AN INCOMPLETE LU PRECONDITIONER FOR PROBLEMS IN ACOUSTICS

MARTIN J. GANDER

Section de Mathématiques, Université de Genève, CP 240, 1211 Genève, Switzerland
Martin.Gander@unige.ch

FRÉDÉRIC NATAF

CMAP, Ecole Polytechnique, Palaiseau, France, CNRS, UMR7641
nataf@cmap.polytechnique.fr

We present an incomplete LU preconditioner for solving discretized Helmholtz problems. The preconditioner is based on an analytic factorization of the Helmholtz operator. This allows us to take the physical properties of the acoustics problem modeled by the Helmholtz equation into account in the preconditioner. We show how the parameters in the preconditioner can be chosen in order to make it effective. Numerical experiments show that the new preconditioner leads to convergent iterative methods even for large wave numbers, and it outperforms classical ILU preconditioners by a large margin.

Keywords: Helmholtz Equation; Preconditioning; ILU; Incomplete Factorization.

1. Introduction

Discretizing an elliptic problem $\mathcal{L}(u) = f$ with a finite element, finite difference or finite volume method, one obtains a matrix equation $A\boldsymbol{u} = \boldsymbol{f}$ where A is large and sparse. Hence Krylov methods are the methods of choice to solve these problems. Unfortunately Krylov methods can have serious convergence problems, especially for problems in acoustics, modeled by the Helmholtz equation [5,6]. These problems remain when matrix based, black box preconditioners like ILU are used. Even a preconditioner which comes close to a factorization, like ILU(1e-2), can not alleviate the situation [5,6].

Motivated by the effectiveness of ILU preconditioners based on the underlying analytic factorization of symmetric positive definite operators [4], we analyze in this paper two approximate factorizations based on the analytic factorization of the Helmholtz operator. Extensive numerical tests for the second factorization were very promising for solving the discretized Helmholtz equation with a preconditioned Krylov method [5,6], and a first analysis indicated how the parameters in the approximate factorization should be chosen for good performance [7].

The analytic factorization of elliptic operators has been of theoretical interest for some time [3,11], and more recently the use of these factorizations to construct preconditioners for convection diffusion problems was advocated [9,10,8]. However, the performance of these

factorizations was not satisfactory, because they were missing a link between the analytic factorization of the continuous operator and the exact block LU decomposition of the discretization of the same operator. Establishing this link [4] led to approximate factorizations of high quality for symmetric, positive definite problems. These factorizations are related to frequency filtering factorizations at the fully discrete level [14,15,12,13,2]. A recursive extension of AILU factorizations to 3d problems is also possible [1].

This paper is organized as follows: in Section 2, we introduce a continuous factorization of the Helmholtz operator. We show that this factorization leads to non-local operators, where the non-local nature corresponds to the fill-in one would obtain in a direct matrix factorization of the corresponding discretized problem. We propose two different sparse approximations of the factors, based on the insight from the continuous factorization, and analyze their performance when the factorization is used as a splitting in a stationary iterative method. While the continuous analysis shows that our approach is general, it does not reveal directly how the factorization can be used as a preconditioner in a discrete setting. We introduce therefore in Section 3 a factorization of the semi-discrete Helmholtz operator. This allows us to build a direct link with the fully discrete matrix factorization and leads to the AILU preconditioner. We analyze again the performance of the preconditioner in the semi-discrete setting and show how the parameters in AILU can be chosen to get good performance. In Section 4, we show extensive numerical experiments to illustrate the performance of AILU for acoustic problems, and to test how well the analysis predicts a good choice of the parameters in AILU. We present our conclusions in Section 5.

2. The Continuous Analytic Factorization

Given an elliptic operator $\mathcal{L}(u)$, we write the operator as a product of two evolution operators,

$$\mathcal{L}(u) = -(\partial_x + \Lambda_1)(\partial_x - \Lambda_2)(u) \tag{1}$$

where Λ_1 and Λ_2 are positive operators up to a compact operator. The first factor represents an evolution operator acting in the positive x direction and the second one an evolution operator acting in the negative x direction. We focus in the sequel on the Helmholtz operator $\mathcal{L} = (-\omega^2 - \Delta)$ as a model for acoustic problems.

