{L}\underset{\sim}{x}^{(m+1)} + \underset{\sim}{W}^{-1}\underset{\sim}{b} \tag{24}$$

Varga has shown[38a] that (24) will converge if and only if $\underset{\sim}{A}$ is a positive definite matrix. One way of insuring[38b] that $\underset{\sim}{A}$ is a positive definite matrix is for it to be strictly diagonally dominant, i.e.,

$$|A_{ii}| > \sum_{\substack{j=1 \\ j \neq 1}}^{N} |A_{ij}| , \quad i = 1,2,\ldots,N \tag{25}$$

If the option to try Gauss-Seidel iteration is chosen for a given step, our program checks (25) and uses the regular Numerov instead if it is not satisfied.

Comparing (12) with (21) and (19) with (23) shows that $\underset{\sim}{A} = \underset{\sim}{I} + \underset{\sim}{F}_{n+1}$, $\underset{\sim}{x} = \underset{\sim}{f}_{n+1}$, $\underset{\sim}{b} = \underset{\sim}{Y}_{n+1}$, $\underset{\sim}{W} = \underset{\sim}{I} + \underset{\sim}{d}_{n+1}$, and $(\underset{\sim}{U} + \underset{\sim}{L})\underset{\sim}{x} = \underset{\sim}{\Delta}_{n+1}$. Recalling (16) and putting these assignments into (24) yields the following Gauss-Seidel iteration procedure

$$\left(\underset{\sim}{f}_{n+1}^{(m+1)}\right)_{ij} = [1 + (\underset{\sim}{d}_{n+1})_{ii}]^{-1} [(\underset{\sim}{Y}_{n+1})_{ij} + \left(\underset{\sim}{\Delta}_{n+1}^{(m)}\right)_{ij}] \tag{26}$$

$$\left(\Delta_{n+1}^{(m)} \right)_{ij} = -(h^2/12) \: [\sum_{k<i} (V_{n+1})_{ik} \left(f_{n+1}^{(m+1)} \right)_{kj} + \sum_{k>i} (V_{n+1})_{ik} \left(f_{n+1}^{(m)} \right)_{kj}] \quad (27)$$

Allison doesn't describe what to use for $f_{n+1}^{(0)}$ in (27). We let

$$f_{n+1}^{(0)} = f_n \tag{28}$$

Then using (27) the first row of $\Delta_{n+1}^{(0)}$ can be calculated and put into (26) to calculate the first row of $f_{n+1}^{(1)}$ which together with (28) can be used in (27) to calculate the second row of $\Delta_{n+1}^{(0)}$ and so on, i.e., row k of $\Delta_{n+1}^{(m)}$ only requires $f_{n+1}^{(m)}$ and the first $(k-1)$ rows of $f_{n+1}^{(m+1)}$ and row k of $f_{n+1}^{(m+1)}$ requires only row k of $\Delta_{n+1}^{(m)}$. Provided $(I + F_{n+1})$ is positive definite, equations (28), (26), and (27) are used in the above manner until

$$\frac{1}{N} \left[\sum_{i=1}^{N} \sum_{j=1}^{N} \left| \left(f_{n+1}^{(m+1)} \right)_{ij} - \left(f_{n+1}^{(m)} \right)_{ij} \right|^2 \right]^{\frac{1}{2}} < EPS \tag{29}$$

where EPS is an input variable. Then putting $f_{n+1} = f_{n+1}^{(m+1)}$ into (16), Δ_{n+1} is evaluated and $Y(r)$ is evaluated at the next grid point using (18). We refer to the above procedure as the "iterative" Numerov method.

The program has three options regarding the use of the regular and iterative Numerov methods. In the first mode of operation, only the regular Numerov method will be used. In the second, the regular Numerov method is used until an r value is reached at which the iterative Numerov method becomes faster. The time comparison between the two methods is made after the first integration step and every KCKth step until the iterative Numerov method becomes faster. KCK is an input variable. In the third mode only the iterative Numerov method should be used. If the matrix $(I + F_{n+1})$ does not satisfy (25) for some r, the regular Numerov method is used for that r

independent of the choice of mode. For modes two and three the program
has an input variable NCONV. NCONV is the maximum number of Gauss-Seidel
iterations made at a mesh point. If convergence is not achieved in NCONV
iterations, the regular Numerov method is used to find $f(r)$ at that mesh
point. If convergence is not reached in NCONV iterations but the time
used is less than used by the ı ılar Numerov method the program automatically
increases NCONV. Usually we set EPS = 10^{-4} and NCONV = 200. As stressed
by Allison,[5] the efficiency of the iterative Numerov method is largely
dependent on one's ability to vary the value of the convergence criterion,
consistent with obtaining the desired accuracy. Allison used a convergence
criterion he called ε and found that $\varepsilon = 10^{-2}$ was sufficient for his calcu-
lations. Since we use the reasonably safe value EPS = 10^{-4} for our conver-
gence criterion, we need not check so carefully for convergence with respect
to EPS, but our efficiency is not optimized. To obtain the best possible
timings one should vary EPS to obtain the value just sufficient for the
desired precision.

Next we consider the method used to change the stepsize h. Blatt[4] has
suggested a method for applications to single differential equations for
determining when h can be doubled or should be halved. A generalization
of his error criterion to coupled differential equations has been included
in our program. The criterion we use for the variable-stepsize runs used
to gather timing information for the present report is to require h be small
enough that

$$REPS(r) < 64\delta \tag{30}$$

where δ is an input variable and the estimated relative error per step is

$$\text{REPS}(r) = 7.2 \left(\frac{h^2}{12} \max_i |D_{ii}(r)| \right)^3 \tag{31}$$

In practice we double the stepsize if $\text{REPS}(r_n) < \delta$ and we half it if $\text{REPS}(r_n) > 64\delta$. There is also a provision for restricting the maximum stepsize h_{max} and the stepsize is not doubled if h exceeds $\frac{1}{2} h_{max}$. Although (31) is based on Blatt's error analysis, an essentially equivalent result can be obtained by the following argument. To obtain reasonable accuracy one would expect to require a certain number of steps per deBroglie wavelength, i.e.,

$$h \leq \delta' \min_i \lambda_i \tag{32}$$

where λ_i is the deBroglie wavelength in channel i and δ' is the reciprocal of the number of steps per deBroglie wavelength. Using (3) and a corresponding criterion for locally closed channels we rewrite (32) as

$$h \leq (\text{const.}) \delta' (\max_i |D_{ii}|)^{-\frac{1}{2}} \tag{33}$$

For comparison we use (31) to rewrite (30) as

$$h \leq (\text{const.}') \delta^{\frac{1}{6}} (\max_i |D_{ii}|)^{-\frac{1}{2}} \tag{34}$$

The effect of (33) is the same as the effect of (34), but the constant has a different name.

Because the stepsize is not changed with $\delta < \text{REPS}(r) < 64\delta$, and because the stepsize is only changed by factors of 2, the actual stepsize used at a given r depends on the initial stepsize h_0 as well as on r. The program can also operate in a fixed-stepsize mode, or used fixed stepsizes at small r and variable stepsizes at large r.

Doubling the stepsize is straightforward. It merely requires using the values of $\underline{f}(r)$ and $\underline{D}(r)$ at r_n and r_{n-2} instead of r_n and r_{n-1} and using 2h instead of h in (1𝒸) to calculate $\underline{f}_{n+2}$ instead of $\underline{f}_{n+1}$.

However, when the stepsize should be halved, $\underline{f}(r)$ has not been calculated at a half step back and, until it is, the integration cannot proceed using half the current stepsize h. The method[4] used to calculate $\underline{f}(r)$ at $r' = r_n - \frac{h}{2}$ from h, $\underline{f}_n$, $\underline{D}_n$, $\underline{f}_{n-1}$, and $\underline{D}_{n-1}$ is as follows. Let $h' = \frac{h}{2}$, i.e., half the current stepsize, let $r_{n+1} = r_{n-1} + 2h' = r_{n-1} + h$, i.e., the old r_n becomes the new r_{n+1} and the old $\underline{f}_n$ and $\underline{D}_n$ become the new $\underline{f}_{n+1}$ and $\underline{D}_{n+1}$, respectively, and the new $r_n = r_{n-1} + h' = r'$. Using this new grid, $\underline{f}(r') = \underline{f}_n$ can be found from (10) and $\underline{D}_n = \underline{D}(r')$ as

$$\underline{f}_n = (\underline{I} + \frac{10}{12}h'^2\underline{D}_n)^{-1}\left[(\underline{I} - \frac{h'^2}{12}\underline{D}_{n+1})\underline{f}_{n+1} + (\underline{I} - \frac{h'^2}{12}\underline{D}_{n-1})\underline{f}_{n-1}\right] \tag{35}$$

after which the integration can proceed as before but with h' replacing h in all equations.

Now we consider the procedure for starting the propagation of the solution vectors. If the elements of $\underline{V}(r)$ have no singularities of order two or higher at the origin, then for small r the solutions of (2) that satisfy (4) are given by

$$f_{ij}(r) \underset{r\to 0}{\sim} c_{ij}r^{\ell_i+1} \tag{36}$$

Since, as explained in the previous subsection, we need merely obtain any linearly independent set of N solutions of the coupled equations we may let $c_{ij} = c_{ii}\delta_{ij}$. For electron-molecule and electron-atom scattering the program calculates the initial grid point r_0 by solving

$$r_0^{\ell_{min}+1} = \text{UTEST} \tag{37}$$

where UTEST is a small number, usually 10^{-12} atomic units, which is an input variable of the program, and

$$\ell_{min} = \min_i \{\ell_i\}_{i=1}^N \tag{38}$$

For atom-molecule scattering we set r_0 to a value sufficiently small so that all channels are strongly closed. In either case the results should be invariant to decreasing UTEST or r_0. Since the Numerov method is a 3-point method, it requires $\underset{\sim}{f}(r)$ at r_0 and $r_1 = r_0 + h_0$ to get started. As discussed above, the calculation of a subblock of the reactance matrix requires that the column vectors of the solution matrix be linearly independent. Therefore the program starts the solution with a linearly independent set of column vectors as

$$\underset{\sim}{f}_0 = \underset{\sim}{0} \tag{39}$$

and

$$\underset{\sim}{f}_1 = \underset{\sim}{I} \tag{40}$$

Starting the solution with a set of linearly independent column vectors does not insure that the computed solution vectors will remain linearly independent until (5) is valid. For example, if the local kinetic energy in the ith channel [which is proportional to $-D_{ii}(r)$] is negative, the ith row of $\underset{\sim}{f}(r)$ will grow exponentially as the integration proceeds and the linear independence of the column vectors will be lost. If this problem occurs one must perform stabilizing transformations, i.e., one must periodi-cally replace the columns of $\underset{\sim}{f}(r)$ by linear combinations of the columns

to insure linear independence is not lost. There is of course more than one way to do this. Riley's method,[39] consisting of periodic reorthogonalization of the columns of $\underset{\sim}{f}(r)$, is particularly easy to apply and was chosen for our program. An alternative method has been presented by Gordon.[12] Riley's method consists of defining a transformed solution $\underset{\sim}{f}^t(r)$ in terms of the solution $\underset{\sim}{f}(r)$ as

$$\underset{\sim}{f}^t_k = \underset{\sim}{f}_k \underset{\sim}{f}^{-1}_n \qquad k = n, n-1, n-2, \ldots, 0 \tag{41}$$

with $\underset{\sim}{f}^t(r)$ replacing $\underset{\sim}{f}(r)$ in all equations using $\underset{\sim}{f}(r)$. The transformation is applied only when at least one of the $D_{ii}(r)$ is positive. The number of integration steps taken between successive applications of (41) in such regions is set by an input parameter, NLINDP, of the program. If NLINDP is set too large, the subblock $\underset{\sim}{R}^J$ of the reactance matrix calculated in the asymptotic region is not symmetric. For a typical case, we find it sufficient to reorthogonalize at every 20th step for which the local kinetic energy in any channel is negative. Riley's method of stablization may also be used as part of a procedure for eliminating closed channels at large r so that the number of channels propagated may be reduced to those that are open asymptotically.[14]

As mentioned above, the solution to (2) is matched in the asymptotic region using (5) to a linear combination of matching functions $\underset{\sim}{S}(r)$ and $\underset{\sim}{C}(r)$. The program uses either of two sets of matching functions as described next.

If the elements of $\underset{\sim}{V}(r)$ all go to zero at large r faster than r^{-2}, then beyond some large value r' of r the quantity $\underset{\sim}{\ell}^2 r^{-2}$ will dominate $\underset{\sim}{V}(r)$

to such an extent that $\underline{V}(r)$ can be dropped from (3). This reduces (2) to the N^2 uncoupled second order differential equations

$$\frac{d^2}{dr^2} f_{ij}(r) - \frac{\ell_i(\ell_i+1)}{r^2} f_{ij}(r) + k_i^2 f_{ij}(r) = 0 \tag{42}$$

with $1 \le i \le N$, $1 \le j \le N$, and $r > r'$. The solution to (42) is just a linear combination of regular and irregular Ricatti-Bessel functions $j_{\ell_i}(k_i r)$ and $n_{\ell_i}(k_i r)$, respectively, for $i = 1,2,\ldots,N$, where the Ricatti-Bessel functions are defined in terms of spherical Bessel functions of the first kind and second kind[40] by

$$j_\ell(x) = x j_\ell(x) \tag{43}$$

$$n_\ell(x) = x n_\ell(x) \tag{44}$$

(Reference 40 uses $y_n(z)$ in denoting the spherical Bessel function of the second kind.) For large r[41]

$$k_i r j_{\ell_i}(k_i r) \underset{r \to \infty}{\sim} \sin(k_i r - \tfrac{1}{2}\ell_i \pi) \tag{45}$$

and

$$k_i r n_{\ell_i}(k_i r) \underset{r \to \infty}{\sim} -\cos(k_i r - \tfrac{1}{2}\ell_i \pi) \tag{46}$$

Comparing (6) and (7) with (45) and (46) yields

$$S_{ij}(r) = k_i^{-\frac{1}{2}} \delta_{ij} j_{\ell_i}(k_i r) \tag{47}$$

and

$$C_{ij}(r) = -k_i^{-\frac{1}{2}} \delta_{ij} n_{\ell_i}(k_i r) \tag{48}$$

as one set of matching functions.

However, if some elements of $\underset{\sim}{V}(r)$ go to zero at large r as r^{-2}, r^{-3}, or r^{-4} (as is the case for electron-molecule scattering with realistic effective potentials), then the r' for which (42) is valid is generally very large and the use of (47) and (48) requires numerical integration of (2) over a very large region which is expensive. Therefore, Burke and Schey[42] and Burke, McVicar, and Smith[43] (BMS) have derived asymptotic solutions to (2) computed from the long-range part of the potential and a computer program for using the BMS solutions as matching functions has been described in detail by Norcross.[44] For open channels, the matching functions are written in terms of asymptotic series as

$$S_{ij}(r) = k_i^{-\frac{1}{2}}[\alpha_{ij}(r) \sin \phi_i(r) + \beta_{ij}(r) \cos \phi_i(r)] \tag{49}$$

and

$$C_{ij}(r) = k_i^{-\frac{1}{2}}[\alpha_{ij}(r) \cos \phi_i(r) - \beta_{ij}(r) \sin \phi_i(r)] \tag{50}$$

where

$$\phi_i(r) = k_i r - \frac{1}{2} \ell_i \pi \tag{51}$$

and

$$\gamma_{ij}(r) = \sum_{p=0}^{P_i} \gamma_{ij}^p r^{-p} \tag{52}$$

for $\gamma = \alpha$ or β where P_i is determined as described below and

$$\alpha_{ij}^0 = \delta_{ij} \qquad \text{and} \quad \beta_{ij}^0 = 0 \tag{53}$$

The remaining coefficients γ_{ij} are determined by the following procedure. The function $\underset{\sim}{V}(r)$ in (2) and (3) is replaced by its long range form

$$V_{ij}(r) \underset{r \to \infty}{\sim} \sum_{m=1}^{M} c_{ij}^{(m)} r^{-m-1} \tag{54}$$

and substituting the $\underline{S}(r)$ of (49) or the $\underline{C}(r)$ of (50) into (2) for $\underline{f}(r)$ yields recursion relations among the coefficients of like powers of r^{-1}. The recursion relations have two different forms, one for when $k_i = k_j$ and the other for when $k_i \neq k_j$. The latter form depends on $(k_i^2 - k_j^2)^{-1}$ and becomes numerically unstable as k_i^2 approaches k_j^2 so a test is made using the input parameter n_1 to decide which form to use. Only if $|k_i^2 - k_j^2| > n_1 k_i^2$ is the latter form used. As compared to Norcross' version[44] we made one additional change in the recursion relation for the case of degenerate energies, i.e., if $|k_i^2 - k_j^2| < n_1 k_i^2$ we approximated both k_i and k_j by their arithmetic mean rather than replacing them both by k_i as Norcross did. From these relations the remaining values of γ_{ij}^p can be calculated. It is well known that the best approximation which can be obtained from an asymptotic series is obtained by summing up to the smallest term and retaining half that term. In this spirit we replace (52) by

$$\gamma_{ij}(r) = \sum_{p=0}^{P_i} \gamma_{ij}^p r^{-p} - \tfrac{1}{2}\gamma_{ij}^{P_i} r^{-P_i} \tag{55}$$

(This is another change from the procedure used by Norcross.) The value of P_i is determined for each channel $i = 1,2,\ldots,N$ using the input parameter τ as follows. For a given p and i, the maximum absolute value of $\alpha_{ij}^p r^{-p}$ or $\beta_{ij}^p r^{-p}$ is set equal to τ and solved for r. The solution is called $r(i,p)$. This is done for each $p \geq 1$ in increasing order of p until $p = p'$ where p' is the smallest p for which $r(i,p+1) > r(i,p)$ or until p is the maximum value P_{max} allowed by the program. If the former, then P_i is set equal to p' and $r(i,P_i)$ is the minimum value of r for which (52) may be used with the smallest term being less than or equal to τ for the ith channel.

(It is interesting to note that Norcross[44] always uses $P_i \geq 4$ in his program.)

If $r(i,p) < r(i,p-1)$ for all $p \leq P_{max}$ then P_i is set equal to P_{max} and $r(i,P_i)$ is set to the value of r for which the maximum absolute value of $\alpha_{ij}^{P_{max}} r^{-P_{max}}$ or $\beta_{ij}^{P_{max}} r^{-P_{max}}$ is equal to 10^{-14}.

An estimate of where (5) with (49) and (50) becomes valid is then made as the maximum of the $r(1,P_i)$ values for the various channels i and the value of r for which (54) is valid.

The above method can also be generalized to include closed channels.[44]

The choice between matching to Ricatti–Bessel functions or BMS functions is controlled by an input variable MMAX. Assuming one has reached an r for which the asymptotic form (5) is valid, the program matches $\underline{f}(r)$ to either matching functions (47) and (48) or (49) and (50) at two grid points r_n and r_{n+1}. After rearranging terms, this yields

$$(\underline{C}_{n+1} - \underline{S}_{n+1} \underline{S}_n^{-1} \underline{C}_n)\underline{Q} = \underline{f}_{n+1} - \underline{S}_{n+1} \underline{S}_n^{-1} \underline{f}_n \tag{56}$$

which is solved for the approximate $\underline{Q}$, called $\underline{Q}_n$, and two approximate $\underline{P}$ matrices, called $\underline{P}_n$ and $\underline{P}_{n+1}$, are then found using

$$\underline{P}_k = \underline{S}_k^{-1}(\underline{f}_k - \underline{C}_k \underline{Q}_n) \qquad k = n, n+1 \tag{57}$$

If

$$norm(\underline{P}_{n+1} - \underline{P}_n) < STEST \tag{58}$$

where STEST is an input variable and the norm of a matrix is defined as

$$norm \ \underline{A} = \frac{1}{N} \left(\sum_{i=1}^{N} \sum_{j=1}^{N} |A_{ij}|^2 \right)^{\frac{1}{2}} \tag{59}$$

then $\underset{\sim}{R}$ is calculated from $\underset{\sim n}{P}$ and $\underset{\sim n}{Q}$ as

$$\underset{\sim n \sim}{P^T} R = \underset{\sim n}{Q^T} \tag{60}$$

In deriving (60) we use the fact[45] that $\underset{\sim}{R} = \underset{\sim}{R^T}$.

  C. R matrix propagation method. The version of the R matrix pro-
pagation method we use is our own adaptation[25,26,28,29] to inelastic scat-
tering problems of the method Light and Walker[24] developed for reactive
scattering problems. The R matrix propagation method has also been adapted
to inelastic scattering problems by Stechel et al.,[27] but their procedures
differ from ours in several respects.

  In the R matrix propagation method the range of the translation co-
ordinate r is subdivided into many sectors. In sector (i) the total wave
function $\Psi(\vec{z}, r)$ is expanded in a "primitive" basis of N orthonormal functions
$X_n(\vec{z})$, here assumed to be the same in every sector, and a set of N close
coupling equations in the primitive representation is derived for each of
the 2N linearly independent translational wave functions $\underline{f}(r)$. We use
the convention that $\underline{A}$ denotes a column vector with component A_i and $\underline{\underline{A}}$ denotes
a matrix each of whose columns is an $\underline{A}$. These equations have the form (2).
An "adiabatic" basis for sector (i) is found by diagonalizing $\underline{\underline{D}}_C^{(i)}$, the
interaction matrix at the center of sector (i). The wave function is ex-
panded in P of the adiabatic-basis functions $Z_n^{(i)}(\vec{z})$, and a set of P close
coupling equations in the adiabatic representation is obtained for each of
the 2P linearly independent translational wave functions $\underline{g}^{(i)}(r)$. The basis
functions $X_n(\vec{z})$ and $Z_n^{(i)}(\vec{z})$ are related by a transformation matrix $\underline{\underline{T}}^{(i)}$.
Similarly, $\underline{\underline{T}}^{(i)}$ is used to relate the translational wave functions $f_n(r)$
and $g_n^{(i)}(r)$ to each other. The adiabatic-representation translational

wave functions and their derivatives are propagated through sector (i) by a propagation matrix $\underset{\sim}{P}^{(i)}(E)$. Then the requirement that $\Psi(\vec{z},r)$ and $\Psi'(\vec{z},r)$, where an apostrophe denotes a derivative, be continuous at the boundary between sectors (i) and (i+1) is used to obtain sector matching conditions. To express the continuity between the wave function in the adiabatic representation in sector (i) and the wave function in the adiabatic representation in sector (i+1), the transformation matrix $\underset{\sim}{T}^{(i)}$ is used to transform from the adiabatic representation to the primitive representation in sector (i), and the matrix $\underset{\sim}{T}^{(i+1)^T}$ is used to transform from the primitive representation in sector (i+1) to the adiabatic representation in sector (i+1). The combined effects of these two steps is expressed in terms of a transformation matrix $\underset{\sim}{T}(i,i+1)$ which relates the adiabatic representation in sector (i) to that in sector (i+1). The propagator $\underset{\sim}{P}^{(i+1)}$ is then used to propagate the adiabatic-representation translational wave functions $\underset{\sim}{g}^{(i+1)}(r)$ and their derivatives through sector (i+1), $\underset{\sim}{T}(i+1,i+2)$ is used to transform to the adiabatic representation in sector (i+2), and so forth. In this way the translational wave functions and their derivatives could be propagated from the strong-interaction region through each sector and across sector boundaries. Rather than propagate the wave function and its derivative though, we propagate the global R matrix $\underset{\sim}{R}^{(i)}$, which relates the matrices $\underset{\sim}{g}^{(i)}[r_L^{(i)}]$ and $\underset{\sim}{g}^{(i)}[r_R^{(i)}]$ of 2P linearly independent wave functions at the left side of the first sector and at the right side of sector (i) to their derivatives at these locations. In each sector this global R matrix $\underset{\sim}{R}^{(i)}$ is computed from $\underset{\sim}{R}^{(i-1)}$ and the sector R matrix $\underset{\sim}{r}^{(i)}$, which relates the adiabatic-basis translational wave functions at the right sides of sectors (i-1) and (i) to their derivatives at these locations. In turn the sector R matrix is obtained from the

transformation matrix $\underset{\sim}{T}(i-1,i)$ and the propagator $\underset{\sim}{P}^{(i)}$ In this way we pro-
pagate the global R matrix from sector to sector until we obtain the global R
matrix in the last sector. This relates the matrix of linearly independent
physical wave functions in the strong-interaction or small-r region and
in the large-r asymptotic region to the matrix of their derivatives. Small-
r and large-r boundary conditions on the wave functions and their derivatives
are then imposed in such a way that the reactance matrix $\underset{\sim}{R}$ can be obtained
in terms of known quantities.

The sector-by-sector propagation is essentially the same as in reference
24. The 2P × 2P sector propagator $\underset{\sim}{P}^{(i)}$ is defined by

$$\underset{\sim}{G}_L^{(i)} = \underset{\sim}{P}^{(i)} \underset{\sim}{G}_R^{(i)} \tag{61}$$

where each column of the 2P × 2P matrix $\underset{\sim}{G}^{(i)}(r)$ is defined by

$$\underline{G}^{(i)}(r) = \left[\begin{array}{c} \underline{g}^{(i)}(r) \\[2mm] \underline{g}'^{(i)}(r) \end{array} \right] \tag{62}$$

where $\underline{g}^{(i)}(r)$ is one of the 2P linearly independent P-component solution
vectors. Here and in the following equations subscripts R, L, and C de-
note quantities evaluated at the right-hand and left-hand sector boundaries
and the center of the sector, respectively, e.g., $g_L^{(i)} = g^{(i)}[r_L^{(i)}]$. The
matrix $\underset{\sim}{P}^{(i)}$ is partitioned into four P × P submatrices

$$\underset{\sim}{P}^{(i)} = \left[\begin{array}{cc} \underset{\sim}{P}_1^{(i)} & \underset{\sim}{P}_2^{(i)} \\[2mm] \underset{\sim}{P}_3^{(i)} & \underset{\sim}{P}_4^{(i)} \end{array} \right] \tag{63}$$

$\underline{P}^{(i)}$ is computed by diagonalizing $\underline{D}_C^{(i)}$ and assuming that the eigenvalues of $\underline{D}$ are independent of r within a sector.[24] This assumption would be true if the interaction potential were constant throughout the sector. The sector R matrix $\underline{r}^{(i)}$ is the matrix which relates the wave functions $g_R^{(i-1)}$ and $g_R^{(i)}$ evaluated at the right-hand side of sectors (i-1) and (i) to their derivatives. It is defined by

$$\underline{r}^{(i)} = \begin{pmatrix} \underline{r}_1^{(i)} & \underline{r}_2^{(i)} \\ \underline{r}_3^{(i)} & \underline{r}_4^{(i)} \end{pmatrix} \tag{64}$$

where

$$\begin{bmatrix} g_R^{(i-1)} \\ g_R^{(i)} \end{bmatrix} = \begin{bmatrix} \underline{r}_1^{(i)} & \underline{r}_2^{(i)} \\ \underline{r}_3^{(i)} & \underline{r}_4^{(i)} \end{bmatrix} \begin{bmatrix} g'^{(i-1)}_R \\ -g'^{(i)}_R \end{bmatrix} \tag{65}$$

The equations for the sector R matrices are

$$\underline{r}_1^{(i)} = \underline{T}(i-1,i)\, \underline{P}_1^{(i)}\, \underline{P}_3^{(i)^{-1}}\, \underline{T}(i-1,i)^{-1} \tag{66}$$

$$\underline{r}_2^{(i)} = \underline{T}(i-1,i)\, \underline{P}_3^{(i)^{-1}} \tag{67}$$

$$\underline{r}_3^{(i)} = \underline{P}_3^{(i)^{-1}}\, \underline{T}(i-1,i)^{-1} \tag{68}$$

$$\underline{r}_4^{(i)} = \underline{P}_3^{(i)^{-1}}\, \underline{P}_4^{(i)} \tag{69}$$

The global R matrix $\underline{R}^{(i)}$ spanning the configuration space from the first sector to sector (i) is defined by

$$\underline{R}^{(i)} = \begin{pmatrix} \underline{R}_1^{(i)} & \underline{R}_2^{(i)} \\ \underline{R}_3^{(i)} & \underline{R}_4^{(i)} \end{pmatrix} \tag{70}$$

$$\begin{bmatrix} g_L^{(1)} \\ g_R^{(1)} \end{bmatrix} = \begin{bmatrix} R_1^{(1)} & R_2^{(1)} \\ R_3^{(1)} & R_4^{(1)} \end{bmatrix} \begin{bmatrix} g_L'^{(1)} \\ -g_R'^{(1)} \end{bmatrix} \tag{71}$$

$R^{(1)}$ is propagated from the first sector to the asymptotic region in which the scattering matrix elements and other physically interesting quantites are calculated.

A real symmetric global R matrix $R^{(1)}$ insures a symmetric reactance matrix and hence a unitary scattering matrix. If $R^{(i-1)}$ and $r^{(1)}$ are symmetric matrices, then $R^{(1)}$ is symmetric. But

$$R_1^{(1)} = P_1^{(1)} P_3^{(1)-1} \tag{72}$$

$$R_2^{(1)} = R_3^{(1)} = P_3^{(1)^{-1}} \tag{73}$$

$$R_4^{(1)} = P_3^{(1)^{-1}} P_4^{(1)} \tag{74}$$

so that $R^{(1)}$ is symmetric. The sector R matrix $r^{(1)}$ is symmetric if the transformation matrix $T(i-1,i)$ is orthogonal. The transformation matrix is orthogonal if and only if the number P of propagated adiabatic channels equals the number N of primitive basis functions. Thus the scattering matrix is automatically unitary if and only if $P = N$.[29] It has been suggested[24,27] that in order to insure a symmetric $R^{(1)}$ the matrix $T(i-1,i)^{-1}$ be replaced by $T(i-1,i)^T$. This also may afford computational advantages. When $P = N$, this is not an approximation. However, for contracted basis sets, i.e., when $P < N$, $T(i-1,i)$ is not orthogonal. For contracted bases replacing $T(i-1,i)^{-1}$ by $T(i-1,i)^T$ in each sector changes the results unless of course the calculations are converged with respect to P and N. For

unconverged calculations we generally obtained more accurate answers by using the inverse and symmetrizing the reactance matrix than by using the transpose. For this reason we have used the inverse for our final production runs; however, the choice in general is still somewhat ambiguous. Both choices should converge to the correct limit as P and N are increased. Since, when $P \neq N$, using the inverse does not automatically produce a symmetric reactance matrix and hence a unitary scattering matrix, we symmetrized our reactance matrix by taking an arithmetic mean with its transpose.

We obtain the reactance matrix $\tilde{R}$ from the global R matrix equation by imposing asymptotic scattering boundary conditions on the adiabatic wave functions in the following way. The P × 2P solution matrix $g^{(c)}(r)$ in the adiabatic representation in sector (c) consists of 2P linearly independent vectors of order P. Because the solutions are linearly independent, the P × P matrix $\underset{\sim}{\phi}^{(c)}(r)$, each column of which is a scattering vector, i.e., one of P linearly independent linear combinations of the former set of vectors which satisfies correct small-r and large-r boundary conditions, satisfies the global R matrix equations for sector (c). Using (71), one can show that

$$
\begin{bmatrix} \phi_L^{(1)} \\ \phi_R^{(c)} \end{bmatrix} = \begin{bmatrix} R_1^{(c)} & R_2^{(c)} \\ R_3^{(c)} & R_4^{(c)} \end{bmatrix} \begin{bmatrix} \phi_L'^{(1)} \\ -\phi_R'^{(c)} \end{bmatrix}
\tag{75}
$$

where $\phi_{mn}^{(i)}(r)$ is the m-th component of the n-th scattering vector and the P × P matrices $\underset{\sim}{\phi}_L^{(1)}$ and $\underset{\sim}{\phi}_R^{(c)}$ consist of the P linearly independent scattering vectors evaluated at the left side of the first sector and the right side of sector (c), respectively. $R_1^{(c)}$, $R_2^{(c)}$, $R_3^{(c)}$, and $R_4^{(c)}$ are the P × P submatrices of the 2P × 2P global R matrix $\underset{\sim}{R}^{(c)}$ which spans

the configuration space from the left side of the first sector to the
right side of sector (c). To obtain R we substitute scattering boundary
conditions for $\underset{\sim}{\phi}_L^{(1)}$ and $\underset{\sim}{\phi}_R^{(c)}$ and their derivatives into (75).

To use (75) to extract the reactance matrix we first require expres-
sions for $\underset{\sim}{\phi}_L^{(1)}$ and $\underset{\sim}{\phi}_L'^{(1)}$. Since only P of the 2P linearly independent
solutions $\underset{\sim}{g}^{(i)}(r)$ satisfy physical boundary conditions in the strong
interaction region, we include in our analysis of (75) only the P
functions which provide physical solutions in the small-r region.

For channels which are closed at $r = r_L^{(1)}$ we use the following expo-
nential functions for the wave functions

$$\left(\underset{\sim}{\phi}_L^{(1)}\right)_{mn} = d_{mn} \exp\left[\left|\kappa_m\left(r_L^{(1)}\right)\right| r_L^{(1)}\right] \qquad \text{m closed} \qquad (76)$$

$$\left(\underset{\sim}{\phi}_L'^{(1)}\right)_{mn} = d_{mn} \left|\kappa_m\left(r_L^{(1)}\right)\right| \exp\left[\left|\kappa_m\left(r_L^{(1)}\right)\right| r_L^{(1)}\right] \qquad (77)$$

where $\kappa_m\left(r_L^{(1)}\right)$ is the local wave number in the m-th channel calculated at
$r_L^{(1)}$ and the coefficients d_{mn} are unknowns. All channels with nonzero
orbital angular momentum are closed at the origin. For channels which are
open in the small-r region the boundary conditions on the wave functions
can be written

$$\left(\underset{\sim}{\phi}_L^{(1)}\right)_{mn} = e_{mn} \sin\left[\left|\kappa_m\left(r_L^{(1)}\right)\right| r_L^{(1)}\right] \qquad \text{m open} \qquad (78)$$

$$\left(\underset{\sim}{\phi}_L'^{(1)}\right)_{mn} = e_{mn} \left|\kappa_m\left(r_L^{(1)}\right)\right| \cos\left[\left|\kappa_m\left(r_L^{(1)}\right)\right| r_L^{(1)}\right] \qquad (79)$$

where the coefficients e_{mn} are not known. To facilitate the calculations
we cast (76)-(77) and (78)-(79) in the same form, so that for both open
and closed channels the boundary conditions we use at $r = r_L^{(1)}$ are

$$\Phi_L^{(1)} = \mathcal{C} \tag{80}$$

$$\phi_L^{'(1)} = \chi^{(1)} \mathcal{C} \tag{81}$$

where

$$x_{mn}^{(1)}(E) = \delta_{mn} \begin{cases} |\kappa_m\left(r_L^{(1)}\right)| \cos\left[|\kappa_m\left(r_L^{(1)}\right)|r_L^{(1)}\right] \bigg/ \sin\left[|\kappa_m\left(r_L^{(1)}\right)|r_L^{(1)}\right] & \text{m open} \\[4mm] |\kappa_m\left(r_L^{(1)}\right)| & \text{m closed} \end{cases} \tag{82}$$

and the coefficients C_{mn} are unknowns.

At large r the close coupling equation becomes

$$\left\{\frac{d^2}{dr^2} - [\lambda_{mn}^{(c)}(r)]^2\right\} g_m^{(c)}(r) = 0 \tag{83}$$

where $[\lambda^{(c)}(r)]^2$ is a diagonal matrix, the elements of which at the center $r = r_C^{(c)}$ of sector (c) are the eigenvalues of the interaction matrix $D_C^{(c)}$.

For the discussion of the large-r boundary conditions we find it useful to introduce the following notation. We define the symptotic interaction matrix D^a by

$$\lim_{r\to\infty} D(r) = D^a \tag{84}$$

We denote the diagonal matrix of the eigenvalues of D^a by $[\lambda^a]^2$, where the square of the diagonal matrix of the asymptotic channel wave numbers is given by

$$[k]^2 = -[\lambda^a]^2 \tag{85}$$

Once the real potential has vanished and the r dependence of $D(r)$ is dominated

by the centrifugal potential, the eigenvalues of the interaction matrix in sector (c) are given by

$$[\lambda_{mm}^{(c)}(r)]^2 = \lambda_{mm}^{a\,2} + \frac{\ell_m(\ell_m+1)\hbar^2}{2\mu r^2} \tag{86}$$

In the large-r region P^0 of the P channels are open, and the channels are ordered in such a way that channel m corresponds to an open channel for $1 \leq m \leq P^0$ and to a closed channel for $P^0 < m \leq P$. In the large-r asymptotic region the boundary conditions for closed-channel wave functions are expressed as linear combinations of exponentially increasing and decreasing functions

$$\left[\phi_{\sim R}^{(c)}\right]_{mn} = \delta_{mn} \, b \, \exp\left[|k_m|r_R^{(c)}\right] + a_{mn}^{(c)} \exp\left[-|k_m|r_R^{(c)}\right] \quad P^0 < m \leq P \tag{87}$$

$$\left[\phi_{\sim R}^{\prime\,(c)}\right]_{mn} = \delta_{mn} \, b \, |k_m| \exp\left[|k_m|r_R^{(c)}\right] - a_{mn}^{(c)} \, |k_m| \exp\left[-|k_m|r_R^{(c)}\right] \tag{88}$$

where b is an unknown coefficient of the exponentially increasing component of the $P - P^0$ closed-channel wave functions.

