Tech Report 94-03

SOME NUMERICAL RESULTS ON ALGORITHMS FOR STURM-LIOUVILLE PROBLEMS

XINGZHI JI

Department of Computer Science

University of Waterloo

Waterloo, Ontario, Canada N2L 3G1

email: xrji@neumann.uwaterloo.ca

KEY WORDS: Eigenvalue, Sturm-Liouville problem, finite difference method, Numerov's method, asymptotic correction, Prüfer method.

AMS(MOS) Subject Classifications. 65L15, 65L12, 65F15.

Running title: **Numerical Results on Algorithms for S-L Problems**

Abstract

For the numerical solution of Sturm-Liouville eigenvalue problems, finite difference methods and Prüfer methods are two kinds of popular methods. However, the conditions under which each method is more efficient are not obvious. The experimental results reported in this paper show: (1) The relative efficiency of the methods depends on the desired accuracy; (2) Finite difference methods, with correction techniques, remain effective even for eigenvalues with higher oscillation numbers.

1 Introduction

It is well known that the Sturm-Liouville (S-L) eigenvalue problem

$$-(p(x)y'(x))' + q(x)y(x) = \lambda s(x)y(x), \qquad a < x < b \tag{1.1}$$

$$y(a)cos\alpha - (py')(a)sin\alpha = 0, \tag{1.2}$$

$$y(b)cos\beta - (py')(b)sin\beta = 0 \tag{1.3}$$

arises in many mathematical physics and engineering applications. If $p', s \in C[a,b], q$ is piecewise continuous and $p > 0, s > 0$, this is a regular problem in the sense that the eigenvalues λ are discrete, simple and can be listed variationally or, equivalently, according to oscillation numbers of the corresponding eigenfunctions.

Concerning numerical solution of the S-L problems, finite difference methods and the Prüfer methods are two kinds of popular methods. Generally speaking, finite difference methods (including those associated with asymptotic correction techniques, cf. [16]) have the advantage of simplicity and programming ease, and have been widely used in practice. If solutions with higher oscillation numbers are required, however, one usually has to use a matrix of very high order. The Prüfer methods are more complicated but give good accuracy, and can calculate any specified eigenvalue without consideration of other eigenvalues. They also have the advantage that a relatively small amount of storage space is required.

It is in general meaningless if we simply ask which method is more efficient, because the efficiency usually depends on the conditions under which the algorithms are implemented. These conditions, however, are not obvious and, even somewhat confusing.

Recently, we have made systematic numerical experiments about a central finite difference method (F-D), F-D with asymptotic correction, Numerov's method and modified Prüfer

methods. For the same problems in our experiments, we first fixed the CPU time to compare accuracies reached by these methods ; then fixed desired accuracy to compare the CPU times spent by these methods. Our numerical results show: When desired eigenvalues, no matter with low or high oscillations, are required high accuracy, the modified Prüfer methods are more efficient; otherwise, preference would be given to finite difference methods. Especially with asymptotic correction technique, finite difference methods maintain effective even for the eigenvalues corresponding to higher oscillation eigenfunctions.

In the following section, the two kinds of methods are briefly described. Their error estimations are given in Section 3. In Section 4, our results of numerical test performed on a CYBER 860 computer are reported.

2 The Method

Consider the S-L problem (1.1–1.3) . To obtain a discrete algebraic eigenvalue problem with a symmetric tridiagonal coefficient matrix, we take the uniform mesh length h and use the Marčuk's integral identities to construct the difference scheme

$$p_{k-\frac{1}{2}}y_{k-1} - (p_{k-\frac{1}{2}} + p_{k+\frac{1}{2}} + h^2 q_k)y_k + p_{k+\frac{1}{2}}y_{k+1} = -\lambda h^2 s_k y_k, \tag{2.1}$$

and similarly deal with the boundary conditions (cf. [4]). The local truncation error is $O(h^2)$.