Taking a Fourier transform of $\mathcal{L}$ in y, we obtain

$$\mathcal{F}_y(-\omega^2 - \Delta) = -\partial_{xx} + k^2 - \omega^2 = -(\partial_x + \sqrt{k^2 - \omega^2})(\partial_x - \sqrt{k^2 - \omega^2}), \tag{2}$$

and thus we have the continuous analytic factorization

$$(-\omega^2 - \Delta) = -(\partial_x + \Lambda_1)(\partial_x - \Lambda_2), \tag{3}$$

where $\Lambda_1 = \Lambda_2 = \mathcal{F}_y^{-1}(\sqrt{k^2 - \omega^2})$. Note that the Λ_j, $j = 1, 2$, are non local operators in y, because of the square root in their symbol $\sqrt{k^2 - \omega^2}$. This non-locality corresponds to the fill-in in an LU factorization of the discretized Helmholtz operator: the LU factorization also produces an operator looking forward, namely the L, and another operator looking

backward, namely the U, if the variables are in the corresponding lexicographic ordering [4].
Hence a local approximation of the nonlocal symbol in the factorization leads to a new type
of ILU preconditioner, which we call AILU, for "Analytic Incomplete LU".

Suppose we approximate the nonlocal operators Λ_j by a local approximation $\Lambda_j^{app} = \mathcal{F}_y^{-1}(p + qk^2)$, $p, q \in \mathbb{C}$, $\Re(q) > 0$. Then the approximate new operator is

$$\mathcal{L}^{app} = \mathcal{F}_y^{-1}(-\partial_{xx} + p^2 + 2pqk^2 + q^2k^4) \tag{4}$$

and an iterative method to solve the original problem would then only solve two local
evolution problems at each step,

$$\mathcal{L}^{app}(u^{n+1}) = (\mathcal{L}^{app} - \mathcal{L})(u^n) + f. \tag{5}$$

The parameters $p, q \in \mathbb{C}$ should be chosen so that the convergence factor $\rho := \mathcal{F}(1 - (\mathcal{L}^{app})^{-1}\mathcal{L})$ is as small as possible, except for possibly a few frequencies which will be taken
into account once the stationary iterative procedure (5) is replaced by a Krylov method. In
Fourier, the convergence factor becomes after some calculations

$$\rho(k, k_x, \omega, p, q) = \frac{(p + qk^2)^2 + w^2 - k^2}{(p + qk^2)^2 + k_x^2}, \tag{6}$$

where we have also transformed the x direction with Fourier parameter k_x.

To choose $p, q \in \mathbb{C}$, we first note that we can restrict our analysis to $k \geq 0$, because ρ
depends on k^2 only. Since the free parameters are complex, we have four real parameters
to use in the optimization process. To simplify the optimization, we make two different
particular choices in the next two subsections. These choices allow us to find good parameters
and to analyze the performance of the method.

2.1. *Low Frequency Approximation*

In this first approach, we choose the parameter p, such that the convergence factor vanishes
at $k = 0$, $p := i\omega$, and the parameter q such that the convergence factor vanishes again
at an intermediate frequency $k_2 > \omega$, $q := \frac{(k_2^2 - \omega^2)^{1/2} - i\omega}{k_2^2}$. This leads to an approximate
factorization which is exact for the zero frequency and also for the frequency k_2. Estimating
the lowest frequency in the x-direction by zero, $k_x = 0$, the modulus of the convergence
factor becomes

$$R(k, \omega, k_2) := |\rho(k, 0, \omega, i\omega, \tfrac{(k_2^2 - \omega^2)^{1/2} - i\omega}{k_2^2})|^2 = \frac{k^4(k - k_2)^2(k + k_2)^2}{(k^4 - 2\omega^2 k^2 + k_2^2\omega^2)^2}. \tag{7}$$

Note that this simplified convergence factor is not a uniform bound on the modulus of the
convergence factor of the method for all frequencies k_x, since p and q are complex numbers
in (6), but it allows us to further study the method for $k_x = 0$ and determine various choices
for the free parameters.

