Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

Lars Eldén and Valeria Simoncini

1 Department of Mathematics, Linköping University, Sweden
E-mail: laeld@math.liu.se

2 Dipartimento di Matematica, Università di Bologna, 40127 - Bologna - Italy
E-mail: valeria@dm.unibo.it

Abstract.

We study the numerical solution of a Cauchy problem for a self-adjoint elliptic partial differential equation $u_{zz} - Lu = 0$ in three space dimensions $(x, y, z)$, where the domain is cylindrical in $z$. Cauchy data are given on the lower boundary and the boundary values on the upper boundary are sought. The problem is severely ill-posed. The formal solution is written as a hyperbolic cosine function in terms of the two-dimensional elliptic operator $L$ (via its eigenfunction expansion), and it is shown that the solution is stabilized (regularized) if the large eigenvalues are cut off. We suggest a numerical procedure based on the rational Krylov method, where the solution is projected onto a subspace generated using the operator $L^{-1}$. This means that in each Krylov step a well-posed two-dimensional elliptic problem involving $L$ is solved. Furthermore, the hyperbolic cosine is evaluated explicitly only for a small symmetric matrix. A stopping criterion for the Krylov recursion is suggested based on the difference between two successive approximate solutions, combined with a check that the residual is small enough. Two numerical examples are given that demonstrate the accuracy of the method and the efficiency of the stopping criterion.

Submitted to: Inverse Problems
1. Introduction: A Cauchy Problem on a Cylindrical Domain

Let $\Omega$ be a connected domain in $\mathbb{R}^2$ with smooth boundary $\partial \Omega$, and assume that $L$ is a linear, self-adjoint, and positive definite elliptic operator defined in $\Omega$. We consider the ill-posed Cauchy problem,

\begin{align*}
  u_{zz} - Lu &= 0, & (x, y, z) &\in \Omega \times [0, z_1], \\
  u(x, y, z) &= 0, & (x, y, z) &\in \partial \Omega \times [0, z_1], \\
  u(x, y, 0) &= g(x, y), & (x, y) &\in \Omega, \\
  u_z(x, y, 0) &= 0, & (x, y) &\in \Omega.
\end{align*}

The problem is to determine the values of $u$ on the upper boundary, $f(x, y) = u(x, y, z_1), (x, y) \in \Omega$.

This is an ill-posed problem in the sense that the solution (if it exists), does not depend continuously on the data. It is a variant of a classical problem considered originally by Hadamard, see e.g. [22], and it is straightforward to analyze it using an eigenfunction expansion. In Appendix B we discuss the ill-posedness of the problem and derive a stability result.

Since the domain is cylindrical with respect to $z$, we can use a separation of variables approach, and write the solution of (1) formally as

\[ u(x, y, z) = \cosh(z\sqrt{L})g. \]

The operator $\cosh(z\sqrt{L})$ can be expressed in terms of the eigenvalue expansion of $L$, cf. Appendix B. Due to the fact that $L$ is unbounded, the computation of $\cosh(z\sqrt{L})$ is unstable and any data errors or rounding errors would be blown up, leading to a meaningless approximation of the solution.

The problem can be stabilized (regularized) if the operator $L$ is replaced by a bounded approximation. In a series of papers [10, 11, 12, 28, 29, 30] this approach has been used for another ill-posed Cauchy problem, the sideways heat equation, where wavelet and spectral methods were used for approximating the unbounded operator (in that case the time derivative). A similar procedure for a Cauchy problem for the Laplace equation was studied in [5]. However, for such an approach to be applicable, it is required that the domain is rectangular or can be mapped conformally to a rectangular region. It is not clear to us how a spectral or wavelet approximation of derivatives can be used in cases when the domain is complicated so that, e.g., a finite element procedure is used for the numerical approximation of the 2-D operator $L$.

Naturally, since it is the large eigenvalues of $L$ (those that tend to infinity) that are associated with the ill-posedness, it is natural to devise the following regularization method:

- Compute approximations of the smallest eigenvalues of $L$ and the corresponding eigenfunctions, and discard the components of the solution (2) that correspond to large eigenvalues.

It is straightforward to prove that such a method is a regularization method in the sense that the solution depends continuously on the data (Theorem 3). However, in
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

The direct implementation of such a method one would use unnecessarily much work to compute eigenvalue-eigenfunction approximations that are not needed for the particular data function \( g \). Thus the main contribution of this paper is a numerical method for approximating the regularized solution that has the following characteristics:

- The solution (2) is approximated by a projection onto a subspace computed by means of a Krylov sequence generated using the operator \( L^{-1} \).
- In each step of the Krylov method a well-posed two-dimensional elliptic problem involving \( L \) is solved. Any standard (black box) elliptic solver, derived from the discretization of \( L \), can be used.
- The hyperbolic cosine of the restriction of the operator \( L^{-1} \) to a low-dimensional subspace is computed.
- The method takes advantage of the fact that the regularized solution operator is applied to the particular data function \( g \).

We will demonstrate that the proposed method requires considerably fewer solutions of two-dimensional elliptic problems, than the approach based on the eigenvalue expansion.

A recent survey of the literature on the Cauchy problem for the Laplace equation is given in [2], see also [4]. There are many engineering applications of ill-posed Cauchy problems, see [34] and the references therein. A standard approach for solving Cauchy problems of this type is to apply an iterative procedure, where a certain energy functional is minimized; a recent example is given in [1]. Very general (non-cylindrical) problems can be handled, but if the procedure from [1] were to be applied to our problem, then at each iteration four well-posed elliptic equations would have to be solved over the whole three-dimensional domain. In contrast, our approach for the cylindrical case requires the solution of only one two-dimensional problem at each iteration‡.

We would also like to stress that we are not aware of any papers in the literature that treat the numerical solution of elliptic Cauchy problems in three dimensions. Those in the reference list of [2] all discuss less general problems.

Krylov methods with explicit regularization have been used before for ill-posed problems. For instance, [24, 6] describe regularized Lanczos (Golub-Kahan style) bidiagonalization procedures for the solution of integral equations of the first kind. Our approach is different in that it uses a Krylov method for approximating the solution operator.

We conclude by noticing that the procedure described in this paper can be generalized in a straightforward way to problems in more than three space dimensions.

The paper is organized as follows. In Section 2 we give a brief review of the ill-posedness and stabilization of the problem. More details of this are given in Appendix B. The Krylov method is described in Section 3. In Section 4 we describe a couple of

‡ In one of our examples the two-dimensional problem had 8065 degrees of freedom and the three-dimensional one had 129040.
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

In Appendix A we show that the assumption that the Cauchy data are \( u_z(x, y, 0) = 0 \) is no restriction: the general case can be transformed to this special case by solving a 3-D well-posed problem.

Throughout we will use an \( L_2(\Omega) \) setting with inner product and norm,
\[
\langle f, g \rangle = \int_{\Omega} f(x, y) g(x, y) \, dx \, dy, \quad \| f \| = \langle f, f \rangle^{1/2},
\]
and their finite-dimensional counterparts.