The asymptotic correction technique for finite difference methods is quite simple [16]. Since true solutions of the problem

$$\begin{cases} -y'' = \lambda y \\ y(a) = y(b) = 0 \end{cases} \tag{2.2}$$

are $\lambda_k = k^2(\frac{\pi}{b-a})^2$, $(k = 1, 2, \cdots)$ and its approximate solutions by (2.1) are $\tilde{\lambda}_k = 4sin^2(\frac{k\pi h}{2(b-a)})/h^2$, it is evident that $\triangle_k = \lambda_k - \tilde{\lambda}_k (k = 1, 2, \cdots)$ are exact correction terms for approximate solutions $\tilde{\lambda}_k$. Furthermore, $\triangle_k$'s are also effective corrections for solutions of the problems that are in the normal Liouville form

$$\begin{cases} -y'' + qy = \lambda y, \\ y(a) = y(b) = 0, \end{cases} \tag{2.3}$$

provided that q is relatively small (comparing with λ_k). In other words, $\triangle_k$'s are generally good correction when λ_k are sufficiently large (It is usually the case when k is sufficiently large). The above technique can be extended to the problems with general boundary conditions [1].

For the regular S-L problem

$$\begin{cases} -y'' + qy = \lambda sy, \\ y(a) = y(b) = 0, \end{cases} \tag{2.4}$$

Numerov's method is also frequently used. Let $\tilde{\lambda}_k$ be an approximation of λ_k and employ uniform grid points $x_i = a + ih (i = 1, 2, \cdots, n), h = (b-a)/(n+1)$, Numerov's method gives

$$-A\mathbf{u} + BQ\mathbf{u} = \tilde{\lambda} BS\mathbf{u} \tag{2.5}$$

where $A := (a_{ij})$ is symmetric tridiagonal with $a_{ii} := -2/h^2$ $(i = 1, \cdots, n)$, $a_{i,i+1} := 1/h^2 (i = 1, \cdots, n-1)$, $B := 12I + h^2 A$, $Q := diag(q(x_1), \cdots, q(x_n))/12$, $S := diag(s(x_1), \cdots, s(x_n))/12$, I is the identity matrix. The local truncation error of formula (2.5) is $O(h^4)$, but BQ, BS are usually nonsymmetric.

The Prüfer methods are another kind of popular methods. By the Prüfer transformation

$y(x) = r(x)sin\theta(x), p(x)y'(x) = r(x)cos\theta(x)$, a pair of first order equations are obtained:

$$\frac{d\theta}{dx} = p^{-1}cos^2\theta - (q - \lambda s)sin^2\theta, \tag{2.6}$$

$$\frac{dr}{dx} = \frac{1}{2}r(p^{-1} + q - \lambda s)sin2\theta, \tag{2.7}$$

$$\theta(a) = \alpha, \quad \theta(b) = \beta + i\pi. \tag{2.8}$$

The system (2.6–2.8) and the original problem (1.1–1.3) are equivalent. In the boundary conditions (2.8), i may take the values $0, 1, 2, \cdots$ and corresponds to the oscillation number of associated eigenfunction.

We note that the equation (2.6) is independent of $r(x)$ and can be solved for λ with boundary conditions (2.8) by shooting methods. As the function $\theta(x, \lambda)$ is strictly monotone about λ (for any fixed x), a starting value for λ can be obtained by the bisection method [7, 9].

To increase the stability of integration in the shooting process, especially when eigenvalues with high oscillation numbers, some scaling techniques are necessary. To this end, rewrite (1.1) as

$$(p(x)y'(x))' + v(x)y(x) = 0 \tag{2.9}$$

and introduce a modified Prüfer transformation by

$$y(x) = r(x)sin\theta(x), \quad p(x)y'(x) = r(x)cos\theta(x)f(v(x)),$$

where $f(v) \neq 0$ is a function of v to be chosen subsequently. Then we get the modified Prüfer equation (see [8])

$$\theta' = \frac{f}{p}cos^2\theta + \frac{v}{f}sin^2\theta + \frac{f'}{2f}sin2\theta. \tag{2.10}$$

In the case $p(x) \equiv 1$ and $v(x) > 0$, we may take $f(v) = v^{1/2}$ to obtain

$$\theta' = v^{1/2} + \frac{v'}{4v} sin2\theta. \tag{2.11}$$

This is modified Prüfer phase equation [9]. If we take $f(v) = c_0 = const$, (2.10) becomes

$$\theta' = \frac{c_0}{p} cos^2\theta + \frac{v}{c_0} sin^2\theta. \tag{2.12}$$

This slight modification avoids evaluating the derivative and the square root of v [5].