Theorem 2.1. *The convergence factor $R(k, \omega, k_2)$ given in (7) satisfies $R(k, \omega, k_2) \leq 1$ for
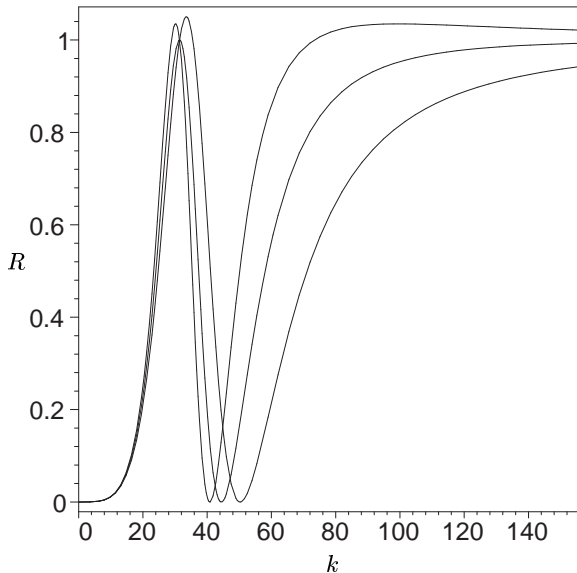all k, if and only if $k_2 = \bar{k}_2 := \sqrt{2}\omega$.*

4   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

Fig. 1. Convergence factor $R(k, \omega, k_2)$ for three values of k_2: one where $k_2 < \bar{k}_2 = \sqrt{2}\omega$, one where $k_2 = \bar{k}_2$, and one where $k_2 > \bar{k}_2$.

Proof. First we note that for $k = \omega$, the convergence factor $R(\omega, \omega, k_2) \equiv 1$, independently of what we choose for the parameter k_2. Now the partial derivative of R with respect to k is

$$\frac{\partial R}{\partial k} = 4 \frac{k^3 (k_2 - k)(k_2 + k)(k_2^4 \omega^2 - k_2^2 k^4 - 2k_2^2 \omega^2 k^2 + 2\omega^2 k^4)}{(k^4 + \omega^2 k_2^2 - 2\omega^2 k^2)^3}, \tag{8}$$

which evaluated at $k = \omega$ gives $4 \frac{k_2^2 - 2\omega^2}{\omega(k_2^2 - \omega^2)}$. Hence, if $k_2 \neq \bar{k}_2 = \sqrt{2}\omega$, then the slope at $k = \omega$ is non-zero, and since at $k = \omega$ we have $R(\omega, \omega, k_2) = 1$, this implies by continuity that $R(k, \omega, k_2) > 1$ for some values of k in the neighborhood of ω, which shows the only if part. For the if part, we find from (8) with $k_2 = \bar{k}_2 = \sqrt{2}\omega$ that the extrema in k are at 0, ω and $\bar{k}_2$. Since $R(k, \omega, \bar{k}_2) = 0$ at $k = 0$ and $k = \bar{k}_2$, these are minima, and in between at $k = \omega$ we find by continuity the only maximum, where $R(\omega, \omega, \bar{k}_2) = 1$. Hence $R(k, \omega, \bar{k}_2)$ has no local maximum bigger than one, and it could thus only exceed one for large k. But $\lim_{k \to \infty} R(k, \omega, \bar{k}_2) = 1$, which shows that $R(k, \omega, \bar{k}_2) \leq 1$ for all k.   □

In Figure 1, we show the convergence factor given in (7) for three different choices of the parameter k_2: one where $k_2 < \bar{k}_2 = \sqrt{2}\omega$, one where $k_2 = \bar{k}_2$, and one where $k_2 > \bar{k}_2$. One can clearly see the implications of Theorem 2.1: the convergence factor is only bounded by one for the choice $k = \bar{k}_2$.