2. Ill-posedness and Stabilization of the Cauchy Problem

Let the eigenvalues and eigenfunctions of the operator \( L \) be \( (\lambda^2_\nu, s_\nu(x, y))_1^\infty \); the eigenfunctions are orthonormal with respect to the inner product (3), and the system of eigenfunctions is complete; see, e.g., [9, XIV.6.25], [14, Chapter 6.5]. Further, we assume that the eigenvalues are ordered as \( 0 < \lambda_1 \leq \lambda_2 \leq \cdots \). In analogy to the case when Fourier analysis can be used, we will refer to the eigenvalues as frequencies.

It is a standard exercise (see Appendix B) in Hilbert space theory to show that the formal solution (2) can be understood as an expansion in terms of eigenfunctions,
\[
\sum_{\nu=1}^{\infty} \cosh(\lambda_\nu z) \langle s_\nu, g \rangle \ s_\nu(x, y).
\]

The unboundedness of the solution operator is evident: a high-frequency perturbation of the data, \( g_m = g + e \), will cause the corresponding solution to blow up.

It is customary in ill-posedness problems to incorporate the data perturbation in the problem formulation and stabilize the problem by assuming that the solution is bounded. Thus we define the stabilized problem,
\[
\begin{align*}
u_{zz} - Lu &= 0, \quad (x, y) \in \Omega, \quad z \in [0, z_1], \\
u(x, y, z) &= 0, \quad (x, y) \in \partial \Omega, \quad z \in [0, z_1], \\
u_z(x, y, 0) &= 0, \quad (x, y) \in \Omega, \\
\| u(\cdot, \cdot, 0) - g_m(\cdot, \cdot) \| &\leq \epsilon, \\
\| u(\cdot, \cdot, z_1) \| &\leq M.
\end{align*}
\]

It is again a standard exercise (and for this reason we relegate it to Appendix B) to demonstrate that the solution of (5)-(9) is stable, but not unique.

**Proposition 1.** Any two solutions, \( u_1 \) and \( u_2 \), of the stabilized problem (5)-(9) satisfy
\[
\| u_1(\cdot, \cdot, z) - u_2(\cdot, \cdot, z) \| \leq 2 e^{1-z/z_1} M^{z/z_1}, \quad 0 \leq z < z_1.
\]
Definition 2. For $\lambda_c > 0$, a regularized solution is given by

$$v(x, y, z) = \sum_{\lambda \leq \lambda_c} \cosh(\lambda \nu z) \langle s_\nu, g_m \rangle s_\nu(x, y). \tag{11}$$

The quantity $\lambda_c$ is referred to as a cut-off frequency. It is easy to show that the function $v$ satisfies an error bound that is optimal in the sense that it is of the same type as that in Proposition 1. A proof is given in Appendix B.3.

Theorem 3. Suppose that $u$ is a solution defined by (4) (with exact data $g$), and that $v$ is a regularized solution (11) with measured data $g_m$, satisfying $\|g - g_m\| \leq \epsilon$. If $\|u(\cdot, \cdot, 1)\| \leq M$, and if we choose $\lambda_c = (1/z_1) \log(M/\epsilon)$, then we have the error bound

$$\|u(\cdot, \cdot, z) - v(\cdot, \cdot, z)\| \leq 3\epsilon^{1-z/z_1} M^{z/z_1}, \quad 0 \leq z \leq z_1. \tag{12}$$

The result above indicates that if we can solve approximately the eigenvalue problem for the operator $L$, i.e. compute good approximations of the eigenvalues and eigenfunctions for $\lambda_\nu \leq \lambda_c$, then we can compute a good approximation of the regularized solution. The solution of the eigenvalue problem for the smallest eigenvalues and eigenfunctions by a modern eigenvalue algorithm for sparse matrices [3] requires us to solve a large number of well-posed 2-D elliptic problems with a discretization of $L$.

If we use the eigenvalue approach then we do not take into account that we actually want to compute not a good approximation of the solution operator itself but rather the solution operator applied to the particular right-hand side $g_m$. We will now show that it is possible to obtain a good approximation of (11) much more cheaply by using a Krylov subspace method initialized with $g_m$.

Remark Theorem 3 only gives continuity in the interior of the interval, $[0, z_1]$. In the theory of ill-posed Cauchy problems one often can obtain continuous dependence on the data for the closed interval $[0, z_1]$ by assuming additional smoothness and using a stronger norm, see e.g. [27, Theorem 3.2]. We are convinced that this can be done also here, but we have not pursued this.

3. A Krylov Subspace Method

From now on we assume that the problem has been discretized with respect to $(x, y)$, and that the operator $L \in \mathbb{R}^{N \times N}$ is a symmetric, positive definite matrix. The details of the discretization are unimportant for our presentation, we only assume that it is fine enough so that the discretization errors are small compared to the uncertainty $\epsilon$ of the data; this means that $L$ is a good approximation of the differential operator, whose unboundedness is reflected in a large norm of $L$. In the following we use small roman letters to denote vectors that are the discrete analogs of the continuous quantities. Thus the solution vector $u(z)$ is a vector-valued function of $z$. 
For a given \( z \), the discrete analogs of the formal and regularized solutions in (2) and in (11) are given by

\[
u(z) = \cosh(z\sqrt{L})g = \sum_{j=1}^{N} (\cosh(z\lambda_j)s_j^\top g)s_j, \tag{13}
\]
\[
v(z) = \sum_{\lambda_j \leq \lambda_2} (\cosh(z\lambda_j)s_j^\top g_m)s_j, \tag{14}
\]
respectively, where \((\lambda_j^2, s_j)\) are the eigenpairs of \( L \), such that \( 0 < \lambda_1^2 \leq \cdots \leq \lambda_N^2 \).

We will now discuss how to compute an approximation of (14) using a Krylov subspace method, which is an iterative method. Error estimates for the Krylov approximation are given in Propositions 4 and 5, which form the bases of a stopping criterion that is derived in Section 3.2.

Krylov subspace approximations of matrix functions have been extensively employed in the solution of certain discretized partial differential equations, see, e.g., \[15, 32, 19, 33, 18, 20, 7\], while more recently attention has been devoted to acceleration procedures, see, e.g., \[8, 23, 21, 26\], where shift-invert type procedures are explored. The standard approach consists in generating the Krylov subspace \( K_k(L, g) = \text{span}\{g, Lg, \ldots, L^{k-1}g\} \) by a step-wise procedure (for details, see Section 3.2). Let \((\hat{q}_i)_{i=1}^k\) be an orthonormal basis of \( K_k(L, g) \), with \( \hat{q}_1 = g/\|g\| \).

Letting \( \hat{Q}_k = (\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_k) \) and \( \hat{T}_k = \hat{Q}_k^\top L\hat{Q}_k \in \mathbb{R}^{k \times k} \) be the symmetric representation of \( L \) onto the space, an approximation to \( u \) in \( K_k(L, g) \) may be obtained by projection,

\[
u_k(z) = \hat{Q}_k \cosh(z\hat{T}_k^{1/2})e_1 \|g\|. \tag{15}
\]
Here and in the following, \( e_j \) denotes the \( j \)'th canonical vector, of appropriate dimension. It may be shown that the error norm satisfies

\[
\|u_k(z) - u(z)\| \approx \frac{\alpha}{2k^2} \exp \left( -\alpha \left( \frac{2k^2}{\alpha^2} + O\left(\frac{2k}{\alpha} \right) \right) \right), \tag{16}
\]
where \( \alpha = z\lambda_{\text{max}} \) and \( \lambda_{\text{max}}^2 \) is the largest eigenvalue of \( L \). Convergence is superlinear, and the quality of the approximation depends on how small \( \lambda_{\text{max}} \) is. An approximation to the stabilized solution (14) in the approximation space \( K_k(L, g_m) \) (note that \( g \) has been replaced by \( g_m \)) may be obtained by accordingly truncating the expansion of the solution \( u_k \) in terms of the eigenpairs of \( \hat{T}_k \).