3 The Error Estimation

To discuss error estimation for the various approaches described above we take a uniform mesh with the step size h. Suppose that $\{\lambda_k\}$ are true eigenvalues of the S-L problem (1.1–1.3) and $\{\lambda_k^{(n)}\}_{k=1}^n$ are approximate eigenvalues of the corresponding finite difference equation, where n is the order of matrices in the difference equation.

Keller [13] has shown that , in a very general sense, the error arising in the $\lambda_k^{(n)}$ is proportional to the local truncation error associated with the particular finite difference approximation used. For example, if central differences are used with sufficiently small h, then

$$\mid \lambda_k - \lambda_k^{(n)} \mid \leq ch^2\lambda_k^2, \tag{3.1}$$

where (here and below) c is a constant independent of k and h.

When the corrected central difference method is applied to the problem (2.3) on $[a, b] = [0, \pi]$, [16] proved that if $q \in C^2[0, \pi]$ and $0 < \alpha < 1$, then there exists a constant $c(\alpha)$ such that for all n and all $k < \alpha\pi/h$, we have

$$\mid \lambda_k - \lambda_k^{(n)} \mid \leq c(\alpha)kh^2. \tag{3.2}$$

For Numerov's method applied to the regular S-L problem (2.4), Andrew [2] has shown that for $q, s \in C^4[a,b], s > 0$, and for all k and sufficiently large n , there is an eigenvalue Λ of the approximating scheme for which

$$| \lambda_k - \Lambda | \leq ck^6h^4. \tag{3.3}$$

For the Prüfer method , Paine and Anderssen [15] shown that, when the classical Runge-Kutta method is used to perform the integration, scaled and modified Prüfer methods satisfy

$$| \lambda_k - \lambda_k^{(n)} | \leq ch^4\lambda_k^{1/2}(\lambda_k^{(n)})^{1/2} \tag{3.4}$$

for suitably small h and $4\lambda_k^{(n)}h^2 \leq \pi^2$.

It should be pointed out that both the finite difference method and the Prüfer method can be employed to regular or singular S-L problems, but the above error estimations usually require enough continuity of the coefficient functions. This is limitation of the theoretical estimate. In spite of this, there is a tendency to judge the efficiency of algorithms by their error estimations if the stability and convergence of the algorithms are presumed. Therefore, the following opinion is quite popular: The Prüfer methods are more efficient than finite difference methods; finite difference methods are only suitable for eigenvalues with lower oscillation numbers. However, this is not true in general. It should be noted that although the above formulae give the error estimations, they do not tell how much work is needed for each method to achieve those accuracies. It is more reasonable to consider the error estimation of a method together with its cost[1]. Because in the same computing time, a "big-error" method may use much smaller h to calculate so that it eventually reach higher

[1]In this paper, we concentrate on time cost.

accuracy than "small-error" methods. As our numerical tests showed, if we do not let the computing time last too long, in other words, if we do not want high accurate solutions, finite difference methods are even more efficient than the Prüfer methods. Besides, these methods are not only simple but also need no starting values. However, if high accurate eigenvalues are needed, it is more efficient to employ the Prüfer methods even for eigenvalues with lower oscillation numbers.

4 Numerical Results

In this section, we present the results of our numerical experiment on a CYBER 860 computer about the methods described in Section 2.