    At first sight, the optimal choice of k_2 seems to be determined by Theorem 2.1, because with the choice $k_2 = \bar{k}_2 = \sqrt{2}\omega$, the method converges for all frequencies except for $k = \omega$ and $k = \infty$. Using Krylov acceleration, these two modes would be taken care of by the Krylov method and the preconditioned method would converge overall. But in a numerical computation, two additional issues come into play: first, a numerical grid can not carry

arbitrary large frequencies, there is a maximum frequency $k_{\max}$ which can be estimated for a grid-size h by $k_{\max} = \pi/h$. Second the domain is bounded, which leads to a discrete spectrum, $k = dn$, where n is an integer and d is some spacing. Hence we do not need $R(k, \omega, k_2) \leq 1$ for all k, we only need to consider $k \in K$, where

$$K := (0, \omega - \delta\omega) \cup (\omega + \delta\omega, k_{\max}), \tag{9}$$

and $\delta\omega$ is a parameter we can choose to determine how many of the modes around $k = \omega$ should be taken care of by the Krylov method. For example choosing $\delta\omega = d/2$, there is at most one mode left for the Krylov method, if we manage to obtain $R(k, \omega, k_2) < 1$ for $k \in K$. We first obtain an important corollary of Theorem 2.1 in this case.

Corollary 2.1. *Let $R(k, \omega, \bar{k}_2)$ be the convergence factor given in (7) with $k_2 = \bar{k}_2$ from Theorem 2.1, and let $k_{\max} = \pi/h$. Then, for $k \in K$ defined in (9), the convergence factor $R(k, \omega, \bar{k}_2)$ is for h small bounded by*

$$R(k, \omega, \bar{k}_2) \leq R(k_{\max}, \omega, \bar{k}_2) = 1 - 4\frac{\omega^4}{\pi^4}h^4 + O(h^6). \tag{10}$$

Proof. For $k_2 = \bar{k}_2 = \sqrt{2}\omega$, the convergence factor $R(k, \omega, \bar{k}_2)$ given in (7) attains its maximum on K at $k_{\max}$ for $k_{\max}$ large, and since $k_{\max} = \pi/h$, we have for h small that $R(k, \omega, \bar{k}_2) \leq R(k_{\max}, \omega, \bar{k}_2)$ for all $k \in K$. The result then follows by expanding $R(\pi/h, \omega, \bar{k}_2)$ for h small. $\qquad \square$

Now the question arises if there is a better choice of the parameter k_2 in the discretized setting, if we need to consider $k \in K$ from (9) only in the convergence factor given in (7). To obtain a fast method, we want to make the convergence factor as small as possible for all $k \in K$, which leads to the min-max problem

$$\min_{\omega+\delta\omega < k_2 < k_{\max}} \left(\max_{k \in K} R(k, \omega, k_2) \right) = \min_{\omega+\delta\omega < k_2 < k_{\max}} \left(\max_{k \in K} \frac{k^4(k - k_2)^2(k + k_2)^2}{(k^4 - 2\omega^2 k^2 + k_2^2\omega^2)^2} \right). \tag{11}$$

Theorem 2.2. *With K given in (9), the solution of the min-max problem (11) is for*

$$\begin{aligned} k_{\max} &\geq \sqrt{\frac{\delta\omega^4 + 4\omega^3\delta\omega + \omega^4 - 2\omega^2\delta\omega^2 + \sqrt{\omega^8 - 2\omega^4\delta\omega^4 + \delta\omega^8 + 16\omega^6\delta\omega^2}}{2(\delta\omega(2\omega - \delta\omega))}} \\ &= \frac{\omega^{3/2}}{\sqrt{2\delta\omega}} + \frac{5\sqrt{2\omega}}{8}\sqrt{\delta\omega} + O((\delta\omega)^{3/2}), \end{aligned} \tag{12}$$

given by

$$\bar{k}_2^* = \frac{\sqrt{2}(\omega + \delta\omega)k_{\max}}{\sqrt{k_{\max}^2 + (\omega + \delta\omega)^2}}. \tag{13}$$

