Solving inverse Sturm–Liouville problems: theory and practice

Alan L. Andrew[1]

Abstract

Theoretical results on the solution of inverse Sturm–Liouville problems generally consider only idealized problems requiring much more data than is available in real applications. Typical theorems describe problems where infinitely many eigenvalues are known exactly, but in most applications we know only approximations of a finite, and usually small, number of eigenvalues. This paper considers how idealized theoretical results may assist practical numerical computation. It also reviews recent progress on a class of numerical methods for inverse Sturm–Liouville problems, it discusses some open questions, and it announces a new convergence result.

Contents

1 Introduction

There are many excellent methods [23] for numerical solution of the *direct*
Sturm–Liouville problem: computing eigenvalues, λ, and eigenfunctions, y,
of

$$- y'' + qy = \lambda y, \tag{1}$$

when the potential, q, and the boundary conditions are known. This paper
addresses the more challenging *inverse* problem of computing q from a
knowledge of eigenvalues and eigenfunctions. For simplicity (and because
solutions to inverse problems on infinite intervals are often approximated
by solutions of problems on finite intervals [3]), we restrict our attention
to problems defined on a finite interval. Without further loss of generality,
we take this to be $[0, \pi]$. We emphasize the case of separated boundary
conditions,

$$\cos(c_1)y(0) + \sin(c_1)y'(0) = \cos(c_2)y(\pi) + \sin(c_2)y'(\pi) = 0, \tag{2}$$

where $c_i \in \mathbb{R}$, though similar considerations apply to problems with non-
separated boundary conditions such as periodic conditions [2, 4]. We denote
the eigenvalues of (1–2), by $\lambda_1(q, c_1, c_2) \leqslant \lambda_2(q, c_1, c_2) \leqslant \ldots$, and the
corresponding eigenfunctions by $y_1(q, c_1, c_2), y_2(q, c_1, c_2), \ldots$.

Some background theory. The theoretical literature is largely concerned with establishing sufficient conditions for desirable outcomes such as existence and uniqueness of a potential q producing the given data, and whether sufficiently small perturbations in the data cause only small perturbations in the potential. These conditions are also important when modeling practical problems, but unfortunately the finite amount of data available in applications is not enough to satisfy the requirements of theoretical uniqueness results. We emphasize three important cases where no information on eigenfunctions is needed to define q uniquely in the Hilbert space $L_2(0, \pi)$.

U.1 The **symmetric** problem [5]. If $c_1 + c_2 = \pi$, and q is known to satisfy $q(x) = q(\pi - x)$ for almost all $x \in (0, \pi)$, then q is uniquely determined in $L_2(0, \pi)$ by the eigenvalues of (1–2).

U.2 The (closely related [7]) **two spectra** problem. If both $\lambda_i(q, c_1, c_2)$ and $\lambda_i(q, c_1, c_3)$ are given for all $i \in \mathbb{N}$, and $\sin(c_2 - c_3) \neq 0$, then q is uniquely determined in $L_2(0, \pi)$.

U.3 The **half inverse** (or Hochstadt–Lieberman) problem [8]. If q is known almost everywhere on $(0, c)$ or $(\pi - c, \pi)$, then q is uniquely determined in $L_2(0, \pi)$ by the eigenvalues of (1–2) when $c = \pi/2$, and by a suitable (infinite) subset of these eigenvalues when $c > \pi/2$.

These conditions give L_2 uniqueness, not pointwise uniqueness. To determine q at the uncountably many points of $(0, \pi)$ from countably many eigenvalues we need a smoothness assumption, such as continuity. Knowing eigenfunctions can make things easier. For the general equation, $(py')' + (q + \lambda r)y = 0$, if we know λ_1, and also r, y_1 and $(py_1')'$ everywhere in $(0, \pi)$, then, q is trivially given by $q = -(py_1')'/y_1 - \lambda_1 r$, everywhere in $(0, \pi)$. In applications, we are unlikely to know more than approximations of y_1, and perhaps y_1', at a finite number of points. Liu [19] suggested solving this problem numerically. His method can be adapted for more challenging problems [20]).