For the large-r boundary conditions on the open-channel wave functions in sector (c) we use

$$\left[\phi_{\sim R}^{(c)}\right]_{mn} = \delta_{mn} \, j_{\ell_m}\left(k_m r_R^{(c)}\right) - a_{mn}^{(c)} \, n_{\ell_m}\left(k_m r_R^{(c)}\right) \quad 1 \leq m \leq P^0 \tag{89}$$

$$\left[\phi_{\sim R}^{\prime\,(c)}\right]_{mn} = k_m\left[\delta_{mn} \, j'_{\ell_m}\left(k_m r_R^{(c)}\right) - a_{mn}^{(c)} \, n'_{\ell_m}\left(k_m r_R^{(c)}\right)\right] \quad 1 \leq m \leq P^0 \tag{90}$$

where the Ricatti-Bessel functions are defined in terms of spherical Bessel functions by (43) and (44). The elements of the reactance matrix are related to the open-channel amplitudes $a_{mn}^{(c)}$ by

$$R_{mn} = |k_m|^{\frac{1}{2}} a_{mn}^{(c)} |k_n|^{-\frac{1}{2}} \qquad\qquad \begin{array}{l} 1 \leq m \leq P^0 \\ 1 \leq n \leq P^0 \end{array} \tag{91}$$

The large-r boundary conditions on the adiabatic wave functions can be written concisely in matrix notation as

$$\underset{\sim R}{\Phi}^{(c)} = \underset{\sim}{B}^{(c)} \underset{\sim}{\Delta} + \underset{\sim}{F}^{(c)} \underset{\sim}{a}^{(c)} \tag{92}$$

$$\underset{\sim R}{\Phi}^{'(c)} = \underset{\sim}{G}^{(c)} \underset{\sim}{\Delta} - \underset{\sim}{H}^{(c)} \underset{\sim}{a}^{(c)} \tag{93}$$

where

$$B_{mn}^{(c)} = \delta_{mn} \begin{cases} j_{\ell_m}\left(k_m r_R^{(c)}\right) & 1 \leq m \leq P^0 \\[2mm] \exp\left(|k_m| r_R^{(c)}\right) & P^0 < m \leq P \end{cases} \tag{94}$$

$$F_{mn}^{(c)} = \delta_{mn} \begin{cases} -n_{\ell_m}\left(k_m r_R^{(c)}\right) & 1 \leq m \leq P^0 \\[2mm] \exp\left(-|k_m| r_R^{(c)}\right) & P^0 < m \leq P \end{cases} \tag{95}$$

$$G_{mn}^{(c)} = \delta_{mn} \begin{cases} k_m j'_{\ell_m}\left(k_m r_R^{(c)}\right) & 1 \leq m \leq P^0 \\[2mm] |k_m| \exp\left(|k_m| r_R^{(c)}\right) & P^0 < m \leq P \end{cases} \tag{96}$$

$$H_{mn}^{(c)} = \delta_{mn} \begin{cases} k_m n'_{\ell_m}\left(k_m r_R^{(c)}\right) & 1 \leq m \leq P^0 \\[2mm] |k_m| \exp\left(-|k_m| r_R^{(c)}\right) & P^0 < m \leq P \end{cases} \tag{97}$$

$$\Delta_{mn} = \delta_{mn} \begin{cases} 1 & 1 \leq m \leq P^0 \\[2mm] b & P^0 < m \leq P \end{cases} \tag{98}$$

and

$$\underset{\sim}{R} = [\underset{\sim}{k}^{00}]^{\frac{1}{2}} \underset{\sim}{a}^{(c)00}(E) [\underset{\sim}{k}^{00}]^{-\frac{1}{2}} \tag{99}$$

where the superscript 00 is used to denote the $P^0 \times P^0$ submatrix which links open channels to open channels.

Substituting (80), (81), (92), and (93) into (75) yields

$$
\left[\begin{array}{c} \underset{\sim}{C} \\[2ex] \underline{B}^{(c)} \underset{\sim}{\Delta} + \underline{F}^{(c)} \underline{a}^{(c)} \end{array} \right] = \left[\begin{array}{cc} \underline{R}_1^{(c)} & \underline{R}_2^{(c)} \\[2ex] \underline{R}_3^{(c)} & \underline{R}_4^{(c)} \end{array} \right] \left[\begin{array}{c} \underset{\sim}{x}^{(1)} \underset{\sim}{C} \\[2ex] -\underline{G}^{(c)} \underset{\sim}{\Delta} + \underline{H}^{(c)} \underline{a}^{(c)} \end{array} \right] \tag{100}
$$

Solving for the $P \times P$ matrix $\underline{a}^{(c)}$ gives

$$
\underline{a}^{(c)} = [-\underline{F}^{(c)} + \underline{M}^{(c)} \underline{H}^{(c)}]^{-1} \; [\underline{B}^{(c)} + \underline{M}^{(c)} \underline{G}^{(c)}] \underline{\Delta} \tag{101}
$$

where

$$
\underline{M}^{(c)} = \underline{R}_4^{(c)} + \underline{R}_3^{(c)} \underline{x}^{(1)} \; [\underline{1} - \underline{R}_1^{(c)} \underline{x}^{(1)}]^{-1} \underline{R}_2^{(c)} \tag{102}
$$

Therefore, the matrix $\underline{a}^{(c)}$ from which we calculate $\underline{R}$ is independent of $\underline{C}$, the matrix of coefficients giving the appropriate linear combinations of functions at $r = r_L^{(1)}$. Further $\underline{a}^{(c)00}$, the only portion of $\underline{a}^{(c)}$ which we used to obtain $\underline{R}$, is independent of the last $P - P^0$ columns of $\underline{\Delta}$ and consequently is independent of b, the coefficient of the exponentially increasing component of the closed-channel wave functions in sector (c). In the computer code b is set equal to zero and $B_{mn}^{(c)}$ and $G_{mn}^{(c)}$ are set equal to zero for $P^0 < m \leq P$.

When all channels are closed in the small-r region one finds[29]

$$
\underline{R}_2^{(1)} \overset{\sim}{=} \underline{0} \tag{103}
$$

and

$$
\underline{R}_3^{(1)} \overset{\sim}{=} \underline{0} \tag{104}
$$

if sector (1) is located deep enough into the classically forbidden region. In such a case $R_2^{(i)}$ and $R_3^{(i)}$ remain small in all subsequent sectors and in particular

$$R_2^{(c)} \cong 0 \tag{105}$$

and

$$R_3^{(c)} \cong 0 \tag{106}$$

In such a case (102) becomes

$$M^{(c)} = R_4^{(c)} \tag{107}$$

Consequently the determination of the reactance matrix from (99) and (101) becomes independent of $R_1^{(c)}$, $R_2^{(c)}$, and $R_3^{(c)}$. Furthermore the propagation equation[24] for $R_4^{(i)}$ is independent of $R_1^{(i-1)}$, $R_2^{(i-1)}$, and $R_3^{(i-1)}$. Thus when all channels are closed at small r we propagate only $R_4^{(i)}$ to save computer time.

Although the asymptotic analysis just presented allows for the inclusion of channels which are closed for large r, the presence of strongly closed channels in the asymptotic analysis sometimes gives rise to numerical difficulties. To eliminate these, the program has two options which can be used to simplify the asymptotic analysis. One option is used to eliminate closed channels from the propagation at large r. We have shown that if all elements of the last row and column, corresponding to the most strongly closed channel of $r_2^{(i)}$ are small, the last channel is uncoupled from the remaining channels and may be dropped from propagation without degrading the accuracy of the results.[29] We have implemented this option in the following way. In the large-r region, if the number P of channels propagated

in sector (i) is greater than the number of open channels, the program
checks whether

$$\left|\left(\underset{\sim}{r}_2^{(i)}\right)_{iP}\right| \leq \text{EPSRED} \qquad i = 1,2,\ldots,P \tag{108}$$

and

$$\left|\left(\underset{\sim}{r}_2^{(i)}\right)_{Pi}\right| \leq \text{EPSRED} \qquad i = 1,2,\ldots,P \tag{109}$$

In these equations, channel P is the most strongly closed propagated channel,
and EPSRED is an input variable. If this criterion is satisfied, channel
P is dropped from propagation and only the remaining $P-1$ channels are
propagated in sector (i+1). As we have implemented this option, at most
one channel is dropped from propagation in any sector. We have also im-
plemented a second method to avoid numerical difficulties associated with
including in the asymptotic analysis channels which are strongly closed
for large r. In this method we eliminate those closed channels from the
asymptotic analysis even if they have been included in the propagation.
The option as we have coded it in our programs is appropriate when the
asymptotic analysis is based on only the $\underset{\sim}{R}_4$ part of the global R matrix.
The procedure we use for deciding whether a closed channel can be eliminated
from the asymptotic analysis is the following. In the large-r region,
if the number of channels propagated is greater than the number P^0 of
channels open asymptotically, the program compares the off-diagonal elements
of $\underset{\sim}{R}_4^{(i)}$ for channel P, the most strongly closed propagated channel, to
EPSDR, an input variable. If

$$\left|\left(\underset{\sim}{R}_4^{(i)}\right)_{iP}\right| \leq \text{EPSDR} \qquad i = 1,2,\ldots,(P-1) \tag{110}$$

- 251 -

and

$$\left|\left(R_4^{(i)}\right)_{Pi}\right| \leq \text{EPSDR} \qquad\qquad i = 1, 2, \ldots, (P-1) \qquad (111)$$

then the number of channels P' to be included in the asymptotic analysis is set to $P' = P - 1$. This procedure is repeated for subsequent closed channels until an element for $R_4^{(i)}$ for some closed channel fails the test of (110) and (111) or until all closed channels have been eliminated and $P' = P^0$.

The stepsize $h^{(i)}$ for sector (i) is defined by

$$h^{(i)} = r_R^{(i)} - r_L^{(i)} \qquad\qquad (112)$$

We set $r_L^{(1)}$, $h^{(1)}$, and $h^{(2)}$ by input variables, and we check that they are sufficiently small that the calculations are converged with respect to them. If one or more channels is open at the origin, then $r_L^{(1)}$ should be close to zero. The determination of the stepsizes for subsequent sectors is crucial to the efficiency of the method. A reasonable stepsize criterion can be obtained for $i > 2$ by requiring that the effect of the lowest-order neglected term in the propagator[46] be small. We simplified this argument to make it computationally more convenient and arrived at the following algorithm which is used by the code for $i \geq 2$:

$$h^{(i+1)} = \min \left\{ \begin{array}{l} \varepsilon^{(1)}\left[\dfrac{1}{m} \displaystyle\sum_{j=1}^{m}\left(\dfrac{dD_{jj}^{(i)}}{dr}\right)^2\right]^{-1/6} \\[20pt] h_{\max} \end{array} \right. \qquad (113)$$

where h_{max} is an input variable, m is the largest value of P to be used in a given run, and the derivative is estimated by a backward difference. The error-control parameter is determined as follows. The range of r is divided into three subranges in each of which $\varepsilon^{(i)}$ is constant, i.e., independent of (i). The values of $\varepsilon^{(i)}$ for a given run are set by input variables, EPSA, EPSB, and EPSC. The calculations must be tested for convergence with respect to decreasing all three values. For most applications it has been sufficient to set all values of $\varepsilon^{(i)}$ equal to each other.

One of the advantages of the piecewise analytic method and the R matrix propagation method is that if calculations are required at several energies with the same potential, calculations at the second and subsequent energy can be performed more rapidly by saving certain information generated in the first calculation. We have not made much use of this feature for electron-molecule scattering because we use energy-dependent potentials to include exchange effects (see, e.g., references 10, 11, and 25 and references therein). However, even for this type of problem, this feature might be useful at higher energies or for large orbital angular momenta where exchange effects can sometimes be neglected. For atom-molecule scattering this feature is very useful, and the dramatic reduction in execution time of R matrix propagation calculations for subsequent energies as compared with the calculation at the first energy is discussed in section V.

Because we are interested in making close coupling calculations more efficient, we also want to demonstrate the success we have had in reducing the size of the close coupling calculations by using adiabatic basis functions in the context of R matrix propagation calculations. We have found for

electron-molecule and atom-molecule scattering calculations that we could generally obtain similar or better accuracy with an adiabatic basis of the same or even significantly smaller propagation dimension P than with a conventional $N = P$ basis. As discussed above we construct a P-function adiabatic basis for sector (i) by diagonalizing the $N \times N$ interaction matrix at the center of sector (i) and taking its P lowest-energy eigenfunctions.

D. Piecewise analytic method. The piecewise analytic method of Gordon is described elsewhere.[12,13] We have used two different programs for calculations with this method. One was obtained originally from Quantum Chemistry Program Exchange[35] (QCPE) and was modified in various ways by two of us (M.A.B. and D.G.T.). The second program was written by Wagner.[14]

E. Integral equations algorithm. For our calculations using the integral equations algorithm we used the computer code of Morrison, Lane, and Collins, which is described in detail by them.[20-22] The integral equations method was first presented by Sams and Kouri.[17] The applications presented here have used a trapezoidal rule quadrature scheme. An important feature of the integral-equations formalism for electron scattering problems is the recent development of an efficient means of including non-local Hartree-Fock exchange operators without increasing the size of the solution matrix over the local-potential case.[47] This feature will not be explored here since the examples presented for electron-molecule scattering involve local, energy-dependent potentials. Another important feature is the truncation procedure by which the number of coupled channels is decreased at a "truncation radius"; this can yield substantial savings of computer time.[21,22,48]

F. Timing comparisons. It is very difficult to make precise timing comparisons for several reasons. The most obvious reason is that one method may be programmed more efficiently than another, or a given compiler may produce more efficient code for one method than another. Another difficulty is that one seldom completely optimizes all the numerical parameters for a given application. It is usually more efficient for production runs to set some or all numerical parameters at safe values which produce more accurate results than are really required for parts of the calculation or even for the final cross sections. A related problem is the efficiency of utility codes, e.g., our R matrix propagation code uses the EISPACK sub-program RSP for matrix diagonalization and University of Minnesota codes for solution of sets of linear equations. Another problem, less significant than those mentioned above, is that the computer time even for an identical run may vary 10% or more depending on the time of day and overall computer load. A question which has no unique answer but depends on the application is how to define accuracy or precision. For the present report we have made special runs designed to approximately determine the minimum computing time required to achieve a given precision for some test cases. By precision we refer to all accuracy criteria except convergence with respect to N and P, i.e., we mean accuracy of the numerical solution for given basis-set sizes. [Recall that N is the order of the close coupling equation (2) and P is the number of channels propagated; for all the methods considered here except R matrix propagation, $P = N$.] Bearing in mind the above caveats, one should not attempt to draw conclusions based on the fine details of the computer times presented in this report. Some overall trends

and general magnitudes are however meaningful and interesting. All computer times given in this report are execution times, excluding compilation. All source codes, except for the University of Minnesota linear equation solver used in both our Numerov and R matrix propagation codes, were written in FORTRAN. For purposes of rough comparison to calculations performed on other computers, Table 1 gives approximate conversion factors.

III. Electron-molecule scattering

We have applied all four methods introduced in section II to electron scattering by N_2. We consider two classes of problems: (A) vibrational-rotational close coupling and (B) rotational close coupling with the rigid rotator approximation.

A. Vibrational-rotational close coupling. The first method we attempted to apply to electron-molecule scattering was the piecewise analytic method. We used our modified version of the QCPE program. Although we had limited success with this method, we found that it was inadequate to complete some of the applications we attempted. These applications[9] involved vibrational-rotational close coupling calculations for electron scattering by N_2 at energies of 5-45 eV. Two difficulties we encountered were: (i) we were unable to calculate accurate small transition probabilities, especially those associated with vibrational transitions, with reasonable stepsizes; (ii) for some problems the results were not converged even with impractically small stepsizes, e.g., 5×10^{-6} a_0. The first difficulty we tentatively associate with the fact that the piecewise analytic solution used is correct for a diagonal linear reference potential, but although the transformation method used diagonalizes the potential at the center of each sector, it does not diagonalize the derivative of the potential. Thus the transformed potential through linear terms is not diagonal in a sector. This or some other aspect of the method causes it to be poorly suited for the accurate calculation of small S matrix elements. In considering difficulty (ii), we note that the success and efficiency of a sophisticated variable-stepsize integrator is highly dependent on the reliability of the stepsize algorithm. In the present case, however, during the course of the solution, the predicted

stepsize sometimes became exceedingly small, e.g., 10^{-12} a_0. It was suggested to us that we just substitute a very small minimum stepsize, e.g., 10^{-3} a_0, and continue to propagate until the stepsize algorithm again predicted stepsizes larger than the minimum. Since the piecewise analytic method requires more expense per step than less sophisticated methods like the Numerov method, its efficiency requires that the stepsizes be fairly large. Nevertheless we tried the minimum stepsize procedure and were disappointed to find some applications where the results were not converged with respect to minimum stepsize even at 5×10^{-6} a_0. The difficulty of calculating small vibrational transition amplitudes by the piecewise analytic method has also been noted elsewhere.[49]

The second method we applied to electron-molecule vibrational-rotational close coupling was the Numerov method.[8,9] We found this method to be accurate and reliable for all cases attempted, even those for which the piecewise analytic method was unacceptable.

B. Rotational close coupling with the rigid rotator approximation. We now consider the electron-N_2 rotational close coupling problem studied in reference 26. In that study the diatom is treated as a rigid rotator with the equilibrium internuclear distance (2.068 a_0), and the rotational close coupling problem is formulated in the laboratory frame using the total angular momentum representation of Arthurs and Dalgarno.[50] Only the ground electronic state is included explicitly, and effects of electronic-charge-cloud polarization and of electron exchange are included by means of an effective potential.[10] The anisotropic electron-molecule interaction potential is expanded as

$$V(r,\chi) = \sum_\lambda V_\lambda(r) \, P_\lambda(\cos \chi) \tag{114}$$

where χ is the colatitude of the scattering electron with respect to the internuclear axis. The individual terms $V_\lambda(r)$ are represented by spline functions that have a cusp at one half the internuclear distance, i.e., at $r = 1.034 \, a_0$. Both our Numerov and our R matrix propagation codes have special provisions for choice of stepsize in the vicinity of the cusp. For the Numerov calculations reported here we did not use this provision; we just ignored the cusp. For the R matrix propagation calculations we shortened the sector before the cusp to put a sector boundary at the cusp.

The calculations of reference 26 used the R matrix propagation scheme, and we attempted to obtain three-significant-figure precision in the elastic and inelastic transition probabilities. In a separate study[10] we had found that $j_{max} = 6$ was required for convergence of the $J = 5$, even-j partial cross sections at 30 eV impact energy where j and J are rotational and total angular momentum quantum numbers in the Arthurs-Dalgarno scheme. A conventional basis for this j_{max} and J consists of 15 channels. The next smaller conventional basis ($j_{max} = 4$) and the ℓ-dominant[51] basis both contain 9 functions. In reference 26 we compared calculations with various conventional, ℓ-dominant, and adiabatic bases. In Table 2 we give for comparison some representative results. These examples show the general result that the ℓ-dominant basis provides significantly more accurate results for elastic and inelastic scattering from the ground rotational state than the conventional basis of the same dimension, but that a 9-function adiabatic basis (consisting of the 9 lowest energy eigenfunctions obtained by taking linear combinations of the 15 total angular momentum eigenfunctions of the

primitive basis in each sector) is considerably more accurate than either of these. In fact, for transitions between excited rotational states even a 6-function adiabatic basis provides better agreement with the converged results than either of the two 9-function diabatic bases.

For the $N = 15$, $P = 15$ and $N = 15$, $P = 9$ cases we have used trace statements to make a detailed study of how much computer time is spent in each subprogram. We then related the computer time spent in each sub-program to the computer time spent on various parts of the calculation. The results of this timing study are shown in Table 3. For the computer runs upon which Table 3 is based, we used a single-energy, single-basis-set version of our code which used no disk reads or writes during the cal-culation. Dimensions for arrays in common blocks were set to accomodate a maximum of 15 channels and 3000 sectors. The field length required for this to run was 101500 (base 8) words. We propagated only R_4 and we used the tranpose rather than the inverse of T. We set all $\varepsilon^{(i)} = 0.07$. Results

the same accuracy can be obtained more efficiently by dividing the propagation range into three or four intervals and optimizing $\varepsilon^{(i)}$ separately in each. This extra optimization was employed for some production runs[26] but not for the timing comparison reported here. We placed the center of the first sector at 0.03 a_0 and used 5×10^{-6} a_0 for the first two stepsizes. The stepsize algorithm yields $h^{(3)} = 4.4 \times 10^{-4}$ a_0. The 15/15 run required 250 steps to reach 1 a_0 where $h^{(1)} = 0.014$ a_0, 50 more steps to reach 2 a_0 where $h^{(1)} = 0.031$ a_0, and 45 more steps to reach 4 a_0, where $h^{(1)} = 0.058$ a_0. The final asymptotic analysis was performed at 101.6 a_0 by which point the stepsize had increased monotonically to 1.4 a_0. The first 486 of the sectors are located at $r < 30$ a_0. Multiplying the first three entries in Table 3

by 486/572 gives an estimate of 65 sec for the execution time of a calcu-
lation which would quit at 30 a_0. A similar correction to 30 a_0 yields
47 sec for the 15/9 calculation. One should keep in mind that the time
required for the calculation of $\underset{\sim}{D}$ is highly dependent on the complexity
of the interaction potential and the efficiency of the potential subprogram.

The most striking result in Table 3 is that most of the computer time
is involved in diagonalizing $\underset{\sim}{D}$. The propagation of the R_4 matrix accounts
for only 52% of the time in the 15/15 run and only 13% of the time in the
15/9 run. Thus, increasing the efficiency of this step by even a factor
of two would result in savings of only 16% and 6% in the two runs,
respectively.

For the 15/15 case, decreasing $\varepsilon^{(i)}$ to 0.05 and moving the asymptotic
analysis to 120.8 a_0 to check the accuracy required 867 sectors and 109.5
sec computer time. Most partial cross sections calculated in this run
agreed to ±1 in the third significant figure with those for the run described
above.

For comparison with these calculations we ran the Numerov code with the
same potential and a similar potential subprogram for the same impact energy
and $J = 5$. We also used the same compiler (MNF) and the same computer
(the CDC Cyber 74). Trial calculations showed that one can obtain about
three significant figures of precision by performing the asymptotic analysis
at 30 a_0 with Ricatti-Bessel functions. Similar accuracy can be obtained
at 10 a_0 using BMS functions, but with our computer program the overall
cost for the present case is greater with BMS functions because of the cost
of computing the BMS functions. The BMS functions are less expensive for
cases with a smaller number of channels. We made several runs designed

to discover the most efficient procedure for solving the 15-channel test problem performing the asymptotic analysis at 30 a_0 using Ricatti-Bessel functions. All these runs began the propagation at 10^{-6} a_0 with $h_0 = 0.0005$ a_0. We used h_0 for the first ten steps. First we used the regular Numerov method with $h_{max} = 0.064$ a_0 and various δ to determine the δ required for three significant figures of accuracy. This yielded $\delta = 10^{-4}$ a_0, 547 steps, and a computer time of 33 sec. With the same numerical parameters the iterative Numerov method with EPS = 10^{-4} required 48 sec. With the same values of h_0, h_{max}, and δ, we ran the option which tests the iterative against the regular Numerov method every 25 steps. With EPS = 10^{-4}, the iterative Numerov method was found to be slower at all distances for this test case; this run also required 33 sec. In the runs just discussed the maximum stepsize of 0.064 a_0 was reached at $r = 0.70$ a_0. We removed the maximum stepsize criterion and again searched for the δ which yields about three significant figures of accuracy. For most partial cross sections this could be achieved with $\delta = 10^{-6}$ or 10^{-5}. Using KCK = 25 and EPS = 10^{-4}, these two calculations required 357 and 268 steps and 28 and 19 sec, respectively. These times may be compared with 65 sec for the R matrix propagation code on the same compiler and same computer.

Then the Numerov calculation with KCK = 25 and $\delta = 10^{-6}$ was rerun with the FTN compiler with optimization levels OPT = 1 and OPT = 2; the execution times decreased to 24 sec and 15 sec, respectively. Retaining KCK = 25, $\delta = 10^{-6}$, and OPT = 2, we increased EPS to 10^{-2}; the regular Numerov method was still faster at every check. For some other electron-molecule scattering calculations we have found that the iterative Numerov method is relatively more efficient at large r and large J.

The dimensions in the Numerov program were set for 15 channels and the field length required to execute was 114500 (octal) words. For 15-channel calculations on the He-HF system considered below the field length required is only 105700 (octal) words. The difference is accounted for by the large number of spline coefficients in the electron-N_2 interaction potential.

It is difficult to compare the computation times for the calculations of reference 23 using the integral equations algorithm to those reported here as obtained with the other programs. The integral equations algorithm has been used for calculations in the body-frame formalism,[21-23] whereas the other calculations discussed in this section use the Arthurs-Dalgarno formalism.[50] The input for the integral equations program consists of the $V_\lambda(r)$ values on the integration mesh, whereas the input for the Numerov and R matrix propagation programs consists of spline fits to the $V_\lambda(r)$. This means that the stepsizes for the integral equations calculations are input variables. For a typical run in reference 23, the stepsize for the first twenty steps was 0.001 a_0, followed by 298 steps with h = 0.01 a_0, 152 steps with h = 0.1 a_0, and 100 steps with h = 0.2 a_0, for a total of 570 ending at 38.2 a_0. Table 4 shows typical computation times for various numbers of channels. These calculations are for Σ_g symmetry at E = 13.6 eV with λ_{max} = 28. No truncation of the number of coupled channels was employed, i.e., the truncation radius is infinite. Thus all N channels were propagated at all distances as in the Numerov calculations and the R matrix propagation calculations with P = N.

The potential used for the R matrix propagation and Numerov test cases is called potential i in reference 10. The results presented above were

presented at the NRCC Workshop in June, 1979. We have also submitted a set of $V_\lambda(r)$ for this potential to L. Thomas who prepared a potential subprogram based on this potential for the Workshop participants as a test case for further study.

IV. Electron-atom scattering

Electron-hydrogen atom scattering provides an interesting test case for close coupling codes.[41a] It is a prototype for electronically inelastic electron-atom scattering in general, but it has the advantage that the interaction potential matrix is known analytically. The interaction potential for electron-atom scattering is qualitatively different from the interaction potential for the other test problems considered in this report in that it tends to $-\infty$ at the origin. The same limit occurs for electron scattering from molecules like CO_2, which has a nucleus at the origin.[52] For comparison of computation times for electron-atom scattering we consider a 2-channel problem: 1s-2s close coupling without exchange and with total angular momentum zero.

To use the piecewise analytic method for electron-hydrogen atom scattering, we had to modify the QCPE program to allow for starting channels which are open at the origin. For various 3- and 4-channel examples, reasonably precise results could be obtained by starting at about $10^{-5} a_0$ and using a stepsize error criterion of 10^{-5} or 10^{-6}. Using the FT3 compiler and a CDC 6600 computer, precise results for the 1s-2s s-wave test case required about 1.6 sec. This corresponds to about 1.5 sec on the CDC Cyber 74. In general it was difficult, as compared to using our Numerov program, to test and obtain convergence with respect to the starting point and the stepsize error criterion.

The R matrix propagation method, propagating R_1, R_2, R_3, and R_4, was applied to the test case and the numerical parameters were adjusted so that the partial cross sections were precise to ±1 in the third significant figure. This yielded $\varepsilon^{(i)} = 0.05$. The calculation was $\backslash$:11

converged at sector 166 at 17.5 a_0. The stepsize became very large at large r, and two more steps brought the calculation to the final sector centered at 25.1 a_0. The computation time, using the MNF compiler and the Cyber 74 computer, was 0.6 sec.

To compare the efficiency of the computer codes we applied the variable-stepsize regular Numerov method with Ricatti-Bessel function boundary conditions to the same problem, again using the MNF compiler and the Cyber 74 computer. Asymptotic analysis was performed at 25 a_0 where it was converged with respect to further propagation to 5 significant figures. We used $h = 5 \times 10^{-4}$ a_0 for the first ten steps. Subsequent stepsizes were determined by increasing δ in successive runs until we obtained only three significant figures of precision in the partial cross sections. This required $\delta = 10^{-6}$ and a computation time of 0.42 sec. In this run the stepsize doubled every step from the eleventh until it reached 0.128 a_0. It then increased to its final value, 0.256 a_0 at 0.52 a_0. The calculation required 116 steps. We repeated the calculation using the option to check every 11th step whether the regular or iterative Numerov method is faster. With EPS = 10^{-4}, the iterative Numerov method became faster at r = 18.2 a_0. This whole calculation required 0.44 sec. The iterative Numerov method is relatively more efficient for cases with centrifugal barriers.

To check the sensitivity to compiler we reran the most efficient of the above calculations with a code compiled on the FTN compiler with optimization levels OPT = 1 and 2. The computer time increased from 0.42 sec to 0.44 and 0.43 sec, respectively.

A class of methods which has been widely applied to electron-atom scattering, and to a lesser extent to electron-molecule scattering, is

the algebraic variational method and related techniques involving expansion
of the translational wave function in a basis set.[53-59] In these methods
the potential is generally represented in a basis set rather than as a
function of the radial coordinate. These methods can often be used to
solve the same problems as are attacked by close coupling codes like the
ones discussed here which rely on numerical integration of coupled dif-
ferential equations. However, they become relatively more efficient and
more useful when nonlocal exchange potentials are included.

Using our original algebraic variational program[60] with the FUN com-
pile: on the CDC 6600 computer, a typical run on the 2-channel test problem
of this section required 26 sec (equivalent to about 24 sec on the CDC
Cyber 74). This run involved 15 and 10 uncontracted basis functions for
the expansion of the translational wave functions in the 1s and 2s channels
respectively. This time can be speeded up by using more efficient proce-
dures for evaluation of the integrals over basis functions. Unfortunately
a computer time is not available using our more efficient integrals
packages[61,62] on this test problem. The computer time can also be decreased
by using contracted basis functions.[63]

V. Vibrationally inelastic atom-diatom scattering

For atom-molecule scattering the interaction potential is independent of energy. When the R matrix propagation method is used to study systems with energy-independent interaction potentials, great savings of computer time can be made since calculations with the same number N of primitive basis functions can be carried out for several energies with the time-consuming diagonalization of the interaction matrix $\underset{\sim}{D}$ carried out only once.[24] The calculation at the first energy is called a reference-mode calculation. In a reference mode calculation, $\underset{\sim}{D}$ is calculated and diagonalized in every sector, the eigenvalues are stored on the disk, and the transformation matrix $\underset{\sim}{T}$ is calculated and stored on the disk. If $\underset{\sim}{T}^{-1}$ is to be used, it too is computed and stored on the disk. Calculations at additional energies may be carried out in a propagation-mode. Additionally, to test convergence with respect to P the propagation-mode can also be used to run calculations at the same energy but with successively smaller values of P.

We here report a detailed study of the timing requirements of the many-energy, many-basis version of our R matrix propagation program for a test problem. The test problem is collinear scattering of He by H_2 with a harmonic oscillator potential for H_2 and an exponential repulsive inter-action potential. The Hamiltonian is the same as used in two published studies[64,65] and corresponds to $m = 2/3$ and $\alpha = 0.314$ in the unitless notation of reference 65. We considered total energies of $8\hbar\omega$ and $7.75 \hbar\omega$ and used harmonic oscillator eigenfunctions for the primitive basis. We used the many-energy, many-basis version of our computer program to run a series of nineteen calculations with various values of N and P in the

range 7-15. By printing out the computation times in various subprograms
we were able to approximately distribute the computation times into four
categories: setup, calculation and diagonalization of $\underline{D}$, propagation, and
asymptotic analysis. We then fit each category of computation time (in
sec) to an empirical function of N or P. The total computation time (in
sec) is called t, and the functions for the four categories are called s,
d, p, and a respectively. For a reference-mode calculation we obtained

$$t \stackrel{\sim}{=} s_1 + d(N) + p_1(P) + a \tag{115}$$

where

$$s_1 \stackrel{\sim}{=} 0.5 \tag{116}$$

$$d(N) \stackrel{\sim}{=} 1.13 \times 10^{-3} N^3 + 0.031 N^2 \tag{117}$$

$$p_1(P) \stackrel{\sim}{=} 0.029 P^2 \tag{118}$$

and

$$a \stackrel{\sim}{=} 0.5 \tag{119}$$

For a propagation-mode calculation we obtained

$$t \stackrel{\sim}{=} s_2 + p_2(P) + a \tag{120}$$

where

$$s_2 \stackrel{\sim}{=} 0.3 \tag{121}$$

$$p_2(P) \stackrel{\sim}{=} 0.021 P^2 (1 + 0.008P) \tag{122}$$

and a is as before. The functional forms in (117), (118), and (122) have
not been fit exactly; they are chosen strictly to provide a simple empirical
fit over the range of N and P examined. In principle other powers of N and

P should appear. U ing these functions we can generate the sample compu-
tation times in Table 5.

For the runs used to generate equations (115)-(122), we propagated
only R_4 and we used the inverse of T. We also made runs in which we used
the transpose of T; for our program there was little difference in the
execution time as compared to using the inverse. We determined that
$c^{(i)} = 0.15$ and $r_L^{(1)} = 1.195 \ a_0$ were just sufficient to get 2% accuracy
for all transition probabilities and 3-significant-figure accuracy for
those greater than 10^{-5}. We found that placing the center of the last
sector at $r_R^{(c)} \geq 6 \ a_0$ was sufficient to ensure that our results were con-
verged with respect to increasing the range of the propagation. To
achieve this convergence for all the runs and to use a fixed number of
sectors to simplify the interpretation of the computation times, we used
150 sectors for all the runs and propagated to 7-24 a_0, depending on N and
P; however, the stepsize becomes large at large r (for the last few steps
$h^{(i)} = h_{max} = 3 \ a_0$), and r = 6 a_0 is not reached earlier than the 142nd
sector. Typical values for the centers of the sectors and the stepsizes
are $r_c^{(i)} = 1.40$, 1.67, and 2.54 a_0 and $h^{(i)} = 0.0096$, 010122, and 0.0261 a_0
for i = 25, 50, and 100, respectively.

Table 5 shows that calculations at second and subsequent energies are
faster by factors of about 3-5 than calculations at the initial energy.
These savings are also achieved when a series of P values is run to test
convergence.

In Tables 6 and 7 we give representative transition probabilities
$P_{vv'} = |S_{vv'}|^2$ calculated for the simple model of vibrationally inelastic

collinear atom-molecule scattering described above.[29] For energies corresponding to $4\hbar\omega$ and $8\hbar\omega$ we see that an adiabatic basis of propagation dimension P can give significantly more accurate results than a conventional basis of the same dimension. For example, at $E = 8\hbar\omega$ converged results for all but transitions involving the highest open channel, $v' = 7$, can be obtained with a 9-function conventional basis which includes harmonic oscillator eigenfunctions corresponding to $v = 0-8$. In contrast the transition probabilities obtained by calculations using the 8-function conventional 'asis which includes only the open channels can have large errors. E.g., as shown in Table 7, the $N = 8$, $P = 8$ calculations of P_{25}, P_{46}, and P_{06} have errors of 12%, 47%, and more than a factor of two, respectively. The $N = 9$, $P = 8$ adiabatic basis, however, gives all but 3 transition probabilities (P_{34}, P_{57}, and P_{67}) to within 1%. Thus we see for this example, that when the results differ, adiabatic bases yield considerably more accurate results than conventional bases of the same or frequently even larger propagation dimension P. This is an important result. To obtain the adiabatic basis extra effort must be expended to diagonalize the N × N interaction matrix $\underline{D}_C^{(i)}$ in each sector. However, when the interaction potential is independent of energy, as it is for the present problem, the adiabatic basis functions in each sector are also independent of energy. Consequently, as we have seen above, significant computational savings can result from obtaining the adiabatic basis in a reference-mode calculation and using it for several energies in propagation-mode calculations.

Table 5 shows that most of the computer time in a reference-mode calculation is spent calculating and diagonalizing $\underline{D}$. Since the interaction

potential is very simple for this test case, the diagonalization step is
the slow one. One idea for a method to reduce the time required to evaluate
the eigenvalues and eigenvectors of $\underline{D}$ is to calculate and diagonalize $\underline{D}$
on a coarse grid, fit the eigenvalues and eigenvectors to spline functions,
and use the spline functions for a propagation-mode calculation on a finer
grid. One would have to converge the calculation with respect to the spline
grid as well as the propagation grid. We tried this for the case $N = 8$,
$P = 8$ with the unconverged coars' grid being about ten times coarser than
the propagation grid. We used the storage-efficient but computer-time
inefficient version of our spline subprograms. The computer time was
7.4 sec (compare 5.5 sec in Table 5). Since the spline version of the
program was not at all optimized, this test indicates that this kind of
idea deserves further consideration. Another possible way to speed up
the diagonalization step is to use an iterative method for the diagonal-
ization. The diagonalization at the previous step would be used to start
the iteration.