In our context, the smallest eigenvalues of \( L \) are the quantities of interest; cf. (14). Since the convergence of the Krylov subspace approximation is faster away from the origin (see, e.g., \[3, \text{Section 4.4.3}\]), a shift-invert procedure is commonly used to speed up convergence to the eigenvalues closest to a target value. More precisely, the spectral approximation is obtained in the Krylov subspace

\[
K_k(L^{-1}, g_m) = \text{span}\{g_m, L^{-1}g_m, \ldots, L^{-(k-1)}g_m\},
\]
or more generally, in \( K_k((L - \tau I)^{-1}, g_m) \) for some well selected value of \( \tau \). For simplicity of exposition, we assume in this section that \( \tau = 0 \), and let the orthonormal columns of
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

\( Q_k \) span such a space. If the Arnoldi process is employed to generate the orthonormal basis, we have the relation (see, e.g., [3])

\[
L^{-1}Q_k = Q_kT_k + \beta_{k+1}q_{k+1}e_{k+1}^T, \quad q_{k+1}^TQ_k = 0.
\]

Let \( ((\theta_j^{(k)})^2, y_j^{(k)}), j = 1, \ldots, k \) be the eigenpairs of \( T_k^{-1} \), so that \( ((\theta_j^{(k)})^2, Q_ky_j^{(k)}), j = 1, \ldots, k \) approximate some of the eigenpairs of \( L \). Using \( \cosh(zT_k^{-1/2}) = \sum_{j=1}^k y_j^{(k)} \cosh(z\theta_j^{(k)})(y_j^{(k)})^T \), the truncated approximation can be obtained as

\[
v_k(z) = Q_k \sum_{\theta_j^{(k)} \leq \lambda_c} y_j^{(k)} \cosh(z\theta_j^{(k)})(y_j^{(k)})^Te_1\|g_m\|.
\]

If our purpose were to first accurately approximate the small eigenvalues of \( L \) and then compute \( v_k(z) \) above, then we would have made the problem considerably harder. Indeed, the convergence rate of eigenvalue and eigenvector approximations is in general significantly slower than that of the matrix function (cf. (16)). Following [25, Th. 12.4.1 and Th. 12.4.3], for each eigenpair \( (\lambda_j^2, s_j) \) of interest, one would obtain

\[
|((\theta_j^{(k)})^2 - \lambda_j^2)| = O(2 \exp(-4k\sqrt{\gamma})), \quad \tan(s_j, K_k(L^{-1}, g_m)) = O(2 \exp(-2k\sqrt{\gamma})),
\]

where \( \gamma \) is related to the gap between the sought after eigenvalues and the rest of the spectrum.

Fortunately, we can merge the two steps of the spectral approximation and the computation of \( v_k(z) \), without first computing accurate eigenpairs. By computing the sought after solution while approximating the eigenpairs, the iterative process can be stopped as soon as the required solution is satisfactorily good (see section 3.2 for a discussion on the stopping criterion). In particular, the number of terms in the sum defining \( v_k(z) \) can be chosen dynamically as \( k \) increases, since the number of eigenvalues \( \theta_j^{(k)} \) less than \( \lambda_c \) may increase as \( k \) grows.

The value of \( \lambda_c \) depends on the data perturbation, (see Theorem 3), and it may be known approximately a priori. However, the number of eigenvalues smaller than \( \lambda_c \) is usually not known. As a consequence, it is not possible to fix a priori the number of summation terms neither in \( v(z) \) (stabilized solution (14)) nor in \( v_k(z) \) (Krylov approximation (17) of the stabilized solution). Clearly, these problems would dramatically penalize an approach that first computes accurate eigenvalues and then obtains \( v_k \).

We would also like to stress that, although the convergence rate of \( v_k(z) \) does depend on the eigenpairs and thus it is slower than that in (16), there is absolutely no need to get accurate spectral approximants; indeed, the final error norm \( \|v_k(z) - u(z)\| \) stagnates at a level that depends on the data perturbation, much before accurate spectral approximation takes place. This fact is investigated in the next section.

3.1. Accuracy of the Stabilized Approximation

As a first motivation for the stopping criterion, we now look at an error estimate for the Krylov subspace solution. Note that it is possible to derive an error estimate of the
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

type (12) also for the problem that is discretized in \( \Omega \). Therefore we want to express the error estimate for the Krylov approximation in similar terms, as much as is possible.

Let \( F(z, \lambda) = \cosh(z\sqrt{\lambda}) \) and let \( L^c \) be the restriction of \( L \) onto the invariant subspace of eigenvalues less than the threshold \( \lambda_c \). Let \( E^c \) be the orthogonal projector associated with the eigenvalues less than the threshold. Define \( S_k = T_k^{-1} \) and adopt the corresponding notation for \( S_k^c \). Given \( u(z) = F(z, L)g \) and \( v_k(z) = Q_k F(z, S_k^c)Q_k^\top g_m \), we want to estimate the error norm \( \|u - v_k\| \) so that we can emphasize the stagnation level. We have

\[
\|u(z) - v_k(z)\| = \|F(z, L)g - Q_k F(z, S_k^c)Q_k^\top g_m\|
\leq \|(F(z, L)g - Q_k F(z, S_k^c)Q_k^\top g)\|
+ \|Q_k F(z, S_k^c)Q_k^\top (g - g_m)\| =: \alpha + \beta.
\]

As in Lemma 8 in Appendix B, \( \beta \) can be bounded as follows:

\[
\beta \leq \|Q_k F(z, S_k^c)Q_k^\top\| \|g - g_m\| \leq \exp(\lambda_c) \epsilon \leq \epsilon^{1-z/z_1} M^{z/z_1}.
\]

Then we can estimate

\[
\alpha = \|(F(z, L) - Q_k F(z, S_k^c)Q_k^\top)g\| \leq \alpha_1 + \alpha_2,
\]

where, for the first term we use \( g = (F(z_1, L))^{-1} f \),

\[
\alpha_1 = \|(I - E^c)(F(z, L) - Q_k F(z, S_k^c)Q_k^\top)(F(z_1, L))^{-1} f)\|
\]

\[
\alpha_2 = \|E^c(F(z, L) - Q_k F(z, S_k^c)Q_k^\top)g\|.
\]

Since \( (I - E^c)F(z, L^c) = 0 \), we have

\[
\alpha_1 \leq \|(I - E^c)F(z, L)(F(z_1, L))^{-1} f)\|
+ \|(I - E^c)(F(z, L^c) - Q_k F(z, S_k^c)Q_k^\top)F(z_1, L))^{-1} f)\|.
\]

The first term (19) can be estimated as in the last part of the proof of Lemma 8, giving

\[
\|(I - E^c)(F(z, L)(F(z_1, L))^{-1})\| \leq 2\epsilon^{1-z/z_1} M^{z/z_1},
\]

while the second term is bounded by \( \|(F(z, L^c) - Q_k F(z, S_k^c)Q_k^\top)\| \). Moreover,

\[
\alpha_2 = \|E^c(F(z, L) - Q_k F(z, S_k^c)Q_k^\top)g\|
= \|E^c(F(z, L^c) - Q_k F(z, S_k^c)Q_k^\top)g\| \leq \|(F(z, L^c) - Q_k F(z, S_k^c)Q_k^\top)g\|.
\]

We have thus proved the following error estimate.