Otherwise, the solution of the min-max problem is

$$\bar{k}_2^* = \frac{\sqrt{3\omega^4 + \delta\omega^4 + \sqrt{\omega^8 - 2\omega^4\delta\omega^4 + \delta\omega^8 + 16\omega^6\delta\omega^2}}}{\sqrt{2}\omega}. \tag{14}$$

6   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

Proof. From the partial derivative (8) of the convergence factor $R(k, \omega, k_2)$ with respect to k, we find that for $k_2 > \sqrt{2}\omega$, $R(k, \omega, k_2)$ has only one maximum in k, namely at

$$k = \hat{k}_1 := \frac{k_2\sqrt{\omega(k_2^2 - 2\omega^2)(\sqrt{k_2^2 - \omega^2} - \omega)}}{k_2^2 - 2\omega^2}. \tag{15}$$

For $\omega + \delta\omega < k_2 < \sqrt{2}\omega$, there is one maximum at $\hat{k}_1$ given in (15), and one at

$$k = \hat{k}_2 := \frac{k_2\sqrt{\omega(2\omega^2 - k_2^2)(\sqrt{k_2^2 - \omega^2} + \omega)}}{2\omega^2 - k_2^2}, \tag{16}$$

which is also a maximum, as one can see from the asymptotic expansion

$$R(k, \omega, k_2) = 1 + 2\frac{2\omega^2 - k_2^2}{k^2} + O(\frac{1}{k^4}), \quad \text{for } k \text{ large}, \tag{17}$$

which shows that $R(k, \omega, k_2)$ approaches one from above for k large, see also Figure 1 as an illustration. The expansion (17) also shows that at the second maximum, $k = \hat{k}_2$, we have $R(\hat{k}_2, \omega, k_2) \geq 1$. From Theorem 2.1, we know furthermore that at the first maximum we have $R(\hat{k}_1, \omega, k_2) \geq 1$ and hence the two maxima must be outside of the numerical frequency range K to obtain a convergence factor less than one for $k \in K$, which is possible, as we know already from the choice given in Theorem 2.1. Since there are no other maxima, the maxima in the solution of the min-max problem (11) can only be attained on the boundaries of the set K, at $k = k_- := \omega - \delta\omega$, $k = k_+ = \omega + \delta\omega$ or $k = k_{\max}$ (not at $k = 0$, because there the convergence factor vanishes). We therefore analyze the dependence of $R(k, \omega, k_2)$ on k_2 at these boundaries of K. Taking a partial derivative with respect to k_2, we find

$$\frac{\partial R}{\partial k_2} = \frac{4k^6 k_2(k_2^2 - k^2)(k^2 - \omega^2)}{(k^4 + \omega^2 k_2^2 - 2\omega^2 k^2)^3}. \tag{18}$$

For $k_2 > \omega$ the denominator is positive, as one can see by completing the square, and for $\omega < k < k_2$ the numerator is also positive, otherwise the numerator is negative. Hence at $k = k_-$ and $k = k_{\max}$ the convergence factor $R(k, \omega, k_2)$ decreases with increasing k_2, whereas at $k = k_+$ the convergence factor $R(k, \omega, k_2)$ increases with increasing k_2. A direct computation shows that for

$$\omega + \delta\omega \leq k_2 < \frac{\sqrt{3\omega^4 + \delta\omega^4 + \sqrt{\omega^8 - 2\omega^4\delta\omega^4 + \delta\omega^8 + 16\omega^6\delta\omega^2}}}{\sqrt{2}\omega}, \tag{19}$$