The piecewise analytic method has also been used for this test problem.
We used the program written by Wagner[14] for this purpose. Compiling this
on the MNF compiler and running it on the CDC 6600 computer required 9.1 sec
computer time (corresponding to about 8.4 sec on the Cyber 74) to obtain
slightly less than three-significant-figure accuracy for a basis with 8
channels; at the energy considered 5 channels were open, and no transition
probabilities were less than 10^{-4}. The piecewise analytic method, like the
R matrix propagation method, has the advantage that additional calculations
at subsequent energies can be performed with reduced cost.

VI. Rotationally Inelastic Atom-Diatom Scattering

We have applied three different methods to rotationally inelastic scattering of an atom by a rigid rotator using the Arthurs-Dalgarno[50] scheme Consider first He-HF scattering with the interaction potential of Collins and Lane.[20,28,66] For a test case we study scattering at total angular momentum $J = 12$ and impact energy 0.05 eV using a conventional 10-function basis with $j_{max} = 3$. Using the R matrix propagation scheme (propagating only R_4) we found that propagating from 3.5 a_0 to 20 a_0 was sufficient to give 1% precision for the real and imaginary parts of all S matrix elements greater than 10^{-2} and three-significant-figure precision for partial cross sections from the ground state. We used the same $\varepsilon^{(i)}$ for every sector and increased it till we just retained this accuracy. This required $\varepsilon^{(i)} = 0.3$, and took 4.8 sec execution time for a program compiled on the MNF compiler and run on the Cyber 74 computer. Using the same compiler and computer, we repeated the calculation with the regular Numerov method, starting at 3.5 a_0 and applying Ricatti-Bessel function boundary conditions at 24 a_0. Using fixed stepsizes, we found that we could satisfy the criteria given above with $h = 0.064 \, a_0$. This calculation required 15 sec computer time. We repeated this calculation using the FTN compiler with optimization level OPT = 2, and the execution time decreased to 10 sec. We then tried the variable-stepsize algorithm and obtained comparable accuracy for $\delta = 7 \times 10^{-8}$; in this run the stepsize increased to 0.064 a_0 at $r = 4.50 \, a_0$ and remained at this value; the execution time was still 10 sec with the FTN compiler and optimization level OPT = 2. Based on Allison's experience[5] one would assume that the execution could be improved by using the iterative Numerov method with a carefully chosen value of EPS.

One advantage of the R matrix propagation method for this problem is that, since the interaction potential is independent of energy, one can perform calculations at additional energies at a reduced cost. Further, by judicious use of contraction, i.e., $P < N$, one can often obtain good accuracy without propagating the full number of functions required in the primitive basis set.[28] To illustrate this we give in Table 8 some representative rotational excitation cross sections for $E = 0.05$ eV for $J = 4$ and 12.[28-29] As shown here, we found that in all cases adiabatic bases of propagation dimension P gave more accurate results than conventional calculations of the same dimension except where accidental cancellation of errors occurred. In Table 8 we see that a 10-function conventional basis including channels with $j \leq j_{max} = 3$ gives converged results for both total angular momenta. In contrast the next smaller conventional basis, the 6-function $j_{max} = 2$ basis, gives errors of 15% and 66% in $\sigma_{0 \rightarrow 1}^{4}$ and $\sigma_{0 \rightarrow 2}^{4}$, respectively, and 12% and 48% in $\sigma_{0 \rightarrow 1}^{12}$ and $\sigma_{0 \rightarrow 2}^{12}$. However, the $N = 10$, $P = 6$ adiabatic basis yields all four cross sections with an accuracy of 9%. In Table 9 we show that the significant increase in accuracy obtained when an adiabatic basis is used instead of a conventional basis is not predicated on a fortuitous cancellation of errors. For this example, the $j_{max} = 2$ basis gives errors of almost a factor of 2 for each partial cross section, while the 6-function adiabatic basis yields all three partial cross sections within 6%.

The piecewise analytic method has been widely used for atom-rigid-rotator collisions, although even for this problem the difficulty of obtaining very precise answers has been noted.[5] We applied the piecewise analytic method to the 16-channel problem of Johnson et al.[67] Using the

numerical parameters of the QCPE test data, the FUN compiler, and the CDC
6600 computer, we obtained less than three significant figures of precision
in 12 sec for a first-energy calculation. This corresponds to about 11 sec
on the CDC Cyber 74 computer. This problem required propagating from 0.73 a_0
to 6.5 a_0. For comparison, the R matrix propagation method for the He-HF
problem discussed above, with $\varepsilon^{(1)} = 0.3$ and propagating from 3.5 to 20.0 a_0,
required 10.3 and 8.4 sec computer time on the CDC Cyber 74 computer for
$N = 15$, $P = 15$ and $N = 15$, $P = 10$ problems, respectively, for first-energy
calculations.

VII. Summary

We have presented some details of our implementation of the Numerov and R matrix propagation methods for inelastic close coupling calculations. We have found that these methods are convenient and reliable for electron-atom, electron-molecule, and atom-molecule collisions involving rotationally, vibrationally, and electronically inelastic scattering where the close coupling equations take the form of coupled differential equations. We have shown that the computer time requirements of these two methods as well as the piecewise analytic method and the integral equations method are similar for a variety of inelastic close coupling calculations. The ultimate choice among these methods should therefore often be based on other considerations, such as ease of programming in the Numerov method or of using adiabatic basis sets in the R matrix propagation method. Both these methods have favorable properties for checking convergence with respect to numerical parameters in calculations performed on a production basis.

VIII. Acknowledgements

We are grateful to Drs. John R. Rumble and Bruce C. Garrett for help with this project. This work was supported in part by grant no. CHE77-27415 from the National Science Foundation.

1. B. Numerov, Publ. Observatoire Central Astrophys. Russ. 2, 188 (1933).

2. R. W. Hamming, Numerical Methods for Scientists and Engineers (McGraw-Hill, New York, 1962), p. 215.

3. L. L. Barnes, N. F. Lane, and C. C. Lin, Phys. Rev. 137, A388 (1965).

4. J. M. Blatt, J. Comput. Phys. 1, 382 (1967).

5. A. C. Allison, Ph.D. thesis, University of Glasgow, Glasgow, Scotland, 1967; A. C. Allison, J. Comput. Phys. 6, 378 (1970).

6. N. F. Lane and S. Geltman, Phys. Rev. 160, 53 (1967).

7. R. J. W. Henry and E. S. Chang, Phys. Rev. A 5, 276 (1972).

8. M. A. Brandt, M.S. thesis, University of Minnesota, Minneapolis, 1975.

9. M. A. Brandt, D. G. Truhlar, and F. A. Van-Catledge, J. Chem. Phys. 64, 4957 (1976); D. G. Truhlar, M. A. Brandt, A. Chutjian, S. K. Srivastava, and S. Trajmar, J. Chem. Phys. 65, 5535 (1976); D. G. Truhlar, M. A. Brandt, A. Chutjian, S. K. Srivastava, and S. Trajmar, J. Chem. Phys. 65, 2962 (1976).

10. K. Onda and D. G. Truhlar, J. Chem. Phys. 69, 1361 (1978), 70, 1681 (1979).

11. K. Onda and D. G. Truhlar, to be published.

12. R. G. Gordon, J. Chem. Phys. 51, 14 (1969).

13. R. G. Gordon, Methods Comput. Phys. 10, 81 (1971).

14. A. F. Wagner and V. McKoy, J. Chem. Phys. 58, 2604 (1973), 58, 5561 (1973).

15. M. H. Alexander, J. Comput. Phys. 20, 248 (1976).

16. R. J. Gordon, J. Chem. Phys. 67, 5923 (1977).

17. W. N. Sams and D. J. Kouri, J. Chem. Phys. 51, 4809 (1969), 51, 4815 (1969).

18. R. A. White and E. F. Hayes, J. Chem. Phys. 57, 2895 (1972), Chem. Phys. Lett. 14, 98 (1972).

19. J. T. Adams, R. L. Smith, and E. F. Hayes, J. Chem. Phys. 61, 2193 (1974).

20. L. A. Collins, Ph.D. Thesis, Rice University, Houston, Texas, 1975.

21. M. A. Morrison, Ph. D. thesis, Rice University, Houston, Texas, 1976.

22. M. A. Morrison, L. A. Collins, and N. F. Lane, Chem. Phys. Lett. $\underline{42}$, 356 (1976), M. A. Morrison, N. F. Lane, and L. A. Collins, Phys. Rev. A $\underline{15}$, 2186 (1977).

23. J. R. Rumble and D. G. Truhlar, J. Chem. Phys. $\underline{70}$, 4101 (1979).

24. J. C. Light and R. B. Walker, J. Chem. Phys. $\underline{65}$, 4272 (1976).

25. D. G. Truhlar and N. A. Mullaney, J. Chem. Phys. $\underline{68}$, 1574 (1978).

26. N. A. Mullaney and D. G. Truhlar, Chem. Phys. Lett. $\underline{58}$, 512 (1978).

27. E. B. Stechel, R. B. Walker, and J. C. Light, J. Chem. Phys. $\underline{69}$, 3518 (1978).

28. N. A. Mullaney and D. G. Truhlar, Chem. Phys. $\underline{39}$, 91 (1979).

29. N. Mullaney Harvey, Ph.D. thesis, University of Minnesota, Minneapolis, 1979.

30. J. W. Cooley, Math. Comp. $\underline{15}$, 363 (1961).

31. I. H. Sloan, J. Comp. Phys. $\underline{2}$, 414 (1968).

32. B. R. Johnson, J. Chem. Phys. $\underline{69}$, 4678 (1978).

33. W. A. Lester, Jr., J. Comput. Phys. $\underline{3}$, 322 (1968); W. A. Lester, Jr., University of Wisconsin Theoretical Chemistry Institute Technical Report WIS-TCI-285, Madison, 1968.

34. R. deVogelaere, J. Research Natl. Bur. Std. $\underline{54}$, 119 (1955); R. deVogelaere, Z. Angew. Math. Phys. $\underline{8}$, 151 (195.).

35. R. G. Gordon, program entitled "Coupled Channel Scattering Matrices", program no. 187 described in Quantum Chemistry Program Exchange Catalog and Procedures, Vol. 10 (Indiana University Chemistry Department, Bloomington, Indiana, 1974), p. 67.

36. W. A. Lester, Jr. and R. B. Bernstein, Chem. Phys. Lett. $\underline{1}$, 207 (1967), erratum: $\underline{1}$, 347 (1967).

37. N. F. Mott and H. W. S. Massey, The Theory of Atomic Collisions, third edition (Oxford University Press, London, 1965), pp. 388ff.

38. R. S. Varga, Matrix Iterative Analysis (Prentice-Hall, New Jersey, 1962), p. 56; (a) p. 78; (b) p. 23.

39. M. E. Riley and A. Kuppermann, Chem. Phys. Lett. $\underline{1}$, 537 (1968); M. E. Riley, Ph.D. thesis, California Institute of Technology, Pasadena, 1968.

40. Handbook of Mathematical Functions, M. Abramowitz and I. A. Stegun, eds. (U.S. Dept. of Commerce, Natl. Bur. Standards Appl. Math Series 55, Washington, D.C., 1964), p. 437.

41. K. Smith, The Calculation of Atomic Collision Processes (Wiley-Interscience, New York, 1971), p. 15, (a) p. 96.

42. P. G. Burke and H. M. Schey, Phys. Rev. 126, 147 (1962).

43. P. G. Burke, D. D. McVicar, and K. Smith, Proc. Phys. Soc. (London) 83, 397 (1964).

44. D. W. Norcross, Comput. Phys. Commun. 1, 88 (1969). This program has been revised in A. T. Chivers, Comput. Phys. Commun. 5, 416 (1973).

45. D. G. Truhlar, C. A. Mead, and M. A. Brandt, Advan. Chem. Phys. 33, 295 (1975).

46. J. C. Light, Methods. Comp. Phys. 10, 111 (1971).

47. L. A. Collins, W. D. Robb, and M. A. Morrison, J. Phys. B 11, L777 (1978).

48. M. A. Morrison, in Electron- and Photon-Molecule Collisions, edited by V. McKoy, T. N. Rescigno, and B. I. Schneider (Plenum, New York), in press.

49. S. Green, J. Chem. Phys. 70, 4686 (1979).

50. A. M. Arthurs and A. Dalgarno, Proc. Roy. Soc., Lond., Ser. A 256, 540 (1960).

51. A. E. dePristo and M. H. Alexander, J. Chem. Phys. 63, 3552, 5327 (1975).

52. K. Onda and D. G. Truhlar, J. Phys. B 12, 283 (1979).

53. F. E. Harris and H. H. Michels, Methods. Comp. Phys. 10, 143 (1971).

54. D. G. Truhlar, J. Abdallah, Jr., and R. L. Smith, Advan. Chem. Phys. 25, 211 (1974).

55. P. G. Burke and W. D. Robb, Advan. At. Molec. Phys. 11, 143 (1975).

56. R. K. Nesbet, Advan. Quantum Chem. 9, 215 (1975).

57. J. Callaway, Phys. Reports 45, 89 (1978).

58. B. I. Schneider, in Electronic and Atomic Collisions, edited by G. Watel (North Holland, Amsterdam, 1978), p. 259.

59. A. Fliflet and V. McKoy, Phys. Rev. A 18, 2107 (1978).

60. R. L. Smith and D. G. Truhlar, Phys. Lett. A 39, 35 (1972).

61. R. L. Smith and D. G. Truhlar, Comput. Phys. Commun. 5, 80 (1973), erratum: 8, 333 (1974).

62. J. Abdallah, Jr. and D. G. Truhlar, Comput. Phys. Commun. 9, 327 (1975).

63. J. Abdallah, Jr. and D. G. Truhlar, J. Chem. Phys. 61, 30 (1976).

64. A. P. Clark and A. S. Dickinson, J. Phys. B 6, 164 (1973).

65. J. W. Duff and D. G. Truhlar, Chem. Phys. 9, 243 (1975).

66. L. A. Collins and N. F. Lane, Phys. Rev. A 14, 1358 (1976).

67. B. R. Johnson, D. Secrest, W. A. Lester, Jr., and R. B. Bernstein, Chem. Phys. Lett. 1, 396 (1967).

Table 1. Approximate conversion factors for computer time

Computer	Factor relative to CDC Cyber 74
CDC 6400	0.35[a]
CDC 6600	0.92
CDC 7600	6.0
IBM 360/75	0.6
IBM 370/168	1.4
IBM 360/91	2.8
IBM 370/195	6.0
Univac 1108	0.4

[a]For example, to convert a computation time obtained with the CDC 6400 computer to an expected computation time on the CDC Cyber 74, multiply by 0.35.

Table 2. Sums of transition probabilities for transitions from the ground state $(j,\ell) = (0,5)$ and from the second rotationally excited state[a] as a function of basis for e-N_2, $J = 5$ scattering at 30 eV.[b]

description of basis	conventional	conventional	ℓ-dominant	adiabatic	adiabatic
N/P	15/15	9/9	9/9	15/9	15/6
j'			$\sum_{\ell'} P_{05j'\ell'}$		
2	2.92(-3)[c]	3.16(-3)	2.70(-3)	2.97(-3)	
4	5.36(-4)	2.39(-4)	4.40(-4)	5.63(-4)	
6	1.99(-4)		1.42(-4)	2.11(-4)	
j'			$\sum_{\ell'} P_{23j'\ell'}$		
4	2.46(-1)	1.04(-1)	2.28(-1)	2.48(-1)	2.47(-1)
6	1.07(-1)		8.08(-2)	1.11(-1)	1.10(-1)

[a] These are dominated by transitions from $(j,\ell) = (2,3)$ channel

[b] See reference 26

[c] The number in parentheses denotes the power of ten by which the entry should be multiplied.

Table 3. Timing study using the R matrix propagation method
for two runs on the electron-N_2 scattering problem
of reference 26.

Function	computation time (sec)[a]	
	$N/P = 15$[b]	$N/P = 15/9$[c]
Calculation of $\underline{D}$	9.1	8.8
Diagonalization of D	40.7	39.9
R matrix propagation	24.1	7.5
Asymptotic analysis	0.6	0.8
Miscellaneous	1.5	1.4
Total	76.0	58.5
	computing time (as a percentage of 76.0 sec)	
Calculation of $\underline{D}$	12	12
Diagonalization of $\underline{D}$	54	52
R matrix propagation	32	10
Asymptotic analysis	1	1
Miscellaneous	2	2
Total	100	77

[a] MNF compiler, Cyber 74 computer.

[b] 572 sectors

[c] 561 sectors

Table 4. Computation times (in sec) for body-frame
close coupling calculations on electron-N_2
scattering using the integral equations
program.

N^a	CDC 7600[b]	Cyber 74[c]
14	12.2	61
9	7.0	35
8	6.3	31.5
7	5.5	27.5

[a] number of coupled channels

[b] NCAR FORTRAN compiler

[c] converted using Table 1

Table 5. Computation times (in sec) generated by equations
(113)-(120) for He $\dotplus$ H$_2$.[a]

N	P	Reference mode	Propagation mode
Calculation and diagonalization of $\underset{\sim}{D}$			
14	14	9.2	0.0
	11	9.2	0.0
	8	9.2	0.0
11	11	5.3	0.0
	8	5.3	0.0
8	8	2.6	0.0
Propagation			
14	14	5.7	4.6
	11	3.5	2.8
	8	1.9	1.4
11	11	3.5	2.8
	8	1.9	1.4
8	8	1.9	1.4
Total			
14	14	15.9	5.4
	11	13.7	?.6
	8	12.1	2.2
11	11	9.8	3.6
	8	7.2	2.2
8	8	5.5	2.2

[a] MNF compiler, Cyber 74 computer

Table 6. Transition probabilities $P_{vv'} = |S_{vv'}|^2$ for the model He-li$_2$ system[a] at $E = 4\hbar\omega$.[b]

description of basis		$v_{max} = 5$	$v_{max} = 3$	adiabatic
N/P		6/6	4/4	6/4
v	v'			
0	1	$1.33(-1)$[c]	$1.30(-1)$	$1.33(-1)$
0	2	$2.04(-3)$	$1.65(-3)$	$2.04(-3)$
0	3	$4.97(-7)$	$4.56(-7)$	$4.9(-7)$
1	2	$5.51(-1)$	$4.61(-2)$	$5.51(-1)$
1	3	$3.09(-5)$	$1.11(-5)$	$3.3(-5)$
2	3	$2.11(-3)$	$7.76(-5)$	$2.2(-3)$

[a] The system is described in references 64 and 65

[b] The results are from reference 29

[c] The number in parentheses denotes the power of ten by which the entry should be multiplied

Table 7. Transition probabilities $R_{vv'} = |S_{vv'}|^2$ for the model
He-H_2 system[a] at $E = 8\hbar\omega$.[b]

description of basis		$v_{max} = 8$	$v_{max} = 7$	adiabatic
N/P		9/9	8/8	9/8
v	v'			
0	6	$1.77(-6)$[c]	$4.03(-6)$	$1.77(-6)$
1	4	$7.33(-2)$	$6.98(-2)$	$7.33(-2)$
1	6	$3.31(-5)$	$5.15(-5)$	$3.32(-5)$
2	5	$2.58(-2)$	$2.26(-2)$	$2.58(-2)$
4	6	$1.90(-2)$	$1.11(-2)$	$1.90(-2)$

[a]The system is described in references 64 and 65

[b]The results are from reference 29

[c]The number in parentheses denotes the power of ten by which
the entry should be multiplied

Table 8. Rotational excitation cross sections $\sigma_{0 \to j}^J$, (a_0^2) for transitions
from the ground rotational state for He-HF at $E = 0.05$ eV.[a]

description of basis	$j_{max} = 4$	$j_{max} = 3$	$j_{max} = 2$	adiabatic
N/P	15/15	10/10	6/6	10/6
j'		$J = 4$		
1	$2.06(-1)$[b]	$2.03(-1)$	$2.36(-1)$	$2.25(-1)$
2	$9.75(-2)$	$9.57(-2)$	$1.62(-1)$	$9.55(-2)$
3	$2.70(-2)$	$2.86(-2)$		
j'		$J = 12$		
1	$4.83(-1)$	$4.82(-1)$	$5.42(-1)$	$4.94(-1)$
2	$1.74(-1)$	$1.75(-1)$	$2.58(-1)$	$1.59(-1)$
3	$3.38(-2)$	$3.56(-2)$		

[a]See references 28 and 29

[b]The number in parentheses denotes the power of ten by which the entry should be multiplied

Table 9. Partial cross sections $\sigma^4_{04 \to 2\ell'}$ (a_0^2) for He–HF for

J = 4 and E = 0.05 eV.[a]

description of basis	$j_{max} = 4$	$j_{max} = 2$	adiabatic
N/P	15/15	6/6	10/6
ℓ'			
2	$2.88(-2)$[b]	$4.82(-2)$	$2.72(-2)$
4	$4.57(-2)$	$4.26(-2)$	$2.49(-2)$
6	$4.30(-2)$	$7.07(-2)$	$4.34(-2)$

[a] See references 28 and 29

[b] The number in parentheses denotes the power of ten by which the entry should be multiplied

REFERENCE POTENTIALS WITH INTEGRAL EQUATIONS

Michael J. Redmon
Chemical Physics Group
BATTELLE
Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201

I.   INTRODUCTION

Due to the rapid accumulation of information concerning the potential surfaces of chemically interesting species, it is desirable to be able to perform coupled-channel calculations within the close-coupling method. Unfortunately, even the most efficient tradiational methods of integrating the coupled equations are so time-consuming that their use for more than a few coupled channels becomes prohibitively expensive. A recent method, developed by Roy Gordon (Go69) and based on a piecewise-analytic approximation to the potential, is a major breakthrough in numerical technology because it requires relatively few steps to integrate the coupled equations. At high energies, it has a real advantage over methods which approximate the wavefunction, such as the Runge-Kutta method or the more recent de Vogelaere (dV55) method, since a proper construction of the wavefunction using these methods requires a large number of steps. Recent methods for solving close-coupled integral equations have been developed by Johnson and Secrest (Jo66) and Sams and Kouri (Sa69). Both of these methods use numerical quadrature procedures, which subdivide integrands into slowly varying partitions and, therefore, must employ a finer integration mesh with increasing energy.

We have developed an integral equations method for solving coupled equations that does not require quadrature procedure (Re74a). Thus, it has the desirable feature (characteristic of the Gordon method) that the integration mesh is relatively insensitive to changes in energy (wavelength). It is based on the Volterra integration equation, as is the Sams and Kouri method (Sa69), but, instead of replacing the integrals by a quadrature, it uses a piecewise-constant reference potential, in the spirit of Roy Gordon's method (Go69). The result is that larger steps can be taken as the equations are integrated into the asympotic region since the integrals for the model potential problem can be evaluated exactly for each interval.

The method has been implemented in a general inelastic scattering code, INSCAT, which has been used to study a variety of atom-diatomic mole-cule (Re73) and atom-atom collisions (Re74b). For a typical atom-rigid rotor interaction the integration takes 60-100 steps to reach the asymptatic region. Accuracy and speed are comparable to the program of Gordon. Execu-tion times increase as N^2 at least up to 25 channels. Stability to growth of closed channels is ensured by using variable pickup points and occasional stabilization.

In part 1 we develop the form of the Volterra integral equation for inelastic scattering used in the method. The use of reference poten-tials is discussed in part 2. In the final section, the code INSCAT is discussed, and comparison is made with results obtained with other methods.

1. The Multi-Channel Volterra Equation

The quantum mechanical coupled equations for scattering can be derived from the differential form of the time-independent Schrödinger equation. We wish to derive an integral equation representation for the wavefunction. We consider the Schrödinger equation

$$(E - H_0 - V)|\vec{k}+\rangle = 0 \tag{1}$$

with the boundary condition

$$\Psi_{\vec{k}}^{+}(r) \xrightarrow[r \to \infty]{} (2\pi)^{-3/2} \exp(i\vec{k} \cdot \vec{r}) + f \frac{\exp(ikr)}{r} . \tag{2}$$

$\Psi_{\vec{k}}^{+}(r)$ is the coordinate representation of the outgoing stationary scattering states $|\vec{k}+\rangle$, and f is the scattering amplitude.

For the stationary states $|\vec{k}+\rangle$, equation (1) can be transformed into an integral equation

$$|\vec{k}+\rangle = |\vec{k}\rangle + (E - H_0 + i\epsilon)^{-1}V|\vec{k}+\rangle \tag{3}$$

where the plane wave $|\vec{k}\rangle$ is a solution to the homogeneous problem

$$(E - H_0)|\vec{k}\rangle = 0. \tag{4}$$

Equation (3) was first derived by Lippmann and Schwinger (Li50) and is often called the Lippmann-Schwinger equation for $|\vec{k}+\rangle$.

In the coordinate representation, (3) can be written

$$\langle \vec{r}|\vec{k}+\rangle = \Psi_{\vec{k}}^{+}(r) = \phi_{\vec{k}}(\vec{r}) + \int d\vec{r}' G(\vec{r}, \vec{r}') V(r') \Psi_{\vec{k}}^{+}(\vec{r}'). \tag{5}$$

It is more convenient, for our purposes, to work in the angular momentum representation. The vectors $|\vec{k}+\rangle$ can be expanded in terms of angular momentum eigenvectors, for which we use the notation $|E,\ell,m+\rangle$ (Ta72). These eigenvectors can be written in the coordinate representation as

$$\langle \vec{r}|E,\ell,m+\rangle = i^\ell \left(\frac{2m}{\pi k}\right)^{1/2} r^{-1} \Psi_{\ell,k}(r) Y_\ell^m(\hat{r}) \tag{6}$$

where the radial functions $\Psi_{\ell,k}(r)$ reduce to spherical Bessel functions $j_\ell(kr)$ when $V \to 0$ and satisfy the radial equation (Ta72)

$$[\frac{d^2}{dr^2} - \frac{\ell(\ell + 1)}{r^2} - \bar{v}(r) + k^2]\Psi_{\ell,k}(r) = 0. \tag{7}$$

The states $|E,\ell,m+\rangle$ satisfy the Lippmann-Schwinger equation

$$|E,\ell,m+\rangle = |E,\ell,m\rangle + (E - H_0 + i\varepsilon)^{-1}V|E,\ell,m+\rangle \tag{8}$$

and, in analogy to (5) an integral equation for the radial wavefunction can be obtained:

$$\Psi_{\ell,k}(r) = j_\ell(kr) + \int_0^\infty dr'G^0_{\ell,k}(r,r')V(r')\Psi_{\ell,k}(r') \tag{9}$$

where the kernel is

$$G^0_{\ell,k} = -\frac{1}{k}j_\ell(kr_<)h^+_\ell(kr_>). \tag{10}$$

h^+_ℓ is a spherical Hankel function. The asymptotic condition is

$$\Psi_{\ell,k}(r) \xrightarrow[r \to \infty]{} j_\ell(kr) + kf_\ell(k)\exp[i(kr - \ell\pi/2)]. \tag{11}$$

Equation (9) can be conveniently rewritten:

$$\Psi_{\ell,k}(r) = j_\ell(kr) + \int_0^r dr'G^0_{\ell,k}(r,r')V(r')\Psi_{\ell,k}(r')$$

$$+ \int_r^\infty dr'G^0_{\ell,k}(r,r')V(r')\Psi_{\ell,k}(r'). \tag{12}$$

Using (10) we can write the integral equation as

$$\Psi_{\ell,k}(r) = j_\ell(kr) - \frac{1}{k}h^+_\ell(kr) \int_0^r dr'j_\ell(kr')V(r')\Psi_{\ell,k}(r')$$

$$+ \frac{1}{k}j_\ell(kr) \int_0^r dr'h^+_\ell(kr')V(r')\Psi_{\ell,k}(r')$$

$$- \frac{1}{k}j_\ell(kr) \int_0^\infty dr'h^+_\ell(kr')V(r')\Psi_{\ell,k}(r'). \tag{13}$$

The last integral is a constant, and, following Sams and Kouri (Sa69),

$$\Psi_{\ell,k}(r) = j_\ell(kr)(1 + C) - \frac{1}{k} h_\ell^+(kr) \int_0^r dr' j_\ell(kr')V(r')\Psi_{\ell,k}(r')$$

$$+ \frac{1}{k} j_\ell(kr) \int_0^r dr' h_\ell^+(kr')V(r')\Psi_{\ell,k}(r') \quad (14)$$

where

$$C = - \frac{1}{k} \int_0^\infty dr' h_\ell^+(kr')V(r')\Psi_{\ell,k}(r'). \quad (15)$$

We note that $(1 + C)$ corresponds to the Jost matrix of multi-channel scattering (Ta72). Since this quantity is a constant, we can renormalize our integral equation by multiplying through by $(1 + C)^{-1}$. This yields the equation [Sams and Kouri (Sa69)]:

$$\Psi_{\ell,k}^0(r) = j_\ell(kr) - \frac{1}{k} h_\ell^+(kr) \int_0^r dr' j_\ell(kr')V(r')\Psi_{\ell,k}^0(r')$$

$$+ \frac{1}{k} j_\ell(kr) \int_0^r dr' h_\ell^+(kr')V(r')\Psi_{\ell,k}^0(r'). \quad (16)$$

This is a Volterra equation of the second kind (Ar68) and has the desirable feature that the limits of integration are finite. This equation has been used as the basis of a computational method by Sams and Kouri (Sa69), who replaced the integrals by quadratures and showed that the equation could be solved by a noniterative procedure, due to the cancellation of the two integrals in the final quadrature terms.

A very important property of Volterra integral equations of the second kind is that such equations are always uniquely solvable, regardless of the magnitude of the kernel (Ta58). This is not true in general for Fredholm integral equations, from which the Volterra

equations may be derived as a special case.[*]

The matrix analog of (16) is

$$\Psi^0(r) = J(r) - H^+(r) \cdot \int_0^r dr' J(r') \cdot V(r') \cdot \Psi^0(r')$$

$$+ J(r) \cdot \int_0^r dr' H^+(r') \cdot V(r') \cdot \Psi^0(r'). \qquad (17)$$

In obtaining (17) we have made use of the commutation of diagonal matrices and have absorbed the wave-vectors into H. The basis used for the matrix representation is suitably chosen for each application. For atom-atom scattering it is convenient to use a set of molecular electronic states.

In the next section we develop a new method, based upon an approximation to $V(r)$, for solving (17)

2. A New Computational Procedure for Solving the Volterra Equation

The Volterra integral equation (16) is very convenient for numerical computation because of the finite limits on the integrals. This means that we can break up the range of integration into a finite number of steps by using a quadrature procedure to represent the integral, as Sams and Kouri (Sa69) have done. An alternative scheme is suggested by the work of Roy Gordon (Go69), who made a piecewise-analytic approximation to the potential in each interval and represented the solution by a linear combination of solutions

[*] Reinhardt and coworkers, in a recent series of papers, have developed a numerical procedure for evaluating the Fredholm determinant, thus leading to a direct matrix solution of the Lippmann-Schwinger equation. Representative references are (Ra70a), (Ra70b), (Ra71), and (He73).

to the model potential problem. We will adopt this procedure and write the single-channel integral equation in the form

$$y(r) = \phi_1(r) - \phi_2(r)\left\{\left[\int_0^{r_1} dr' + \int_{r_1}^{r_2} dr' + \ldots + \int_{r_n}^{r} dr'\right]\phi_1(r')V(r')y(r')\right\}$$

$$+ \phi_1(r)\left\{\left[\int_0^{r_1} dr' + \int_{r_1}^{r_2} dr' + \ldots + \int_{r_n}^{r} dr'\right]\phi_2(r')V(r')y(r')\right\} \quad (18)$$

or

$$y(r) = \phi_1(r) - \phi_2(r)F(r) + \phi_1(r)G(r) \quad (19)$$

where $F(r)$ and $G(r)$ are the net values of the two integrals up to r. The functions ϕ_1 and ϕ_2 are solutions to our homogeneous problem, being sines and cosines if we choose H_o to be the kinetic energy and absorb the centrifugal potential into V. The derivative of y is easily evaluated:

$$\dot{y}(r) = \dot{\phi}_1(r) - \dot{\phi}_2(r)F(r) + \dot{\phi}_1(r)G(r). \quad (20)$$

In each interval (a,b) we make the approximation

$$V(r) = U_0 \qquad (a<r<b) \quad (21)$$

and represent the wavefunction in this interval by the local approximation

$$y(r) = A\chi_1(r) + B\chi_2(r) \qquad (a<r<b) \quad (22)$$

where A and B are linearly independent solutions for the constant potential problem; i.e., for a classically allowed region

$$\chi_1(r) = \sin(\alpha r) \quad (23)$$

$$\chi_2(r) = \cos(\alpha r) \quad (24)$$

and for a forbidden region

$$\chi_1(r) = \exp(-\alpha r) \quad (25)$$

$$\chi_2(r) = \exp(\alpha r). \quad (26)$$

For convenience, we let $V(r)$ be the effective potential

$$V(r) \longrightarrow V(r) + \frac{\ell(\ell+1)}{r^2} \, . \tag{27}$$

Then the functions ϕ_1 and ϕ_2 are solutions to the free particle problem; i.e.,

$$\phi_1(r) = \sin(kr) \tag{28}$$

and

$$\phi_2(r) = \frac{\cos(kr)}{k}. \tag{29}$$

With these choices for the various functions, the integrals may be written for a classical region, as

$$F(b,a) = \int_a^b \sin(kr')U_0[A\sin(\alpha r') + B\cos(\alpha r')]dr' \tag{30}$$

and

$$G(b,a) = \int_a^b \frac{\cos(kr')}{k} U_0[A\sin(\alpha r') + B\cos(\alpha r')]dr'. \tag{31}$$

Similarly, for a forbidden region, we use the notation

$$F(b,a) = I_1(b,a)A + I_2(b,a)B \tag{32}$$

$$G(b,a) = I_3(b,a)A + I_4(b,a)B \tag{33}$$

where, for example,

$$I_1(b,a) = U_0 \int_a^b \sin(kr')\sin(\alpha r')dr'. \tag{34}$$

All of these integrals can be evaluated analytically.

The calculational scheme consists of assigning initial values of y and $\dot{y}$, determining A and B for an interval by inverting the local expressions for y and $\dot{y}$, and then propagating the solution across the interval (a,b) by use of the integral equation. The explicit forms

for A and B, obtained from (22) and an analogous expression for $\dot{y}$, are

$$A = W^{-1}(\chi_2 y - \chi_2 \dot{y}) \tag{35}$$

and

$$B = W^{-1}(\chi_1 \dot{y} - \dot{\chi_1} y) \tag{36}$$

where W is the Wronskian of χ_1 and χ_2. The multi-channel Volterra equation (17) may be written, using our notation, as

$$y(r) = \phi_1(r) \cdot [1 + G(r)] - \phi_2(r) \cdot F(r) \tag{37}$$

where ϕ_1 and ϕ_2 are diagonal matrices. The matrices F and G are

$$F(r) = I_1 \cdot A + I_2 \cdot B + F(a) \tag{38}$$

and

$$G(r) = I_3 \cdot A + I_4 \cdot B + G(a) \tag{39}$$

where, for example,

$$I_1 = \int_a^r \phi_1(r') \cdot U_0(r') \cdot \chi_1(r') dr'. \tag{40}$$

(The other integral matrices I_i are given by similar expressions.) F(a) and G(a) are the accumulated values of F and G up to $r = a$.

If we have one or more closed channels, we have different boundary conditions which ϕ_1 and ϕ_2 must satisfy: ϕ_1 must be zero at the origin, and ϕ_2 must decay exponentially for large r. Clearly, two linearly independent functions satisfying these criteria are a hyperbolic sine and a decaying exponential. A proper closed-channel Green's function is

$$G^c(r_<, r_>) = -\frac{1}{\kappa} \phi_1^c(r_<) \phi_2^c(r_>) = \begin{cases} -\frac{1}{\kappa} \sinh(\kappa r) \exp(-\kappa r') & \text{(for } r > r') \\ & \tag{41} \\ -\frac{1}{\kappa} \sinh(\kappa r') \exp(-\kappa r) & \text{(for } r < r'). \end{cases}$$

With this choice of G^c, the integrals must be re-evaluated, but the

computational procedure is the same as before.

The reason for making a potential approximation and performing all of the integrals is that we should be able to take large steps as we approach the asymptotic region. In order to vary the step-size efficiently, we need a local measure of the error we are introducing at each step. Roy Gordon has given an expression for choosing the step-size:

$$h_{n+1} = h_n \left[\frac{T_0}{T}\right]^{1/3} \tag{42}$$

where T_0 is the desired error, and T is an estimate of the error incurred in a step. The power of 1/3 arises because the first-order correction to the wavefunction for a linear reference potential is cubic in the step-size h (Go71). We have found that this relation, when used for choosing the step-size for our constant reference potential, maintains the error produced in each interval at a value very near the desired error T_0.