Scope of this paper. Of the many available methods [9] for solving inverse Sturm–Liouville problems (ISLPs), we concentrate on those using the

technique of 'asymptotic correction' [4] (sometimes called the AAdHP correction [23]). This allows us to use a method such as Numerov's method to replace the ISLP by a more easily solved matrix inverse eigenvalue problem (MIEVP) [5]. The idea is that, before solving the MIEVP, we add to each known eigenvalue a correction calculated so that, when q is constant, the MIEVP produces the same constant solution. This eliminates the source of failure of earlier attempts to solve ISLPs by finite difference methods: the asymptotic difference between the continuous and discrete eigenvalues [5, 6, 7, 8, 9]. In many important cases [4, 8, 9], the correction is known in closed form. Originally proposed [22] for centered finite difference solution of the direct problem (1–2) with $c_1 = c_2 = 0$, asymptotic correction was soon extended to higher order methods [11, 25] and to more general differential operators [4, 6, 25].

The next paragraph summarizes important recent work not covered in the next two sections or in earlier surveys [6, 9]. Section 2 discusses some observed properties of numerical results in the light of known theoretical results. Section 3 reviews recent results on sufficient conditions for the computed potential to converge to the true potential, q, as the number of data points increases. It also announces a new result on the convergence of Numerov's method [7] for the two spectra problem, and discusses possible extensions.

Numerov's method [5, 7, 11] has the advantage that the matrices whose eigenvalues are required are tridiagonal, but the disadvantage that the mesh length is fixed by the number of available data points [5, 7, 8]. Gao et al. [14] showed that this limitation could be overcome by using interpolation to refine the mesh. An older approach allowing indefinite mesh refinement seeks a potential of the form $\sum_{i=1}^{n} c_i \phi_i$, for some specified ϕ_i [7]. Asymptotic correction has proved useful for this also [1, 8, 15]. These methods are much more computationally expensive than the simple Numerov's method [5, 7, 8], but are useful when available accuracy is limited by availability of data rather than by computational costs. Comparing results obtained by different methods can further increase the information obtained from limited data [5].

Other recent work on the use of asymptotic correction for inverse problems concerns its use with higher order methods [1, 15]. Also, although the theory of asymptotic correction [11, 22] concerns equations in the Liouville normal form (1), recent numerical results [13, 16] show that it can also improve results for equations in the impedance form $(ay')' + \lambda ay = 0$.

2   Clues from theoretical results

The uniqueness conditions, U.1, U.2, U.3, require infinitely many eigenvalues, but, for sufficiently smooth q of limited variation, many methods (including those using asymptotic correction [6, 9]) have produced remarkably accurate results from limited data. Less smooth q are more difficult to compute. The computed values of q share some properties with truncated Fourier expansions, exhibiting something like a Gibbs phenomenon at points of discontinuity and (sometimes) at the boundaries. Asymptotic expansions have been used [5, pp. 228–229 and references] to explain numerical results. Savchuk and Shkalikov [28], strengthened these asymptotic results, obtaining expansions for eigenvalues of (1) with various boundary conditions, for all q in the Sobolev space $W_2^{\theta-1}, \theta \geqslant 0$. (When $\theta < 1$, q is a distribution.) For smoother q, they obtained stronger results. For $\theta \in \mathbb{N}$, when q has Sobolev norm $\|q\|_{\theta,2} = \left(\int_0^\pi \sum_{i=0}^\theta (q^{(i)}(x))^2) dx \right)^{1/2}$, their result [28, Theorem 4.1], and some known results for low θ [27], may be summarized as follows.