In the computation, we employed the routine F02BFF in the NAG Library [11] to compute the selected eigenvalues of a real symmetric tridiagonal matrix. For Numerov's method, we adopted the bisection method developed by Peters and Wilkinson [17], which takes the advantage of sparse property of the problem (2.5). We treated BQ as a symmetric matrix. The caused error is neglectable when h is very small (cf. [12]). We also used the routine D02KDF of scaled Prüfer method in the NAG (cf. [10, 18]), which can find a specified eigenvalue of a regular or singular self-adjoint second-order S-L problem with permitting of discontinuities in the coefficient functions or their derivatives. Combining this routine with slightly modified Prüfer transformation (see (2.12)) usually gave better results. We point out, there is also the SLEIGN Software for the S-L problems available from Sandia Laboratories [6, 7, 14].

The first example we considered is

$$\begin{cases} -y'' = 200 sgnx \cdot y - \lambda y, \\ y(-1) = y(+1) = 0. \end{cases} \tag{4.1}$$

By the two kinds of methods, we calculated the three groups of its eigenvalues: For the first group, the oscillation number k is from 0 to 4; for the second, k from 40 to 49; the third, k from 100 to 109. We compared the accuracies of obtained solutions in the almost equal computing periods (not including compiling times), which are coincide with each other till $\frac{1}{100}$ second. There are many factors to affect the CPU time such as the order of a matrix or the adopted tolerance for an algorithm. In order to make the conclusion acceptable, one must properly adjust these factors at the same time.

The computing CPU times we set were as follows: 0.9 and 1.7 seconds for the first group eigenvalues, 1.9 and 3.1 seconds for the second group, and 1.45 and 4.6 seconds for the third group . In each case, the absolute errors of computed eigenvalues are listed in Table 1 $\sim$ 6. In which, F-D and F-D-C represent center difference method and its correcting version, respectively; PRU-1 and PRU-2 represent the scaled Prüfer method(i.e., D02KDF in the NAG) and this method plussing slight modification, respectively; MAX-ERR and MAX-RE-ERR denote maximum error and maximum relative error, respectively. The true solutions of 4 decimal places were obtained by the routine D02KDF with Tol (tolerance) $= 10^{-12}$, and the corresponding starting values were produced by F-D with the order of matrices $n = 10,001$.

Remark 1. Owing to the coefficient function $signx$, taking the point $x = 0$ as one of the uniform mesh points is significant for accuracy. If we set $n = 3001$ for F-D method, for instance, the maximum absolute error is 0.0015281 as in Table 1. While the maximum absolute error will go to 0.13843 if we set $n = 3000$ (other conditions remain the same).

Remark 2. We used F-D method with $n = 10$ to produce starting values for Prüfer methods. The CPU time spent was included in corresponding Prüfer methods.

It is clear, by our computational results, that the efficiency of the algorithms is not determined by the size of eigenvalue's oscillation number, but by the accuracy one desired. In other wards, Prüfer methods are more efficient for high accurate solutions while finite difference methods are more efficient for not so high accurate solutions, because the CPU time grows rapidly with the increasing of order of coefficient matrices but relatively slower with the decreasing of tolerance error in the Prüfer methods. The behavior of Numerov's method, which is of local truncation error $O(h^4)$, was not as satisfactory as expected, for more work involved in calculation of associated generalized eigenvalue problems. However, Numerov's method combining with the correction technique will give better results, but we didn't carry out that test. The interested readers may refer to [3].

The second example we considered is

$$\begin{cases} y'' = \lambda y, \\ y(-1) = y(1) = 0. \end{cases} \tag{4.2}$$

The computed results are reported in Table 7. It also supports the above conclusions.

For the problems (4.1) and (4.2) , we also compared the computing CPU times spent by these methods by fixing the accuracy of results. As expected, the conclusions remain the same.

At last, we should point out, for the problems with complicated coefficient functions the Prüfer methods will not be as efficient as supposed to be, because there are thousands of times of evaluating these functions in implementation of the Prüfer methods. In some practical situations, a single eigenvalue is needed. In these cases, however, more preference

will be given to the Prüfer methods.