The formula requires an estimate (T) of the actual error incurred in each step. Following Gordon (Go69), we assume that A in (35) is a function of r:

$$\frac{dA(r)}{dr} = W^{-1}(\ddot{\chi}_2 y - \chi_2 \ddot{y}). \tag{43}$$

The Wronskian W is constant because the functions χ_1 and χ_2 are local solutions to a second-order differential equation with missing first derivative. Both χ_2 and y satisfy second-order differential equations, so we have

$$\ddot{\chi}_2 = - [k^2 - U_0]\chi_2 \tag{44}$$

and

$$\ddot{y} = - [k^2 - V(r)]y. \tag{45}$$

Using these expressions, (43) can be written

$$\frac{dA(r)}{dr} = - W^{-1}[V(r) - U_0]\chi_2 y. \tag{46}$$

We can obtain the variation of $A(r)$ over an interval (a,b) by integrating

$$A(b) - A(a) = - W^{-1} \int_a^b \chi_2(r')[V(r') - U_0]y(r')dr'. \tag{47}$$

Inserting our local approximation for $y(r)$ and dividing by the step-size h (Go69), we have an estimate of the average variation of $A(r)$ over the interval (a,b):

$$\frac{\Delta A(r)}{h} \approx - (hW)^{-1} \int_a^b \chi_2(r')[V(r') - U_0][A_a\chi_1(r') + B_a\chi_2(r')]dr'. \tag{48}$$

We take for A_a and B_a their values at $r = a$. To evaluate this integral, we expand $V(r)$ in a Taylor series about the midpoint of (a,b) and keep only the first two terms. This gives

$$\frac{\Delta A(\bar{r})}{h} \approx -(hW)^{-1}a_1 \int_a^b \chi_2(r')(r' - \bar{r})[A\chi_1(r') + B\chi_2(r')]dr' \tag{49}$$

where a_1 is the potential slope [See Appendix G.]

$$a_1 = \left(\frac{dV(r)}{dr}\right)_{r = \bar{r}}. \tag{50}$$

We can also obtain an expression for the average variation of $B(r)$ over an interval (a,b):

$$\frac{\Delta B(r)}{h} \approx (hW)^{-1}a_1 \int_a^b \chi_1(r')(r' - \bar{r})[A\chi_1(r') + B\chi_2(r')]dr'. \tag{51}$$

The integrals appearing in (49) and (51) are easily evaluated since χ_1 and χ_2 are transcendental functions. It is worth mentioning that computationally convenient expressions are obtained by performing

the integration over the interval $(-\frac{h}{2}, \frac{h}{2})$ and taking advantage of the parity of the resulting integrand.

Once the integrals in (49) and (51) are evaluated, the first-order correction to the local wavefunction is estimated by writing

$$A(r) = A_a + \left(\frac{\Delta A}{h} \right)(r - a) \tag{52}$$

$$B(r) = B_a + \left(\frac{\Delta B}{h} \right)(r - a) \tag{53}$$

so that, for $r = b$

$$y^1(b) = \Delta A \cdot \chi_1(b) + \Delta B \cdot \chi_2(b). \tag{54}$$

Using this estimate for the correction to y at the right-hand side of an interval, Gordon (Go71) defines the error introduced in the interval (a,b) as

$$T = \frac{\left| y^1(b) \right| + \alpha^{-1} \left| \dot{y}^1(b) \right|}{\left| y^0(b) \right| + \alpha^{-1} \left| \dot{y}^0(b) \right|} . \tag{55}$$

where the zero-order function is given by (22), and α is the local wave number. Since the corrections to A and B vanish at the midpoint of an interval, (55) is a very stable measure of the error introduced in one step.

The step-size selector (42) requires an estimate of the local error introduced in an interval before the next step-size can be determined. This error estimate, T, requires the evaluation of the "perturbation integrals" (49) and (50) . Since we are approximating the potential, these integrals are slowly varying functions of energy, and in practice it is found that one need only evaluate these integrals and determine a set of step-sizes at one energy. It is then possible to use these same step-sizes at other energies.

This desirable feature of embedding with a reference potential has
been discussed by Gordon (Go69). In our implementation, this prac-
tice results in savings of about one-third in computation time in
comparison with runs in which the step-size is computed. A good
procedure to follow is to divide a set of energies into partitions
such that the lowest energy of one overlaps with the highest energy
of another. This makes possible e check of the assumption that the
step-size is independent of energy.

3. A Multichannel Scattering Program

We have implemented the computational procedure developed in Section 2
into a coupled-channel program INSCAT for solving the coupled integral
equations which arise in scattering theory. We will discuss the program
arrangement and give selected results in this section.

The main program, INSCAT, reads in the required input and calls
either INROT or ATMATM to solve the coupled equations for the rota-
tional inelastic problem or atom-atom collision problem. These subroutines
are called once for each value of the total angular momentum. INROT calls
the subroutine F2INIT and F2 to set up the F2-matrix, sets up the potential,
and then calls MIKINT to integrate the coupled equations. ATMATM sets up
the potential for the atom-atom problem and calls MIKINT to perform the
integration.

When the coupled equations have been integrated for a particular
value of J, XSECTN is called from INSCAT to compute the partial
cross-sections. After the equations have been integrated for all
values of J, ANGDIS is called to compute the angular distribution.

All input is in atomic units (Hartrees and Bohrs), but the
printed output lists the cross-sections in both square Bohrs and

square Angstroms, and the energies in Hartrees and electron volts.
Defaults are available for most parameters, and the program is
documented with detailed information concerning the ·choice of
parameters. For most rotational inelastic scattering calculations,
an initial step-size of .05 Bohrs and a value of T = .001 are
satisfactory. For the atom-atom collisions we have studied, we found
that an initial step-size of .025 Bohrs and a value of T = .0005 were
required to obtain acceptable results.

The spherical Bessel functions were calculated by downward
recursion for $J < x$, and by a procedure recommended by Wills (Wi71a)
for $J > X$.

INSCAT is presently dimensioned to allow one to compute the
step-sizes at one energy and use them for calculations at nine other
energies. Partial cross-sections and angular distributions may be
calculated for each energy. Phase-shifts and R-matrices can be
written on a data set for later use.

Several standard potentials are already programmed into the
potential subroutine POT and can be accessed by simply setting a
flag. In addition, a set of matrices defining a set of potentials
may be read in and a spline-fit used to allow interpolation between
the points.

The value of 1/3 used for the exponent in (42) was chosen by
a trial-and-error procedure for single-channel calculations, and
appears to work reasonably well for the problems we have considered.
We have actually plotted the calculated error (55) as a function
of the radial coordinate, and found that it oscillated around the
preselected optimum value.

The potential slope required in obtaining the variations of A and B in (49) and (51) is calculated from the difference in the potential at the midpoint of two successive intervals. This can be a poor approximation if large steps are taken too soon in a radial integration. This is not a very severe restriction since the use of a constant reference potential precludes the use of very large steps before the potentials begin to flatten out.

A normalization factor is used to insure that the solutions start to grow at a sufficiently small rate; for rotational inelastic problems, it was sufficient to set $N = 1$, but this procedure would not be adequate if the channel wave-numbers were very different from one another.

One difficulty which often arises in integrating coupled differential or integral equations (with initial value methods) is the gradual appearance of linear dependencies caused by the exponential growth of solutions in classically forbidden regions. Roy Gordon (Go69) has given an elegant discussion of the problem and has pointed out that the situation arises quite naturally from the finite precision of machine arithmetic and has nothing to do with an inherent instability of the coupled equations. Gordon shows that it is possible to stabilize the solutions by unitary transformation to upper-triangular form, thus insuring that the diagonal elements of the solution matrix have the most rapid growth.

The problematic growth of solutions is most troublesome when highly closed channels are involved. However, we have found that it is often not necessary to stabilize, even for closed channels, if the channel wave numbers do not differ appreciably from one another. When stabilization is required, it is in practice only necessary to transform

the solution matrix periodically to upper-triangular form, and not necessarily by constructing a unitary transformation. Calculations have been performed for rotational-inelastic scattering of an atom colliding with a rigid rotor. The formalism used was that of Arthurs and Dalgarno (Ar60), and low-energy collisions were studied for $(He + H_2)$, $(He + N_2)$, and $(Hg + H_2)$. Our choice of these systems was based on the fact that published data for close-coupling calculations are available, and the potentials involved are sufficiently varied to permit a useful analysis of the method.

The calculations chosen for comparison were the $(He + H_2)$ cross-sections of McGuire and Micha (MG72), the $(Hg + H_2)$ R-matrices of Lester and Bernstein (Le67a, Le67b, and Le68a) and Lester (Le71a), and the extensive calculations of Erlewein et al. (Er68) and von Seggern and Toennies (vS69) for both open and closed channels.

Two other programs were used to generate cross-sections. These were the de Vogelaere differential equation program of Paul McGuire (CLSCPL) and the Volterra integral equation program of Neal Sams, Charles Wells, and Donald Kouri (SPIE). The version of CLSCPL available to us did not have the capability of including closed channels in the basis, but the integral equation program, SPIE, did have this option.

We have made four-channel calculations, with each of the three programs, on the three potentials mentioned above. The related speeds of the three methods depend very much on the potential. INSCAT takes roughly as much time as SPIE for $(He + H_2)$ collisions at 0.2 eV, and CLSCPL is slightly faster than either of the others. This is not surprising since the $(He = H_2)$ potential of Krauss and Mies (Kr65) is

exponential and our constant potential approximation is rather severe. For potentials with wells and long-range interactions, such as $(He + N_2)$ or $(Hg + H_2)$, we are from four to ten times as efficient as the other two methods, depending on the relative energy (both of the other programs are wavelength dependent). We did not carry our calculations above 1 eV; however, at higher energies we would expect to be significantly faster than the other methods. These statements should be taken as a rough comparison of the relative speeds of the three programs.

In Table 1 we present a comparison of our calculated R-matrix for $(He + N_2)$ with one obtained from calculations using the program of Kouri. The symmetry in the R-matrix is an indication of the stability of the numerical procedure, although it does not guarantee accuracy in the calculated cross-sections. In Figure 1 we present the $(He + N_2)$ compound-state resonance at about 1.066×10^{-3} eV, which was reported by von Seggern and Toennies (vS69). The resonance curve was constructed from calculations of the total cross-section, including 25 partial waves, at ten energies. The computations required about one minute of CPU time (on the University of Florida IBM 370/165 computer) to integrate the coupled equations and to determine the total cross-sections and angular distributions (not reported here) for all ten energies. Each integration required about 80 steps to reach the asymptotic region $(60a_0)$. This represents a good closed-channel test of INSCAT since we have only one channel open at this energy. We found that it was not necessary to stabilize to obtain these results.

Two planned improvements in the code are the use of a simple (fast) quadrature procedure to integrate through regions of steep (repulsive) potentials near the origin, and relaxation of the rigid-rotor constraint to allow for vibrational transitions. The present code is very portable, and has been used in IBM, CDC, UNIVAC, and VAX computers.

REFERENCES

(Ar68) G. Arfken, Mathematical Methods for Physicists, Academic Press,
        N.Y. (1968).

(Ar60) A. M. Arthurs and A. Dalgarno, Proc. Roy. Soc. (London) A256,
        540 (1960).

(dV55) R. de Vogelaere, J. Res. N. B. S. 54, 119 (1955).

(Er68) W. Erlewein, M. von Seggern and J. P. Toennies, Z. Physik. 211,
        35 (1968).

(Go69) R. G. Gordon, J. Chem. Phys. 51, 14 (1969).

(Go71) R. G. Gordon, Meth. Comp. Phys. 10, 81 (1971).

(He73) E. J. Heller and W. P. Reinhardt, Phys. Rev. A7, 365 (1973).

(Jo66) B. R. Johnson and D. Secrest, J. Math. Phys. 7, 2187 (1966).

(Le67a) W. A. Lester, Jr., and R. B. Bernstein, Chem. Phys. Lett. 1,
        207 (1967).

(Le67b) W. A. Lester, Jr., and R. B. Bernstein, Chem. Phys. Lett. 1,
        347 (1967).

(Le68a) W. A. Lester, Jr., and R. B. Bernstein, J. Chem. Phys. 48,
        4896 (1968).

(Le71a) W. A. Lester, Jr., J. Chem. Phys. 54, 3171 (1971).

(Le71b) W. A. Lester, Jr., Meth. Comp. Phys. 10, 211 (1971).

(Li50) B. A. Lippmann and J. Schwinger, Phys. Rev. 79, 469 (1950).

(MG72) P. McGuire and D. Micha, Int. J. Quantum. Chem. 6, 111 (1972).

(Re70a) W. P. Reinhardt and A. Szabo, Phys. Rev. A1, 1162 (1970).

(Re70b) W. P. Reinhardt, Phys. Rev. A2, 1767 (1970).

(Re71) W. P. Reinhardt, Phys. Rev. A4, 429 (1971).

(Re73) M. J. Redmon, Dissertation, University of Florida, 1973
        (University Microfilms).

(Re74a) M. J. Redmon and D. A. Mich, Chem. Phys. Lett. 28, 341 (1974).

(Re74b) M. J. Redmon and D. A. Micha, Int. J. Quantum. Chem. Symp. 8,
        253 (1974).

(Sa69)   W. N. Sams and D. J. Kouri, <u>J</u>. <u>Chem</u>. <u>Phys</u>. <u>51</u>, 4809 (1969).

(Ta58)   A. E. Taylor, <u>Introduction</u> <u>to</u> <u>Functional</u> <u>Analysis</u>, John Wiley
         and Sons, N.Y. (1958).

(Ta72)   J. R. Taylor, <u>Scattering</u> <u>Theory</u>, John Wiley and Sons, N.Y. (1972).

(vS69)   M. von Seggern and J. P. Toennies, <u>Z</u>. <u>Physik</u> <u>218</u>, 341 (1969).

(Wi71a) J. G. Wills, <u>J</u>. <u>Comp</u>. <u>Phys</u>. <u>8</u>, 162 (1971).
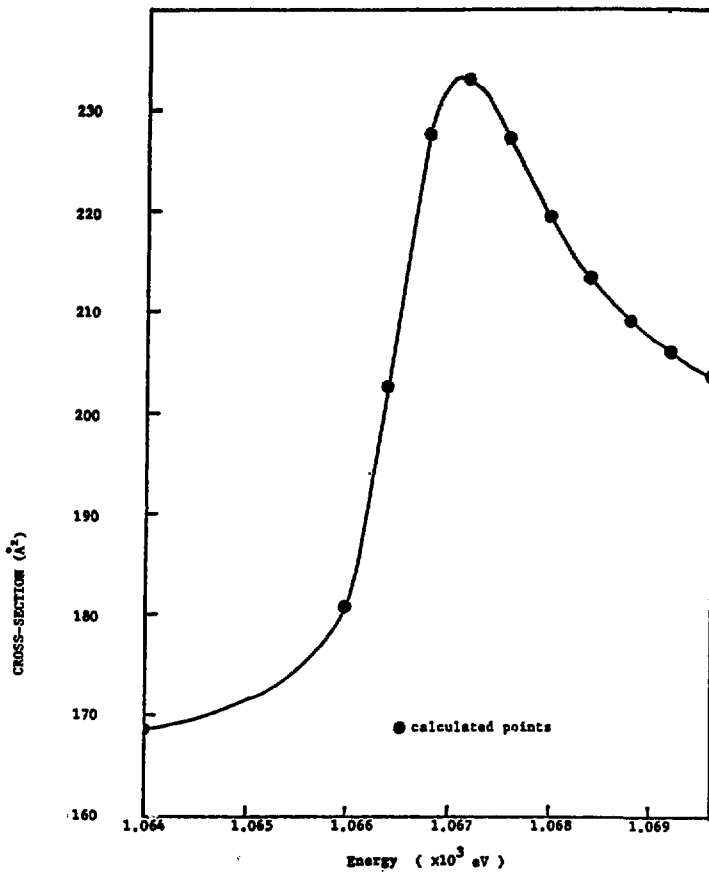
Figure 1. Ne+N$_2$ Compound-State Resonance at 1.066×10^{-3} eV.

Table 1. A Four-Channel R-Matrix

	Col 1	Col 2	Col 3	Col 4
Row 1	1.5296[*]	0.1749	-0.0717	0.1060
	(1.5412)[†]	(0.1773)	(-0.0734)	(0.1073)
Row 2	0.1747	-2.5740	0.0107	0.0304
	(0.1773)	(-2.5423)	(0.0120)	(0.0307)
Row 3	-0.0717	0.0110	0.1350	-0.0040
	(-0.0734)	(0.0120)	(0.1389)	(-0.0040)
Row 4	0.1057	0.0304	-0.0040	0.2230
	(0.1073)	(0.0307)	(-0.0040)	(0.2242)

INSCAT parameters: $He + N_2$ for $J = 4$.

$x0 = 4.72$	$E = 8.085054D-5$	$V1 = 1.08046D-4$
$xf = 67.3$	$Rm = 3.5022$	$V3 = 6.65196$
time = 0.3 sec	$H0 = 0.025$	$V4 = 0.375$
no. steps = 72	$T = 0.001$	$V5 = 0.176$

[*]INSCAT results

[†]SPIE results

SOME THOUGHTS ON THE SOLUTION OF A CLASS
OF LINEAR DIFFERENTIAL EQUATIONS*

Fred T. Krogh
California Institute of Technology
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91103

The problem of interest, which arises from scattering theory, can be written

$$(1) \quad Y'' = \frac{d^2 Y}{dt^2} = A(t)Y, \quad Y(t_0) = Y_0$$

where Y is an $N \times N$ matrix, A is a symmetric $N \times N$ matrix; and $Y'(t_0)$ is to be determined from certain asymptotic conditions. There are two regions of interest with very different characteristics. In the first region $A(t)$ is quite dense and changing rapidly. It also has large positive eigenvalues which leads to the need for periodic reorthogonalization of the solution in order to keep the linear independence required for the final determination of Y'_0. In the second region $A(t)$ is relatively sparse, and slowly varying.

These very different characteristics have led to the development of two classes of methods for the solution of this problem. We believe it likely that no single method will be highly effective in both regions, and thus, that a general purpose method for solving (1) should be a hybrid consisting of a method for the first region, a method for the second region, and some mechanism for deciding when to switch. We believe that this view is generally held by people working in this area, although as yet no such hybrid code has been written.

In this note, we consider only methods for the first region.

Methods for the First Region

In the first region, there are a large number of plausible methods which correspond to methods in use for general nonlinear equations. Two points are worth mentioning. Most important is that methods designed specifically for the solution of second order equations will perform significantly better than related methods applied to the equivalent first order system. The other point is that because (1) is a matrix equation and because $A(t)$ is symmetric, the cost of using implicit methods is significantly less than is generally the case. This second point may not be as significant as might appear at first glance. At least for multistep methods, the primary advantage of implicit methods is their superior stability, and it happens that explicit methods for second order systems without first derivatives are significantly more stable that explicit methods for the equivalent first order system.

Recent studies [1], [2] have found that when evaluation of derivatives is expensive, variable order Adams codes are best; when evaluation of derivatives is cheap, Runge-Kutta methods are best at low accuracy, and extrapolation methods at high accuracy. Thus for large N (experiment is necessary to define large), we believe the method for second order equations

derived in [3] will be most effective. Although this type of method is usually used in PECE mode (Predict, Evaluate derivatives, Correct, Evaluate derivatives), for the problem at hand we believe one should either use the PEC mode or the implicit mode. Using the notation of [3]; except with capital letters for matrix variables, the PEC mode is defined by

$$
(2) \quad P_{n+1} = Y_n + hY'_n + h^2 \sum_{i \equiv 0}^{q-1} g_{i,2} \phi^*_i(n)
$$

$$
P'_{n+1} = Y'_n + h \sum_{i=0}^{q-1} g_{i,1} \phi^*_i(n)
$$

$$
Y_{n+1} = P_{n+1} + h^2 g_{q,1} [A(t_{n+1}) P_{n+1} - \phi_0^{(e)}(n+1)]
$$

$$
(3) \quad Y'_{n+1} = P'_{n+1} + h g_{q,1} [A(t_{n+1}) P_{n+1} - \phi_0^{(e)}(n+1)]
$$

The only N^3 process is the matrix multiply, $A(t_{n+1}) P_{n+1}$. The most efficient implementation of the implicit mode replaces eqs. (3) with

$$
(4) \quad [I - h^2 g_{q,2} A(t_{n+1})] Y_{n+1} = P_{n+1} - h^2 g_{q,2} \phi_0^{(e)}(n+1)
$$

$$
Y'_{n+1} = P'_{n+1} + \frac{g_{q,1}}{h g_{q,2}} (Y_{n+1} - P_{n+1})
$$

where the only N^3 process is the solution of the top equation in (4) for Y_{n+1}. If the full generality in [3] is used, $g_{i,j}$ is a matrix, but multiplication of a matrix by $g_{i,j}$ involves multiplying corresponding matrix elements, not a full matrix multiply. If integration orders, q, are all the same or just the same over individual columns, then a slight reduction in arithmetic operations is obtained by multiplying the top equation in (4) by $1 / h^2 g_{q,2}$.

Closely related to the implicit method given above, is Cowell's method, [4, p.292]

$$(5) \quad Y_{n+1} = 2Y_n - Y_{n-1} + h^2 \sum_{i=0}^{q} \sigma_i^* \nabla^i (A(t_{n+1}) Y_{n+1})$$

$$= 2Y_n - Y_{n-1} + h^2 \{A(t_{n+1}) Y_{n+1} + \sum_{i=1}^{q} \sigma_i^* [A(t_{n+1}) Y_{n+1} - \sum_{j=0}^{i-1} \nabla^j (A(t_n) Y_n)]\}$$

For q=2 and 3 this is the Numerov formula

$$(6) \quad (I - \frac{1}{12} h^2 A(t_{n+1}) Y_{n+1} = 2Y_n - Y_{n-1} + \frac{1}{12} h^2 [10A(t_n) Y_n + A(t_{n-1}) Y_{n-1}]$$

which may be an attractive choice when integration overhead dominates computing time in the above method. A reduction in operation count is obtained with the substitution

$$(7) \quad Z_n = (I - \frac{1}{12} h^2 A(t_n)) Y_n$$

Using the identity $(I + 5A) (I - A)^{-1} = 5I + 6 (I - A)^{-1}$ it is easy to obtain from eqs. (6) and (7)

$$(8) \quad Z_{n+1} - 2 Z_n + Z_{n-1} = -12 [I - (I - \frac{1}{12} h^2 A(t_n))^{-1}] Z_n$$

It is well known, see e.g. [4,p.327],that a double summation of eq.(8) gives an algorithm with better round-off characteristics. Thus

$$(9) \quad Z_{n+1} = -12 \nabla^{-2} \{[I - (I - \frac{1}{12} h^2 A(t_n))^{-1}] Z_n\}$$

The only N^3 operation required by this algorithm is the formation of

$$(10) \quad (I - \frac{1}{12} h^2 A(t_n))^{-1} Z_n \equiv Y_n$$

and thus computationally this algorithm is quite efficient. Unfortunately there may be excessive cancellation in (9) when subtracting (10) from Z_n. If this should be the case, one may prefer to replace (9) with the equivalent.

$$(11) \quad Z_{n+1} = h^2 \nabla^{-2} \{ A(t_n)[I - \frac{1}{12} h^2 A(t_n)]^{-1} Z_n \}$$

which requires nearly twice as much work (13/6 vs. 7/6 N^3) for large N. Changing stepsize when using the summed form of these formulas requires adjustments to the first and second sums. Details for the cases of halving and doubling the stepsize can be found in [5, pp. 957-958].

Other possibilities for the case when integration overhead dominates the time for the variable order Adams methods are Runge-Kutta-Nystrom methods developed by Fehlberg, [6], and Horn, [7], and extrapolation methods based on the Numerov formula.

References

1. Shampine, L.F., Watts, H. A., and Davenport, S. M., "Solving Nonstiff Ordinary Differential Equations - The State of the Art," SIAM Review 18, (1976), pp. 376-411.

2. Hull, T. E., Enright, W. H., Fellen, B. M., and Sedgwick, A. E., "Comparing Numerical Methods for Ordinary Differential Equations," SIAM J. Numer. Anal. 9 (1972), pp. 603-637.

3. Krogh, Fred T., "Changing Stepsize in the Integration of Differential Equations Using Modified Divided Differences," Lecture Notes in Mathematics, Volume 362 (Proceedings of the Conference on the Numerical Solution of Ordinary Differential Equations, 19, 20 October 1972, the University of Texas at Austin), Springer-Verlag, 1974, pp. 22-71.

4. Henrici, Peter, "Discrete Variable Methods in Ordinary Differential Equations," John Wiley and Sons, 1962.

5. Krogh, Fred T., "Algorithms for Changing the Step Size," SIAM J. Numerical Analysis 10, (1973), pp. 949-965.

6. Fehlberg, E., "Classical Eighth-and Lower-Order Runge-Kutta-Nystrom Formulas with a New Step Size Control Procedure for Special Second-Order Differential Equations," NASA TR R-410, June, 1973.

7. Horn, Mary Kathleen, "Developments in High-Order Runge-Kutta-Nystrom Formulas," Texas Institute for Computational Mechanics Report 77-11, The University of Texas at Austin, Texas, October 1977.

CHANGING STEPSIZE IN THE INTEGRATION
OF DIFFERENTIAL EQUATIONS USING
MODIFIED DIVIDED DIFFERENCES*

Fred T. Krogh

California Institute of Technology
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91103

Abstract

Multistep methods for solving differential equations based on numerical integration formulas or numerical differentiation formulas (for stiff equations) require special provision for changing the stepsize. New algorithms are given which make the use of modified divided differences an attractive way to carry out the change in stepsize for such methods. Error estimation and some of the important factors in stepsize selection and the selection of integration order are also considered.

1. Introduction

We have considered a number of methods for changing the stepsize of multistep methods in [1], and the use of modified divided differences in particular in [2]. The algorithm proposed here is different than that given in [2] in that the stepsize is not changed on every step, and the procedure for computing integration coefficients is designed to take advantage of this situation. Even when the stepsize is changed on every step, the new algorithm computes the required coefficients more efficiently than the algorithms in [2] - [7].

However, if there are more than approximately 2 first order stiff equations and the stepsize is changed on every step, then the overall algorithm is slightly less efficient than a carefully organized Lagrangian formulation. (See [7] for example) Reference [7], which gives the only algorithm of the type considered here for stiff equations, reactivated our interest in this area by giving an algorithm with a computational cost that only goes up linearly with the integration order, as opposed to quadratically for methods based on numerical integration formulas.

The use of a completely variable stepsize gives more flexibility than what we propose here, and other methods considered in [1] require less computation. We believe the new method is a good compromise between the conflicting goals of flexibility, computational economy, and stability and reliability.

The following section gives algorithms for computing integration, interpolation, and differentiation coefficients in a framework useful for the step-by-step integration of ordinary differential equations of arbitrary order. The interpolation and differentiation formulas are useful for the case of stiff equations.

In section 3, details connected with the implementation of these algorithms in a program for integrating differential equations are considered. Computational details associated with implementing the algorithms in an efficient way are considered in section 4.

Section 5 considers the problem of interpolating to points which do not coincide with the end of an integration step. The paper concludes with a discussion of some details such as selection of integration order and stepsize, which are difficult to make rigorous statements about, but which are very important in making an integration program efficient.

2. Algorithms for Computing Coefficients and Updating Differences

Let $w(t)$ be a function given at discrete points, t_i, with $t_{i+1} > t_i$ for all i, and consider the polynomial interpolating w at the points t_n, $t_{n-1}, \cdots, t_{n-q+1}$ given by the Newton divided difference formula

$$P_{q-1,n}(t) = w[t_n] + (t-t_n)w[t_n, t_{n-1}] + \cdots$$

$$+ (t-t_n)(t-t_{n-1})\cdots(t-t_{n-q+2})w[t_n, t_{n-1}, \ldots, t_{n-q+1}]$$

(2.1)

where

$$w[t_n, t_{n-1}, \ldots, t_{n-i}] = \begin{cases} w(t_n) & i=0 \\ \dfrac{w[t_n, \ldots, t_{n-i+1}] - w[t_{n-1}, \ldots, t_{n-i}]}{t_n - t_{n-i}} & i=1, 2, \ldots \end{cases}$$

(2.2)

Also consider the polynomial

$$P_{q,n}^{*}(t) = P_{q-1,n}(t) + (t-t_n)(t-t_{n-1})\cdots(t-t_{n-q+1})w[t_{n+1}, t_n, \ldots, t_{n-q+1}]$$

(2.3)

which interpolates w at $t_{n+1}, t_n, \ldots, t_{n-q+1}$. Although apparently $P_{q,n}^{*}$ $= P_{q,n+1}$, later P is associated with a predictor formula and P^{*} with a corrector formula, and the divided difference $w[t_{n+1}, \ldots, t_{n-q+1}]$ in equation (2.3) is computed using a predicted value of $w(t_{n+1})$.

We are interested in the problem of efficiently integrating, interpolating, and differentiating the polynomials $P_{q-1,n}$ and $P_{q,n}^{*}$ at $t = t_{n+1}$ and then obtaining the polynomials $P_{q-1,n+1}$ and $P_{q,n+1}^{*}$.

As in [2] we make a change of variable and introduce some additional notation. Let

$$h_i = t_i - t_{i-1}$$

$$\tau = (t-t_n)/h_{n+1}$$

$$\xi_i(n+1) = h_{n+1} + h_n + \cdots + h_{n+1-i}$$

$$\alpha_i(n+1) = h_{n+1}/\xi_i(n+1)$$

$$\beta_0(n+1) = 1$$

$$\beta_i(n+1) = [\xi_0(n+1)\xi_1(n+1)\cdots\xi_{i-1}(n+1)]/[\xi_0(n)\cdots\xi_{i-1}(n)] \qquad (2.4)$$

$$\varphi_0(n) = w(t_n)$$

$$\varphi_i(n) = \xi_0(n)\xi_1(n)\cdots\xi_{i-1}(n)w[t_n, t_{n-1}, \ldots, t_{n-i}]$$

$$H_k(n+1) = h_{n+1}/k, \quad H_0(n+1) = h_{n+1}$$

$$s_k(n+1) = \begin{cases} H_0(n+1)H_1(n+1)\cdots H_{k-1}(n+1) & k > 0 \\ 1 & k = 0 \\ [H_1(n+1)H_2(n+1)\cdots H_{-k}(n+1)]^{-1} & k < 0 \end{cases}$$

The (n) or (n+1) following a variable name may be dropped if the value to be assigned to n is obvious. From equations (2.2) and (2.4) it is easy to obtain

$$\eta_{i+1}(n+1) = \varphi_i(n+1) - \beta_i(n+1)\varphi_i(n) \qquad (2.5)$$

from which it is apparent that if $h_{n+1} = h_n = \cdots = h_{n+2-i}$ then the modified divided difference $\varphi_i(n+1)$ is equivalent to the i-th backward difference of w at $t = t_{n+1}$. Later we shall find it useful to estimate $\varphi_i(n+1)$ from the values of $\varphi_i(n)$. These estimates are obtained using equation (2.5) and estimating that $\varphi_q = 0$. Thus

$$\varphi_q^{(e)}(n+1) = 0$$

$$\varphi_i^{(e)}(n+1) = \varphi_{i+1}^{(e)}(n+1) + \beta_i(n+1)\varphi_i(n), \quad i = q-1, \ q-2, \ldots, 0. \qquad (2.6)$$

Clearly

$$\varphi_i(n+1) = \varphi_i^{(e)}(n+1) + [w(t_{n+1}) - \varphi_0^{(e)}(n+1)], \quad i=0, 1, \ldots, q. \tag{2.7}$$

The procedure used to update the modified divided differences from t_n to t_{n+1} is given by equations (2.6) and (2.7). The expression $\beta_i(n+1)\varphi_i(n)$ which is required in equation (2.6) is also used in the formulas for integration, interpolation and differentiation, since its use permits a more economical evaluation of the required coefficients. Thus we introduce

$$\varphi_i^*(n) = \beta_i(n+1)\varphi_i(n) \tag{2.8}$$

From equations (2.1), (2.3), (2.4), (2.6), (2.7), and (2.8), simple substitution gives

$$P_{q-1, n}(t) = P_{q-1, n}(t_n + h_{n+1}\tau) = \sum_{i=0}^{q-1} c_{i, n}(\tau)\varphi_i^*(n) \tag{2.9}$$

and

$$P_{q, n}^*(t) = P_{q-1, n}(t) + c_{q, n}(\tau)[w(t_{n+1}) - \varphi_0^{(e)}(n+1)] \tag{2.10}$$

where

$$c_{i, n}(\tau) = \begin{cases} 1 & i = 0 \\ \left(\dfrac{h_{n+1}\tau}{\xi_0(n+1)}\right) \left(\dfrac{h_{n+1}\tau + \xi_0(n)}{\xi_1(n+1)}\right) \cdots \left(\dfrac{h_{n+1}\tau + \xi_{i-2}(n)}{\xi_{i-1}(n+1)}\right) & i \geq 1 \end{cases} \tag{2.11}$$

$$c_{i,\,n}(\tau) = \begin{cases} 1 & i = 0 \\ \alpha_0(n+1)\tau = \tau & i = 1 \\ [\alpha_{i-1}(n+1)\tau + (\xi_{i-2}(n)/\xi_{i-1}(n+1))]c_{i-1,\,n} & i \ge 2 \end{cases} \qquad (2.12)$$

To treat the case of extrapolating to $t = t_{n+1}$, set $\tau = 1$ and observe that

$$\alpha_{i-1}(n+1) + (\xi_{i-2}(n)/\xi_{i-1}(n+1)) = 1 \quad (i \ge 2) \qquad (2.13)$$

and hence $c_{i,\,n}(1) = 1$.

The formula for the k-th derivative of $\mathbb{P}$ with respect to t is obtained by computing the k-th derivative of $c_{i,\,n}(\tau)$ with respect to τ. Thus from equation (2.12)

$$c'_{i,\,n}(\tau) = \alpha_{i-1}(n+1)c_{i-1,\,n}(\tau) + [\alpha_{i-1}(n+1)\tau + \frac{\xi_{i-2}(n)}{\xi_{i-1}(n+1)}]c'_{i-1,\,n}(\tau)$$
$$\vdots$$
$$c^{(k)}_{i,\,n}(\tau) = k\alpha_{i-1}(n+1)c^{(k-1)}_{i-1,\,n}(\tau) + [\alpha_{i-1}(n+1)\tau + \frac{\xi_{i-2}(n)}{\xi_{i-1}(n+1)}]c^{(k)}_{i-1,\,n}(\tau)$$

Setting $\tau = 1$ and using equation (2.13)

$$c^{(k)}_{i,\,n}(1) = k\alpha_{i-1}(n+1)c^{(k-1)}_{i-1,\,n}(1) + c^{(k)}_{i-1,\,n}(1) \qquad (2.14)$$

Define

$$d_{i,\,k} = \frac{1}{k!} c^{(k)}_{i,\,n}(1) \qquad (2.15)$$

Then from the above $(d_{i,\,k} = 0$ for $i < k)$

$$
d_{i,k} = \begin{cases} 1 & k=0, i=0, 1, \ldots, q \\ \alpha_{i-1}(n+1)+d_{i-1,1} & k=1, i=1, 2, \ldots, q \\ \alpha_{i-1}(n+1)d_{i-1,k-1}+d_{i-1,k} & k=2, 3, \ldots; i=k, k+1, \ldots, q \end{cases} \tag{2.16}
$$

From equations (2.4), (2.9), (2.10), and (2.15) there follows

$$
\left[\frac{d^k}{dt^k} P_{q-1,n}(t)\right]_{t=t_{n+1}} = s_{-k}(n+1) \sum_{i=k}^{q-1} d_{i,k} \varphi_i^*(n) \tag{2.17}
$$

$$
\left[\frac{d^k}{dt^k} P_{q,n}^*(t)\right]_{t=t_{n+1}} = \left[\frac{d^k}{dt^k} P_{q-1,n}(t)\right]_{t=t_{n+1}} \\
+ s_{-k}(n+1)d_{q,k}[w(t_{n+1})-\varphi_0^{(e)}(n+1)] \tag{2.18}
$$

where the $d_{i,k}$ can be determined as indicated in equation (2.16).

For the case of integration we proceed much as we did for differentiation, except that now integration by parts is used. Let $c_{i,n}^{(-k)}(\tau)$ denote the k-fold integral $\int_0^\tau \int_0^\tau \cdots \int_0^\tau c_{i,n}(\tau)d\tau$. Then (starting with equation (2.12))

$$
c_{i,n}^{(-1)}(\tau) = (\alpha_{i-1}(n+1)\tau + \frac{\xi_{i-2}(n)}{\xi_{i-1}(n+1)})c_{i-1,n}^{(-1)}(\tau)-\alpha_{i-1}(n+1)c_{i-1,n}^{(-2)}(\tau)
$$
$$
\vdots
$$
$$
c_{i,n}^{(-k)}(\tau) = (\alpha_{i-1}(n+1)\tau + \frac{\xi_{i-2}(n)}{\xi_{i-1}(n+1)}) c_{i-1,n}^{(-k)}(\tau)-k\alpha_{i-1}(n+1)c_{i-1,n}^{(-k-1)}(\tau)
$$

$$c_{i,n}^{(-k)}(1) = c_{i-1,n}^{(-k)}(1) - k\,s_{i-1}(n+1)c_{i-1,n}^{(-k-1)}(1)$$

Define

$$g_{i,k} = (k-1)!\,c_{i,n}^{(-k)}(1) \tag{2.19}$$

From equation (2.11) and the above

$$g_{i,k} = \begin{cases} 1/k & i = 0 \\[1mm] 1/[k\cdot(k+1)] & i = 1 \\[1mm] g_{i-1,k} - \sigma_{i-1}(n+1)g_{i-1,k+1} & i = 2, 3, \ldots, q;\ k = q+d-i, \ldots, 1 \end{cases} \tag{2.20}$$

where d is the largest value of k for which one desires $g_{q,k}$. From equations (2.4), (2.9), (2.10), (2.19), and the definition of $c_{i,n}^{(-k)}$ it follows that the k-fold integrals of P and P^* are given by

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^{t} \cdots \int_{t_n}^{t} P_{q-1,n}(t)dt = s_k(n+1) \sum_{i=0}^{q-1} g_{i,k}\varphi_i^*(n) \tag{2.21}$$

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^{t} \cdots \int_{t_n}^{t} P_{q,n}^*(t)dt = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} \cdots \int_{t_n}^{t} P_{q-1,n}(t)dt$$
$$+ s_k(n+1)g_{q,k}[w(t_{n+1}) - \varphi_0^{(e)}(n+1)] \tag{2.22}$$

where the $g_{i,k}$ can be determined as indicated in equation (2.20).