Theorem 1. *For all $R > 0$, there exists a constant $C(R)$ such that, for all $q \in W_2^{\theta-1}$ with $\|q\|_{\theta,2} < R$, there exists a sequence $\{\alpha_\ell\}_1^\infty \in \ell_2$ with $\sum_{i=1}^\infty \alpha_\ell^2 < C(R)$, such that, when $\theta = 2s + 1$ is an odd integer,*

$$\lambda_k^{1/2}(q,0,0) = k + \sum_{i=0}^s \frac{h_i(q)}{(2k)^{2i+1}} - (-1)^s \frac{a_{2k}(q)}{2(2k)^{2s+1}} + \frac{\alpha_{2k}}{k^{2s+2}},$$

$$\lambda_k^{1/2}(q,0,\pi/2) = k - \frac{1}{2} + \sum_{i=0}^s \frac{g_i(q)}{(2k-1)^{2i+1}} - (-1)^s \frac{a_{2k-1}(q)}{2(2k-1)^{2s+1}} + \frac{\alpha_{2k-1}}{k^{2s+2}},$$

and, when $\theta = 2s$ *is an even integer,*

$$\lambda_k^{1/2}(q,0,0) = k + \sum_{i=0}^{s} \frac{h_i(q)}{(2k)^{2i+1}} - (-1)^s \frac{b_{2k}(q)}{2(2k)^{2s}} + \frac{\alpha_{2k}}{k^{2s+1}},$$

$$\lambda_k^{1/2}(q,0,\pi/2) = k - \frac{1}{2} + \sum_{i=0}^{s} \frac{g_i(q)}{(2k-1)^{2i+1}} - (-1)^s \frac{b_{2k-1}(q)}{2(2k-1)^{2s}} + \frac{\alpha_{2k-1}}{k^{2s+1}}.$$

Here

$$a_i(q) = \frac{2}{\pi} \int_0^\pi q^{(\theta-1)}(t)\cos(it)dt, \quad b_i(q) = \frac{2}{\pi} \int_0^\pi q^{(\theta-1)}(t)\sin(it)dt,$$

for $i = 1,2,\ldots$, *while* h_i , g_i , $0 \leqslant i \leqslant s$, *are continuous bounded functionals on* $W_2^{\theta-1}$, *with* $h_0(q) = g_0(q) = \pi^{-1}\int_0^\pi q(t)dt$, *and, for* $1 \leqslant i < s$ *(and for* $i = s$ *when* $\theta = 2s+1$*), the linear parts,* h_i^0 , g_i^0 *of* $h_i(q)$ *and* $g_i(q)$ *are*

$$h_i^0 = (-1)^i \pi^{-1}[q^{(2i-1)}(0) - q^{(2i-1)}(\pi)],$$
$$g_i^0 = (-1)^i \pi^{-1}[q^{(2i-1)}(0) + q^{(2i-1)}(\pi)].$$

Also $h_1(q) = \pi^{-1}[q'(0) - q'(\pi) + \int_0^\pi q^2(t)dt] - 2(h_0(q))^2$, *and the linear parts are generally dominant for small* q . *If* $\theta = 2s$, *then* $h_s^0 = g_s^0 = 0$.

Discussion. The a_i and b_i terms give the cos and sin terms in the Fourier expansion of $q^{(\theta-1)}$. The linear parts (dominant for sufficiently small $\|q\|$) of $g_i + h_i$ and $g_i - h_i$ give the odd terms in the Taylor expansions of q about 0 and π respectively. Also, $\alpha_i \to 0$ as $\|q\|_{\theta,2} \to 0$. As $\theta \to \infty$ (q smoother), the final (α_i) terms $\to 0$ (exponentially), and more Taylor coefficients are involved. This helps explain why q is usually easier to compute when it has high θ , low $\|q\|_{\theta,2}$, and rapidly converging Taylor and Fourier approximations.

Using $\lambda_i(q,0,0)$, $i = 1,\ldots,n$, and $\lambda_i(q,0,\pi/2)$, $i = 1,\ldots m$, for the two spectra problem, Numerov's method [7] led to well-conditioned equations when $m \leqslant n \leqslant m+1$, but produced a singular Jacobian matrix when $n < m$

or $n > m + 1$ [7]. For the half-inverse problem [8] with various boundary conditions, using the first n eigenvalues of (1), and the values of q at the first (or last) m grid points, Numerov's method [8] produced well-conditioned equations when $|m - n| \leqslant 1$, but produced a singular Jacobian matrix when $m < n - 1$. When $m > n + 1$, the Jacobian matrix was still nonsingular, but ill-conditioning increased dramatically as $m - n$ increased.