**Proposition 4.** Let \( u \) be defined by (13) and assume that hypotheses corresponding to those in Theorem 3 hold. Let \( v_k \) be defined by (17). Then

\[
\|u(z) - v_k(z)\| \leq 3\epsilon^{1-z/z_1} M^{z/z_1} + 2\|F(z, L^c) - Q_k F(z, S_k^c)Q_k^\top g\|.
\]

The two terms in the upper bound of Proposition 4 emphasize different stages of the convergence history. The error \( \|u(z) - v_k(z)\| \) may be large as long as the approximate low frequencies are not accurate. Once this accuracy has improved sufficiently, then the error \( \|u(z) - v_k(z)\| \) is dominated by the “intrinsic error”, due to the data perturbation. This behavior is confirmed by our numerical experiments; see Section 4.
3.2. Implementation aspects

The matrix \( Q_k \), whose columns \( q_1, \ldots, q_k \) span the Krylov subspace \( K_k(L^{-1}, g_m) \), may be obtained one vector at the time by means of the Lanczos procedure. Starting with \( q_0 = 0 \) and \( q_1 = g_m/\| g_m \| \), this process generates the subsequent columns \( q_2, q_3, \ldots \) by means of the following short-term recurrence,

\[
L^{-1}q_k = q_{k-1}^T \beta_{k-1} + q_k \alpha_k + q_{k+1}^T \beta_k, \quad k = 1, 2, \ldots,
\]

with \( \alpha_k = q_k^T L^{-1} q_k \) and \( \beta_k = q_{k+1}^T L^{-1} q_k \); see, e.g. [25, 3]. An analogous recurrence is derived when the shift-inverted matrix \((L - \tau I)^{-1}\) is employed. These coefficients form the entries of the tridiagonal symmetric matrix \( T_k \), that is \( T_k = \text{tridiag}(\beta_{k-1}, \alpha_k, \beta_k) \), with the \( \alpha_k \)'s on the main diagonal. At each iteration \( k \), the eigenpairs of \( T_k \) are computed, and the approximate solution \( v_k \) in (17) could be derived. An approximation to the theoretical quantity \( \lambda \) is determined a-priori (see Theorem 3), so that the partial sum in (17) is readily obtained. The process is stopped when the approximate solution is sufficiently accurate. In the absence of a stopping criterion based on the true error, we consider the difference between consecutive solutions as a stopping strategy, i.e.,

\[
\text{if } \|v_{k+1} - v_k\| < \text{tol} \text{ then stop.}
\]

This difference may be computed without first performing the expensive multiplication by \( Q_k \). Indeed, for \( v_j = Q_j w_j, j = k, k + 1 \), with \( w_j \in \mathbb{R}^l \), we have

\[
\|v_{k+1} - v_k\| = \left\|w_{k+1} - \begin{bmatrix} w_k \\ 0 \end{bmatrix} \right\|.
\]

In the next proposition we bound this norm in a way that emphasizes the dependence on the spectral accuracy.

**Proposition 5.** Let \( d > 0 \) be an natural number. Then for \( k \) and \( d \) large enough,

\[
\|v_{k+d}(z) - v_k(z)\| \lesssim 6e^{1-z/1}M^{z/1} + 2\|(F(z, L^c) - Q_k F(z, S_k^c) Q_k^T)g\|.(24)
\]

**Proof.** Using Proposition 4, we have

\[
\|v_{k+d}(z) - v_k(z)\| \leq \|v_{k+d}(z) - u(z)\| + \|v_k(z) - u(z)\|
\]

\[
\leq 6e^{1-z/1}M^{z/1} + 2\|(F(z, L^c) - Q_k F(z, S_k^c) Q_k^T)g\|
\]

\[
+ 2\|(F(z, L^c) - Q_{k+d} F(z, S_{k+d}^c) Q_{k+d}^T)g\|.
\]

If \( k \) is large enough,

\[
\|(F(z, L^c) - Q_{k+d} F(z, S_{k+d}^c) Q_{k+d}^T)g\| \ll \|(F(z, L^c) - Q_k F(z, S_k^c) Q_k^T)g\|,
\]

and the final estimate follows. \(\square\)

Proposition 5 shows that the difference between two subsequent estimates depends on the quality of the approximation to the low frequencies and on the data perturbation. Clearly, the quantity \( \|v_{k+d}(z) - v_k(z)\| \) may be small without the two right-hand side terms in (24) being small. However, our numerical experience suggests that premature
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

stagnation in the approximation is rarely encountered, and a small \( \|v_{k+d}(z) - v_k(z)\| \), even for \( d = 1 \), is usually associated with the final stagnation of \( v_k \), at the level of the data perturbation.

To ensure that the stopping criterion is also in agreement with the standard inverse problem framework, we compute the residual

\[
r_k = K v_k - g,
\]

where \( K \) is the operator that maps the function \( f(x, y, z_1) = u(x, y, z_1) \), with data \( u_z(x, y, 0) = 0 \), and homogeneous boundary values on the lateral boundary \( \partial \Omega \times [0, z_1] \), to the function values at the lower boundary, \( u(x, y, 0) \). This is related to the discrepancy principle [13, p. 84], [17, p. 179], but our criterion is different in one important aspect.

A stopping criterion for an iterative regularization procedure, based on the discrepancy principle is usually of the type “stop iterating as soon as \( \|r_k\| \leq C \epsilon \)”, where \( C \) is of the order 2, say, and \( \epsilon = \|g - g_m\| \) (the safety factor \( C \) is used to make up for the possible uncertainty concerning the knowledge of \( \epsilon \)). There the number of iterations is the regularization parameter. In our method the cut-off value \( \lambda_c \) is the regularization parameter (Theorem 3), and the stopping criterion only determines when the numerical approximation of the regularized solution is good enough.

In our problem, the computation of the residual requires the solution of a 3-D elliptic boundary value problem, which is much more costly than solving the 2-D elliptic problems. Therefore we only compute the residual in (25) when \( \|v_{k+1} - v_k\| \) is so small that we have a reasonable chance that the residual stopping criterion is satisfied.

The overall algorithm can be summarized as follows.