3. Implementation of the Algorithms

Consider the single d-th order differential equation

$$Y^{(d)} = f(t, Y, Y', \ldots, Y^{(d-1)}), \quad Y^{(k)}(t_0) = Y_0^{(k)}, \quad k=0, 1, \ldots, d-1 \tag{3.1}$$

and let

$$y_n = y(t_n) = \text{computed approximation to } Y(t_n)$$

$$p_n = \text{predicted value of } y_n \tag{3.2}$$

Except where noted to the contrary, the extension to systems of differential equations of what is done in this section is a simple matter of applying what is done for the single equation to each equation in the system. Permitting different values for d and/or q (the integration order) for different equations in the system can also be done in an obvious way.

The formulas of the previous section permit a numerical solution of equation (3.1) with d+1 different choices of w. Let

$$w = y^{(d-j)} \tag{3.3}$$

where j takes one of the values $0, 1, \ldots, d$. The smaller the value of j, the more accurate the formulas for solving equation (3.1). But for some problems (stiff equations), too small a value for j will unduly restrict the stepsize because of stability problems. If $j > 0$, then in the general case an approximate solution to a system of nonlinear equations must be obtained on every step. Thus the case j = 0 is to be preferred if the stepsize is not thereby unduly restricted. This choice gives the well known Adams-Bashforth-Moulton method when d = 1 and the stepsize is constant, and to methods which we shall simply call Adams methods for all d. The case j = 1, d = 1 gives a class of formulas first

suggested by Curtiss and Hirschfelder [8] and recently popularized by Gear [9], [10]. If the $Y^{(d)}$ in equation (3.1) is replaced by 0, then equation (3.1) is an algebraic equation if $d = 1$, and an implicit differential equation of order d-1 for $d > 1$. Including such equations, which we shall call implicit for all d, is a simple matter although one is restricted to $j \geq 1$ for this case.

Although equation (3.1) could be broken up into d first order equations, thereby simplifying much of what follows, the extra complexity of dealing with a d-th order equation directly is well worth while in many cases. The most obvious advantage is that only one set of differences is required, thus saving on both storage and the computation of differences. In the case of the Adams methods we have found that integrating 2-nd order equations directly sometimes permits a larger stepsize or enables one to integrate efficiently with a PEC (Predict-Evaluate derivatives-Correct) method instead of the more usual PECE method. The best choice depends on both the differential equations and the initial conditions. For example, in terms of derivative evaluations to obtain a given accuracy on $x'' = -x/r$, $y'' = -y/r$, $r = (x^2+y^2)^{1/2}$ (a simple 2-body problem), if the motion is circular $(PECE)_2$ is approximately twice as efficient as either $(PECE)_1$ or $(PEC)_2$, where the subscript 2 indicates the integration of the above two 2-nd order equations, and 1 the integration of the equivalent first order equations $x' = u$, $u' = -x/r^3$, $y' = v$, $v' = -y/r^3$. On the other hand if the motion is elliptic with eccentricity .6, then $(PEC)_2$ is approximately twice as efficient as $(PECE)_2$ or $(PECE)_1$. Examples can also be given where reduction to a system of first order equations is best; see e.g. [2] or [11]. The advantages or disadvantages of integrating d-th order stiff equations directly is as far as we know an open question. One advantage as we shall show below is that the direct integration permits a reduction in the effective number of equations which must be solved at each step. The reader who is not interested in the general case may find it advantageous to substitute specific values for d and j in the following text. If this is done, any equation with a final index less than the starting index should be deleted.

For predictor formulas, equations (3.3), (2.21), (2.17), (2.9), and (2.4) give

$$
\left.
\begin{aligned}
p_{n+1}^{(d-j-1)} &= y_n^{(d-j-1)} + h \sum_{i=0}^{q-1} s_{i,1} \varphi_i^*(n), \\
p_{n+1}^{(d-j-k)} &= y_n^{(d-j-k)} + H_1[y_n^{(d-j-k+1)} + H_2[y_n^{(d-j-k+2)} + \dots \\
&\quad + H_{k-1}[y^{(d-j-1)} + h \sum_{i=0}^{q-1} s_{i,k} \varphi_i^*(n)]] \dots], \quad k=2,3,\dots,d-j
\end{aligned}
\right\}
\tag{3.4}
$$

$$
p_{n+1}^{(d-j+k)} = s_{-k} \sum_{i=k}^{q-1} d_{i,k} \varphi_i^*(n), \quad k=1,2,\dots,j, \tag{3.5}
$$

and, of course,

$$
p_{n+1}^{(d-j)} = \varphi_0^{(e)}(n+1) = \sum_{i=0}^{q-1} \varphi_i^*(n) \tag{3.6}
$$

For the case $j = 0$, the corrector formulas are

$$
y_{n+1}^{(d-k)} = p_{n+1}^{(d-k)} + s_k s_{q,k}[f(t_{n+1}, p_{n+1}p_{n+1}', \dots, p_{n+1}^{(d-1)}) - \varphi_0^{(e)}(n+1)], \tag{3.7}
$$
$$
k = 1, 2, \dots, d.
$$

$$
y_{n+1}^{(d)} = f(t_{n+1}, y_{n+1}, y_{n+1}', \dots, y_{n+1}^{(d-1)}) \tag{3.8}
$$

The case $j > 0$ requires the (approximate) solution of the following system of equations.

$$y_{n+1}^{(d-j-k)} = p_{n+1}^{(d-j-k)} + s_k s_{q,k}[y_{n+1}^{(d-j)} - \varphi_0^{(e)}(n+1)], \quad k=1,2,\ldots,d-j \tag{3.9}$$

$$y_{n+1}^{(d-j+k)} = p_{n+1}^{(d-j+k)} + s_{-k} d_{q,k}[y_{n+1}^{(d-j)} - \varphi_0^{(e)}(n+1)], \quad k=1,2,\ldots,j-1 \tag{3.10}$$

$$f(t_{n+1}, y_{n+1}, y'_{n+1}, \ldots, y_{n+1}^{(d-1)}) = p_{n+1}^{(d)} + s_{-j} d_{q,j}[y_{n+1}^{(d-j)} - \varphi_0^{(e)}(n+1)] \tag{3.11}$$

The expression

$$e = [y_{n+1}^{(d-j)} - \varphi_0^{(e)}(n+1)] = \varphi_q(n+1) \tag{3.12}$$

which appears in equations (3.9) and (3.10) is also required for updating the difference table. For the control of round-off error <u>it is essential that e be solved for directly</u>, rather than computing e from equation (3.12) after obtaining $y_{n+1}^{(d-j)}$. Substituting equations (3.9) and (3.10) into equation (3.11), the problem of solving equations (3.9)-(3.11) for e is reduced to

$$f(t_{n+1}, p_{n+1} + s_{-j} s_{q,d-j} e, \ldots, p_{n+1}^{(d-j-1)} + s_1 s_{q,1} e, p_{n+1}^{(d-j)} + e, p_{n+1}^{(d-j+1)}$$

$$+ s_{-1} d_{q,1} e, \ldots, p_{n+1}^{(d-1)} + s_{-j+1} d_{q,j-1} e) = p_{n+1}^{(d)} + s_{-j} d_{q,j} e \tag{3.13}$$

Of course, in the general case equation (3.13) is a system of equations with order equal to the total number of differential equations in the system. There are many ways that one might obtain an approximate solution to this system; see e.g.[12]. For general purpose use a two iteration constant slope Newton method appears to be a good choice; see [13]. The (1×1) matrix required to apply the Newton method to equation (3.13) is given by

$$A = \left\{ (\frac{\partial f}{\partial y})s_{d-j}\beta_{q,\,d-j} + \cdots + (\frac{\partial f}{\partial y^{(d-j-1)}})s_j\beta_{q,\,1} + (\frac{\partial f}{\partial y^{(d-j)}}) \right.$$

$$\left. + (\frac{\partial f}{\partial y^{(d-j+1)}})s_{-1}d_{q,\,1} + \cdots + (\frac{\partial f}{\partial y^{(d-1)}})s_{-j+1}d_{q,\,j-1} \right\} - s_{-j}d_{q,\,j} \tag{3.14}$$

Although it is a little awkward notationally, it is completely straightforward to extend A to the case of a system of equations, even if different values of d, j, and/or q are used for different equations. The same matrix can and should be used over several steps. This is not important for a single equation, but for large systems a significant gain in efficiency results from using the same factorization of A over as many steps as possible. Of course, the partial derivatives of f need not be computed on any step that A is left unchanged.

The two stage iteration process proceeds as follows. Solve

$$Ae_1 = p_{n+1}^{(d)} - f(t_{n+1}, p_{n+1}, p'_{n+1}, \ldots, p_{n+1}^{(d-1)}) \tag{3.15}$$

for e_1. Compute $c_{n+1}, c'_{n+1}, \ldots, c_{n+1}^{(d-1)}$ using equations (3.9) and (3.10) with y_{n+1} replaced by c_{n+1} and $[y_{n+1}^{(d-j)} - \varphi_0^{(e)}(n+1)]$ replaced by e_1. $(c_{n+1}^{(d-j)} = p_{n+1}^{(d-j)} + e_1)$ Then solve

$$Ae_2 = p_{n+1}^{(d)} + s_{-j}d_{q,\,j}e_1 - f(t_{n+1}, c_{n+1}, c'_{n+1}, \ldots, c_{n+1}^{(d-1)}) \tag{3.16}$$

for e_2. Substituting c_{n+1} for p_{n+1} and e_2 for $[y_{n+1}^{(d-1)} - \varphi_0^{(e)}(n+1)]$ in equations (3.9) and (3.10) the final values of $y_{n+1}, \ldots, y_{n+1}^{(d-1)}$ are obtained. The difference table is updated using $e = e_1 + e_2$, and the ratio $\|e_2\|/\|e_1\|$ gives an

indication of the convergence of the iteration and thus is useful in deciding how frequently a new matrix A should be introduced for the iteration. Note that two evaluations of f are required, the same as is required in equations (3.7) and (3.8). For implicit equations, simply replace $p_{n+1}^{(d)}$ and $s_{-j}d_{q,j}$ with 0 in equations (3.11), (3.13)-(3.16).

For the purpose of obtaining error estimates, observe that equations (2.6) and (2.8) give

$$\sum_{i=0}^{q-1} s_{i,k}\phi_i^*(n) = \sum_{i=0}^{q-1} s_{i,k}[\phi_i^{(e)}(n+1)-\phi_{i+1}^{(e)}(n+1)] = \sum_{i=0}^{q-1} s_{i,k}^*\phi_i^{(e)}(n+1) \qquad (3.17)$$

where

$$s_{i,k}^* = \begin{cases} s_{i,k} & i = 0 \\ s_{i,k}-s_{i-1,k} & i > 0 \end{cases} \qquad (3.18)$$

Since $s_{q,k} = s_{0,k}^* + s_{1,k}^* + \cdots + s_{q,k}^*$, equations (2.7) and (3.17) permit us to write

$$\sum_{i=0}^{q-1} s_{i,k}\phi_i^* + s_{q,k}[w(t_{n+1})-\phi_0^{(e)}(n+1)] = \sum_{i=0}^{q} s_{i,k}^*\phi_i(n+1) \qquad (3.19)$$

Clearly the same type formulas can be obtained with $d_{i,k}$ substituted for $s_{i,k}$. and

$$d_{i,k}^{*} = \begin{cases} d_{i,k} & i = 0 \\ d_{i,k} - d_{i-1,k} & i > 0 \end{cases} \qquad (3.20)$$

It is also clear that replacing $g_{q,k}$ in equation (3.19) with $g_{q-1,k}$ will simply change the upper limit of the sum on the right side of equation (3.19) from q to q-1. The corrector formulas as given in equations (3.7), (3.9)-(3.11) have an order one greater than the predictor formulas in equations (3.4) and (3.5). If $g_{q,k}, g_{q,k}, d_{q,k}$ and $d_{q,j}$ in equations (3.7), (3.9), (3.10), and (3.11) respectively were replaced by $g_{q-1,k}, g_{q-1,k}, d_{q-1,k}, d_{q-1,j}$ then the correctors would have the same order as the predictors. We have given our algorithm with correctors of higher order than the predictors because in the case of a constant stepsize and d = 1, the Adams methods for q = 1, 2, ..., 12 have significantly better stability characteristics when the corrector has an order one greater than the predictor. (For q = 13, ..., 19, as far as we have checked, the opposite is true.) For the ·case j = 1, d = 1, Klopfenstein [13] has shown that the method using a corrector with order one greater than the order of the predictor has the same region of asymptotic absolute stability as the method using the same order corrector. (This means that for h sufficiently large, the methods are equally sensitive to error in the matrix A of equation (3.14).) In both cases one also has the advantage of using a corrector which will give better accuracy on most problems. (As far as we know, no effort has been made to compare the algorithm which uses a corrector with order one greater than that of the predictor to the more usual one for the case d > 1 or for the case when the stepsize is not constant.)

It is frequently stated that for the purpose of error estimation the predictor and corrector should be of the same order. If the usual error estimate is added to the corrector with the same order as the predictor, one increases the order of the corrector by one, obtaining the type of algorithm we recommend. We suggest using the same error estimate for the case when the corrector presumably

.

is more accurate. Also see Shampine and Gordon [:∗] where this matter is considered in more detail. For a given stepsize, if the integration order is selected to minimize the local error, then by necessity any error estimate will be quite crude. (Despite the crude error estimates one gets with this policy of selecting the integration order, we believe it is the best policy since it tends to reduce global errors for a given amount of work, and since reasonable estimates of the global error are usually difficult to obtain from local errors even if they are known with high precision.)

Let

$$E_{n+1}^{(d-k)} = y_{n+1}^{(d-k)} - \hat{y}_{n+1}^{(d-k)} \tag{3.21}$$

where $\hat{y}_{n+1}^{(d-k)}$ is the result of using a corrector with the same order as the predictor to compute $y_{n+1}^{(d-k)}$, and $E_{n+1}^{(d-k)}$ is to serve as an indication of the error in $y_{n+1}^{(d-k)}$.

Clearly for the case j=0, equations (3.7), (3.18), and (3.21) give

$$E_{n+1}^{(d-k)} = s_k s_{q,k}^* [f(t_{n+1}, p_{n+1}, p_{n+1}'), \ldots, p_{n+1}^{(d-1)}) - \varphi_0^{(e)}(n+1)] \tag{3.22}$$

Although the local error in $y_{n+1}^{(d-k)}$ is of higher order in h_{n+1} the larger the value of k, this is not true for the global error; see [11]. The global error in all cases has order one less than the order of the local error in computing $y^{(d-1)}$.

Computing good theoretical error bounds for the case j > 0 is more work than can be justified. Thus we suggest estimating the change in e due to using correctors of different orders by the change in $e_1 (= e_1 - \hat{e}_1)$. This gives a good approximation if $\|e_1\|$ is considerably larger than $\|e_2\|$, which should be the case since ordinarily one will want to recompute A if $\|e_2\| > \alpha \|e_1\|$, where $\alpha \approx 1/8$. With $\hat{A}$ defi. d as is A in equation (3.14), except with every q replaced by q-1,

it is easy to obtain from equations (3.15), (3.15) with Ae_1 replaced by $\hat{A}\hat{e}_1$, (3.20), and (3.21)

$$\hat{A}\, E_{n+1}^{(d-j)} = s_{-j} d_{q,\,j}^* e_1 \qquad (3.23)$$

By appropriately bounding the stepsize (to prevent A and $\hat{A}$ from becoming too poorly conditioned), one can guarantee that $E_{n+1}^{(d-j)}$ will be changed very little if $\hat{A}$ is replaced by A in equation (3.23). This is what we recommend to the cautious user. Those who regard the error estimate primarily as a means to reasonable stepsize control may want to simplify equation (3.23) by assuming the partial derivatives are all zero (always a reasonable assumption for suf-ficiently small stepsize), obtaining

$$E_{n+1}^{(d-j)} \approx -\frac{d_{q,\,j}^*}{d_{q-1,\,j}}\, e_1 = -\frac{d_{q,\,j}^*}{d_{q-1,\,j}}\{c_{n+1}^{(d-j)} - p_{n+1}^{(d-j)}\} \qquad (3.24)$$

The use of equation (3.24) for large stepsizes can be justified for some types of problems. For example, it provides safe error bounds if A is diagonally dominant with all negative elements on the diagonal. Given $E_{n+1}^{(d-j)}$, equations (3.9), (3.10), (3.12), (3.18), (3.20) and (3.21) clearly yield

$$E_{n+1}^{(d-j-k)} = s_k(g_{q-1,\,k} E_{n+1}^{(d-j)} + g_{q,\,k}^* e_1),\quad k=1,2,\ldots,d-j \qquad (3.25)$$

$$E_{n+1}^{(d-j+k)} = s_{-k}(d_{q-1,\,k} E_{n+1}^{(d-j)} + d_{q,\,k}^* e_1),\quad k=1,2,\ldots,j-1 \qquad (3.26)$$

Almost as important as obtaining an estimate of the local error is estimating the effect that various strategies of selecting the stepsize will have on future estimated errors. As a first step consider the case when the divided difference (and hence the corresponding derivative) which is in the error estimate is constant,

and $h_{n+k} = h_n$, $k \geq 1$. If there has been a recent change in stepsize, then local error estimates will change from step to step until q steps without a stepsize change have occurred. This is due to the factors $\zeta_0 \zeta_1 \cdots \zeta_{q-1}$ which multiply $w[t_n, t_{n-1}, \ldots, t_{n-q}]$ to form $\varphi_q(n)$; see equation (2.4). For this simple case, after q steps without a change in the stepsize, the error estimate will have changed from its current value by a factor of (approximately)

$$\sigma_q = \frac{h_n (2h_n) \cdots (qh_n)}{h_n (h_n + h_{n-1}) \cdots (h_n + h_{n-1} + \cdots + h_{n-q+1})} \tag{3.27}$$

In order to limit the frequency with which the stepsize is changed and to reduce the work required to decide how much to change the stepsize, we suggest giving the user the option of specifying the two parameters

$\rho_i(\rho_i > 1)$ = the basic factor by which the stepsize is to be increased.

$\rho_d(\rho_d < 1)$ = the basic factor by which the stepsize is to be decreased. $\tag{3.28}$

The closer to one these parameters are selected, the more frequent changes in the stepsize will be, and thus the more overhead that is required for computing integration coefficients and difference tables. (See the next section.) At the same time, increased flexibility in selecting the stepsize (within reason) enables the solution to be computed to a given accuracy with fewer derivative evaluations. A reasonable choice for these parameters in most applications is $\rho_i = 2$, $\rho_d = 1/2$; but for problems with extremely expensive derivative evaluations, values as close to one as $\rho_i = 1.1$, $\rho_d = .9$ may prove useful.

We propose the following strategy for selecting the stepsize.

1. After computing error estimates, but before the second derivative evaluation of the step, check to see if the estimated error is "too" big. If so, go back to the beginning of the current step and try again with the stepsize reduced by a factor of $\min\{1/2, \rho_d\}$.

2. At the completion of the step, estimate what the error would be on the next step if the stepsize were held constant. If this estimated error is "too" big, reduce the stepsize by a factor of ρ_d before starting the next step.

The "too" big in the test for redoing a step should be at least twice as large as the "too" big in the test for simply reducing the step. With such a policy a step will require being repeated only rarely, thus saving the derivative evaluation that is wasted in such cases and also some of the overhead associated with backing up.

If no reduction in the stepsize is required, then increase the stepsize by a factor of ρ_i^k, where k is the smallest integer for which (any estimated error)$\cdot$(max $\{\sigma_q, 1\}) \cdot (\rho_i^{q_i})^{k+i}$ is "too" big. The "too" big used here should be no bigger than one tenth the "too" big used in the test for deciding if the stepsize should be reduced at the end of the step unless there has been a fairly long and consistent history of the error decreasing from one step to the next, in which case it pays to gradually increase the tolerance used here until it is the same size as that used for decreasing the stepsize. We have found it prudent to restrict k (in ρ_i^k) so that $\rho_i^k \leq \max(2, \rho_i)\rho_i^j$, where h increased on the previous step by ρ_i^j. The factors ρ_i^q used in estimating the growth in the error should be stored during the initialization procedure for all values of q which may be used.

4. Computational Details

In presenting the algorithms below, the following notation is used.

n_e number of differential equations.

$d(\ell)$ order of the ℓ-th equation.

$q(\ell)$ integration order used on the ℓ-th equation.

$j(\ell)$ the $(d(\ell)-j(\ell))$-th derivative of $y(\ell)$ is used in forming the differences $\varphi(i, \ell)$, $i=0, 1, \ldots$

$y^{(k)}(\ell)$ current value of $y^{(k)}$ for the ℓ-th equation, $k=0, 1, \ldots, d(\ell)-1$.

$\hat{y}^{(k)}(\ell)$ value of $y^{(k)}(\ell)$ from the previous step.

$f(\ell)$ current value of $f(t, y, y', \ldots, y^{(d-1)})$ for the ℓ-th equation.

$\varphi(i, \ell)$ i-th modified divided difference for the ℓ-th equation.

$e(\ell)$ for stiff equations, the value of e for the ℓ-th equation. (See equations (3.12), (3.15), (3.16).)

h current value of the stepsize.

$\zeta(k), \alpha(k), \theta(k), H(k), s(k)$ current values of $\zeta_k, \alpha_k, \theta_k, H_k,$ and s_k. (See equation (2.4).)

$g(i, k), d(i, k)$ coefficients for integration and differentiation formulas. Same as $g_{i, k}$, $d_{i, k}$. (See equations (2.16), (2.20), (3.4), and (3.5).)

$g^*(i, k), d^*(i, k)$ coefficients required for error estimation. (See equations (3.18), (3.20), (3.22), and (3.23).)

$\sigma(k)$ Same as σ_k in equation (3.27).

q_I $\{\ell : j(\ell) < d(\ell)\}^{\max} \{q(\ell)\}=$ maximum order integration formula used.

q_D $\{\ell : j(\ell) > 0\}^{\max} \{q(\ell)\}=$ maximum order differentiation formula used.

m_I $\max \{d(\ell)-j(\ell)\}=$ maximum number of repeated integrations.

m_D $\max \{j(\ell)\}=$ maximum number of repeated differentiations.

n_h number of steps that h has been constant (not counting the current step).

$$\left. \right\} \quad (4.1)$$

We have found that permitting different values for d, q, and j for the different equations in a system, and permitting any q to change from one step to the next, provides a useful flexibility. Different values for d can give a more

efficient integration if equations of different orders are being integrated; different values for q can make for more efficient integrations and makes available valuable information for diagnostic purposes; and different values for j can significantly reduce the size of the system of nonlinear equations which must be solved on every step if only a few equations in a large system cause the stiffness. When implemented as described here, this flexibility is obtained with little cost over what is required when the same values are used for each equation. Where it is not obvious, we indicate the simplifications that can be obtained when one or another of d, q, or j is fixed. In order that the implementation be as efficient as possible, we require $d(\ell) \leq 4$. The extension to larger values of d is trivial, but in practice larger values of d are rarely used. (Such equations can always be broken up into lower order equations.)

The description of the algorithm for computing integration, differentiation, and related coefficients, assumes that the following initial values have been assigned as indicated. These coefficients are never changed by the algorithm.

$$
\left.
\begin{aligned}
&\alpha(0) = \beta(0) = s(0) = \sigma(1) = 1. \\
&d(1,1) = d^{*}(1,1) = 1. \\
&d(i,k) = d^{*}(i,k) = 0, \quad i = 1, 2, \ldots, k-1; \ k = 2, 3, \text{ and } 4. \\
&g(0,k) = 1/k, \quad g(1,k) = 1/[k(k+1)], \quad k = 1, 2, 3, \text{ and } 4.
\end{aligned}
\right\} \quad (4.2)
$$

The following variables are used internal to the algorithm.

$$
\left.
\begin{array}{ll}
q_s & \max\{q_1, q_D, 2\} = \text{step number of the method.} \\[4pt]
n_s & \text{number of steps for which } \alpha, \beta, \sigma, \text{ and } \zeta \text{ coefficients} \\
& \text{determined by a constant stepsize have been computed.} \\[4pt]
n_1(n_D) & \text{number of steps for which integration (differentiation)} \\
& \text{coefficients determined by a constant stepsize have been} \\
& \text{computed.} \\[4pt]
\tau_1, \tau_2 & \text{locations used for temporary storage when computing } \zeta.
\end{array}
\right\} \quad (4.3)
$$

(4.3 cont'd.)

q_I^* value of q_I the last time integration coefficients were computed. (Initially $q_I^* = 0.$)

$\hat{q}_I$ $= q_I^* + m_I.$

$B(k)$ $= 1/k(k+1), k=1, 2, \ldots, \max\{q_I + m_I - 1\}$ (used to initialize $V(k)$).

$V(k)$ $= g(n_I - 1, k)$ if $k \le \hat{q}_I - n_I$; $= g(\hat{q}_I - k, k)$, for $k = \hat{q}_I - n_I + 1, \ldots, \hat{q}_I - 1$ (used to initialize $W(k)$).

$W(k)$ $= g(n, k)$ (a: $C20$ in the algorithm below) $k = 1, 2, \ldots, \hat{q}_I - n$).

    (4.3)

For use in selecting the order, it is useful to carry along one more difference than is required by the integration (or differentiation) formula. Since the value of ß used in forming this last difference is not very critical, a simple extrapolation formula is used to obtain the last ß (see C13). In order to make good decisions on when to increase the order we have found it necessary to examine differences of at least 4 different orders. In order to have enough differences for the order selection then it is necessary to restrict q_s to be ≥ 2. With such a policy it is necessary to set $\xi(0)$ equal to the starting stepsize when starting an integration. (We also use a different method for selecting the order on the first few steps.) For best efficiency the place to go at statements C20 and C28 should be set initially (once per integration if m_I and m_D are constant) based on the values of m_I and m_D. (In FORTRAN this is best done with the assigned GO TO statement.) It is assumed that if $m_I = 0$ then $q_I = 0$, if $m_D = 0$ then $q_D = 0$, and that neither q_I nor q_D can be increased by more than one on any step. We have left out the calculation of $g^*(n, k)$ for $k \ge 2$ at C21-C23 since we never estimate errors in anything but $y^{(d-1)}$. Error estimates for $y^{(d-j)}(j > 1)$ should not be used for stepsize selection since the error estimates tend to be much too small when starting. (Due to the small stepsize required by the low order.)

Algorithm for Computing Coefficients

C1. [Set q_s=step number of method.] $q_s \leftarrow \max[q_I, q_D, 2]$.

C2. [Test if stepsize did not change.]   if $n_h \neq 0$, go to C5.

C3. [Set new values for H and s.] $H(0) \leftarrow h$; $H(k) \leftarrow h/k$,

$s(k) \leftarrow H(k-1)s(k-1)$, $k=1, 2, \ldots, m_I-1$;

$s(-k) \leftarrow s(1-k)/H(k)$, $k=1, 2, \ldots, m_D$.

C4. [Set variables indicating a step change.] $n_s \leftarrow 1$; $n \leftarrow 1$; $n_I \leftarrow 1$; $n_D \leftarrow 1$;

$\tau_1 \leftarrow h$; go to C11.

C5. [Test if integration order did not increase.]   if $q_I \leq q_I^{\ast}$, go to C7.

C6. [Compute new V's required by the increase in q_I.] $V(\hat{q}_I) \leftarrow B(\hat{q}_I)$; if $n_I=2$, go to C7.

$V(k) \leftarrow V(k) - \alpha(\hat{q}_I-k)V(k+1)$,

$k = \hat{q}_I-1, \hat{q}_I-2, \ldots, \hat{q}_I-n_I+2$.

C7. [Test if stepsize has been constant long enough.]   if $q_s \leq n_s$, go to C14.

C8. [Update n_s and set the index n.] $n_s \leftarrow n_s+1$; $n \leftarrow n_s$.

C9. [Compute precisely, those coefficients which remain fixed if h is held constant.] $\beta(n_s-1) \leftarrow 1$; $\sigma(n_s) \leftarrow 1$;

$\alpha(n_s-1) \leftarrow 1/n_s$; $\tau_1 \leftarrow n_s h$.

C10. [Test if step has been constant for q_s steps.]   if $n \geq q_s$, go to C13.

C11. [Compute coefficients which will change on next step, even if h is held constant.] $\tau_2 \leftarrow \xi(n-1)$; $\xi(n-1) \leftarrow \tau_1$;

$\beta(n) \leftarrow \beta(n-1)\tau_1/\tau_2$; $\tau_1 \leftarrow \tau_2+h$;

$\alpha(n) \leftarrow h/\tau_1$; $\sigma(n+1) \leftarrow (n+1)\sigma(n)\alpha(n)$.

C12. [Test if more coefficients need be computed.] $n \leftarrow n+1$; if $n < q_s$, go to C11.

C13. [Set $\xi(q_s-1)$ and approximate $\beta(q_s)$.] $\xi(q_s-1) \leftarrow \tau_1$; $\beta(q_s) \leftarrow \beta^2(q_s-1)/\beta(q_s-2)$

C14. [Test if no integ. coeff. are required.]   if $n_I > q_I$, go to C26.

C15. [Set indices for computing integ. coefficient.] $n \leftarrow n_I$; $n_I \leftarrow n_I+1$; $q_I^{\ast} \leftarrow q_I$;

$\hat{q}_I \leftarrow q_I+m_I$; $j \leftarrow \hat{q}_I-n$.

C16. [Test if stepsize did not change.]   if $n > 1$, go to C18.

C17. ⌈Initialize $V(k)$ and $W(k)$.⌉ $V(k) \leftarrow B(k)$, $W(k) \leftarrow V(k)$,

    $k=1, 2, \ldots, j$; go to C25.

C18. [Update $V(k)$ (and initialize $W(k)$).] $V(k) \leftarrow V(k) - \alpha(n-1)V(k+1)$,

    $W(k) \leftarrow V(k)$, $k=1, 2, \ldots, j$;

    go to C20.

C19. [Inner loop for computing integration $j \leftarrow j-1$; $W(k) \leftarrow W(k) - \alpha(n-1)W(k+1)$,

    coefficients.] $k=1, 2, \ldots, j$.

C20. [Go store integration coefficients.]   go to C25-m_I.

C21. [C21 = C25-4.] $g(n, 4) \leftarrow W(4)$.

C22. [C22 = C25-3.] $g(n, 3) \leftarrow W(3)$.

C23. [C23 = C25-2.] $g(n, 2) \leftarrow W(2)$.

C24. [C24 = C25-1.] $g(n, 1) \leftarrow W(1)$; $g^*(n, 1) \leftarrow g(n, 1) - g(n-1, 1)$

C25. [Test if more integ. coeff. required.] $n \leftarrow n+1$; if $n \leq q_I$, go to C19.

C26. [Test if no differentiation coeff.   if $n_D > q_D$, go to C34.

    required.]

C27. [Set indices for computing diff. coeff.] $n \leftarrow n_D$; $n_D \leftarrow n_D+1$; if $n=1$, go to C33.

C28. ⌈Go compute and store diff. coeff.⌉   go to C33-m_D.

C29. [C29 = C33-4.] $d^*(n, 4) \leftarrow \alpha(n-1)d(n-1, 3)$;

    $d(n, 4) \leftarrow d^*(n, 4) + d(n-1, 4)$.

C30. [C30 = C33-3.] $d^*(n, 3) \leftarrow \alpha(n-1)d(n-1, 2)$;

    $d(n, 3) \leftarrow d^*(n, 3) + d(n-1, 3)$.

C31. [C31 = C33-2.] $d^*(n, 2) \leftarrow \alpha(n-1)d(n-1, 1)$;

    $d(n, 2) \leftarrow d^*(n, 2) + d(n-1, 2)$.

C32. [C32 = C33-1.] $d^*(n, 1) \leftarrow \alpha(n-1)$

    $d(n, 1) \leftarrow d^*(n, 1) + d(n-1, 1)$.

C33. [Test if more diff. coeff. required.] $n \leftarrow n+1$; if $n \leq q_D$ go to C28.

C34. [End of computing coefficients.]   Exit.

A straightforward implementation of the difference formulation of the Adams method involves accessing the difference tables in three different loops: to form the predicted values, to form the differences from predicted derivative values in order to obtain corrected values, and to form the difference tables at the end of the step from the final corrected derivative values. Each of these loops involves overhead associated with initializing indices and with the loop itself, and each must be passed through for each equation. An algorithm is given below which requires accessing the difference tables in only one loop. (The difference tables also must be accessed when correcting, estimating errors, and selecting orders, but the entire difference table is not required for these operations.) The capability of treating equations with different orders is obtained using preassigned transfers (similar to what was done in the algorithm for computing integration and differentiation formula coefficients) rather than a loop which would require additional overhead. The variable order Adams program DVDQ [15] uses three loops for operations on the difference tables, and loops on the order (for both predicting and correction) to permit equations of different orders. Thus a similar program based on the approach used here should require significantly less overhead than is reported for DVDQ in [16] and [17]. (Note, DVDQ uses a different method for changing stepsize, see [1] and it has no provision for stiff equations.)

The algorithm given below should be executed just after computing integration coefficients, which in turn is the first thing done on a step. The algorithm includes among the jobs it does:

1. An updating of the difference table based on $[\, y_n^{(d-j)} - \varphi_0^{(e)}(n)\,]$ from the previous step.

2. The calculation of predicted values for the differences on the next step, $\varphi_i^{(e)}(n+1)$, to be used for job 1 on the next step.

3. The calculation of predicted values for the dependent variables.

These are two situations when job 1 will have been done previous to the execution of this algorithm. If a step is being repeated, it is easiest to return

the difference tables to the point they would be in just after step 1 using the
formula

$$\varphi_i(n) = [\varphi_i^{(e)}(n+1) - \varphi_{i+1}^{(e)}(n+1)]/\beta_i(n+1), \quad i=0, 1, \ldots \tag{4.4}$$

which follows immediately from equation (2.6). If an interpolation to an off-step
point is required at the end of a step, then it is best to convert the $\varphi_i^{(e)}$'s to $\overline{\varphi}_i$'s
using equation (2.7) before doing the interpolation. We examine this point in
more detail in the next section. Thus we introduce

$$\nu = \begin{cases} 0 \text{ if no update has occurred } (\varphi \text{ contains } \varphi^{(e)'}s) \\ 1 \text{ if there has been an update } (\varphi \text{ contains } \varphi's) \end{cases} \tag{4.5}$$

The statement $T_{51} \leftarrow P64-I$ (for example) means there is a "go to" at P_{51}, and
this "go to" is now to indicate a transfer to the line labeled with a P followed by
the integer 64-I. The "go to" at P51 will then contain in braces the possible
transfers and the conditions which determine the actual transfer to be used.

Additional notation used in the algorithm includes

$$\left. \begin{array}{l} I \quad = d(\ell) - j(\ell) \text{ for the current value of } \ell. \\ \Sigma(i) = \text{sum used in the formula for predicting } y^{(1+i)} \\ \tau_1 = y_n^{(d-j)} - \varphi_0^{(e)}(n) \text{ for the current equation if } \nu \text{ '}; \text{ and } = 0 \text{ otherwise} \\ \tau_2 = \text{used to contain } \varphi_k^*(n+1) \text{ (see equation 2.8) for the current equation} \\ \tau_3 = \text{used to contain } \varphi_k^{(e)}(n+1) \text{ (see equation 2.6) for the current equation} \end{array} \right\} \tag{4.6}$$

For $q(\ell) > 1$, $\varphi_q^{(e)}$ is zero as far as the computation of the other differences is
concerned (since τ_3 is initially set to 0), but at P23, $\varphi(q(\ell), \ell)$ is set equal to $\varphi_q^{\wedge}$
for use in the order selection algorithm later. (It is useful to have a difference
with order one greater than is used in the corrector for purposes of order
selection.) It is assumed that $\varphi(q(\ell), \ell)$ is set equal to 0 before coming back to

this algorithm on the next step. This means that when beginning this algorithm (when $v = 0$)

$\varphi(q(\ell), \ell) = 0$ if the order was not changed on the last step.