Since q' is odd when q is even, the dominant terms in the asymptotic expansions of $\lambda_k(q, 0, 0))$ and $\lambda_k(q, 0, \pi/2)$ give information on the even and odd parts respectively of q about $\pi/2$. As having similar amounts of information on both parts is usually an advantage, this partially explains why best results were obtained with m close to n. (In U.2 and U.3, both types of data have similar importance.) Nevertheless, some questions remain open. Why were the restrictions [7, 8] on $|m - n|$ so severe? To what extent can these restrictions be overcome by the use of a nonuniform grid [8], or by using an alternative method to solve the main nonlinear equation [8, eqn. 8]?

3   Convergence questions

Given N eigenvalues, numerical methods typically estimate q as $f(\mathbf{q})$, where the known function f depends on the method, and $\mathbf{q} \in \mathbb{R}^N$ is obtained by solving the N nonlinear equations $\Lambda_i(\mathbf{q}) = \lambda_i$, where $\Lambda_i(\mathbf{q})$ is the method's estimate of the given eigenvalue λ_i. (With Numerov's method [5], the components of $\mathbf{q}$ are the computed values of q at the grid points.) Two convergence questions generally arise. Does the iterative method used to solve the N equations $\Lambda_i(\mathbf{q}) = \lambda_i$ converge? Does $f(\mathbf{q}) \to q$ as $N \to \infty$? Until recently [17], only the first question had been answered for Numerov's method [5, 7, 8] or for most other methods. Much work on numerical methods addresses neither convergence question. We discuss recent work on the more difficult second convergence question, and suggest some possible extensions.

Gao et al. [17] used a result of Hald [18, Theorem 3], and a known error

bound [11, Theorem 1] for the direct (Dirichlet) problem, to provide an affirmative answer to the second convergence question for the corrected Numerov method [5] for the symmetric inverse Sturm–Liouville problem. (Their proof can be simplified by noting that the nonsingularity of the matrix D in their Lemma 2.4 follows immediately from the independently proved fact that all eigenvalues of DD^T are non-zero.)

Gao et al. also showed that a conjecture of Andrew [4, p. 363], which they called Conjecture A, would, if proved, establish a corresponding convergence result for the Numerov method [7] for the two spectra problem. Subsequently, Andrew has shown that, for q satisfying an appropriate smoothness condition, that conjecture can be deduced from the previously cited result [11, Theorem 1] for the direct Dirichlet problem, by using the known relationship [7, Section 2] between the $\lambda_k(q, 0, \pi/2)$ and the Dirichlet eigenvalues on an extended interval. The proof uses symmetry properties of the Numerov equations [11] and a property of centrosymmetric matrices [5]. The result of Gao et al. then establishes convergence of the Numerov method [7] for the two spectra problem. Details will be given elsewhere [10].

Some open questions. Hald's result [18, Theorem 3], and consequently also the results of Gao et al. [15, 17], apply only for q 'sufficiently close' to a constant, but mathematical models often involve potentials of the form $q = q_1 + q_2$, where $q_1 \in C^\infty[0, \pi]$ and q_2 is a simple step function. Numerical results of Raffler and Böckmann [24] show that, for q of this form, with known q_2, performance of an important classical method [26] is greatly improved if the constant reference potential used in that method is replaced by q_2. Other methods, including those using asymptotic correction, can, in principle, be similarly modified [9], though such modifications have not yet been tested numerically or studied theoretically. In the important case when q_2 has only one discontinuity in $(0, \pi)$, Efremova and Freiling [12] proposed a simple and effective method for computing q_2, provided only that q_1 is absolutely continuous in $[0, \pi]$. This adds a new challenge to those already mentioned [8, 9]. Since the degree of uncertainty may be more important than the variation of q in limiting how well we can approximate q from limited