5 Further Questions

The Figure 1 gives the sketch of relative efficiency of finite difference methods and the Prüfer methods. Evidently, h_0 is a important value in application. However, how to determine h_0 for a specific question is an open problem.

Generally, given a stable algorithm **A** for a class of problems . Regardless of the amount of storage space, the efficiency of **A**, $E_f(\mathbf{A})$, is determined by both the accuracy $e(h)$ of the result and the computing time $t(h)$ (or more precisely, the number of operations) to reach that accuracy:

$$E_f(\mathbf{A}) = f(e(h), t(h)) = g(h),$$

where h is the step size used, $h^{-1} = n$ can be viewed as the size of the problems . Usually $g(h)$ is not a constant, therefore the efficiency of the algorithm varies with h. At present, we do not know what is proper form of the function $E_f(\mathbf{A})$.

Acknowledgments

This work was completed when the author visited the University of Calgary . He is grateful to Professor P.A. Binding and Professor P. J. Browne for their hospitality and many stimulating discussions.

References

[1] R. S. Anderssen and F. R. de Hoog. On the correction of finite difference eigenvalue approximations for Sturm-Liouville problems with general boundary conditions. *BIT*, 24:401–412, 1984.

[2] A. L. Andrew. The accuracy of Numerov's method for eigenvalues. *BIT*, 26:251–253, 1986.

[3] A. L. Andrew and J. W. Paine. Correction of Numerov's eigenvalue estimates. *Numer. Math.*, 47:289–300, 1985.

[4] I. Babuška, M. Práger, and E. Vitásek. *Numerical Processes in Differential Equations.* Interscience Publishers, London, 1966.

[5] P.B. Bailey. A slightly modified Prüfer transformation useful for calculating Sturm-Liouville eigenvalues. *J. Comput. Phys.*, 29(2):306–310, 1978.

[6] P.B. Bailey, B. S. Garbow, H. G. Kaper, and A. Zettl. Eigenvalue and eigenfunction computations for Sturm-Liouville problems. *ACM Trans. Math. Softw.*, 17(4):491–499, 1991.

[7] P.B. Bailey, M. K. Gordon, and L. F. Shampine. Automatic solution of the Sturm-Liouville problem. *ACM Trans. Math. Softw.*, 4(3):193–208, 1978.

[8] P.A. Binding, P.J. Browne, and Xingzhi Ji. A numerical method using the Prüfer transformation for the calculation of eigenpairs of two parameter Sturm-Liouville problems. accepted by IMA J. Numer. Anal., 1992.

[9] G. Birkhoff and G. C. Rota. *Ordinary Differential equations, 2nd ed.* Blaisdell, Toronto, 1969.

[10] D02KDF-NAG FORTRAN Library Routine Document.

[11] F02BFF-NAG FORTRAN Library Routine Document.

[12] Xingzhi Ji. The pseudosymmetric tridiagonalization of an arbitrary real matrix. *Linear Algebra Appl.*, 129:1–12, 1990.

[13] H. B. Keller. *Numerical Methods for two-point boundary-value problems.* Blaisdell, Waltham, 1968.

[14] M. Marletta. Certification of algorithm 700: Numerical tests of the SLEIGN software for Sturm-Liouville problems. *ACM Trans. Math. Softw.*, 17(4):481–490, 1991.

[15] J. W. Paine and R. S. Anderssen. Uniformly valid approximation of eigenvalues of Sturm-Liouville problems in geophysics. *Geophys. J. R. astr. Soc.*, 63:441–465, 1980.

[16] J.W. Paine, F. R. de Hoog, and R. S. Anderssen. On the correction of finite difference eigenvalue approximation for Sturm-Liouville problems. *Computing*, 26:123–139, 1981.

[17] G. Peters and J.H. Wilkinson. Eigenvalues of $Ax = \lambda Bx$ with band symmetric A and B. *Computer Journal*, 12:398–404, 1969.

[18] J. D. Pryce. Error control of phase-function shooting methods for Sturm-Liouville problems. *IMA J. Numer. Anal.*, 6:103–123, 1986.