we have $R(k_-, k_2, \omega) > R(k_+, k_2, \omega)$. Thus the maximum is at k_- or at $k_{\max}$ and can be reduced by increasing k_2. When the convergence factor at k_- and k_+ are balanced, $R(k_-, \omega, k_2) = R(k_+, \omega, k_2)$, which implies by (19) that $k_2 = \bar{k}_2^*$ given in (14), we could be at the optimum, provided the convergence factor at $k_{\max}$ is smaller than the balanced convergence factor at k_- and k_+, which by a direct computation implies that (12) is not satisfied. If however (12) holds, then the maximum is still at $k_{\max}$ and can be further

decreased by increasing k_2, while $R(k_+, \omega, k_2)$ increases, and $R(k_-, \omega, k_2)$ decreases further and becomes irrelevant. By continuity, the optimum is in that case attained when $R(k_+, \omega, k_2) = R(k_{\max}, \omega, k_2)$, which leads to $k_2 = \bar{k}_2^*$ given in (13). $\qquad\square$

Corollary 2.2. *If $k_2 = \bar{k}_2^*$ according to Theorem 2.2, and $k_{\max} = \pi/h$, then the convergence factor $R(k, \omega, \bar{k}_2^*)$ given in (7) is for h small and $k \in K$ defined in (9) bounded by*

$$R(k, \omega, \bar{k}_2^*) \le R(k_{\max}, \omega, \bar{k}_2^*) = 1 - 4\frac{\pi + 2\omega}{\pi}h^2 + O(h^4). \tag{20}$$

Proof. For $k_{\max}$ large, we are in the first case of Theorem 2.2, and therefore the maximum of the solution of the min-max problem is attained at $k_{\max}$. An expansion of $R(\pi/h, \omega, \bar{k}_2^*)$ for h small then leads to the result of the corollary. $\qquad\square$

We see that the optimized parameter $\bar{k}_2^*$ given by (13) leads to a superior asymptotic performance of the method when h becomes small than the fixed parameter $k_2 = \bar{k}_2$ from Theorem 2.1, which leads to the asymptotic performance shown in Corollary 2.1. However, the asymptotic convergence rate with the optimized parameter $\bar{k}_2^*$ is not better than the one of an unpreconditioned diffusion problem. We will show in Section 3, how a refined analysis in a semi-discrete setting leads to a substantially improved asymptotic convergence rate. But first, we introduce a more balanced approximate continuous factorization.

2.2. *A More Balanced Approximation*

We now choose the parameters p and q such that the convergence factor vanishes at two intermediate frequencies, at $k_1 < \omega$ and at $k_2 > \omega$. This leads after a short calculation to the parameters

$$p = -\frac{k_1^2\sqrt{k_2^2 - \omega^2} + ik_2^2\sqrt{\omega^2 - k_1^2}}{k_2^2 - k_1^2}, \qquad q = \frac{\sqrt{k_2^2 - \omega^2} + i\sqrt{\omega^2 - k_1^2}}{k_2^2 - k_1^2}. \tag{21}$$

With this choice, the new approximate factorization is exact for both the frequencies k_1 and k_2. If we estimate again the lowest frequency in the x-direction by zero, $k_x = 0$, the modulus of the convergence factor (6) becomes after simplification

$$R(k, \omega, k_1, k_2) := \frac{(k^2 - k_1^2)^2(k^2 - k_2^2)^2}{(k^4 - k_1^2 k_2^2 + (k_1^2 + k_2^2 - 2k^2)\omega^2)^2}. \tag{22}$$

Theorem 2.3. *The convergence factor $R(k, \omega, k_1, k_2)$ given in (22) is bounded by one for all $k \in \mathbb{R}$, if and only if*