$\varphi(q(\ell)-1, \ell)= 0$ if the order was increased on the last step.   (4.7)

In addition, when the order is increased, $\varphi(\text{new } q(\ell), \ell)$ should be set equal to $-\varphi(\text{old } q(\ell), \ell)$ before setting $\varphi(\text{old } q(\ell), \ell)$ to zero in order that the correct value for $\varphi(q(\ell), \ell)$ be obtained by the algorithm below. Note that condition (4.7) can be used to detect if the order was increased prior to repeating a step. (Clearly, the order can not be allowed to increase on a step that is rejected. One must also replace $\xi(k)$ with $\xi(k+1)-h$, $k = n_h$, $n_h+1, \ldots$, before executing the algorithm for obtaining new coefficients, when a step is being repeated.)

The case $q(\ell) = 1$ is treated the same as $q(\ell) = 2$ as far as computing other differences is concerned, but only $\varphi(0, \ell)$ is included in the computation of the sums. As indicated earlier, $q(\ell) = 1$ is treated as a special case in order to have an additional difference available to assist in order selection. Obvious simplifications can be made if this extra difference is not required. As given, the algorithm uses a first order predictor for $y_n^{(d-j)}$ when $q(\ell) = 1$, $j(\ell) > 0$, contrary to what is given in equation (3.6). (Also note that at P58 $\hat{y}^{(1)}(\ell)$ $+\varphi(1, \ell) = \varphi(0, \ell)$. The former is used because some implementations will want to carry y and $\hat{y}$ to more precision than φ.) The corrector formulas (3.7) and (3.11) should have $\varphi_0^{(e)}$ replaced by $\varphi_0^{(e)} - \varphi_1^{(e)}$ when $q(\ell) = 1$, if $q(\ell) = 1$ is being treated as a special case.

Finally, note that k_{min} should be set to 0 initially. (This can be done with the DATA initialization statement in FORTRAN.)

Algorithm for Predicting and Updating Differences

P1.   [Test if differences not updated yet.]   if $v=0$, go to P3.

P2.   [Set $\tau_1=0$, and transfers at P26 and P50.] $\tau_1 \leftarrow 0$: $T_{50} \leftarrow T_{26} \leftarrow P_{35}$: go to P4.

P3. [Set transfers at P26 and P50.] $\qquad$ $T_{50} \leftarrow T_{26} \leftarrow P37$.

P4. [Initialise the equation counter.] $\qquad$ $\ell \leftarrow 1$.

P5. [Set max. no. of repeated integrations.] $\qquad$ $1 + d(\ell) - j(\ell)$.

P6. [Test if equation is stiff.] $\qquad$ if $j(\ell) = 0$, go to P15.

P7. [Set τ_1 if $v = 0$.] $\qquad$ if $v = 0$, $\tau_1 \leftarrow e(\ell)$.

P8. [Set transfers at P30, P33, P36, P39, $\qquad$ $T_{39} \leftarrow T_{36} \leftarrow T_{33} \leftarrow T_{30} \leftarrow P44 - j(\ell)$;
P44, P51, and P59.] $\qquad$ $T_{44} \leftarrow P49 - I$; $T51 \leftarrow P58 - j(\ell)$;
$\qquad$ $T_{59} \leftarrow P64 - I$.

P9. [Go store 0 in diff. formula sums.] $\qquad$ go to $P14 - j(\ell)$.

P10. $\qquad$ $\Sigma(-4) \leftarrow 0$.

P11. $\qquad$ $\Sigma(-3) \leftarrow 0$.

P12. $\qquad$ $\Sigma(-2) \leftarrow 0$.

P13. $\qquad$ $\Sigma(-1) \leftarrow 0$.

P14. [Go store 0 in integ. formula sums.] $\qquad$ go to $P22 - I$.

P15. [Set τ_1 if $v = 0$.] $\qquad$ if $v = 0$, $\tau_1 = f(\ell) - \varphi(0, \ell)$.

P16. [Set transfers at P30, P33, P36, P39, and $\qquad$ $T_{39} \leftarrow T_{36} \leftarrow T_{33} \leftarrow T_{30} \leftarrow P49 - I$;
P51.] $\qquad$ $T51 \leftarrow P64 - I$.

P17. [Go store 0 in integ. formula sums.] $\qquad$ go to $P22 - I$.

P18. $\qquad$ $\Sigma(4) \leftarrow 0$.

P19. $\qquad$ $\Sigma(3) \leftarrow 0$.

P20. $\qquad$ $\Sigma(2) \leftarrow 0$.

P21. $\qquad$ $\Sigma(1) \leftarrow 0$.

P22. [Set index for use in loop below.] $\qquad$ $k \leftarrow q(\ell) - 1$.

P23. [Store φ_{k+1}^* (eq. (2.9)) into $\varphi(k+1, \ell)$.] $\qquad$ $\varphi(k+1, \ell) \leftarrow [\varphi(k+1, \ell) + \tau_1] \beta(k+1)$.

P24. [Test if usual case of $q(\ell) > 1$.] $\qquad$ if $k > 0$, go to P26.

P25. [Do special calculations when $q(\ell) = 1$.] $\qquad$ $\tau_3 \leftarrow \varphi(1, \ell); \varphi(2, \ell) \leftarrow [\varphi(2, \ell) + \tau_1]$
$\qquad$ $\beta(2)$; go to P38.

P26. [Initialise τ_3, and test if only backward $\qquad$ $\tau_3 \leftarrow 0$; if $k \leq n_h$, go to [P37 if
differences are required.] $\qquad$ $v = 0$; P35 if $v \neq 0$].

P27. [Set stopping index for mod. div. diffs. $\qquad$ $k_{min} \leftarrow n_h + 1$; if $v = 0$, go to P31.
and test if τ_1 needed.]

P28. [Set transfer at P50.] $T_{50} \leftarrow$ P29.

P29. [Compute φ_k^*, $\nu \neq 0$, and $k > n_h$.] $\tau_2 \leftarrow \varphi(k, \ell)\beta(k)$.

P30. [Go form sums.]   go to {P44-j(ℓ) if j(ℓ) > 0;
P49-1 otherwise}.

P31. [Set transfer at P50.] $T_{50} \leftarrow$ P32.

P32. [Compute φ_k^*, $\nu = 0$, and $k > n_h$.] $\tau_2 \leftarrow \lceil \varphi(k, \ell) + \tau_1 \rceil \beta(k)$.

P33. [Go form sums.]   go to {P44-j(ℓ) if j(ℓ) > 0,
P49-1 otherwise}.

P34. [Set transfer at P50.] $T_{50} \leftarrow$ P35.

P35. [Compute φ_k^*, $\nu \neq 0$, h constant.] $\tau_2 \leftarrow \varphi(k, \ell)$.

P36. [Go form sums.]   go to {P44-j(ℓ) if j(ℓ) > 0;
P49-1 otherwise}.

P37. [Set transfer at P50.] $T_{50} \leftarrow$ P38.

P38. [Compute φ_k^*, $\nu = 0$, h constant (usual case?)] $\tau_2 \leftarrow \varphi(k, \ell) + \tau_1$.

P39. [Go form sums.]   go to {P44-j(ℓ) if j(ℓ) > 0;
P49-1 otherwise}.

P40. [Form sums for differentiation $\Sigma(-4) \leftarrow \Sigma(-4) + d(k, 4)\tau_2$.

P41. formulas.] $\Sigma(-3) \leftarrow \Sigma(-3) + d(k, 3)\tau_2$.

P42. $\Sigma(-2) \leftarrow \Sigma(-2) + d(k, 2)\tau_2$.

P43. $\Sigma(-1) \leftarrow \Sigma(-1) + d(k+1)\tau_2$.

P44. [Go form sums for integrations (if any).]   go to {P49-1}.

P45. [Form sums for integration $\Sigma(4) \leftarrow \Sigma(4) + g(k, 4)\tau_2$.

P46. formulas.] $\Sigma(3) \leftarrow \Sigma(3) + g(k, 3)\tau_2$.

P47. $\Sigma(2) \leftarrow \Sigma(2) + g(k, 2)\tau_2$.

P48. $\Sigma(1) \leftarrow \Sigma(1) + g(k, 1)\tau_2$.

P49. [Compute $\varphi_k^{(e)}$, see equation (2.6).] $\tau_3 \leftarrow \tau_2 + \tau_3$; $\varphi(k, \ell) \leftarrow \tau_3$.

P50. [Test for end of forming differences or $k \leftarrow k-1$; if $k \geq k_{min}$, go to
end of forming mod. div. differences.]   {P29 if $\nu \neq 0$, $k_{min} \neq 0$;
P32 if $\nu = 0$, $k_{min} \neq 0$; P35 if $\nu \neq 0$
$k_{min} = 0$; P38 if $\nu = 0$, $k_{min} = 0$}.

P51. [Test if done forming differences and   if $k_{min}=0$, go to [P64-I if

      sums.] $j(\ell)=0$; P58-$j(\ell)$ otherwise].

P52. [Set up to compute differences based $k_{min} \leftarrow 0$; if $v=0$, go to P37.

P53. on constant stepsize.]   go to P34.

P54. [Compute y's using differentiation formulas.] $y^{(I+4)}(\ell) \leftarrow s(-4)\Sigma(-4)$.

P55. $y^{(I+3)}(\ell) \leftarrow s(-3)\Sigma(-3)$.

P56. $y^{(I+2)}(\ell) \leftarrow s (-2)\Sigma(-2)$.

P57. $y^{(I+1)}(\ell) \leftarrow s(-1)\Sigma(-1)$.

P58. $y^{(I)}(\ell) \leftarrow \hat{y}^{(I)}(\ell) + \varphi(1,\ell)$

                                              go to [P64-1].

P60. [Compute y's using integration formulas.] $y^{(I-4)}(\ell) \leftarrow \hat{y}^{(I-4)}(\ell) +$

                                              $H(1) [\hat{y}^{(I-3)}(\ell) +$

                                              $H(2)[\hat{y}^{(I-2)}(\ell) +$

                                              $H(3)(\hat{y}^{(I-1)}(\ell) + h\Sigma(4))]$.

P61. $y^{(I-3)}(\ell) \leftarrow \hat{y}^{(I-3)}(\ell) +$

                                              $H(1) [\hat{y}^{(I-2)}(\ell) +$

                                              $H(2)[\hat{y}^{(I-1)}(\ell) + h\Sigma(3)]\}$.

P62. $y^{(I-2)}(\ell) \leftarrow \hat{y}^{(I-2)}(\ell) +$

                                              $H(1)[\hat{y}^{(I-1)}(\ell) + h\Sigma(2)]$.

P63. $y^{(I-1)}(\ell) \leftarrow \hat{y}^{(I-1)}(\ell) + h\Sigma(1)$.

P64. [Test if more equations to be processed.] $\ell \leftarrow I+1$; if $\ell \leq n_e$, go to P5.

P65. [Set v to indicate that φ is not updated.] $v \leftarrow 0$.

P66. [End of predicting and updating diff.]   Exit.

5. Interpolation to Off-Step Points

     A significant advantage of multistep methods over one-step methods is that a multistep method has sufficient information stored to enable one to get the solution at any point passed during the integration: with the same accuracy as is obtained at the end of the individual steps, without interfering with the integration process in any way, and without requiring any additional derivative

evaluations.

Ordinarily one is interpolating to a point t which satisfies $t'_{n-1} < t \le t_n$ where t_n is the value of t at the end of the current step. Sometimes it is necessary to extrapolate the solution ($t_n < t \le t_{n+1}$) because derivatives are impossible to compute at $t = t_{n+1}$. Finally, if the solution is being saved for later use, it is useful to know that the algorithm gives reasonable accuracy for $t_{n-q} \le t \le t_{n+1}$, where $q = \min\{q(\ell)\}$. It is assumed that values of $y^{(k)}$ from the current step have been stored in $\hat{y}^{(k)}$ (see equation 4.1) and that the interpolated values are to be stored in $y^{(k)}$, and in f.

In order that full accuracy be obtained, the $\varphi_i^{(e)}(n)$'s should be replaced by $\varphi_i(n)$'s before doing the interpolation. This can be done using equation (2.7) (with n+1 replaced by n); it should be done only if $\nu=0$, see equation (4.5); and if it is done, ν should be set equal 1.

Let

$$h_I = t - t_n$$
$$\rho = h_I / h_n.$$

(5.1)

then h_I plays the same role as h_{n+1} in the predictor formulas (3.4)-(3.6), if the interpolation is looked at as as just taking a new step. However the recursions (2.16) and (2.20) (with h_{n+1} replaced by h_I) can not be used since arbitrarily large values of α_1 occur as $h_I \to -h_n$. These recursions wouldn't give the desired coefficients anyway since from equations (2.8), (3.4)-(3.6) and the fact that we are now using $\varphi_i(n)$, it is clear that we should introduce

$$g_{i,k}^{(1)} = \beta_i^{(1)} \tilde{a}_{i,k}$$
$$d_{i,k}^{(1)} = \beta_i^{(1)} d_{i,k}$$

(5.2)

for use in the interpolation, where $\beta_i^{(I)}$ is defined as in equation (2.4) with h_{n+1} replaced by h_I. Then equations (3.4)-(3.6) can be used for the interpolation if H and s are computed using h_I, φ_i^* is replaced by φ_i, g and d are replaced by $g^{(I)}$ and $d^{(I)}$, and it is understood that $d_{i,0}'$ implicitly multiplied φ_i^c in equation (3.6). (We recommend using q=2 in the case that q(ℓ) = 1.)

To avoid potential overflow in computing $s_{-k} = k!/h_I^k$, we suggest defining

$$\hat{d}_{i,k}^{(I)} = (k!/h_I^k)d_{i,k}^{(I)} \tag{5.3}$$

and using equation (3.5) with s_{-k} removed, and $d_{i,k}$ replaced by $\hat{d}_{i,k}^{(I)}$. (It is not a bad idea to simply compute $s_{-k}d_{i,k}$ and $s_k g_{i,k}$ instead of $d_{i,k}$ and $g_{i,k}$ when computing the coefficients for continuing an integration.)

With

$$\eta_i = h_I/\xi_i(n) \tag{5.4}$$

$$\gamma_i = \begin{cases} \rho & i = 0 \\ [h_I + \xi_{i-1}(n)]/\xi_i(n) & i > 0 \end{cases} \tag{5.5}$$

there follows immediately from equations (2.16), (2.20), and (5.2)-(5.5)

$$\hat{d}_{i,k}^{(I)} = \begin{cases} 0 & i < k \\ 1 & i=0, k=0 \\ \gamma_{i-1}\hat{d}_{i-1,0}^{(I)} & \begin{matrix} i=1, 2, \ldots, q-1; \\ k=0 \end{matrix} \\ (\dfrac{k}{\xi_{i-1}(n)})\hat{d}_{i-1,k-1}^{(I)} + \gamma_{i-1}\hat{d}_{i-1,k}^{(I)} & \begin{matrix} i=k, k+1, \ldots, q-1; \\ k=1, 2, \ldots \end{matrix} \end{cases} \tag{5.6}$$

$$g_{i,k}^{(I)} = \begin{cases} \gamma_{i-1} g_{i-1,k}^{(I)} - \eta_{i-1} g_{i-1,k+1}^{(I)} & i=q-1, q-2, \ldots, 1; \ k=m-1, \ldots, 1 \\ 1/k & k=m-1, \ldots, 1 \end{cases} \quad (5.7)$$

Of course, equation (5.7) is valid for any value of m, but we derive below a more efficient algorithm for computing $g_{i,m}^{(I)}$, where m is the maximum value of d-j (d and j defined as in equation (3.3)). If j=0, then $\hat{d}_{i,0}^{(I)}$ need not be computed, and f $(=y^{(d)})$ can be computed using the formula

$$f = \varphi_0 + \gamma_0[\varphi_1 + \gamma_1[\varphi_2 + \cdots \gamma_{q-2}\varphi_{q-1}]] \cdots] \quad (5.8)$$

The recursion to compute $g_{i,m}^{(I)}$ is obtained starting from the coefficients of the interpolating polynomial. Define

$$\varphi_{i,k}(n) = \begin{cases} \xi_0^k(n) \xi_0(n) \xi_1(n) \cdots \xi_{i-k-1}(n) w[t_n, \ldots, t_n, t_{n-1}, \ldots, t_{n-i+k}] & i \geq k+1 \\ \xi_0^k w[t_n, \ldots, t_n] & i = k \end{cases} \quad (5.9)$$

where the t_n's in the divided differences are repeated k+1 times. Clearly $\varphi_{i,0}(n) = \varphi_i(n)$ as defined in equation (2.4), and $\varphi_{i,i}$ gives scaled coefficients of the interpolating polynomial, so that

$$P_{q-1,n}(t) = \sum_{i=0}^{q-1} \varphi_{i,i}(n)[(t-t_n)/h_n]^i \quad (5.10)$$

From equations (2.2) and (5.9) it is easy to obtain

$$\varphi_{i,k}(n) = [h_n/\xi_{i-k}(n)]\varphi_{i,k-1}(n) + \varphi_{i+1,k}(n), \quad k=1,2,\ldots,i; \ i=q-1,\ldots,1 \quad (5.11)$$

where $\varphi_{q, k}(n) = 0$.

(If one wants to integrate/interpolate/differentiate $P_{q-1, n}(t)$ for many different values of t, then it is most efficient to compute the $\varphi_{i, i}$ from the $\sigma_{i, 0}$'s, and then to compute the desired result using equation (5.10). With $\varphi_{i, i}(n)$ $= (h_n^i / i!) y_n^{(i)}$, $i = 0, 1, \ldots, q+d-2$, one gets a method like ours in the Nordsieck formulation if one writes

$$\varphi_{i, i}(n+1) = \varphi_{i, i}^{(e)}(n+1) + (c_{q, n}^{(j-d+i)}(1))[(d-j)!/i!][\frac{h_{n+1}^{(d-j)}}{(d-j)!} y_{n+1}^{(d-j)} - \varphi_{d-j, d-j}^{(e)}(n+1)] \qquad (5.12)$$

where $c_{i, n}^{(k)}$ is defined as in section 2. Recursions for $\varphi_{i, i}^{(e)}(n+1)$ follow from considering $w[t_n, t_n, \ldots, t_n]$, $w[t_n, \ldots, t_n, t_{n+1}]$, $w[t_n, \ldots, t_n, t_{n+1}, t_{n+1}]$, $\ldots$, $w[t_n, t_{n+1}, \ldots, t_{n+1}]$, $w[t_{n+1}, \ldots, t_{n+1}]$; the recursions obtained are identical to the computational shortcut due to Gear, [18]. Also see Thomas, [19]. In order for this Nordsieck type formulation to give a method equivalent to what we have described, some care is required when changing the order, see [14]. If many interpolations are to be performed during the integration, this Nordsieck formulation may be preferred to the use of modified divided differences if one is not interested in using different integration orders for different equations, and not too many equations are being integrated.)

The m-th integral of P is given by

$$\int_{t_n}^t \cdots \int_{t_n}^t P_{q-1, n}(t) = \sum_{i=0}^{q-1} a_{i, k}^{(m)}(n) \varphi_{i, k}(n) \qquad (5.13)$$

where it is clear from equation (5.10) that

$$a_{i, i}^{(m)} = h_n^m \rho^{i+m}/(i+m)\cdots(i+1) = h_i^m \rho^i/(i+m)\cdots(i+1) \qquad (5.14)$$

and from equations (5.11) and (5.13)

$$a_{i,k}^{(m)} = \begin{cases} \dfrac{h_n}{\xi_{j-k-1}(n)} \, a_{i,k+1}^{(m)} & k=0; \ i=1, 2, \ldots, q-1 \\[4mm] \dfrac{h_n}{\xi_{j-k-1}(n)} a_{i,k+1}^{(m)} + a_{i-1,k}^{(m)} & k=i-j; \ i=q-1, \ldots, j+1; \ j=1, \ldots, q-2 \end{cases} \tag{5.15}$$

With the definition

$$\hat{a}_{i,k}^{(m)} = h_I^{-m}(m-1)! \, \rho^{-k} a_{i,k}^{(m)} \tag{5.16}$$

we have from equations (5.2), (5.4), (5.14)-(5.16)

$$\hat{a}_{i,k}^{(m)} = \begin{cases} \dfrac{(m-1)!}{(i+m)\cdots(i+1)} & k=i; \ i=0, 1, \ldots, q-1 \\[3mm] \eta_{j-k-1} \hat{a}_{i,k+1}^{(m)} + \hat{a}_{i-1,k}^{(m)} & k=i-j; \ i=q-1, \ldots, j+1; \ j=1, \ldots, q-2 \\[3mm] \eta_{j-k-1} \hat{a}_{i,k+1}^{(m)} & k=0; \ i=1, 2, \ldots, q-1 \end{cases} \tag{5.17}$$

and

$$g_{i,m}^{(I)} = \hat{a}_{i,0}^{(m)}, \qquad i=0, 1, \ldots, q-1 \tag{5.18}$$

(Note that $a_{i,i}^{(m)}$ need be computed at most one time if m is not changed. Calculations can be arranged so that $d_{i,k}$, $g_{i,k}$, and $a_{i,k}$ ($k < i$) all occupy the same vector in storage.)

6.   Concerning Some of the Messy Details

If one computes the difference table of a function with sufficient precision, for sufficiently small h, he will most likely find that for a given order q, differences of order q+k tend to decrease in magnitude as k takes the values 0, 1, ... Since in practice we almost always select the order q in such a way that the differences behave in the opposite way, it is reasonable to suspect that any theory based on results for h → 0 and ignoring the effect of round-off error will be of limited value. We believe there is greater danger in attempting to apply rigorous mathematics to problems which do not satisfy the underlying assumptions, than there is in careful inductive reasoning from results on a selection of problems which individually are simple enough to understand, and collectively cover the types of difficulties found in real problems. Thus we have taken primarily an empirical approach. Most of our effort has been spent in poring over difference tables generated in the solution of a variety of problems while using a variety of algorithms for selecting the integration order. Most of what follows is either trivial or mere opinion, yet much of it is important in determining the effectiveness of a variable order algorithm for solving differential equations. The specific algorithms presented give an idea of what we have done, they are not intended as recommendations. For results from a variable order Adams method which makes use of some of the ideas below, see [16]; for a comparison with other methods, see [17].

6.1 General Design

The interface between algorithms and users, despite its importance in determining the effectiveness of an algorithm, has been given minimal attention in the numerical analysis literature. Because of the problems inherent in the addition of features not considered in the original design (probably by a person unfamiliar with the code) we believe it is a good idea to design for maximum flexibility if an algorithm is meant for general purpose use. Complex applications will require the flexibility, and it is a relatively simple matter to insert

a flexible program into a package to be used by the unsophisticated user. Since code of wide generality is liable to contain much that is superfluous to some applications, we think it is also a good idea to make it easy for a user to remove that code he is not interested in using. A procedure for handling code which consists of many versions is described in [20].

We have found it difficult to anticipate user needs. Users can't know what they should have available when they don't know what the possibilities are, and thus they are not as much help in this area as one might think. The integrator in [15] is reasonably flexible and all of its features have been heavily used. But it proved to be insufficiently flexible for some users, leading to the design [21]. This in turn has had to be modified and extended to meet user needs that have surfaced since it was written. The results of a survey on the importance of various factors in a program for solving differential equations can be found in [22].

6.2 Some General Comments on Differences

In the case of the Adams method we have made it a practice to correct y, estimate errors, and select the integration order in the same block of code, thus making multiple use of the differences formed from predicted derivative values. Especially at low integration orders, we prefer such differences to those formed from corrected derivative values, since the former tend to converge less rapidly, giving a more conservative algorithm. For $q \leq 2$ (d=1, j=0) it is possible for the difference table formed from corrected derivative values to converge nicely, while at the same time the numerical solution is diverging rapidly from an acceptable solution. Such problems do not occur when using predicted derivative values. Let

$$\varphi_k = \begin{cases} f(t_{n+1}, p_{n+1}, \ldots, p_{n+1}^{(d-1)}) - \varphi_0^{(e)}(n+1) & k = q \\ \varphi_k^{(e)}(n+1) + \varphi_q & k = q-1, q-2, \ldots \end{cases}$$

$$(6.2.1)$$

and let φ_{q+1} be computed as φ_q - [What is stored in the q-th difference location by the algorithm for predicting; see P23.], where $q = \max\{2$, integration order of the predictor]. Ordinarily, only the differences φ_{q-2}, φ_{q-1}, φ_q, and φ_{q+1} need be computed since they almost always (see below) provide sufficient information for selecting the order.

We have found the linear equation $y' = Gy$ (G a constant matrix) helpful in organizing some of our thoughts. For this case, if h is constant and G has distinct eigenvalues, $y_n = \Sigma c_i r_i^n$. Some of the r_i are extraneous to approximating the desired solution (See [23] for some background on this.), and for the cases $d=1$, $j=0$, $q>3$; $d=1$, $j=1$, $q>1$ it can be shown that the largest extraneous r_i has a negative real part. Since the k-th backward difference of r^n is given by $r^n[(r-1)/r]^k$, differences of an extraneous root tend to increase. By selecting the order at about that point where the differences start to increase, one is stopping at about that order where the influence of an extraneous root is starting to dominate the differences, which in turn guarantees a stable method.

A given difference may be small because the error is small, or because it happens to be passing through zero, or just as a fluke. Any decision made on the basis of one difference being small has a good chance of giving the wrong result. (In the sense for example that the order might be increased when it should not be; such wrong results do not necessarily do significant harm to the solution.) On the other hand, two small differences in succession will only rarely be misleading.

In judging the convergence of the differences, their signs are also important. For given magnitudes, alternating signs indicate the most rapid convergence.

6.3 Starting the Integration

Variable order methods do not require any special logic to start an integration, but still there are advantages to treating the start in a special way. Efficiency in the starting process can be critical for problems with frequent discontinuities. Because of their ability to find the proper stepsize quickly, we believe that variable order methods will frequently prove superior to good one-step methods

on very short integrations. The Adams method starting procedure outlined
below does not require much additional code, despite its apparent complexity.
Initially, of course $q = 1$.

Let p denote the predicted values at t_0+h, and c_1, c_2, c_3 successive corrected
values. After $p^{(d)}$ is computed, compute c_1 and the estimated error on the first
step. If the estimated error is too large (in the sense that such an error would
ordinarily cause the step to be repeated), a new starting stepsize is selected on
the assumption that the estimated error is proportional to the square of the
stepsize. The resulting stepsize will usually give an acceptable error, and thus
an initial value for the stepsize that is too large usually costs only one extra
derivative evaluation. It is rare for more than two evaluations to be required.
In estimating the error on the first step, we multiply our usual estimate by $1/4$
since estimated errors would otherwise have a tendency to be much too large.

After obtaining a satisfactory c_1, compute $c_1^{(d)}$ and c_2. Estimate the error
in $c_2^{(d-1)}$ by $4h(p^{(d)}-c_1^{(d)})$ and compute $s_1(l) = |h[p^{(d)}(l) - c_1^{(d)}(l)]/[p^{(d-1)}(l)$
$- c_1^{(d-1)}(l)]|$. If the estimated error is too large (rarely happens) reduce the
stepsize as in the preceding paragraph and start over. Otherwise check if
$s_1 < 1/16$ for all equations, and if it is proceed to the next paragraph. If
$s_1(l) > 1/16$ for some l, then compute $c_2^{(d)}$ and c_3. Compute a new error and
$s_2(l)$ using formulas like those above with p replaced by c_2. If the error is too
large, start over; if not and $\min\{s_1(l), s_2(l)\} > 1/16$ for some l then end the
the starting phase. If none of these, then continue to the end of the first step
immediately below.

Set $y_1 =$ the final corrected value, and increase the order to 2. From this
point until the end of the starting phase, no derivatives are computed after
computing corrected values. (A PEC method is used.) This is justified since
the second derivative evaluation is required primarily to improve the stability
properties of the method. For low orders and small stepsizes, instability is
usually not a problem. At the end of the second step increase the order to 3.
At the end of the third step leave the order at 3, making φ_{q+1} available (and
thus the usual order selection process possible) for the first time on the fourth

step. If at any time the stepsize must be reduced, or the order must be reduced
and the estimated error is too large to permit an increase in the stepsize, then
end the starting phase. One can gain a little in the starting process by making
it easier for the order to increase than would be prudent after getting started.

6.4   Determining if Too Much Precision Has Been Requested

We believe that some test for unreasonable accuracy requests should be
considered an intrinsic part of any general purpose integration program. Many
users have come to us because our integrator gave them a diagnostic to the
effect that it could not get the accuracy they desired. These users appreciated
being warned of a problem they would not have been aware of otherwise. Users
of [15] rarely question us about this type of diagnostic now. We presume they
have learned to select reasonable error tolerances, or to trust the diagnostic
when they get it. Since a missed diagnostic of this type will usually only result
in a less efficient integration, and since unjustified diagnostics make for bad
relations with users, we recommend that one be conservative in the test which
results in the diagnostic.

Clearly the absolute accuracy that can be obtained depends on the number
of significant digits to which f is computed, and on the size of f. Thus in solving
$y'' = -y$, $y(0) \neq 0$, $y'(0) = c$, an accuracy request that is impossible to meet for
$c = 1$, may be easily met for $c = 10^{-20}$.

The test used in [15] assumes that round-off errors in f get magnified by
2^k in φ_k. (This is the magnification one gets if errors are of equal magnitude
and alternate in sign from one step to the next. This assumption appears to
work best, even though smaller magnification factors would appear more
reasonable.) Thus, with

$$R_1 = [\,|\varphi_q| + |\varphi_{q+1}|\,]/\lceil 2^q |f| 10^{-d_f} \rceil \qquad (6.4.1)$$

where d_f is the number of significant digits in f, $R_1 < 1$ is an indication that round-off errors are limiting the accuracy. To a first approximation, the test in [15] gives a diagnostic if the estimated error is too large and $R_1 < 1$, where R_1 is computed on the assumption that d_f is approximately the number of significant digits in the computer's number system. We believe that it is too much of an imposition on the user to ask him to supply d_f; but even if it weren't, it is hard to justify trusting him for d_f when he cannot be trusted to supply a reasonable error tolerance. If round-off error is the primary component of the estimated error, then the estimated error for a larger stepsize tends to be too large. Thus in [15] the estimated error at twice the stepsize is reduced whenever $R_1 < 1$.

We have been experimenting recently with a test which is impossible to justify, and for which examples could be constructed which cause it to give unjustified diagnostics. It makes no assumptions about the precision to which f is computed, but rather makes some implicit assumptions about the regularity with which difference tables converge. An integrator using this test has been used on a few applications, and has resulted in three diagnostics. In two cases (once when the user thought everything was done in double precision) some calculations done in single precision, should have been done in double precision, and in the other case results were being interpolated from a table with insufficient accuracy. *The integrator* [15] *would not detect these problems.* This new test should not be trusted unless the order selection algorithm does a very good job of selecting the order, it should not be applied in the starting phase of an integration, and $q \leqslant 2$ should be treated as a special case.

Let

$$R_2 = [\, |\varphi_q| + |\varphi_{q+1} | \qquad {}^{\cdot}q)[\, |\varphi_1^{(e)}| + |\varphi_2^{(e)}|\,]\,] \qquad (6.4.2)$$

where

$$r(q) = \begin{cases} .1 & 3 \leq q \leq 4 \\ 2^{-[(q-1)^2 - 3]/3} & q > 4 \end{cases} \tag{6.4.3}$$

and let n_R be the number of steps for which the equation with the largest ratio of estimated error to requested error has had an $R_2 < 1$. If the estimated error is too large, and $n_R \geq 4$, then a diagnostic is indicated. The estimated error with the stepsize increased is reduced by the factor (See equations (3.28) and (4.1) for notation.)

$$\max\{R_2, \rho_i^{-k}\}, \text{ where } k = \max\{n_h, n_R\} - 3] \tag{6.4.4}$$

whenever $R_2 < 1$. When $R_2 < 1$, the integration order is increased if

$$|\varphi_{q+1}| < \min\{|\varphi_{q-1}|, |\varphi_q|\}. \tag{6.4.5}$$

6.5 Selection of Integration Order

The most important point to be made here is that a reasonably good job of selecting the order can be done with very little effort. For example, the algorithm used in [15] for $q > 2$ is given by the following.

Q1. $\tau_1 \leftarrow q+2$

Q2. If $|\varphi_{q-2}| < \tau_1 |\varphi_q|$, go to Q5

Q3. If $|\varphi_{q-1}| < \tau_1 |\varphi_{q+1}|$, go to Q8

Q4. increase q by one, and go to Q8

Q5. $\tau_1 \leftarrow .25 \tau_1$

Q6. If $|\varphi_{q-2}| > \tau_1 |\varphi_q|$ or $|\varphi_{q-1}| > \tau_1 |\varphi_{q+1}|$, go to Q8

Q7.   decrease q by 1

Q8.   end of order selection

Note that the above algorithm uses 4 different differences and requires either 2 or 3 comparisons. If one does not want the order to oscillate unduly, then we believe that at least 4 differences must be examined.

In order to have a check for discontinuities it is sometimes necessary to check more than 4 differences. We now do this as follows. If

$$|\varphi_{q-2}| \le |\varphi_{q-1}| \le |\varphi_q| \le |\varphi_{q+1}|, \text{ and} \qquad (6.5.1)$$

$$.875|\varphi_q - \varphi_{q-1}| < |\varphi_{q+1}| - |\varphi_q| \qquad (6.5.2)$$

then q is reduced by 1; φ_{k+1} is replaced by φ_k for k = q-2, q-1; a new φ_{q-2} is computed; and the test above passed through again. The conditions for reducing q by more than 1 must be very stringent. For example, on a simple 2-body with eccentricity .6, using just the condition (6.5.1) for this test, we have seen the order reduced by about 6 over several steps, when a reduction of at most one was appropriate. If q starts out ≥ 5, is reduced to 3 and the differences still converge slowly, then a discontinuity is indicated and the integration restarted. Otherwise the final q is treated just as if it were the original as far as order selection is concerned.

The more factors that are taken into consideration, the better job one can do in selecting the order. For example, it helps to bias the test towards a low order when errors are decreasing, and vice versa when errors are increasing. (The former may increase the current error estimate, while decreasing the estimate of what the error would be with the stepsize increased.) Since the selection of integration order influences both the integration efficiency and tests such as discussed in 6.4 and 6.8, it is not easy to decide at what point additional

effort does not justify the return. We think it is better to use a simple scheme such as the one in [15] on every step, rather than some more complicated scheme less frequently, and that at a minimum, some test for decreasing the order be made on every step.

The algorithm we are currently using works reasonably well, but is probably unduly complicated. It uses as a measure of the convergence of the difference table at φ_q (with appropriate safeguards to prevent overflow):

$$\theta_q = \begin{cases} |\varphi_{q+1}|/|\varphi_{q-1} - \varphi_q| & \text{if } \varphi_{q+1}\varphi_{q-1} \geq 0 \\ |\varphi_{q+1}|/|\varphi_{q-1}| & \text{if } \varphi_{q+1}\varphi_{q-1} < 0. \end{cases} \tag{6.5.3}$$

To a first approximation, the order is increased if

$$\theta_q + \theta_{q-1} < \begin{cases} .75 \text{ if } h \text{ is closer to being decreased} \\ \rho_i^{-2} \text{ if } h \text{ is closer to being increased,} \end{cases} \text{ and } \theta_q^2 < \theta_{q-1}, \text{ and } \theta_q < 2\theta_{q-1} \tag{6.5.3}$$

and it is decreased if

$$|\varphi_{q+1}| > |\varphi_q| > |\varphi_{q-1}| \tag{6.5.5}$$

Note that these tests differ from those used in [15] in that signs of the differences influence the tests, the order need not be increased when θ_q and θ_{q-1} are both very small, φ_{q-2} is not used in the test for decreasing the order, and the tests to not have an explicit dependence on q.

6.6 Error Estimation and the Selection of Stepsize

Many of our ideas in this area are given at the end of section 3. We currently estimate the error in $y^{(d-1)}$ for the Adams method by

$$|hg_{q,1}^{*}|[\,|\varphi_q| + |\varphi_{q+1}|\,]/(1 - \bar{\theta}) \qquad (6.6.1)$$

where $\bar{\theta} = \min\{\theta_q, \theta_{q-1}, .75\}$, see equation (6.5.3). This reduces to equation (3.22) if φ_{q+1} and $\bar{\theta}$ are replaced by 0. The inclusion of $|\varphi_{q+1}|$ is for reasons discussed in 6.2, and the factor involving $\bar{\theta}$ is included because otherwise error estimates tend to be a little low when the differences decrease slowly. We currently use

$$|hg_{q,1}^{*}|\rho^q \max\{\sigma_q, 1\}[\,|\varphi_{q-1}| + |\varphi_q| + |\varphi_{q+1}|\,]/(1-\bar{\theta})^2 \qquad (6.6.2)$$

for the estimated error with the stepsize increased by the factor ρ. The role of $\bar{\theta}$ in (6.6.2) is quite important since a slowly convergent difference table is liable to require a significantly lower order when the stepsize is increased, thus giving a larger error estimate.

As mentioned at the end of section 3, we think it is a good idea to vary the parameter one uses for making decisions on increasing the stepsize as a function of the increase or decrease of the error estimates. The way we do this is outlined below. Numbers in parentheses give the actual values of the parameters as currently implemented.

    E   largest value in any equation of (estimated error)/(requested
             error).