spectral data, it should be possible to extend known convergence results, for
q with small Sobolev norm, to q of the form $q = q_1 + q_2$, where q_1 has small
Sobolev norm, and q_2 is a step function with one or two discontinuities. There
are several results which may help. Some asymptotic results are available
for eigenvalues corresponding to q with simple discontinuities [5]. Not all
analogues of Hald's result have the same restrictions on q. The paper by
Savchuck [27] is a good source of references.

4   Conclusions

It is hoped that research students and others working on the numerical solution
of inverse Sturm–Liouville problems will find at least the following features of
this paper useful. (i) The critical assessment of recent contributions using the
asymptotic correction technique, and their relationship to work using other
methods. (ii) The discussion on the role of asymptotic expansions, which
complements earlier work on the information available from finitely many
eigenvalues [5, 21, and references]. (iii) The suggestions for further work,
which complement suggestions made in earlier papers [8, 9].

References

[1] L. Aceto, P. Ghelardoni and C. Magherini. Boundary value methods for
    the reconstruction of Sturm–Liouville potentials. *Appl. Math. Comp.*,
    219:2960–2974, 2012. doi:10.1016/j.amc.2012.09.021 C127, C128

[2] A. M. Akhtyamov, V. A. Sadovnichy and Ya. T. Sultanaev.
    Generalizations of Borg's uniqueness theorem to the case of non-
    separated boundary conditions. *Eurasian Math. J.*, **3(4)**:10–22, 2012.
    http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=
    emj&paperid=101&option_lang=eng C125

[3] H. Altundağ, C. Böckmann and H. Taşeli. Inverse Sturm–Liouville problems with pseudospectral methods. *Int. J. Comput. Math.*, **92**:1373–1384, 2015. doi:10.1080/00207160.2014.939646 C125

[4] A. L. Andrew. Asymptotic correction of Numerov's eigenvalue estimates with natural boundary conditions. *J. Comp. Appl. Math.*, **125**:359–366, 2000. doi:10.1016/S0377-0427(00)00479-9 C125, C127, C131

[5] A. L. Andrew. Numerov's method for inverse Sturm–Liouville problems. *Inverse Problems*, **21**:223–238, 2005. doi:10.1088/0266-5611/21/1/014 C126, C127, C128, C130, C131, C132

[6] A. L. Andrew. Asymptotic correction and inverse eigenvalue problems: an overview. *ANZIAM J.*, **46(E)**:C1–C14, 2005. doi:10.21914/anziamj.v46i0.943 C127, C128

[7] A. L. Andrew. Computing Sturm–Liouville potentials from two spectra. *Inverse Problems* **22**:2069–2081, 2006. doi:10.1088/0266-5611/22/6/010 C126, C127, C129, C130, C131

[8] A. L. Andrew. Finite difference methods for half inverse Sturm–Liouville problems. *Appl. Math. Comp.*, **218**:445–457, 2011. doi:10.1016/j.amc.2011.05.085 C126, C127, C130, C131, C132

[9] A. L. Andrew. Inverse Sturm–Liouville problems: some recent developments. *ANZIAM J.*, **52(E)**:C287–C302, 2011. doi:10.21914/anziamj.v52i0.3888 C126, C127, C128, C131, C132

[10] A. L. Andrew. Convergence of Numerov's method for an inverse Sturm–Liouville problem. In preparation. C131

[11] A. L. Andrew and J. W. Paine. Correction of Numerov's eigenvalue estimates. *Numer. Math.*, **47**:289–300, 1985. MR 86j:65101. doi:10.1007/BF01389712 C127, C128, C131

[12] L. Efremova and G. Freiling. Numerical solution of inverse problems for Sturm–Liouville operators with discontinuous potentials. *Cent. Eur. J. Math.* **11**:2044–2051, 2013. doi:10.2478/s11533-013-0301-1 C131