**Algorithm.** Given \( L, z, g_m, \text{tol}, \text{maxit}, \lambda_c \),

\[
q_0 = 0, q_1 = g_m/\|g_m\|, Q_1 = q_1, \beta_0 = 0
\]

for \( k = 1, 2, \ldots, \text{maxit} \)

\begin{align*}
&\text{Compute } \hat{q} = L^{-1} q_k - q_{k-1} \beta_{k-1} \\
&\text{Compute } \alpha_k = \hat{q}_k \hat{q} \\
&\text{Compute } \hat{q} = \hat{q} - q_k \alpha_k \\
&\text{Compute } \beta_k = \|\hat{q}\| \text{ and } q_{k+1} = \hat{q}/\beta_k \\
&\text{Expand } T_k \\
&\text{Compute eigenpairs } ((\theta_j^{(k)})^2, y_j^{(k)}) \text{ of } T_k^{-1} \\
&\text{Compute } w_k = \sum_{\theta_j^{(k)} \leq \lambda_c} y_j^{(k)} \cosh(z \theta_j^{(k)})(y_j^{(k)})^\top e_1 \|g_m\| \\
&\text{If } (k > 1 \text{ and } \|w_k - [w_{k-1}; 0]\| < \text{tol}) \text{ then} \\
&\quad \text{Compute the residual } r_k \\
&\quad \text{If } \|r_k\| \leq C \text{tol then Compute } v_k = Q_k w_k \text{ and stop} \\
&\text{endif} \\
&\text{Compute } Q_{k+1} = [Q_k, q_{k+1}] \\
&\text{endfor} \\
&\text{Compute } v_k = Q_k w_k
\end{align*}

The recurrence above generates the new basis vector by means of a coupled
two-term recurrence, which is known to have better stability properties than the three-term recurrence (22); see, e.g., [3, section 4.4]. In a practical implementation, additional safeguard strategies such as partial or selective reorthogonalization are commonly implemented to avoid well known loss of orthogonality problems in the Lanczos recurrence [3, section 4.4.4].

3.3. Dealing with the generalized problem

When procedures such as the finite element method are used to discretize the given equation over the space variables, equation (5) becomes

$$Hu_{zz} - Lu = 0,$$

where $H$ is the $N \times N$ symmetric and positive definite matrix associated with the employed inner product; it is usually called the mass matrix. Clearly, using the Cholesky factorization of $H$, i.e. $H = R^{\top}R$, the equation in (26) may be reformulated in the original way as $\tilde{u}_{zz} - \tilde{L}\tilde{u} = 0$, where $\tilde{L} = R^{-\top}LR^{-1}$, and $\tilde{u} = Ru$. Such procedure entails performing the factorization of $H$ and applying the factors and their inverses, whenever the matrix $\tilde{L}$ is employed.

To avoid the explicit use of the factorization of $H$, one can rewrite (26) as

$$u_{zz} + H^{-1}Lu = 0.$$

Since both $H$ and $L$ are symmetric and positive definite, the eigenvalues of $H^{-1}L$ are all real and equal to those of $\tilde{L}$. Moreover, the eigenvectors $s_j$ of $H^{-1}L$ are $H$-orthogonal, and can be made to be $H$-orthonormal by a scaling, that is, $\tilde{s}_j = s_j/\sqrt{s_j^{\top}Hs_j}$. Therefore, setting $\tilde{S} = [\tilde{s}_1, \ldots, \tilde{s}_N]$ so that $H^{-1}L = \tilde{S}\Lambda\tilde{S}^{-1}$ and $\tilde{S}^{-1} = \tilde{S}^{\top}H$, we have that

$$u = \cosh(z\sqrt{H^{-1}L})g = \tilde{S}\cosh(z\Lambda)\tilde{S}^{-1}g = \tilde{S}\cosh(z\Lambda)\tilde{S}^{\top}Hg = \sum_{j=1}^{N}(\cosh(z\lambda_j)\tilde{s}_j^{\top}Hg)\tilde{s}_j.$$

Hence, the stabilized approximation may be obtained by truncating the eigenvalue sum.

The approximation with the Lanczos algorithm may be adapted similarly. Following a procedure that is typical in the generalized eigenvalue context, see, e.g., [3, Chapter 5], the approximation to the stabilized solution may be sought after in the Krylov subspace $K_k((L - \tau H)^{-1}, g_m)$. The basis vectors are computed so as to satisfy an $H$-orthogonality condition, that is $q_{k+1}^{\top}HQ_k = 0$; see, e.g., [3, section 5.5].

It is important to remark that the use of the mass matrix also affects the norm employed throughout the analysis, and in particular, in the determination of the perturbation tolerance $\epsilon$ associated with the measured data $g_m$; see Theorem 3. More precisely, we assume that $g_m$ satisfies

$$\|g - g_m\|_{H}^2 := (g - g_m)^{\top}H(g - g_m) \leq \epsilon^2,$$

and the error is measured in the same norm.
4. Numerical Experiments

Example 1. In our numerical experiments we used MATLAB 7.5. In the first example we chose the region $\Omega$ to be the unit square $[0, 1] \times [0, 1]$, and the operator the Laplace operator. Thus the Cauchy problem was

\[ u_{zz} + \Delta u = 0, \quad (x, y, z) \in \Omega \times [0, 0.1], \]
\[ u(x, y, z) = 0, \quad (x, y, z) \in \partial \Omega \times [0, 0.1], \]
\[ u(x, y, 0) = g(x, y), \quad (x, y) \in \Omega, \]
\[ u_z(x, y, 0) = 0, \quad (x, y) \in \Omega. \]

We wanted to determine the values at the upper boundary, that is $f(x, y) = u(x, y, 0.1), (x, y) \in \Omega$.

In constructing our data we chose a solution, $f(x, y) = 30x(1-x)^6y(1-y)$; with this solution the Cauchy problem is not easy to solve, as $|\partial f / \partial x|$ is relatively large along $x = 0$. We computed the data function $g(x, y)$ by solving the well posed problem with boundary values $u(x, y, 0.1) = f(x, y)$ and $u_z(x, y, 0) = 0$, and evaluating the solution at the lower boundary $z = 0$. That solution was taken as the “exact data” $g$. The well-posed problem was solved by separation of variables: trigonometric interpolation of $f$ (relative interpolation error of the order of the machine precision) on a grid with $h_x = h_y = 0.01$, and numerical evaluation of the hyperbolic cosine (i.e. MATLAB’s $\text{cosh}$). In Figure 1 we give the solution and the data function.

![Figure 1. Example 1. True solution (left) and data function (right).](image)

We then perturbed the data, and added normally distributed noise to each component, giving $g_m$. The relative data perturbation $\|g - g_m\| / \|g\|$ was of the order 0.0085. From the singular value expansion (B.3) we can deduce that the condition number of the discrete problem is $\cosh(\lambda_{\text{max}1}) / \cosh(\lambda_{\text{min}1})$, where $\lambda_{\text{max}1}$ and $\lambda_{\text{min}1}$ are the largest and smallest eigenvalues of the matrix $L$. Here the condition number is $8.7 \cdot 10^{11}$ (the largest and smallest eigenvalues of the matrix $L$ are 80000 and 20, respectively), and therefore computing an unregularized solution with $g_m$ is completely meaningless.