    ϵ_R(.3)   if $E > \epsilon_R$ the step is repeated with h reduced by the factor
             $\min\{1/2, \rho_d\}$.

E_A average value of E. Computed as $(E+E_A)/2$ at the end of every step, with an initial value of ε_R.

$\hat{E}_A$ value of E_A just before current E_A was computed.

E_x estimated value for E on the next step. $E_x = E\min\{2, E/\hat{E}_A\}$.

$\varepsilon_x(.1)$ if $E_x > \varepsilon_x$ the stepsize is reduced by the factor ρ_d.

$\overline{E}_I(.2)$ upper bound on E_I, see below.

$\underline{E}_I(.01)$ lower bound on E_I, see below.

E_I step is increased if in so doing the estimated E at the new stepsize is less than E_I. With an initial value of ε_x it is computed using

$$E_I = \begin{cases} \min\{\overline{E}_I, E_I\hat{E}_A/E_A\} & \text{if } E \leq E_A \\ [E_I + \underline{E}_I]/2 & \text{if } E > E_A \end{cases}$$

Note that except for starting the integration, the stepsize is never reduced by more than $\min\{1/2, \rho_d\}$ no matter how large the error. If the step has to be repeated more than once we restart the integration. If the stepsize has been decreased, then increased, and is about to be decreased again, we reduce h by the factor $(1+\rho_d)/2$; this tends to reduce oscillations in the stepsize.

6.7 Treating $q(\ell) = 1$ As a Special Case When Using the Adams Method

Frequently when the order selection algorithm picks an order of 1, there is a stability problem. In this case the stability of the method can be improved by introducing a parameter ω into the corrector formula. Thus

$$y_{n+i}^{(d-k)} = p_{n+i}^{(d-k)} + \omega s_k \delta_{1,k} \nabla_i, \qquad 1/4 \leq \omega \leq 1. \tag{6.7.1}$$

When solving $y' = \lambda y$ with $\omega = 1$, the above method is absolutely stable for $s = h\lambda \in [-2, 0]$. The largest interval of absolute stability is obtained for $\omega = 1/4$, which gives a method absolutely stable for $s \in [-8, 0]$. We have, with mixed success, attempted to approximate s, and for $s \leq -2$ set $\omega = \max\{1/4, -2(1+s)/s^2\}$. This choice for ω zeros the root of the characteristic equation if $-4-2\sqrt{2} \leq s \leq -2$, assuming the equation is linear, and that s is approximated correctly.

Since $|g_{2,1}^*|$ is so much smaller than $|g_{1,1}^*|$ (1/12 vs. 1/2 for h constant), there is a tendency to overestimate the error when using (6.6.1) and (6.6.2). We have been substituting 1/8 for $\gamma_{1,1}^*$ in these expressions. We also increase the error estimates a little if $|s|$ is large or $\bar{\theta}$ is large. (A larger increase is used for (6.6.2) than for (6.6.1).) The increase in the estimate (6.6.1) is prudent; the increase in (6.6.2) quite important if one has a mildly stiff equation, the order has dropped to one, and one wants to avoid oscillations in the stepsize and minor irregularities in the solution.

6.8   Testing for Stiffness

It is frequently the case that only some of the differential equations in a large system are the cause of stiffness. Since the solution of a large stiff system requires considerable storage and computation for the iteration process and since users frequently do not know which, if any, of their equations are stiff, it is desirable to have a test for stiffness. With equations separated into those which are stiff and those which are not, fewer partial derivatives and smaller matrices are required. Since suggesting automatic classification of equations in [12], we have tried off and on without success to find a stiffness test which does not require computing partial derivatives. A test to tell when an equation being integrated as a stiff equation, could be integrated with the Adams method would also be useful, but should not be as difficult since in this case the Jacobian matrix is available.

Sometimes it is possible to say in advance which equations are stiff and which are not, in which case no tests are required. One problem at JPL was integrated with 9 equations treated as stiff and 5 as non-stiff. In this case, in addition to the savings in the iterations there was the bonus that partials connected with the stiff equations were easy to compute, while the other partials were extremely complicated.

It is worth noting that tests based on ratios such as $[f(y)-f(p)]/[y-p]$, where y and p are corrected and predicted values at the same time point, are not likely to be successful. For $y' = Gy$ as in 6.2,

$$y(t) \approx \sum_i c_{i,n} e^{\lambda_i(t-t_n)} \qquad (t \text{ near } t_n)$$

and with the onset of stiffness, the $c_{i,n}$ associated with λ_i's which have large negative real parts become very small. Then if the method is stable, which it will tend to be because of the way the integration order is selected, there is not enough of the solution in the direction of the eigenvectors associated with the large negative λ_i's to give useful ratios.

When integrating an equation which is stiff with an Adams method it is best if the stepsize is such that the method is absolutely stable, but not relatively stable. The order tends to come down gradually because of the relative instability, but there is no growth in the error since the method is absolutely stable. One problem with halving and doubling the stepsize with the Adams method is that as a result of doubling the stepsize it is possible to move from inside the region of relative stability to outside the region of absolute stability. When this occurs, the order is reduced, there is rapid error growth, and the stepsize must be halved. The algorithm is then relatively stable, the order is increased, the error estimates decrease and the cycle is repeated.

We think that monitoring R_2 of equation (6.4.2) (with suitable (?)

adjustments when q or h is changed) together with the flexibility in changing

stepsize available through the use of modified divided differences, may provide

what is needed to detect stiffness. There is of course also the possibility that

we are attempting to extract information from difference tables that isn't there

to begin with.

References

[1] Krogh, F. T.: Algorithms for Changing the Stepsize Used by a Multi-
    step Method. JPL Internal Document, Section 314 TM No. 275, Jet
    Propulsion Laboratory, Pasadena, Calif. (Oct. 1970). (To appear in
    the SIAM Journal on Numerical Analysis.)

[2] Krogh, F. T.: A Variable Step Variable Order Multistep Method for
    the Numerical Solution of Ordinary Differential Equations. Information
    Processing 68 (Proceeding of the IFIP Congress 1968.) 194-199.
    North Holland Publishing Co., Amsterdam (1969).

[3] Forrington, C. V. D.: Extensions of the Predictor-Corrector Method
    for the Solution of Systems of Ordinary Differential Equations. Computer
    Journal 4 (1961-62), 80-84.

[4] Gabel, G. F.: A Predictor-Corrector Method Using Divided Differences.
    University of Toronto Department of Computer Science Technical Report
    No. 5 (October 1968).

[5] Piotrowski, P.: Stability, Consistency and Convergence of Variable k-
    Step Methods for Numerical Integration of Large Systems of Ordinary
    Differential Equations. Conference on the Numerical Solution of
    Differential Equations, Lecture Notes in Mathematics 109, 221-227.
    Springer-Verlag, Berlin (1969).

[6] Van Wyk, _ · Variable Mesh Multistep Methods for Ordinary Differen-
    tial Equations. J. Comp. Physics 5 (1970), 244-264.

[7] Brayton, R. K., Gustavson, F. G. and Hachtel, G. D.: A New Algorithm
    for Solving Differential-Algebraic Systems Using Implicit Backward
    Differentiation Formulae. Proceedings of the IEEE 60 (1972), 98-108.

[8] Curtiss, C. F. and Hirschfelder, J. O.: Integration of Stiff Equations.
    Proc. of the Nat. Acad. of Sci. 38 (1952), 235-243.

[9] Gear, C. W.: The Automatic Integration of Stiff Ordinary Differential
    Equations. Information Processing 68 (Proceeding of the IFIP Congress
    1968) 187-193. North Holland Publishing Co., Amsterdam (1969).

[10] Gear, C. W.: The Automatic Integration of Ordinary Differential
     Equations. Comm. of the ACM 14 (1971), 176-179.

[11] Rutishauser, H.: Bemerkungen zur Numerischen Integration Gewöhnlicher
     Differential Gleichunger n-ter Ordnung. Numerische Mathematik 2 (1960),
     263-279.

[12] Krogh, F. T.: The Numerical Integration of Stiff Differential Equations. TRW Report No. 99900-6573-R000, TRW Systems, Redondo Beach, Calif. (March 1968).

[13] Klopfenstein, R. W.: Numerical Differentiation Formulas for Stiff Systems of Ordinary Differential Equations. RCA Review 32, (1971), 447-462.

[14] Shampine, L. F. and Gordon, M.K.: Local Error and Variable Order, Variable Step Adams Codes. (Submitted to the SIAM Journal on Numerical Analysis.)

[15] Krogh, F. T.: VODQ/SVDQ/DVDQ - Variable Order Integrators for the Numerical Solution of Ordinary Differential Equations. TU Doc. No. CP-2308, NPO-11643, Jet Propulsion Laboratory, Pasadena, Calif. (May 1969).

[16] Krogh, F. T.: On Testing a Subroutine for the Numerical Integration of Ordinary Differential Equations. JPL Internal Document, Section 314 TM No. 217 (revised), Jet Propulsion Laboratory, Pasadena, Calif. (Oct 1970). (To appear in the Journal of the ACM.)

[17] Hull, T. E., Enright, W. H., Fellen, B. M. and Sedgwick, A. E.: Comparing Numerical Methods for Ordinary Differential Equations. SIAM J. Numerical Analysis 9 (1972), 603-637.

[18] Gear, C. W.: The Numerical Integration of Ordinary Differential Equations. Math. of Comp. 21 (1967), 146-156.

[19] Thomas, L. H.: The Integration of Ordinary Differential Systems. The Ohio State University Engineering Experiment Station News 24 (1952), 8-9, 31-32.

[20] Krogh, F. T.: A Method for Simplifying the Maintenance of Software Which Consists of Many Versions. JPL Internal Document, Section 914 TM No. 314, Jet Propulsion Laboratory, Pasadena, Calif. (Sept. 1972).

[21] Krogh, F. T.: An Integrator Design. JPL TM No. 33-479, Jet Propulsion Laboratory, Pasadena, Calif. (May 1971).

[22] Krogh, F. T.: Opinions on Matters Connected with the Evaluation of Programs and Methods for Integrating Ordinary Differential Equations. SIGNUM Newsletter 7 No. 3 (Oct. 1972).

[23] Krogh, F. T.: A Test for Instability in the Numerical Solution of Ordinary Differential Equations. Journal of the ACM 14 (1967), 351-354.

SUMMARY OF TEST RESULTS WITH VARIANTS
OF A VARIABLE ORDER ADAMS METHOD*

Fred T. Krogh
Jet Propulsion Laboratory
Pasadena, California
91103

A variable order Adams method can be implemented and/or used in many differ-
ent ways. In the process of developing a variable order Adams code which uses
modified divided differences for changing stepsize, see [1] or [2], we have had
occasion to explore the effect of four parameters on the performance of the pro-
gram. The effect of these parameters was studied for test problems 8 and 9 of
[3]. Both are simple two-body problems, one with circular motion, and the other
traces out an ellipse with eccentricity .6. For each of these problems we have
examined all possible combinations of two distinct possibilities for each of the
parameters at requested error tolerances of 10, 1, .1, ..., 10^{-20} (2x16x22
cases). The choices examined are given below.

I.   Use a PECE or a PEC Adams method, always with a corrector that
        has order one greater than the predictor. (It is easy to show for such methods
        that a PEC method is equivalent to a PE method, where the P is the same order
        as the C in the PEC method.)

II.   Use either (2, 1/2) or (9/8, 7/8) as nominal factors for changing the
        stepsize. The program allows other values, but these values are probably
        extreme cases. The program does not restrict itself to the nominal factors
        when changing the stepsize, but factors closer to one do tend to cause more
        frequent changes in stepsize, more integration overhead, and fewer function
        evaluations.

III.   Solve the 2nd order equations directly, or solve them by breaking them
        up into an equivalent system of first order equations.

IV.   Use the same integration orders for all equations or select the inte-
        gration orders independently for each equation.

The results given for DVDQ on these problems in [3] were obtained using a PECE method, (2, 1/2) as the only factors for changing the stepsize, solving the 2nd order equations directly, and selecting integration orders separately for different equations. In its current state, and with options as close to DVDQ as possible, the integrator used for the comparisons here was of approximately the same efficiency for the case of circular motion (although the global error was a more regular function of the local error tolerance as a result of not restricting changes in the stepsize to the nominal factors), and approximately 10-20% more efficient in the case of elliptic motion (which we attribute primarily to the different procedure for changing stepsize). Performance of the new integrator would have been somewhat better if we had allowed it as large a maximum integration order as DVDQ. We plan to investigate this (and some other things) in more detail at a later time.

We summarize below the results of our tests. Generalizations to other problems are risky of course.

1.   There was no appreciable difference between using the same and using different integration orders. We believe the primary advantage of using different integration orders comes when integrating equations with different characteristics. And we have encountered problems where one is better off requiring the integration orders to be the same on certain equations.

2.   Integrating the second order equations directly is always best (of course, problems are known for which this is not true, see e.g. [4]); in the case of circular motion better by a factor of over two, except for the PECE methods at low accuracy where the factor is about 1.5. In the case of the elliptic motion results are somewhat similar, except that for the PECE methods the direct integration offers only a small advantage. We attribute the results to the better stability characteristics of the direct integration of second order equations on these problems. The cases where the differences are not great are those where discretization error is the primary factor limiting the stepsize.

3.   It is not as well known as it should be that the primary advantage of doing two derivative evaluations per step is due to improved stability characteristics and not to a smaller error term. Thus on the elliptic problem when integrating 2nd order equations directly, the PEC method is nearly twice as efficient as the PECE method. (The implications of this for practical problems is not highly significant, since in most problems of this type a predict-partial-correct scheme should be used. Such a scheme computes the main part of the derivative twice

per step, and perturbing forces, which typically are significantly more complicated, only after predicting. The values computed after predicting are used in the computation of the derivatives after correcting.) But there are implications for the comparisons of methods, since many methods don't offer the partial correct option. If a comparison is to be made with the best Adams method on such problems, one should compare with a PEC method integrating the second order equations directly. We have obtained similar results on the restricted 3-body problem in [3].

4.   Changing stepsize by factors close to one results in a definite reduction (10-20%) in the number of derivative evaluations for the problem which calls for a wide variation in the stepsize. When derivatives evaluations are expensive, factors close to one appear to be a good idea, but when derivative evaluations are cheap, the factors (2, 1/2) seem to be a reasonable choice to reduce the integration overhead. There is not much difference for the case of circular motion, but there is also not much difference in overhead since, except for extreme values of the error tolerance, the integrator tends to use a constant stepsize ultimately. (This would not be true for a problem which is best integrated with slight variation in the stepsize, however.)

References

1.   Krogh, Fred T., Changing stepsize in the integration of differential equations using modified divided differences. Proceedings of the Conference on the Numerical Solution of Ordinary Differential Equations, pp. 22-71, Lecture Notes in Mathematics No. 362, Springer-Verlag, New York, 1974.

2.   Shampine, L. F. and Gordon, M. K., Computer Solution of Ordinary Differential Equations, The Initial Value Problem. W. H. Freeman and Co., San Francisco, 1975.

3.   Krogh, Fred T., On testing a subroutine for the numerical integration of ordinary differential equations. J. ACM 20 (1973), pp. 545-562.

4.   Krogh, Fred T., A variable step variable order multistep method for the numerical solution of ordinary differential equations. Information Processing 68, pp. 194-199, North Holland Publishing Co., Amsterdam, 1969.

APPENDIX: NUMERICAL RESULTS

TWO-BODY PROBLEM, CIRCULAR MOTION.
WITH AN ABSOLUTE ERROR TEST
RESULTS ON THE INTERVAL (0, 16 PI)

LOG OF ERROR TOLERANCE				-6			-9			-12	
ND/		EQ	INT	ABSOLUTE			ABSOLUTE			ABSOLUTE	
STEP	HINC	ORD	ORD	ERROR	ND		ERROR	ND		ERROR	ND
2	2	2	SAME	(-5)2 0	299		(-8)2 0	409		(-12)6 7	626
2	2	2	DIFF	(-5)2 7	303		(-9)3 6	455		(-13)4 8	682
2	2	1	SAME	(-4)2 5	491		(-8)2 5	811		(-10)1 0	1347
2	2	1	DIFF	(-4)3 1	493		(-8)5 3	811		(-10)2 4	1229
2	9/8	2	SAME	(-5)4 3	279		(-8)5 0	409		(-12)8 0	610
2	9/8	2	DIFF	(-5)3 3	303		(-8)2 1	409		(-12)8 6	612
2	9/8	1	SAME	(-4)4 0	461		(-8)2 5	811		(-10)2 7	1235
2	9/8	1	DIFF	(-4)3 5	485		(-8)4 6	779		(-10)1 6	1258
1	2	2	SAME	(-6)9 0	269		(-8): 6	410		(-12)2 7	666
1	2	2	DIFF	(-6)1 5	301		(-8)2 3	404		(-12)9 3	612
1	2	1	SAME	(-4)2 0	777		(-7)2 6	1476		(-10)4 4	2515
1	2	1	DIFF	(-4)1 8	743		(-8)6 9	1422		(-10)2 2	2503
1	9/8	2	SAME	(-5)1 3	252		(-8)3 2	413		(-12)2 0	619
1	9/8	2	DIFF	(-5)2 2	248		(-8)2 9	390		(-12)9 4	607
1	9/8	1	SAME	(-5)8 0	807		(-7)1 8	1505		(-9)1 5	2488
1	9/8	1	DIFF	(-4)1 6	746		(-8)3 3	1441		(-10)2 9	2440
DVDQ				(-8)4 0	411		(-9)2 0	469		(-14)4 5	857
DVDQ (@-5,-5,-11)				(-4)3 4	317		(-9)2 2	442		(-10)1 0	726

TWO-BODY PROBLEM, ECCENTRICITY= 6.
WITH AN ABSOLUTE ERROR TEST
RESULTS ON THE INTERVAL (0, 16 PI)

LOG OF ERROR TOLERANCE				-6			-9			-12	
ND/		EQ	INT	ABSOLUTE			ABSOLUTE			ABSOLUTE	
STEP	HINC	ORD	ORD	ERROR	ND		ERROR	ND		ERROR	ND
2	2	2	SAME	(-3)1 5	1557		(-7)8 5	2655		(-10)5 2	3935
2	2	2	DIFF	(-4)2 9	1637		(-7)5 1	2698		(-10)3 4	3909
2	2	1	SAME	(-3)3 1	3002		(-7)9 3	2776		(-10)7 7	4623
2	2	1	DIFF	(-3)1 4	1670		(-6)1 0	2768		(-9)1 2	4388
2	9/8	2	SAME	(-3)2 7	1413		(-6)1 6	2150		(-10)3 5	3393
2	9/8	2	DIFF	(-3)1 8	1456		(-6)1 2	2195		(-10)4 9	3401
2	9/8	1	SAME	(-3)3 3	1416		(-6)2 4	2372		(-9)1 4	3817
2	9/8	1	DIFF	(-3)2 4	1500		(-6)1 5	2415		(-10)7 7	3730
1	2	2	SAME	(-4)9 5	819		(-7)1 9	1379		(-10)2 1	2127
1	2	2	DIFF	(-4)2 2	843		(-7)4 6	1425		(-10)3 5	2111
1	2	1	SAME	(-4)4 9	1318		(-7)9 5	2664		(-10)2 3	4865
1	2	1	DIFF	(-3)1 9	1221		(-6)1 6	2729		(-9)2 0	4716
1	9/8	2	SAME	(-?)1 1	722		(-7)4 7	1177		(-11)5 5	1820
1	9/8	2	DIFF	(-4)2 0	732		(-7)6 6	1209		(-10)1 5	1828
1	9/8	1	SAME	(-4)5 9	1284		(-6)1 3	2661		(-10)3 3	4935
1	9/8	1	DIFF	(-4)8 6	1179		(-6)1 8	2426		(-9)1 6	4435
DVDQ				(-4)2 3	2021		(-7)2 8	3131		(-10)6 3	4306

PROB. 8.   TWO BODY PROBLEM, CIRCULAR MOTION,
           WITH AN ABSOLUTE ERROR TEST

PECE METHOD, HF (2, 1, 2), ODE ORDERS 2, SAME INTEG. ORDERS

PROB. 8   TWO-BODY PROBLEM, CIRCULAR MOTION,
          WITH AN ABSOLUTE ERROR TEST

PROB. 8   TWO-BODY PROBLEM, CIRCULAR MOTION,
          WITH AN ABSOLUTE ERROR TEST

PROB. 8   TWO-BODY PROBLEM, CIRCULAR MOTION,
          WITH AN ABSOLUTE ERROR TEST

PROB. 8   TWO-BODY PROBLEM, CIRCULAR MOTION,
          WITH AN ABSOLUTE ERROR TEST

PROB. 8   TWO-BODY PROBLEM, CIRCULAR MOTION,
          WITH AN ABSOLUTE ERROR TEST

PROB. 8   TWO-BODY PROBLEM, CIRCULAR MOTION,
          WITH AN ABSOLUTE ERROR TEST

TWO-BODY PROBLEM, CIRCULAR MOTION
WITH AN ABSOLUTE ERROR TEST

TWO-BODY PROBLEM, CIRCULAR MOTION
WITH AN ABSOLUTE ERROR TEST

Left column

PECE METHOD. NF=4+6.7/8). ODE ORDERS=1. SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 8

TWO-BODY PROBLEM, CIRCULAR MOTION.
WITH AN ABSOLUTE ERROR TEST

PECE METHOD. HF=(9/8,7/8). ODE ORDERS=2. DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 8

TWO-BODY PROBLEM, CIRCULAR MOTION.
WITH AN ABSOLUTE ERROR TEST

PEC METHOD. NF=(9/8,7/8). ODE ORDERS=1. SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 8

TWO-BODY PROBLEM, CIRCULAR MOTION.
WITH AN ABSOLUTE ERROR TEST

PEC METHOD. HF=(9/8,7/8). ODE ORDERS=2. DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 8

Right column

PECE METHOD. NF=(9/8,7/8). ODE ORDERS=1. SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 8

TWO-BODY PROBLEM, CIRCULAR MOTION.
WITH AN ABSOLUTE ERROR TEST

PECE METHOD. NF=(9/8,7/8). ODE ORDERS=2. DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 8

TWO-BODY PROBLEM, CIRCULAR MOTION.
WITH AN ABSOLUTE ERROR TEST

PEC METHOD. NF=(9/8,7/8). ODE ORDERS=1. SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 8

TWO-BODY PROBLEM, CIRCULAR MOTION.
WITH AN ABSOLUTE ERROR TEST

PEC METHOD. HF=(9/8,7/8). ODE ORDERS=1. DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

WITH AN ABSOLUTE ERROR TEST

PECE METHOD, NF=(2, 1/3), ODE ORDERS=2, SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = 6,
WITH AN ABSOLUTE ERROR TEST

PECE METHOD, NF=(2, 1/2), ODE ORDERS=2, DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 9.   TWO-BODY PROBLEM, ECCENTRICITY = 6,
WITH AN ABSOLUTE ERROR TEST

PEC METHOD, NF=(2, 1/2), ODE ORDERS=2, SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 9.   TWO-BODY PROBLEM, ECCENTRICITY = 6,
WITH AN ABSOLUTE ERROR TEST

PEC METHOD, NF=(2, 1/2), ODE ORDERS=2, DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

WITH AN ABSOLUTE ERROR TEST

PECE METHOD, NF=(2, 1/2), ODE ORDERS=1, SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = 6,
WITH AN ABSOLUTE ERROR TEST

PECE METHOD, NF=(2, 1/2), ODE ORDERS=1, DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = 6,
WITH AN ABSOLUTE ERROR TEST

PEC METHOD, NF=(2, 1/2), ODE ORDERS=1, SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = 6,
WITH AN ABSOLUTE ERROR TEST

PEC METHOD, NF=(2, 1/2), ODE ORDERS=1, DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,6 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND

Left column

PROB 9   TWO-BODY PROBLEM; ECCENTRICITY = .6.
WITH AN ABSOLUTE ERROR TEST   -8-

PECE METHOD, NF=(9/8,7/8), ODE ORDERS=2, SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,4 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND
1	(111.4	11	(134.1	15	(231.1	19
0	(111.1	14	(135.6	20	(231.5	38
-1	(231.2	50	(032.6	105	(032.6	434
-2	(-131.0	95	(-135.0	207	(012.5	714
-3	(-119.2	105	(-135.3	309	(031.2	828
-4	(-415.5	124	(-234.9	389	(-132.1	908
-5	(-516.8	158	(-335.5	442	(-235.1	1178
-6	(-813.5	182	(-435.5	530	(-332.7	1413
-7	(-711.5	211	(-532.3	617	(-639.7	1630
-8	(-811.4	245	(-739.0	712	(-632.0	1874
-9	(-1014.8	279	(-719.2	809	(-611.6	2150
-10	(-1014.8	325	(-935.3	905	(-833.3	2497
-11	(-711.2	380	(-1033.0	1105	(-935.2	2915
-12	(-1213.0	441	(-1135.3	1283	(-1013.5	3219
-13	(-1312.5	517	(-1235.0	1494	(-1235.7	3941
-14	(-1413.0	592	(-1435.0	1724	(-1321.5	4558
-15	(-1413.4	680	(-1435.5	2014	(-1418.0	5288
-16	(-1413.8	932	(-1434.5	1994	(-1531.5	5144
-17	(-1412.5	933	(-1434.4	2069	(-1413.5	5418
-18	(-1412.2	741	(-1531.1	1943	(-1235.5	4905
-19	(-1411.5	793	(-1435.1	1819	(-1415.5	5305
-20	(-1412.8	750	(-1436.0	1940	(-1312.5	4946

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = .6.
WITH AN ABSOLUTE ERROR TEST

PECE METHOD; NF=(9/8,7/8), ODE ORDERS=2; DIFF. INTEG. ORDERS

INTERVAL	(0,2 PI)		(0,4 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND
1	(131.6	11	(134.1	13	(231.1	19
0	(135.6	14	(135.6	20	(231.5	30
-1	(031.5	49	(032.5	203	(035.3	402
-2	(-238.5	79	(011.5	201	(012.5	777
-3	(-531.3	119	(-234.8	309	(-118.4	827
-4	(-435.7	123	(-231.7	389	(-135.2	1060
-5	(-535.1	149	(-412.0	447	(-434.5	1190
-6	(-735.0	191	(-431.9	536	(-331.6	1496
-7	(-735.4	209	(-531.5	619	(-635.6	1685
-8	(-935.3	245	(-732.3	725	(-639.7	1917
-9	(-938.7	284	(-738.8	807	(-631.8	2195
-10	(-1030.6	331	(-938.0	901	(-938.3	2547
-11	(-1235.0	380	(-1139.0	1190	(-934.3	2988
-12	(-1234.6	445	(-1235.5	1289	(-1014.0	3401
-13	(-1331.6	517	(-1235.9	1489	(-1134.5	3935
-14	(-1435.8	587	(-1435.8	1719	(-1334.6	4557
-15	(-1431.5	679	(-1434.8	1989	(-1331.8	5289
-16	(-1438.5	765	(-1432.5	2053	(-1338.0	5985
-17	(-1431.8	745	(-1531.5	2030	(-1334.6	5285
-18	(-1431.6	735	(-1531.4	1981	(-1334.9	4989
-19	(-1431.1	721	(-1412.0	2085	(-1411.9	5411
-20	(-1431.8	710	(-1531.8	1914	(-1333.5	4982

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = .6.
WITH AN ABSOLUTE ERROR TEST

PEC METHOD, NF=(9/8,7/8), ODE ORDERS=2; SAME INTEG. ORDERS

INTERVAL	(0,2 PI)		(0,4 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND
1	(131.4	8	(134.7	15	(231.1	12
0	(131.0	11	(135.6	14	(231.5	19
-1	(031.2	34	(031.3	1203	(032.3	4195
-2	(-234.2	43	(031.5	132	(032.5	302
-3	(-537.2	56	(-136.5	161	(031.7	451
-4	(-331.1	70	(-234.7	197	(-131.5	568
-5	(-431.5	85	(-335.1	215	(-338.9	685
-6	(-832.2	98	(-412.5	275	(-331.1	728
-7	(-534.8	135	(-535.6	329	(-431.4	861
-8	(-814.9	135	(-638.9	383	(-618.5	1009
-9	(-1016.4	180	(-833.6	450	(-514.7	1197
-10	(-1011.0	194	(-933.0	482	(-918.8	1378
-11	(-1133.0	214	(-1014.6	607	(-814.9	1589
-12	(-1215.5	245	(-1239.9	695	(-1115.5	1828
-13	(-1332.1	268	(-1234.3	795	(-1218.4	2081
-14	(-1418.3	314	(-1338.5	692	(-1831.5	2538
-15	(-1431.5	364	(-1531.1	1028	(-1535.0	2680
-16	(-1435.9	398	(-1414.0	1012	(-1415.1	2587
-17	(-1412.1	366	(-1412.2	900	(-1531.3	2518
-18	(-1432.8	416	(-1417.3	1198	(-1532.0	3046
-19	(-1438.5	394	(-1413.2	1034	(-1432.8	2684
-20	(-1431.5	368	(-1331.1	1003	(-1331.5	2136

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = .6.
WITH AN ABSOLUTE ERROR TEST

PEC METHOD; NF=(9/8,7/8), ODE ORDERS=2; DIFF. INTEG. ORDERS

INTERVAL	(0,2 PI)		(0,4 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND
1	(131.4	8	(134.7	15	(211.1	12
0	(131.0	11	(135.6	14	(211.5	19
-1	(011.7	35	(012.6	119	(012.3	401
-2	(-718.0	47	(031.3	139	(012.6	401
-3	(-312.5	55	(-131.3	154	(-111.7	411
-4	(-412.5	70	(-315.7	205	(-311.7	540
-5	(-538.4	93	(-331.4	254	(-319.7	652
-6	(-615.9	98	(-431.9	276	(-418.0	758
-7	(-737.3	115	(-637.2	327	(-815.3	899
-8	(-938.5	137	(-612.1	398	(-515.1	1024
-9	(-534.8	160	(-815.7	460	(-716.6	1001
-10	(-1025.5	186	(-914.9	530	(-817.5	1354
-11	(-1135.4	815	(-1015.5	612	(-811.4	1591
-12	(-1231.9	244	(-1412.0	691	(-1011.4	1820
-13	(-1335.1	258	(-1332.0	797	(-1815.5	2083
-14	(-1331.0	318	(-1312.6	834	(-1519.5	2541
-15	(-1438.9	371	(-1416.8	945	(-1535.2	2548
-16	(-1438.9	391	(-1416.0	1040	(-1533.0	2733
-17	(-1432.7	381	(-1416.4	1068	(-1490.0	2733
-18	(-1432.6	369	(-1416.3	1057	(-1531.5	2751
-19	(-1432.8	398	(-1413.5	1054	(-1512.5	2705

Right column

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = .6.
WITH AN ABSOLUTE ERROR TEST

PECE METHOD, NF=(9/8,7/8), ODE ORDERS=1, SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,4 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND
1	(030.7	35	(135.4	63	(136.5	163
0	(011.9	85	(032.5	303	(012.5	859
-1	(-114.3	92	(012.7	212	(032.9	1200
-2	(-234.1	88	(011.3	257	(012.0	620
-3	(-312.7	103	(-114.1	301	(031.9	858
-4	(-435.1	126	(-231.0	369	(-131.3	981
-5	(-538.3	154	(-331.3	451	(-223.0	1194
-6	(-638.7	187	(-438.5	534	(-333.5	1416
-7	(-734.6	224	(-513.5	642	(-432.9	1583
-8	(-639.9	266	(-613.3	758	(-532.6	1990
-9	(-918.4	316	(-712.8	906	(-632.9	2372
-10	(-1030.9	370	(-812.0	1060	(-712.0	2782
-11	(-1113.2	437	(-813.6	1215	(-811.5	3870
-12	(-1210.7	509	(-1011.1	1437	(-931.6	3819
-13	(-1314.1	589	(-1111.7	1695	(-1019.5	4461
-14	(-1412.5	666	(-1211.1	1988	(-1239.4	5212
-15	(-1413.8	984	(-1416.6	2384	(-1413.5	6100
-16	(-1411.9	889	(-1435.2	2949	(-1336.5	6107
-17	(-1432.8	857	(-1432.2	2418	(-1831.1	6200
-18	(-1431.5	831	(-1431.9	2324	(-1915.5	5809
-19	(-1513.0	909	(-1409.0	2289	(-1333.6	5751
-20	(-1431.9	889	(-1493.6	2394	(-1539.0	6302

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = .6.
WITH AN ABSOLUTE ERROR TEST

PECE METHOD; NF=(9/8,7/8), ODE ORDERS=1, DIFF INTEG ORDERS

INTERVAL	(0,2 PI)		(0,4 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND
1	(030.7	21	(132.6	37	(130.8	95
0	(031.5	58	(031.9	204	(035.8	454
-1	(011.9	36	(031.3	142	(012.5	712
-2	(-817.0	100	(-817.0	339	(011.5	873
-3	(-311.5	142	(-216.4	411	(-112.1	1058
-4	(-331.0	173	(-3.0	483	(-219.6	1240
-5	(-432.6	196	(-415.8	503	(-312.0	1500
-6	(-531.5	230	(-311.5	660	(-431.4	1725
-7	(-938.6	263	(-735.0	967	(-532.3	2065
-8	(-932.6	338	(-736.9	871	(-611.5	2415
-9	(-635.7	368	(-832.0	976	(-711.4	2902
-10	(-1039.9	448	(-914.4	1240	(-819.0	3307
-11	(-1078.4	557	(-1038.5	1436	(-1037.7	3550
-12	(-1111.5	604	(-1133.1	1670	(-1133.2	4356
-13	(-1519.0	686	(-1231.6	1934	(-1219.5	5030
-14	(-1518.7	795	(-1135.6	2251	(-1531.0	5869
-15	(-1435.6	930	(-1433.0	2230	(-1431.4	5761
-16	(-1431.2	868	(-1331.1	2329	(-1539.0	5042
-17	(-1431.4	835	(-1531.4	2501	(-1531.0	5623
-18	(-1431.4	799	(-1331.4	1835	(-1531.5	5817
-19	(-1536.8	404	(-1332.1	2150	(-1218.2	5540

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = .6.
WITH AN ABSOLUTE ERROR TEST

PEC METHOD; NF=(9/8,7/8); ODE ORDERS=1; DIFF. INTEG ORDERS

INTERVAL	(0,2 PI)		(0,4 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND
1	(030.6	15	(132.6	17	(136.7	43
0	(032.9	25	(038.0	30	(114.8	46
-1	(-130.5	35	(035.3	142	(032.6	592
-2	(-216.9	81	(-132.2	309	(031.6	565
-3	(-815.7	98	(-315.5	259	(-116.1	90
-4	(-511.0	123	(-331.6	349	(-211.4	918
-5	(-412.0	155	(-434.9	450	(-418.0	1179
-6	(-519.3	202	(-613.7	575	(-615.9	1504
-7	(-912.6	353	(-635.0	730	(-615.8	1986
-8	(-931.3	307	(-713.0	927	(-611.0	2306
-9	(-911.5	434	(-833.3	373	(-418.0	3025
-10	(-1019.9	511	(-931.9	1393	(-818.7	3723
-11	(-1032.3	507	(-1035.7	1474	(-911.0	4033
-12	(-1136.6	712	(-1132.5	2067	(-1114.3	5594
-13	(-1316.6	832	(-1338.0	2445	(-1111.9	6458
-14	(-1312.5	988	(-1538.7	2893	(-1831.1	5759
-15	(-1413.5	1154	(-1331.7	3413	(-1312.3	5040
-16	(-1413.5	1105	(-1331.1	3254	(-1515.1	5625
-17	(-1413.7	1080	(-1531.9	3260	(-1515.5	5838
-18	(-1413.1	1097	(-1439.0	3232	(-1312.5	5579
-19	(-1413.5	1080	(-1416.9	3205	(-1933.7	5507

PROB 9   TWO-BODY PROBLEM, ECCENTRICITY = .6.
WITH AN ABSOLUTE ERROR TEST

PEC METHOD; NF=(9/8,7/8), ODE ORDERS=1; SAME INTEG ORDERS

INTERVAL	(0,2 PI)		(0,4 PI)		(0,16 PI)	
	ABSOLUTE		ABSOLUTE		ABSOLUTE	
LOG(TOL)	ERROR	ND	ERROR	ND	ERROR	ND
1	(030.6	16	(132.4	35	(136.8	53
0	(-138.1	56	(032.5	107	(035.6	489
-1	(-138.1	44	(-119.3	150	(035.3	614
-2	(-131.6	51	(-132.9	188	(012.9	428
-3	(-818.0	60	(-134.3	199	(032.1	538
-4	(-514.3	82	(-238.8	267	(-118.5	700
-5	(-435.6	121	(-436.6	349	(-811.5	914
-6	(-438.1	190	(-538.5	344	(-413.5	1204
-7	(-632.8	154	(-512.8	417	(-615.7	1501
-8	(-711.7	375	(-638.2	775	(-315.5	2101
-9	(-534.6	160	(-813.1	468	(-518.0	2401
-10	(-1025.5	203	(-915.6	586	(-711.9	3126
-11	(-1013.7	494	(-1017.1	1445	(-1019.9	3846
-12	(-1133.2	626	(-1140.0	1965	(-1033.3	4598
-13	(-1335.5	745	(-1279.2	2205	(-1139.0	5501
-14	(-1312.8	560	(-1436.7	2487	(-1933.3	3822
-15	(-1413.8	1059	(-1435.7	3715	(-1..0	10110
-16	(-1417.7	1363	(-1431.7	3560	(-1..5	9651
-17	(-1413.3	1183	(-1439.3	3364	(-1338.2	9058
-18	(-1413.9	1229	(-1417.5	3608	(-1933.7	9609
-19	(-1431.9	1823	(-1412.3	3400	(-1935.7	9597

AN L^2 APPROACH TO R-MATRIX PROPAGATION

Robert B. Walker and Barry I. Schneider
Theoretical Division
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87545

ABSTRACT

Standard R-matrix methods are used to compute a series of Wigner R-matrices each covering a small region of the scattering coordinate for a model atom-diatom collinear vibrational excitation problem. The R-matrices so obtained are then combined using the R-matrix propagation algorithm. We present here additional data on the performance of this method.