[13] Q. Gao. Decent [descent] flow methods for inverse Sturm–Liouville problem. *Appl. Math. Modelling* **36**:4452–4465, 2012. doi:10.1016/j.apm.2011.11.070 C128

[14] Q. Gao, X. Cheng and Z. Huang. Modified Numerov's method for inverse Sturm–Liouville problems. *J. Comp. Appl. Math.* **253**:181–199, 2013. doi:10.1016/j.cam.2013.04.025 C127

[15] Q. Gao, X. Cheng and Z. Huang. On a boundary value method for computing Sturm–Liouville potentials from two spectra. *Internat. J. Comp. Math.* **91**:490–513, 2014. doi:10.1080/00207160.2013.790535 C127, C128, C131

[16] Q. Gao, Z. Huang and A. Cheng. A finite difference method for an inverse Sturm–Liouville problem in impedance form. *Numer. Algorithms* **70**:669–690, 2015. doi:10.1007/s11075-015-9968-7 C128

[17] Q. Gao, Q. Zhao, X. Zheng and Y. Liung. Convergence of Numerov's method for inverse Sturm–Liouville problems. *Appl. Math. Comp.*, **293**:1–17, 2017. doi:10.1016/j.amc.2016.08.007 C130, C131

[18] O. H. Hald. The inverse Sturm–Liouville problem and the Rayleigh–Ritz method. *Math. Comp.*, **32**:687–705, 1978. doi:10.1090/S0025-5718-1978-0501963-2 C130, C131

[19] C.-S. Liu. Solving an inverse Sturm–Liouville problem by a Lie-group method. *Bound. Value Probl.* **2008**:749865, 2008. doi:10.1155/2008/749865 C126

[20] C.-S. Liu and S. N. Atluri. A novel fictitious time integration method for solving the discretized inverse Sturm–Liouville problems, for specified eigenvalues. *CMES Comput. Model. Eng. Sci.* **36**:261–285, 2008. doi:10.3970/cmes.2008.036.261 C126

[21] B. D. Lowe, M. Pilant and W. Rundell. The recovery of potentials from finite spectral data. *SIAM J. Math. Anal.* **23**:482–504, 1992. doi:10.1137/0523023 C132

[22] J. W. Paine, F. R. de Hoog and R. S. Anderssen. On the correction of finite difference eigenvalue approximations. *Computing* **26**:123–139, 1981. doi:10.1007/BF02241779 C127, C128

[23] J. D. Pryce. *Numerical solution of Sturm–Liouville problems.* Oxford University Press, 1993. C125, C127

[24] M. Rafler and C. Böckmann. Reconstructive method for inverse Sturm–Liouville problems with discontinuous potentials. *Inverse Problems* **23**:933–946, 2007. doi:10.1088/0266-5611/23/3/006 C131

[25] A. Rattana and C. Böckmann. Matrix methods for computing eigenvalues of Sturm–Liouville problems of order four. *J. Comp. Appl. Math.* **249**:144–156, 2013. doi:10.1016/j.cam.2013.02.024 C127

[26] W. Rundell and P. E. Sacks. Reconstruction techniques for classical Sturm–Liouville problems. *Math. Comp.* **58**:161–183, 1992. doi:10.1090/S0025-5718-1992-1106979-0 C131

[27] A. M. Savchuk. Reconstruction of the potential of the Sturm–Liouville operator from a finite set of eigenvalues and normalizing constants. *Math. Notes* **99**:715–228, 2016. doi:10.1134/S0001434616050102 C128, C132

[28] A. M. Savchuk and A. A. Shkalikov. On the eigenvalues of the Sturm–Liouville operator with potentials from Sobolev spaces. *Math. Notes* **80**:814–832, 2006. doi:10.1007/s11006-006-0204-6 C128

Author address

1. **Alan L. Andrew**, Dept. Mathematics, La Trobe University, Victoria 3086, AUSTRALIA.
mailto:a.andrew@latrobe.edu.au