Using the actual value of the norm of the solution $f$, the cut-off frequency defined in Theorem 3 was 50, approximately. The shift-invert Lanczos procedure was used as
described in the algorithm of section 3.2, with shift parameter $\tau$ equal to half the cut-off frequency chosen (see discussion below). To study the efficiency and the reliability of the stopping criterion, we plotted the norm of the true error, of the change in the solution, and of the residual $r_k$, as functions of the iteration number $k$. It turned out that when we used the cut-off level $\lambda_c$ as prescribed by Theorem 3, the error and the residual norms did not reach a stagnation level during the first 40 iterations. However, when we chose the cut-off level to be $0.75\lambda_c$, which provided us with equally good results for this solution, the curves leveled off (on the average) after about 20 steps. We illustrate this in Figure 2. This interesting behavior may be explained by looking at the approximation process of the required frequencies. For a cut-off value of $0.75\lambda_c$ fewer eigenpairs need to be approximated. More precisely, a large cut-off value forces the method to approximate more eigenvalues that are farther away from the shift $\tau$. As a consequence, it takes more iterations before the corresponding second term in the right-hand side of (21) stops dominating; see also the discussion after Proposition 4.

![Figure 2. Example 1. Norm of relative solution difference (dashed with ◦), true relative error (solid line), relative residual (dashed), relative spectral error (dashed-dotted with +), as functions of the iteration $k$. Left: cut-off $\lambda_c$ according to the theory; Right: cut-off value equal to $0.75\lambda_c$.](image)

We then tested the stopping criterion in the two cases. For both the test of the difference between consecutive solutions and that of the residual, we used the tolerance

$$\frac{1.5\|g - g_m\|}{\|g_m\|},$$

which we can compute in the present situation. For a non-synthetic problem, the tolerance should be replaced by an estimate of the relative data perturbation.

In Figure 3 we give the computed solution evaluated at $y = 1/2$ using the two values of the cut-off level. In the smoother case (right plot), the behavior of the exact solution is fully captured. The left plot confirms what we observed in the left plot of Figure 2, that is a final (acceptable) approximate solution has not yet been reached, although the

§ For this example there were 98 eigenvalues satisfying $\lambda_i < 0.75\lambda_c$. 
lack of smoothness is not at all dramatic. We may also say that the solution is slightly under-regularized.

**Figure 3.** Example 1. The true solution (solid) and the approximate (dashed), evaluated at $y = 0.5$. Left: Cut-off $\lambda_c$ according to the theory. The stopping criterion was satisfied after 20 steps. Right: Cut-off $0.75\lambda_c$. The stopping criterion was satisfied after 25 steps. In both cases the residual stopping criterion was satisfied the first time it was tested.

**Example 2** Our second example illustrates the use of a finite element discretization, and our computations are based on the codes from the book [16]]. The region was defined as

$$\Omega = \{(x, y, z) \mid x^2 + y^2/4 \leq 1, \ 0 \leq z \leq z_1 = 0.6\}.$$

The operator $L$ was defined

$$L = -(k(x, y)u_x)_x - (k(x, y)u_y)_y, \quad k(x, y) = 1 + 0.25x^2y,$$

and the two-dimensional problem was discretized using linear elements and six mesh refinements, giving mass and stiffness matrices of dimension 8065. We prescribed the solution $u(x, y, z_1) = f(x, y) = (1 - x^2 - y^2/4) \exp(-(x - 0.2)^2 - y^2)$ on the upper boundary $z = z_1$, and $u_z(x, y, 0) = 0$ on the lower boundary. To generate the “exact data function” we solved the 3D problem, discretized in the $z$ direction using a central difference, with a step length $z_1/15$ (a problem with 129040 unknowns; this boundary value problem was solved using the MATLAB function `pcg` with an incomplete Cholesky preconditioner with drop tolerance $10^{-3}$). The exact solution and the unperturbed data function $g(x, y) = u(x, y, 0)$ are illustrated in Figure 4.

We then perturbed the data function by adding a normally distributed perturbation with standard deviation 0.03 giving the data function $g_m$ illustrated in Figure 5.

We computed the rational Krylov solution as in Example 1, with the modifications outlined in Section 3.3. Here we chose the cut-off level to be $0.5\lambda_c$. In Figure 6 we illustrate the convergence history and the computed solution.

The codes are available at [http://www.math.mtu.edu/~msgocken/fembook/](http://www.math.mtu.edu/~msgocken/fembook/)
Figure 4. Example 2. True solution (left) and data function (right).

Figure 5. Example 2. Perturbed data function, standard deviation 0.03.

Figure 6. Example 2. Convergence history (left): relative change in the solution between two consecutive Krylov steps (dashed with ◦), true relative error (solid), and the relative residual (dashed). The right plot shows the computed solution after twelve Krylov steps.

We also tested the stopping criterion, using the same parameters as in Example 1. When after twelve steps the solution difference criterion was satisfied, the residual criterion was checked, and it too was satisfied. To compare the number of elliptic solves to that in the solution based on the explicit eigencomputation in the eigenvalue
expansion, it is worth saying that for approximating the 11 eigenvalues below the cut-off level, \texttt{eigs} would require 57 elliptic solves. It is also important to realize that to ensure that all wanted eigenvalues are actually found, a larger number of eigenvalues should be sought after. In this particular case, the computation of the first 15 eigenvalues with \texttt{eigs} required 66 elliptic solves.

To see if the problem was sufficiently ill-conditioned to be interesting as a test example, we computed the condition number, and it was equal to $7 \cdot 10^{69}$ (which means that in IEEE double precision it qualifies as a “a discrete ill-posed problem”). We also solved the same problem with cut-off at $1.5\lambda_c$. The result after 12 Krylov steps is given in Figure 7. Clearly too many high frequencies are included in the solution approximation.

Figure 7. Example 2. The solution obtained with cut-off at $1.5\lambda_c$ and 12 Krylov steps.

5. Conclusions

We have proposed a truncated eigenfunction expansion method for solving an ill-posed Cauchy problem for an elliptic PDE in three space dimensions. The method approximates the explicit solution, involving a hyperbolic function on a low-dimensional subspace, by means of a rational Krylov method. A crucial part of the algorithm is to determine when to stop the iteration that increases the dimension of the Krylov subspace. We suggest a stopping criterion based on the relative change of the approximate solution, combined with a final stage check of the residual for the three-dimensional problem. The criterion reflects the accuracy of the approximation of the required components in the solution rather than the accuracy of all the eigenvalues that are smaller than the cut-off value. As a consequence, the procedure dynamically improves the accuracy of the sought after solution, with no a-priori knowledge on the number of involved eigenpairs. This represents a particular feature of this method, because no spectral information on the problem is required. Our preliminary experiments are very promising, and we plan to also adapt the strategy to more general problems.
6. Acknowledgements

We are indebted to Xiaoli Feng for useful literature hints.

Appendix

Appendix A. Transforming a General Cauchy Problem

Consider the Cauchy problem,

\[ u_{zz} - Lu = 0, \quad (x, y, z) \in \Omega \times [0, z_1], \]
\[ u(x, y, z) = b(x, y, z), \quad (x, y, z) \in \partial\Omega \times [0, z_1], \]
\[ u(x, y, 0) = g(x, y), \quad (x, y) \in \Omega, \]
\[ u_z(x, y, 0) = h(x, y), \quad (x, y) \in \Omega. \]  

(A.1)

The problem is to determine the values of \( u \) on the upper boundary, \( f(x, y) = u(x, y, z_1), \ (x, y) \in \Omega \).