$$k_2 = \bar{\bar{k}}_2(k_1) := \sqrt{2\omega^2 - k_1^2}. \tag{23}$$

Proof. First, we note again that at $k = \omega$, the convergence factor satisfies $R(\omega, \omega, k_1, k_2) \equiv 1$, independently of what we choose for the parameters $k_1 < \omega$ and $k_2 > \omega$. Taking a partial
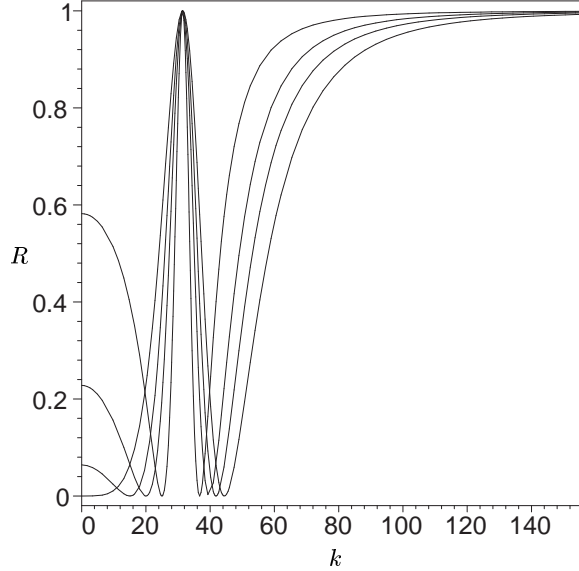
8   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

Fig. 2. Uniformly bounded convergence factor $R(k, \omega, k_1, \bar{\bar{k}}_2(k_1))$ for various choices $k_1 = 0, 15, 20, 25$.

derivative of $R(k, \omega, k_1, k_2)$ with respect to k, and evaluating at $k = \omega$, we find

$$\left.\frac{\partial R}{\partial k}\right|_{k=\omega} = \frac{4\omega(2\omega^2 - k_1^2 - k_2^2)}{(\omega^2 - k_1^2)(\omega^2 - k_2^2)}. \tag{24}$$

Hence, if $k_1^2 + k_2^2 \neq 2\omega^2$, then the partial derivative at $k = \omega$ is non-zero, and since we have $R(\omega, \omega, k_1, k_2) = 1$, by continuity we must have $R(k, \omega, k_1, k_2) > 1$ in the neighborhood of $k = \omega$, which shows the only if part. For the if part, we find from the partial derivative of $R(k, \omega, k_1, k_2)$ with respect to k and assuming that $k_2 = \bar{\bar{k}}_2(k_1)$ given in (23) that the only extrema are at $k = k_1$, $k = \omega$ and $k = \bar{\bar{k}}_2$. Since $R(k_1, \omega, k_1, \bar{\bar{k}}_2) = 0$ and $R(\bar{\bar{k}}_2, \omega, k_1, \bar{\bar{k}}_2) = 0$ by construction, k_1 and $\bar{\bar{k}}_2$ are minima, and in between, at $k = \omega$, we find the only maximum, where $R(\omega, \omega, k_1, \bar{\bar{k}}_2) = 1$. Now at $k = 0$, we have

$$R(0, \omega, k_1, \bar{\bar{k}}_2(k_1)) = \frac{(k_1^2(2\omega^2 - k_1^2))^2}{(2\omega^4 - k_1^2(2\omega^2 - k_1^2))^2} < 1,$$

as one can see by subtracting the numerator from the denominator. Finally, because we have from (22) that $\lim_{k \to \infty} R(k, \omega, k_1, \bar{\bar{k}}_2) = 1$, it follows that $R(k, \omega, k_1, \bar{\bar{k}}_2) \leq 1$ for all k. □

Theorem 2.3 gives us a new free parameter: the method is convergent for all frequencies, except for $k = \omega$ and $k = \infty$, provided that $k_2 = \bar{\bar{k}}_2(k_1) = \sqrt{2\omega^2 - k_1^2}$, where $0 \leq k_1 < \omega$ can be chosen arbitrarily. In Figure 2, we show the convergence factor $R(k, \omega, k_1, \bar{\bar{k}}_2)$ as a function of k for various choices of $k_1 \in [0, \omega)$. One can see that all curves stay bounded by one as predicted in Theorem 2.3. Note also that the choice $k_1 = 0$ leads again to the special case treated already in Subsection 2.1.