I.   INTRODUCTION

In an earlier publication,[1] we described an L^2 approach to the solution of a heavy particle scattering problem; normally, such problems are attacked through the solution of the so-called close coupled equations. These equations are obtained when a basis set expansion is introduced to described all degrees of freedom in the scattering wavefunction except one; motion in the final degree of freedom, represented by the scattering coordinate, is determined by the numerical solution of the close coupled equations.

The method we presented differs from a standard close coupling approach in that the final degree of freedom is also represented by a basis set expansion in the total scattering wavefunction. This method is based on the R-matrix formalism introduced over thirty years ago by Wigner.[2] In this approach, one imagines enclosing configuration space in a "box". This box has to be big enough so that outside the box there is no residual interaction between the scattering particles. Inside the box, the wavefunction is determined by L^2 methods, and outside the box, the wavefunction is forced to fit an appropriate asymptotic form.

As described, the R-matrix approach has not been used extensively to treat the scattering of heavy particles in molecular force fields. The R-matrix method suffers from a practical disadvantage -- the range of the interaction potential is so large that the R-matrix box must be made big enough the encompass many be Broglie wavelengths associated with the scattering degree of freedom. Furthermore, in molecular scattering problems, there are typically many strongly coupled channels. Consequently,

the basis set expansion of the entire wavefunction is required to be impractically large. A step in the right direction was taken by Zvijac and Light,[3] when they presented equations which showed how two R-matrix boxes, convering adjoining regions of configuration space, could be "joined" by matching the scattering information implicitly contained on the surface common to the two boxes. The ability to piece together the information from two separate (but adjacent) boxes effectively solves the basis set problem associated with the translational degree of freedom -- because now the size of each R-matrix box is no longer constrained to encompass the full range of the interaction potential. This advantage has been exploited by Light and Walker,[4] who incorporated these R-matrix propagation equations into an algorithm for solving close-coupled equations. In their approach, which has subsequently undergone further refinement[5] and development,[6] the size of each R-matrix box is dictated by the form of the interaction potential within the box. The box is taken to be small enough so that higher order terms in the (power series) expansion of the interaction potential are negligible. In fact, in their simplest approach, the box is taken to be small enough that the interaction potential may be regarded as a constant within the box. For this simple form of interaction, the R-matrix associated with each box can be determined analytically. Once the R-matrix for each box (commonly referred to as a sector) is known, the sector R-matrices are assembled and a global R-matrix obtained. The global R-matrix is then used in the enforcement of scattering boundary conditions just as if it had been obtained in a standard L^2 approach.

In the L^2 R-matrix propagation approach, we follow the Light-Walker strategy in that the whole of configuration space is subdivided into smaller regions in each of which we determine an R-matrix. This step still solves the de Broglie wavelength problem. However, in the L^2 approach, the size of each R-matrix sector is not constrained by any approximation to the inter-action potential -- the full interaction potential is determined and used in a standard L^2 evaluation of the sector R-matrix. This approach has an advantage over the Light-Walker propagation scheme. Because we treat the full interaction potential exactly, we may use larger R-matrix boxes in regions of configuration space where the Light-Walker method is forced (by the form of the potential) to take small R-matrix sectors. In particular, we are thinking of the steep repulsive wall associated with all inelastic scattering systems. In this region, the rapid variation of the potential implies that the constant coupling approximation used in the simplest form of the Light-Walker method will be justified only if a large number of very small R-matrix sectors are used. One of the objectives of this project is to determine the circumstances under which we may expect the L^2 method to be competetive with the Light-Walker method.

In this paper, we extend the results of the previous paper (paper I) to more fully characterize the properties of the L^2 propagation approach. We review the results of paper I and present additional data on the dependence of the L^2 method on the number of channel functions included in the total wavefunction expansion. This data is presented in the next section, Section II. In Section II, we draw some tentative conclusions about the current approach.

II. PROPERTIES OF THE L^2 METHOD

A. Review of Paper I

We investigated the model collinear vibration excitation problem defined by Secrest and Johnson.[7] The parameters used were given in paper I. The interaction region was divided into equal sized R-matrix boxes (17 boxes from x=0 to x=51; each box has a width of 3). In each box a basis set consisting of products of channel functions and translational functions is constructed. The channel functions are the same as would be used in a standard close coupling expansion (harmonic oscillators). The translational functions are polynomials in x. In each box we construct and diagonalize the sum of the Hamiltonian and Bloch operators. For simplicity in the characterization of this method, each channel function has associated with it the same set of translational functions. In such a case, the matrix to be diagonalized is a real $(N_c N_{TF}) \times (N_c N_{TF})$ symmetric matrix, where N_c is the number of coupled channels, and N_{TF} is the number of translational functions per channel. It is the diagonalization of this matrix which requires the bulk of the computational time for the L^2 propagation method.

Figures 1 and 2 of this paper (which were Figs. 1 and 3 of I) present the essential features of the L^2 method, for a 5-channel example problem $(N_c=5)$. Figure 1 shows the CPU time required to apply the method as a function of the number of translational functions (N_{TF}) associated with each channel. The upper curve of Fig. 1 shows the "setup" time required; this is the time required to construct and diagonalize the Hamiltonian matrix. Once the eigenfunctions and eigenvalues associated with each box are known, we compute and assemble the sector R-matrices to form the global R-matrix. This procedure is repeated at each scattering energy; the time it requires is shown in Fig. 1 as the lower curve (which has been scaled by a factor of ten).

There are two significant features to notice about Fig. 1. First, compared to the R-matrix assembly time at each energy, the initial setup time is very large. Second, both times increase as the number of translational functions increases. The setup time increases more severely -- because of the matrix operations involved in the initial diagonalization stage of the calculation, we expect the setup time to eventually go as N_{TF}^3. On the other hand, the propagation time per energy tends to increase only linearly with N_{TF}.

It should be clear from Fig. 1 that our objective is to minimize number of translational functions which must be carried with each channel function, if we are to speed up the L^2 method. However, as shown by Fig. 2, the quality of the scattering information obtained depends severely on the number of translational functions. The higher in scattering energy we want to go, the more translational functions per channel are required. For example, Fig. 2 shows that with $N_{TF} = 5$, the scattering results are good only to about $E = 7$. With $N_{TF} = 6$, the results are satisfactory up to $E = 8$. To obtain results good to $E = 16$, we must have $N_{TF} = 9$. The conclusion is obvious -- as the scattering energy increases, the number of de Broglie wavelengths in each box also increases. In order to describe this behavior satisfactorily, we must increase N_{TF} accordingly.

Up to this point, we have reviewed the results of paper I. The remainder of this paper will consider additional properties of the L^2 method. In particular, we consider two additional effects -- the effect upon computation times of increasing the number of channel functions, and the effect of varying the widths of the R-matrix boxes.

B. Channel Size Effects

We have repeated calculations as in Fig. 1 but based upon 8, 10, and 16-channel wavefunction expansions. The number of translational functions required per channel depends only on the scattering energy and the size of the R-matrix box, and not on the strength of the potential or the number of strongly coupled channels. Consequently, the number of translational functions required per box is typically the same as in the 5-channel study presented in I. The results of this study are presented in Table I. As we would expect, the CPU effort increases as the number of channel functions increases. Because the majority of the computation effort is involved in the Hamiltonian diagonalization, we once again expect the setup time to go approximately as N_c^3 (with N_{TF} fixed). This point is emphasized in Fig. 3, where we plot (on a log-log scale) the product $N_c N_{TF}$ vs. setup CPU time per box. The figure supports the idea that all the data are effectively described by a single curve -- and the slope of the curve suggests that the setup time goes approximately as the 2.3 power of $N_c N_{TF}$.

C. Effect of Box Size

We repeated the calculations of the 5-channel problem, except that we doubled the total number of R-matrix boxes by halving each box width. In Fig. 4 we show how the scattering data now converge with respect to the number of translational functions per channel (in the smaller boxes). The figure shows clearly that the data converges much more rapidly with N_{TF} than before. Whereas it previously required $N_{TF} = 9$ to obtain good scattering data up to $E = 16$, we now obtain satisfactory results with N_{TF} between 5 and 6. Because the new calculation is required to diagonalize 34 (25x25) matrices

instead of 17 (45x45) matrices, the setup time associated with the current calculation is considerably smaller (2.2. CPU sec. instead of 3.9 CPU sec.). On the other hand, the time required for propagation at each scattering energy is larger (0.093 sec vs. 0.080 sec) because there are now a larger number of R-matrix boxes to be assembled. We therefore have offsetting effects in evaluating the total time required to complete an entire calculation — the setup time is smaller but the scattering times per energy are larger. However, because the setup time is so high in comparison to the scattering time, one can obtain scattering results at quite a large number of energies before the 34 box calculation costs as much as the 17 box calculation (130 energies).

The L^2 R-matrix propagation scheme is clearly capable of providing high quality dynamical information. The method has several strengths and weaknesses, which we will touch upon briefly.

To date we have determined only two significant weaknesses with this method -- (1) the method is typically more expensive than others, and (2) it does not lend itself to calculations requiring a very large number of coupled channels. We dealt here primarily with characteristics of the L^2 method affecting the first weakness. Looking at Fig. 2 of paper I, we ascertain that the L^2 method (at 5 channels) must be applied at approximately 12 energies before the time it requires equals the time required to solve the corresponding 5 coupled channel equations 12 times. We are reading off Fig. 2 of paper I where $N_{TF} = 9$, because we have determined in practic that $N_{TF} = 9$, is required in order to obtain good scattering data at all the energies of interest. Of course, if we had wanted transition probabilities at even higher energies, probably N_{TF} will increase and the L^2 method will compare even less favorably with the analytic method. Fixing $N_{TF} = 9$, we may also determine from Table I (and from Fig. 5 of Ref. 5) that this breakeven number increases as N_c increases (for the 17 box calculation). As N_c increases (N_c=5,8,10,16) the breakeven number increases also (N_B=12,14,16,18). This weakness of the L^2 method is further emphasized by the point alluded to at the end of the preceding section, where we determined that the L^2 method is actually faster if we do a 34 box calculation with $N_{TF} = 5$ instead of a 17 box, $N_{TF} = 9$ calculation. Pushing this strategy to its limit reduces to the analytic R-matrix scheme, where there are many boxes, and only one translational function per channel per box.

The second weakness of the L^2 method pertains to its behavior as N_c increases. The value of N_{TF} depends primarily on the box width, and so if we choose the box widths so that $N_{TF} \approx 5$ for accurate results, then we are committed to a method which must handle matrices 5 times larger than conventional close coupling methods handle. This restriction complicates the coding problem if very large systems (with many channels) are to be handled. For example, the L^2 method might require as much core for a 20 channel problem as a close coupling method would require for a 100-channel problem. As it is, the L^2 method is also expensive in IO charges, since all the box eigenfunctions must be read in from disk at each scattering energy. The IO expense will increase significantly as N_c increases.

Without being too negative about it, we should also point out the strengths of this method. Primarily, it does the one thing we at first wanted it to do -- it handles all regions of the potential with equal ease--including the steep repulsive wall. The L^2 method is actually very powerful because it is flexible. There are few built-in assumptions. The potential is treated exactly. It is easy to study the effect produced by varying the number or type of channel expansion being used. One can easily vary the number of translational functions used, and the type of translational basis expansion used. Probably, it is not necessary to have as many translational functions associated with deeply closed channels as there are with open channels. The use of a mixed trigonometric/polynomial basis should be studied, etc. Finally, the method could be an ideal tool for determining the relative importance of various perturbative corrections to the analytic R-matrix method.

REFERENCES

1. B. I. Schneider and R. B. Walker, "The Coupled Channel R-Matrix Propagation Method", J. Chem. Phys. 70, 2466 (1979).

2. E. P. Wigner, Phys. Rev. 70, 15, 606 (1946); E. P. Wigner and L. Eisenbud, ibid., 72, 29 (1947).

3. D. J. Zvijac and J. C. Light, Chem. Phys. 12, 237 (1976).

4. J. C. Light and R. B. Walker, J. Chem. Phys. 65, 4272 (1976).

5. E. B. Stechel, R. B. Walker, and J. C. Light, J. Chem. Phys. 69, 3518 (1978).

6. T. G. Schmalz, G. A. Parker, J. C. Light, to be published.

7. D. Secrest and B. R. Johnson, J. Chem. Phys. 45, 4556 (1966).
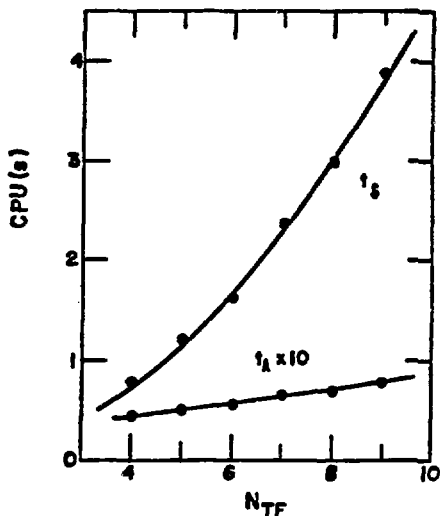
Fig. 1. Fig. 1 of Ref. 1. Approximate CPU time required for a five-channel target expansion, as a function of the number of translational basis functions N_{TF} used per channel function. The heavy line t_s shows the setup time required for the calculation, which consists of the time required to construct and diagonalize the Hamiltonian matrix in all the R-matrix boxes. The thin line t_A shows the additional CPU time required to construct the sector R matrices and assemble them into the global R matrix. Note that the actual CPU times for assembly are a factor of 10 smaller than shown in the graph.
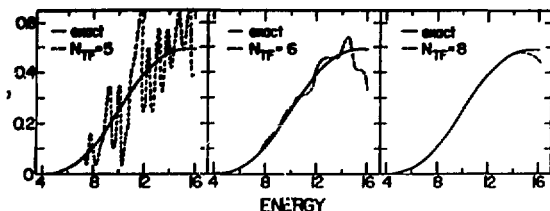
Fig. 2. Fig. 3 of Ref. 1. Effect of the number translational functions used per channel function (N_{TF}) on the accuracy of the current method as a function of scattering energy E. The $0 \rightarrow 1$ vibrational excitation probability is plotted on the ordinate. In each panel, the solid curve shows the exact calculation, obtained by the current method using $N_{TF} = 9$. The three panels show calculations using $N_{TF} = 5, 6,$ and 8. The current method gives better accuracy at higher energies as N_{TF} increases.
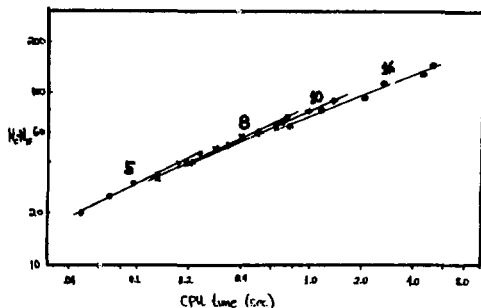

Fig. 3. A log-log plot of the dependence of the CPU setup time per box upon the product of N_c and N_{TF}. The four curves are labelled with the 5,8,10, and 16-channel calculations. The proximity of the curves suggests they are all effectively fit by the same function.
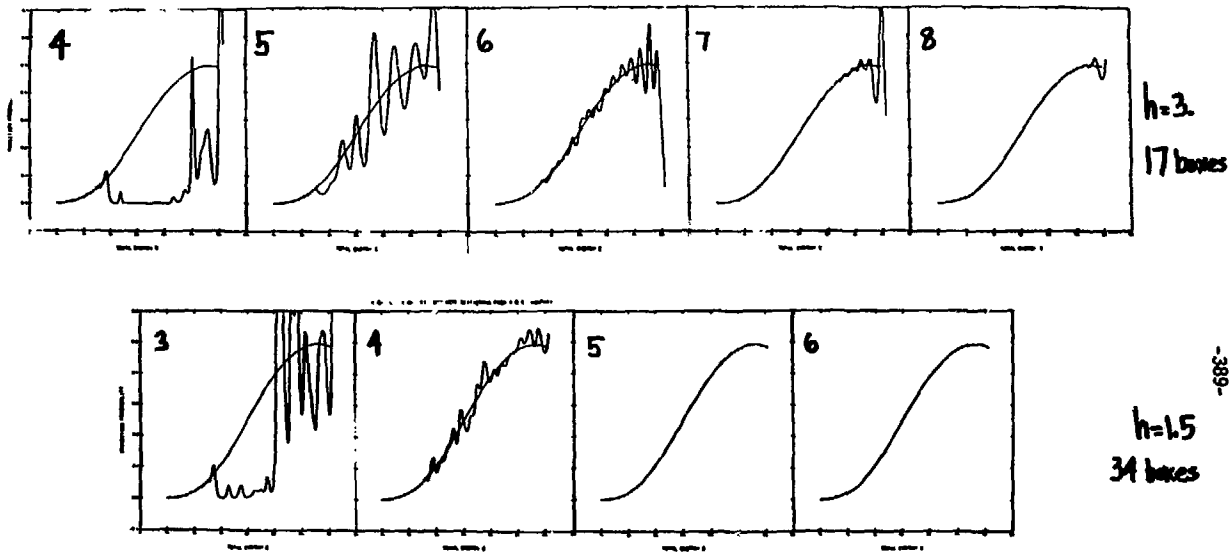
Fig. 4. Same as Fig. 2 above, except that the lower panel compares the
convergence with N_{TF} of the 34 box (h=1.5) calculation with the 17 box
(h=3.0) calculation. The 34 box N_{TF} = 5 calculation is only slightly poorer
in quality than the 17 box N_{TF} = 9 calculation.

Table I. Timings[*] for the L^2 R-Matrix Propagation Method

N_c N_{TF}	5	8	10	16
4	0.047(0.0025)	0.131(0.0071)	0.206(0.0122)	0.759(0.0439)
5	0.069(0.0029)	0.193(0.0081)	0.332(0.0140)	1.140(0.0509)
6	0.096(0.0033)	0.286(0.0090)	0.502(0.0164)	2.039(0.0546)
7	0.140(0.0038)	0.398(0.0100)	0.708(0.0178)	2.625(0.0616)
8	0.176(0.0042)	0.630(0.0110)	0.996(0.0197)	4.449(0.0680)
9	0.229(0.0047)	0.740(0.0123)	1.335(0.0214)	5.173(0.0752)

[*]All times are LASL 7600 CPU sec per R-matrix box. The primary entry is the setup time; the propagation time per box per energy is in parentheses.

A VARIATION-ITERATION METHOD FOR A
SINGLE COLUMN OF THE S-MATRIX

Lowell D. Thomas
NRCC
Lawrence Berkeley Laboratory
Berkeley, California 94720

I. Introduction

The details of this method and the results of a sample calculation with it have been recently published [1], and these will not be repeated here. I will instead discuss some of the motivation for this work, the original implementation of the Kato variational method [2], and finally the future plans for the method.

A comparison of the numerical effort involved in a quantum and a quasiclassical scattering calculation for atom-molecule scattering yields some interesting insights. Consider for the moment rotational excitation of a linear rigid rotor by collision with a structureless particle [3,4,5]. If the rotational levels of the rotor up to j=50 are retained in the expansion of the wave function, then the number of coupled equations, N, is

$$N \simeq 1/2 \ j^2 = 1250 \ .$$

If the potential energy function is expanded in λ_{max} Legendre polynomials, and we work in the body-fixed frame, then the interaction matrix coupling the radial Schroedinger equations will be a banded diagonal matrix which has at most a band width of $2\lambda_{max}+1$. If we choose $\lambda_{max}=2$ (often not enough, but often used in the literature), then for each column of the matrix of wave functions we must solve 1250 second-order differential equations coupled 5 at a time. By this I mean that even though all 1250 wave functions are indirectly coupled to one another, the matrix multiplication of the

interaction matrix times the solution matrix involves only 5 multiplications and 5 additions per element.

Let us now consider the analogous classical trajectory calculation. If solved in Cartesian coordinates, there are 6 independent variables for each trajectory. For a fixed impact parameter one would typically run 200 trajectories. Lagrange's equations for the 6 independent variables give 6 coupled second-order differential equations. The 200 trajectories therefore require the solution of 1200 second-order differential equations coupled 6 at a time. A simple counting of the number and couplings of the differential equations yields roughly the same numbers for the classical and quantum calculations. Of course there are many other differences between the two methods.

First of all it may be argued that the solutions in the two cases have very different characters. The classical trajectories are smooth while the wave functions are oscillatory. As one goes to higher energy this is of course true, but in fact at intermediate energies in the eV range, both the classical trajectories and the wave functions will require something like one- to a few-hundred integration steps.

It may also be argued that classical trajectories can be done one at a time, making them simpler to calculate. This does make them simpler to program and the programs require less computer storage space, but this fact does not reduce the actual number of numerical operations. One could, if one liked, run all 200 trajectories simultaneously with little increase in total computer time.

The real difference between the two methods seems to me to be the fact that the classical calculation can be done for a single initial state. All

trajectories begin with the rotor in the same initial state. The quantum
analogue of this would be to solve the coupled equations for only a single
column of the wave function matrix, and this is not possible in a direct
fashion due to the double-ended boundary conditions which must be imposed.

This kind of reasoning implies that close coupling calculations could
be done for most or at least many of the systems for which only classical
calculations have been possible, if we could somehow solve for a single
column of the wave function matrix. The equation to be solved is,

$$\underline{h}\vec{u} = \underline{V}\vec{u} , \qquad (1)$$

where $\underline{h}$ is a diagonal matrix of second-order differential operators, $\underline{V}$ is
the interaction coupling matrix, and $\vec{u}$ is the vector of solutions. The
numerical work in general increases as N^2, where N is the number of coupled
equations, due to the matrix multiplication. The standard procedure is to
solve for N linearly independent solution vectors and then take linear
combinations to get the proper boundary conditions. Hence, the overall
work increases with N^3.

An alternative approach would be to attempt to solve Eq. (1) iterative-
ly. That is, guess a solution vector, test to see if it satisfies Eq. (1),
improve the guess if not, and so on. This way the numerical work would
increase with MN^2, where M is the number of iterations. If a rapidly
convergent procedure could be found such that $M \ll N$, this would result in
a considerable savings. There are several good reasons to believe that
such a procedure can be found. The coupling matrix, as argued above, is
relatively sparse. Also, several approximate methods, most notably the
IOS [6] and coupled states [7] methods, which ignore much of the coupling

matrix have been very successful. One intuitively feels that there is not much physics left to be accounted for, and that if properly done only a few iterations would be necessary.

The next section is essentially a walk-through of the ideas originally used to develop this approach.

II. The Iterative Method

An iterative approach to the problem is much more transparent when Eq. (1) is written in its equivalent integral form,

$$u_\lambda(r) = x_{\lambda_0}(r)\delta_{\lambda\lambda_0} + \int_0^\infty G_\lambda(r,r') \sum_{\lambda'} V_{\lambda\lambda'}(r')u_{\lambda'}(r')dr' \quad (2)$$

where

$$h_\lambda \left\{ \begin{matrix} x_\lambda \\ y_\lambda \end{matrix} \right\} = 0, \quad (3a)$$

$$\lim_{r\to 0} x_\lambda(r) = 0 , \quad (3b)$$

$$x_\lambda y_\lambda' - x_\lambda' y_\lambda = 1 , \quad (3c)$$

$$h_\lambda G_\lambda(r,r') = \delta(r-r') , \quad (3d)$$

$$G_\lambda(r,r') = \begin{cases} x_\lambda(r)y_\lambda(r'), & r < r' , \\ x_\lambda(r')y_\lambda(r), & r > r' . \end{cases} \quad (3e)$$

Or, in matrix form,

$$\vec{u} = \underset{\sim}{x}\vec{\Delta} + \int \underset{\sim}{G}\underset{\sim}{V}\vec{u} \; dr' \tag{4}$$

where $\underset{\sim}{x}$, $\underset{\sim}{y}$, and $\underset{\sim}{G}$ are diagonal matrices and $\Delta_\lambda = \delta_{\lambda\lambda_0}$.

The idea now is to guess a trial function $\vec{g}^n$, use it to evaluate the right hand side of Eq. (4),

$$\vec{u}^n = \underset{\sim}{x}\vec{\Delta} + \int \underset{\sim}{G}\underset{\sim}{V}\vec{g}^n dr' \; , \tag{5}$$

and then to use $\vec{u}^n$ in some way to obtain an improved guess $\vec{g}^{n+1}$. Four ways have been tried and these will now be discussed. The test problem used for these experiments was the Ar-N_2 problem published by De Pristo and Alexander [8].

The oldest and most obvious iterative method is the Born series [9]. That is, let

$$\vec{g}^0 = \underset{\sim}{x}\vec{\Delta} \tag{6a}$$

and

$$\vec{g}^{n+1} = \vec{u}^n \; . \tag{6b}$$

This series did not converge for the test problem. Since this was a small problem with relatively weak coupling it is to be expected that the Born series will in general not converge for atom-molecule collision problems.

Another possibility suggests itself if we rewrite Eq. (2) as

$$u_\lambda(r) = x_\lambda(r)\{\delta_{\lambda\lambda_0} + A_\lambda(r)\} + y_\lambda(r)\{B_\lambda(0) - B_\lambda(r)\} \tag{7}$$

where

$$A_\lambda(r) = \int_r^\infty y_\lambda(r')W_\lambda(r')dr' \; , \tag{8}$$

$$B_\lambda(r) = \int_r^\infty x_\lambda(r')W_\lambda(r')dr' \ , \tag{9}$$

$$W_\lambda = \sum_{\lambda'} V_{\lambda\lambda'}u_{\lambda'} \ . \tag{10}$$

It was first pointed out by Sams and Kouri [10] that when a quadrature formula is introduced for the integrals of Eqs. (8) and (9), the terms involving $u_\lambda(r)$ at the end points of the integration exactly cancel. The unknown, $u_\lambda(r)$, may be evaluated from the known values, $u_\lambda(r_i)$, $r < r_i$. Therefore, given $\vec{B}(0)$ we can find the exact wave function $\vec{u}$. But unless $\vec{B}(0)$ is just the one we are looking for, $\vec{u}$ is not regular at the origin and the numerical integration becomes unstable in the non-classical region. What we want to try now is to guess the correct $\vec{B}(0)$ rather than the correct wave function.

One of the beauties of Eq. (5) is that the trial function, $\vec{g}^n$, need not have any special boundary conditions nor, for that matter, need it be continuous. As long as it is properly normalized (and hence regular at the origin) the computed function, $\vec{u}^n$, will be continuous and have the proper form (but not values) for the boundary conditions.

The second experiment takes the following form. Let

$$\vec{B}^n = \int_0^\infty \underset{\sim}{x}\underset{\sim}{V}\ \vec{g}^n \ dr \ , \tag{11}$$

$$\vec{v}^n = \underset{\sim}{x}\{\ \vec{\lambda} + \int_r^\infty \underset{\sim}{y}\underset{\sim}{V}\vec{v}^n dr\ \} + \underset{\sim}{y}\{B^{n-1} - \int_r^\infty \underset{\sim}{x}\underset{\sim}{V}\vec{v}^n dr\ \} \ , \tag{12}$$

$$\vec{g}^0 = \vec{0} \ , \tag{13}$$

$$g_\lambda^{n+1} = \begin{cases} v_\lambda^{n+1} \ , & r > r_\lambda^t \\ u_\lambda^n \ , & r < r_\lambda^t \ , \end{cases} \tag{14}$$

where r_λ^t is the classical turning point for the channel λ. What is being done is this. Guess $\vec{B}^0 = \vec{0}$, then integrate inwards for the exact wave function $\vec{v}^0$ (with the wrong boundary condition) until the classical turning point is reached, then set the wave function to zero. We now have a new value for $\vec{B}(0)$,

$$B_\lambda^1 = \int_{r_\lambda^t}^\infty \sum_{\lambda'} x_\lambda V_{\lambda\lambda'} v_{\lambda'}^0 \, dr \ , \tag{15}$$

and we start over again. For the second and higher iterations, however, instead of setting the wave function to zero at the classical turning point, we use u_λ^n from Eq. (5). Unfortunately, this rather elaborate scheme fares no better than the Born series and in retrospect probably a little worse.

It does, however, suggest another experiment and this one is successful. We still feel at this point that integrating inward from a trial $\vec{B}(0)$ is a good idea if we only had a better way to handle the non-classical region. Expansion in a basis set has a well-proven track record in these circumstances, and R-matrix theory [11] comes immediately to mind. It is, however, out of the question. The R-matrix is related to the S-matrix in a nonlinear way and, hence, a single column would do us no good. Besides that recall, we are thinking in terms of eventually solving a 1200 channel problem. Even with one basis function per channel, the eigenvalue problem associated with the R-matrix method becomes a formidable obstacle.

Instead, we look for a variational integral to help us out. In the early literature on the Schwinger variational integral [2,12,13,14] is a variational method by Kato [2] which is well suited to the problem at hand.

Kato's method is to use the wave functions from the previous iterations as basis functions. Therefore, we use the previous scheme except that in the non-classical region we use

$$g_\lambda^{n+1} = \sum_{i=0}^{n} c_i u_\lambda^n , \qquad r < r_\lambda^t . \tag{16}$$

The constraint

$$\sum_i c_i = 1 \tag{17}$$

is also necessary for $\vec{g}^{n+1}$ to be properly normalized. The Schwinger variational integral can only be applied in this case to the elastic, initial channel S- or T-matrix element [9]. Therefore, we relax the requirement that the variational integral be stationary to arbitrary variations in the wave function, and require instead only that it be stationary with respect to variations in the expansion coefficients, c_i, subject to the normalization constraint, Eq. (17). A convenient integral to work with is

$$I[\vec{u}] = \int_0^\infty |(\underline{h} - \underline{V})\vec{u}|^2 \, dr . \tag{18}$$

Therefore,

$$I[\vec{g}^{n+1}] = \sum_{ij} c_i^* J_{ij} c_j , \tag{19}$$

where

$$J_{ij} = \int (\underline{h} - \underline{V})\vec{u}^{i*} \cdot (\underline{h} - \underline{V})\vec{u}^j \, dr , \tag{20}$$

leading to

$$C_i = IC_i' ,\tag{21}$$

where

$$C_i' = \sum_j (\mathfrak{z}^{-1})_{ij}\tag{22}$$

and

$$I = \left(\sum_i C_i' \right)^{-1} .\tag{23}$$

Note that because of Eq. (14) the integration really extends only over the non-classical region.

This is in fact a convergent procedure for the Ar-N_2 test problem. Note, however, that this variational method could just as well have been applied straight away to the Born series. That is, let

$$\vec{g}^{n+1} = \sum_{i=0}^{n} C_i \vec{u}^i , \qquad 0 < r < \infty .\tag{24}$$

This then is the fourth method and it does converge more rapidly than the previous scheme. It is also much easier to write down and explain, but the difference in the computer codes between these last two methods really only amounts to a change in about four statements.

III. Future Plans

Kato's variational method is the essential ingredient to a convergent iterative method for a single column of the wave function matrix. It is by itself, however, not enough. The Ar-N_2 test case is a 16 channel problem. The number of iterations necessary for reasonable convergence of the variationally modified Born series is 22. In addition it is necessary to

do complex arithmetic when solving for a single column of the S-matrix. Therefore, 22 iterations involves something more than twice the numerical work needed for a standard close coupling calculation. Unless it is possible to accelerate the convergence considerably, no gain has been made.

One way to do this is to reduce the strength of the coupling matrix by including its diagonal elements in the definition of $\underline{h}$ and hence in the Green function. The details for implementing this are given in Ref. [1]. This has a dramatic effect on convergence, which is reduced from 22 to 4 iterations for the Ar-N_2 test case. This then points the way to possible further improvements to the method. It is possible to further reduce the strength of the coupling matrix by using a matrix Green function. Of course this means more work in computing the Green function and at some point this extra work will stop paying for itself through fewer iterations. Exactly where that point is remains to be seen, but surely it has not yet been reached.

The matrix Green function is given by

$$\underset{\sim}{G}(r,r') = \begin{cases} \underset{\sim}{x}(r) \ \underset{\sim}{y}^t(r') & r < r' \\ \underset{\sim}{y}(r) \ \underset{\sim}{x}^t(r') & r > r' \end{cases} \tag{25}$$

where

$$\underset{\sim}{h}\underset{\sim}{G}(r,r') = \delta(r-r')\underset{\sim}{I}, \tag{26}$$

$$\underset{\sim}{h}\underset{\sim}{x} = \underset{\sim}{0} , \tag{27}$$

$$\underset{\sim}{h}\underset{\sim}{y} = \underset{\sim}{0} , \tag{28}$$

and

$$\underset{\sim}{x}^t \underset{\sim}{y}' - \underset{\sim}{x}^{t'} \underset{\sim}{y} = \underset{\sim}{I} .$$ (29)

The superscript t indicates the matrix transpose.

Therefore, to calculate the matrix Green function we need to solve for the matrix scattering wave function $\underset{\sim}{x}$ in the standard way and in addition solve for an irregular matrix wave function, $\underset{\sim}{y}$. Calculating the matrix Green function requires roughly twice the numerical work as calculating the scattering wave function. For small matrices of order less than, say 50, this will be a small part of the overall calculation.

There are two matrix choices which seem worth trying first. The first is simply to include all of the coupling matrix elements which directly couple to the initial state. For the 210 channel problem of Ref. [1] there are only about 15 channels which couple directly to the ground state. Therefore, in addition to the diagonal Green function already computed, there would be one 15x15 matrix Green function to calculate. It is expected that the reduction in the number of iterations will more than pay for the additional work involved with the Green function.

A second choice, which may actually involve more work than necessary, is to transform to the body-fixed frame and include everything but the off-diagonal centrifugal terms in the Green function. That is, simply compute a coupled states Green function. All that would remain in the matrix $\underset{\sim}{V}$ of Eq. (4) would be a bidiagonal of centrifugal terms. This is appealing not only because convergence should be reached in a very small number of iterations, but also because it seems the logical way to use the coupled states method [7], which is already known to be in many cases a good

approximation, as a starting point for the full calculation.

Another idea for future work is the applicaton of the method to electron-molecule scattering. A variational method similar to Kato's was developed by Saraph and Seaton [15] in the early days of electron-atom scattering to deal with the problem of electron exchange. Their work and that of others [16] was in many cases successful, but since they were interested in the full matrix of solutions it was abandoned for other non-iterative approaches. In light of the present developments, however, it may be worthwhile to revive the iterative approach. For one thing, there was apparently no attempt at that time to include any part of the coupling matrix other than the centrifugal terms in the Green function. The present approach can very likely be used to kill two birds with one stone; that is, to reduce the problem to a single column of the solution matrix and to eliminate the exchange problem at the same time.

Work on these ideas is currently underway.

References

[1]   L. D. Thomas, J. Chem. Phys., 70, 2979 (1979).

[2]   T. Kato, Prog. Theor. Phys., 6, 295 (1951).

[3]   A. M. Arthurs and A. Dalgarno, Proc. Roy. Soc. (London), Ser. A, 256, 540 (1960).

[4]   R. T. Pack. J. Chem. Phys., 60, 633 (1974).

[5]   R. B. Walker and J. C. Light, Chem. Phys., 7, 84 (1975).

[6]   T. P. Tsien and R. T. Pack, Chem. Phys. Lett., 6, 54 (1970).

[7]   P. M. McGuire and D. J. Kouri, J. Chem. Phys., 60, 2488 (1974).

[8]   A. E. DePristo and M. H. Alexander, J. Chem. Phys., 63, 3552 (1975).

[9]   R. G. Newton, "Scattering Theory of Waves and Particles," (McGraw-Hill, New York, 1966).

[10]   N. W. Sams and D. J. Kouri, J. Chem. Phys., 51, 4809 (1969).

[11]   See, J. C. Light and R. B. Walker, J. Chem. Phys., 65, 4272 (1976) and references therein.

[12]   J. Schwinger, Phys. Rev. 72, 472 (1947).

[13]   J. Schwinger, Phys. Rev. 78, 135 (1950).

[14]   J. M. Blatt and J. D. Jackson, Phys. Rev., 76, 18 (1949).

[15]   H. E. Saraph and M. J. Seaton, Proc. Phys. Soc. London, 80, 1057 (1962).

[16]   For a review, see P. G. Burke and M. J. Seaton, "Methods in Computational Physics," edited by B. Alder, S. Fernbach, and M. Rotenberg (Academic Press, New York, 1971) Vol. 10, pg. 1.