We can transform this problem to a simpler one by using linearity. Let \( u_1 \) satisfy the well-posed problem

\[ u_{zz} - Lu = 0, \quad (x, y, z) \in \Omega \times [0, z_1], \]
\[ u(x, y, z) = b(x, y, z), \quad (x, y, z) \in \partial\Omega \times [0, z_1], \]
\[ u(x, y, z_1) = 0, \quad (x, y) \in \Omega, \]
\[ u_z(x, y, 0) = h(x, y), \quad (x, y) \in \Omega. \]  

Then, let \( u_2 \) be an approximate solution of the ill-posed Cauchy problem,

\[ u_{zz} - Lu = 0, \quad (x, y, z) \in \Omega \times [0, z_1], \]
\[ u(x, y, z) = 0, \quad (x, y, z) \in \partial\Omega \times [0, z_1], \]
\[ u(x, y, 0) = g(x, y) - u_1(x, y, 0), \quad (x, y) \in \Omega, \]
\[ u_z(x, y, 0) = 0, \quad (x, y) \in \Omega. \]  

(A.2)

Obviously, \( u = u_1 + u_2 \) is an approximate solution of the original ill-posed problem (A.1). Therefore, since, in principle, we can solve the well-posed problem with arbitrarily high accuracy, and since the stability analysis is usually considered as an asymptotic analysis as the data errors tend to zero, it makes sense to analyze the ill-posedness of the original problem in terms of the simplified problem (A.2).

Appendix B. Ill-Posedness and Regularization

Appendix B.1. Singular Value Analysis

In order to study the ill-posedness of the Cauchy problem (1) we will first write it in operator form as

\[ Kf = g, \]
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces 18

for some (compact) operator $K$, and then determine the singular value expansion of $K$.

Consider the well-posed problem

$$u_{zz} - Lu = 0, \quad (x, y) \in \Omega, \quad z \in [0, z_1],$$

$$u(x, y, z) = 0, \quad (x, y) \in \partial\Omega, \quad z \in [0, z_1],$$

$$u(x, y, z_1) = f(x, y), \quad (x, y) \in \Omega,$$

$$u_y(x, y, 0) = 0, \quad (x, y) \in \Omega.$$  

With the separation of variables ansatz

$$u(x, y, z) = \sum_{k=1}^{\infty} w^{(k)}(z) s_k(x, y),$$

where $s_k$ are the (orthonormal) eigenfunctions of $L$ (with zero boundary values), the equation $u_{zz} - Lu = 0$ becomes

$$\sum_{k=1}^{\infty} w_{zz}^{(k)}(z) s_k(x, y) - \sum_{k=1}^{\infty} \lambda_k^2 w^{(k)}(z) s_k(x, y) = 0,$$

where $\lambda_k^2$ are the eigenvalues of $L$.

Expanding the boundary values at $z = z_1$,

$$f(x, y) = \sum_{k=1}^{\infty} \langle s_k, f \rangle s_k(x, y),$$

we get a boundary value problem for an ordinary differential equation for each value of $k$,

$$w_{zz}^{(k)} = \lambda_k^2 w^{(k)}, \quad w^{(k)}(z_1) = \langle s_k, f \rangle, \quad w_z^{(k)}(0) = 0,$$

with the unique solution

$$w^{(k)}(z) = \frac{\cosh(\lambda_k z)}{\cosh(\lambda_k z_1)} \langle s_k, f \rangle, \quad 0 \leq z \leq z_1.$$

Thus we can write the solution of the elliptic equation

$$u(x, y, z) = \sum_{k=1}^{\infty} \frac{\cosh(\lambda_k z)}{\cosh(\lambda_k z_1)} \langle s_k, f \rangle s_k(x, y), \quad 0 \leq z \leq z_1, \quad (x, y) \in \Omega.$$  \hspace{1cm} (B.1)

Now consider the solution at the lower boundary,

$$g(x, y) = u(x, y, 0) = \sum_{k=1}^{\infty} \frac{1}{\cosh(\lambda_k z_1)} \langle s_k, f \rangle s_k(x, y).$$  \hspace{1cm} (B.2)

Summarizing the derivation above, we see that the Cauchy problem (1) can be written as an integral equation of the first kind $g = Kf$, where the integral operator is defined in terms of the eigenvalue expansion,

$$g(x, y) = \sum_{k=1}^{\infty} \frac{1}{\cosh(\lambda_k z_1)} \langle s_k, f \rangle s_k(x, y).$$  \hspace{1cm} (B.3)
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

Obviously this is the singular value expansion of the operator, and since the operator is self-adjoint this is the same as the eigenvalue expansion. Thus, the singular values and singular functions are

$$\sigma_k = \frac{1}{\cosh(\lambda_k z_1)}, \quad u_k = v_k = s_k.$$  

The $\lambda_k$ will also be referred to as frequencies. The eigenvalues $\lambda_k^2$ of a self-adjoint elliptic operator satisfy $\lambda_k \to \infty$ as $k \to \infty$. Therefore we have exponential decay of the singular values to zero and the problem is severely ill-posed.

Appendix B.2. Stability in a $z-$Cylinder

We can use the concept of logarithmic convexity to prove a stability result for the Cauchy problem. Put

$$F(z) = \int_\Omega |u(x, y, z)|^2 \, dxdy = \int_\Omega \left| \sum_k \frac{\cosh(\lambda_k z)}{\cosh(\lambda_k z_1)} \alpha_k s_k(x, y) \right|^2 \, dxdy$$

$$= \sum_k \frac{\cosh^2(\lambda_k z)}{\cosh^2(\lambda_k z_1)} \alpha_k^2,$$

where $\alpha_k = \langle s_k, f \rangle$, and where we have used the orthonormality of the eigenfunctions.

We will show that this function is log-convex. The first and second derivatives are

$$F'(z) = 2 \sum_k \frac{\cosh(\lambda_k z) \sinh(\lambda_k z)}{\cosh^2(\lambda_k z_1)} \lambda_k \alpha_k^2,$$

and

$$F''(z) = 2 \sum_k \frac{\sinh^2(\lambda_k z) + \cosh^2(\lambda_k z)}{\cosh^2(\lambda_k z_1)} \lambda_k^2 \alpha_k^2 \geq 4 \sum_k \frac{\sinh^2(\lambda_k z)}{\cosh^2(\lambda_k z_1)} \lambda_k^2 \alpha_k^2.$$

Then it follows that

$$F F'' - (F')^2 \geq$$

$$\geq 4 \left( \sum_k \frac{\cosh^2(\lambda_k z)}{\cosh^2(\lambda_k z_1)} \alpha_k^2 \right) \left( \sum_k \frac{\sinh^2(\lambda_k z)}{\cosh^2(\lambda_k z_1)} \lambda_k^2 \alpha_k^2 \right)$$

$$- 4 \left( \sum_k \frac{\cosh(\lambda_k z) \sinh(\lambda_k z)}{\cosh^2(\lambda_k z_1)} \lambda_k \alpha_k^2 \right)^2 \geq 0$$

by the Cauchy-Schwarz inequality. This implies that log $F$ is convex.

Now consider the stabilized problem,

$$u_{zz} - Lu = 0, \quad (x, y) \in \Omega, \quad z \in [0, z_1],$$

$$u(x, y, z) = 0, \quad (x, y) \in \partial \Omega, \quad z \in [0, z_1],$$

$$u_z(x, y, 0) = 0, \quad (x, y) \in \Omega,$$

$$\|u(\cdot, \cdot, 0) - g_m(\cdot, \cdot)\| \leq \epsilon,$$

¶ When $L$ is the 1D Laplace operator on the interval $[0, \pi]$, then $\lambda_k = k$ and $s_k(x) = \sin(kx)$.
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

\[ \|u(\cdot, \cdot, z_1)\| \leq M. \quad (B.4) \]

From logarithmic convexity it now follows that solutions of the stabilized problem depend continuously on the data (in an \( L_2(\Omega) \) sense) for \( 0 \leq z < z_1 \).

**Proposition 6.** Any two solutions, \( u_1 \) and \( u_2 \), of the stabilized problem satisfy

\[ \|u_1(\cdot, \cdot, z) - u_2(\cdot, \cdot, z)\| \leq 2e^{1-z/z_1} M^{z/z_1}, \quad 0 \leq z < z_1. \quad (B.5) \]

**Proof.** Put \( F(z) = \|u_1(\cdot, \cdot, z) - u_2(\cdot, \cdot, z)\| \). Since \( u_1 - u_2 \) satisfies the differential equation \( u_{zz} - Lu = 0 \), with the Cauchy data, \( F(z) \) is logarithmic convex. This implies that

\[ \log F(z) \leq (1 - z/z_1) \log F(0) + z/z_1 \log F(z_1), \]

or, equivalently,

\[ F(z) \leq F(0)^{1-z/z_1} F(1)^{z/z_1}, \]

Using the triangle inequality and the bounds in (B.4) we obtain (B.5). \( \square \)

**Appendix B.3. Regularization by Cutting off High Frequencies**

Taking the inner product with respect to \( s_k \) in the expansion (B.2) we get

\[ \langle s_k, f \rangle = \cosh(\lambda_k z_1) \langle s_k, g \rangle, \]

and therefore, using (B.1), we can write the solution of the Cauchy problem with exact data \( g \) formally as

\[ u(x, y, z) = \sum_{k=1}^{\infty} \cosh(\lambda_k z) \langle s_k, g \rangle s_k(x, y). \quad (B.6) \]

This does not work for inexact data \( g_m \) (nor for numerical computations with exact data), since high frequency noise (including floating-point round-off) will be blown up arbitrarily. However, we can obtain a useful approximate solution by cutting off high frequencies. Thus, define

\[ v(x, y, z) = \sum_{\lambda_k \leq \lambda_c} \cosh(\lambda_k z) \langle s_k, g_m \rangle s_k(x, y), \quad (B.7) \]

where \( \lambda_c \) is the cut-off frequency. We will call such a solution a regularized solution.

We will now show that a regularized solution satisfies an almost optimal error bound of the type (B.5). A couple of lemmas are needed.

**Lemma 7.** Assume that \( v_1 \) and \( v_2 \) are two regularized solutions defined by (B.7), with data \( g_1 \) and \( g_2 \), respectively, satisfying \( \|g_1 - g_2\| \leq \epsilon \). If we select \( \lambda_c = (1/z_1) \log(M/\epsilon) \), then we have the bound

\[ \|v_1(\cdot, \cdot, z) - v_2(\cdot, \cdot, z)\| \leq \epsilon^{1-z/z_1} M^{z/z_1}, \quad 0 \leq z \leq z_1. \quad (B.8) \]

**Proof.** Using the orthonormality of the eigenfunctions we have

\[ \|v_1 - v_2\|^2 = \sum_{\lambda_k \leq \lambda_c} (\cosh(\lambda_k z) \langle g_1 - g_2, s_k \rangle)^2 \]
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

\[ \leq (\cosh(\lambda_c z))^2 \sum_{\lambda_k \leq \lambda_c} (\langle g_1 - g_2, s_k \rangle)^2 \leq (\cosh(\lambda_c z))^2 \| g_1 - g_2 \|^2 \]

\[ \leq \exp(2\lambda_c z) \epsilon^2. \]

The inequality (B.8) now follows using \( \lambda_c = (1/z_1) \log(M/\epsilon). \)

**Lemma 8.** Let \( u \) be a solution defined by (B.1), and let \( v \) be a regularized solution with the same exact data \( g \). Suppose that \( \| u(\cdot, \cdot, 1) \| \leq M \). Then, if \( \lambda_c = (1/z_1) \log(M/\epsilon) \),

\[ \| u(\cdot, \cdot, z) - v(\cdot, \cdot, z) \| \leq 2\epsilon^{1-z/z_1} M^{z/z_1}, \quad 0 \leq z \leq z_1. \quad \text{(B.9)} \]

**Proof.** Using (B.2) we have

\[ v(x, y, z) = \sum_{\lambda_k \leq \lambda_c} \cosh(\lambda_k z) \langle s_k, g \rangle s_k(x, y) = \sum_{\lambda_k \leq \lambda_c} \frac{\cosh(\lambda_k z)}{\cosh(\lambda_k z_1)} \langle s_k, f \rangle s_k(x, y), \]

where \( f(x, y) = u(x, y, 1) \). Thus, from (B.1) we see that

\[ \| u - v \|^2 = \sum_{\lambda_k > \lambda_c} \left( \frac{\cosh(\lambda_k z)}{\cosh(\lambda_k z_1)} \langle s_k, f \rangle \right)^2 \leq 4 \exp(-2\lambda_c(z_1 - z)) \sum_{\lambda_k > \lambda_c} (\langle s_k, f \rangle)^2, \]

where we have used the elementary inequality

\[ \frac{1 + e^{-\lambda z}}{1 + e^{-\lambda z_1}} \leq 2, \quad \lambda \geq 0, \quad 0 \leq z \leq z_1. \]

The inequality (B.9) now follows from the assumptions \( \| f \| \leq M \) and \( \lambda_c = (1/z_1) \log(M/\epsilon). \)

Now the main error estimate can be proved.

**Theorem 9.** Suppose that \( u \) is a solution defined by (B.1), and that \( v \) is a regularized solution with measured data \( g_m \), satisfying \( \| g - g_m \| \leq \epsilon \). Then, if \( \| u(\cdot, \cdot, 1) \| \leq M \) and we choose \( \lambda_c = (1/z_1) \log(M/\epsilon) \), we have the error bound

\[ \| u(\cdot, \cdot, z) - v(\cdot, \cdot, z) \| \leq 3\epsilon^{1-z/z_1} M^{z/z_1}, \quad 0 \leq z \leq z_1. \]

**Proof.** Let \( v_1 \) be a regularized solution with exact data \( g \). Then using the two previous lemmas we get

\[ \| u - v \| \leq \| u - v_1 \| + \| v_1 - v \| \leq 3\epsilon^{1-z/z_1} M^{z/z_1}. \]
Numerical Solution of a Cauchy Problem for an Elliptic Equation by Krylov Subspaces

References