Unfortunately, the remaining free parameter k_1 can not really be used for an optimization for high frequencies, $k >> 1$, since the condition $k_1^2 + \bar{\bar{k}}_2^2 = 2\omega^2$ implies $\bar{\bar{k}}_2 \leq \sqrt{2}\omega$, and thus $\bar{\bar{k}}_2$ can not be effective to diminish the convergence factor for large k. This is also evident in Figure 2, where one can see that the convergence factor increases for large k, when k_1 moves away from zero. Nevertheless, the free parameter gives an entire family of preconditioners, which can be exact for two frequencies and still non-divergent for all the others.

If we allow however the convergence factor to become slightly larger than one, like in Subsection 2.1, to regain back two parameters in the optimization, one could also treat high frequencies. One would then try to determine two frequencies k_1, k_2, where the factorization is exact, such that the convergence factor is minimized over all $k \in K$, where K is the set of numerically relevant frequencies defined in (9). This leads to the new min-max problem

$$\min_{0 \leq k_1 < \omega < k_2} \left(\max_{k \in K} R(k, \omega, k_1, k_2) \right) = \min_{0 \leq k_1 < \omega < k_2} \left(\max_{k \in K} \frac{(k^2 - k_1^2)^2 (k^2 - k_2^2)^2}{(k^4 - k_1^2 k_2^2 + (k_1^2 + k_1^2 - 2k^2)\omega^2)^2} \right). \quad (25)$$

We do however not pursue this approach further for the moment, and turn our attention now to the semi-discrete case.

3. The Semi-Discrete Analytic Factorization

To relate the analytic factorization to the exact block LU decomposition of the discrete matrix operator, we discretize the operator $(-\omega^2 - \Delta)$ in the x direction [5] and construct the analytic factorization (2) for the semi-discrete operator $(-\omega^2 - \Delta_h)$, where $\Delta_h = D_x^- D_x^+ + \partial_{yy}$ with $D_x^+(u_j) := (u_{j+1} - u_j)/h$ and $D_x^-(u_j) := (u_j - u_{j-1})/h$ representing the discrete derivatives of a grid function u_j on a given mesh. Using a Fourier transform in y of $-\omega^2 - \Delta_h$ as in the continuous case, the semi-discrete analytic factorization is found to be [5]

$$\mathcal{F}_y(-\omega^2 - \Delta_h) = -\left(D_x^- + (\tau h - \frac{1}{h}) \right) \frac{1}{h^2 \tau} \left(D_x^+ - (\tau h - \frac{1}{h}) \right), \quad (26)$$

where the quantity τ is given by

$$\tau = \frac{1}{h^2} + \frac{-\omega^2 + k^2}{2} + \frac{1}{2h}\sqrt{(-\omega^2 + k^2)^2 h^2 + 4(-\omega^2 + k^2)}. \quad (27)$$

Note that, as we take the limit for $h \longrightarrow 0$ in (26), we recover again the continuous analytic factorization (2), the term in the middle of (26) disappears in the limit. It turns out however that this term is important to obtain a good quality approximate factorization [4,5].

As in the continuous analytic factorization, we replace the nonlocal operator represented by its symbol τ in (26) by a local approximation of the form

$$\tau_{app} = \frac{1}{h^2} + \frac{-\omega^2 + k^2}{2} + \frac{1}{2h}(p + qk^2), \quad p, q \in \mathbb{C}, \ \Re(q) > 0, \quad (28)$$

which leads to two local evolution problems in the factorization. We insert the approximation τ_{app} into the factorization (26) and obtain the operator resulting from the approximate

10   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

factorization of $-\omega^2 + k^2 - D_x^+ D_x^-$ in the form

$$\mathcal{L}_{app} = \mathcal{F}_y^{-1}(-D^- D^+ + \tau_{app} + \frac{1}{\tau_{app} h^4} - \frac{2}{h^2}). \tag{29}$$

The complex numbers p and q are to be chosen so that $\mathcal{L}_{app}^{-1}\mathcal{L}$ is as close as possible to the identity except for a few frequencies which will be taken into account by the Krylov method. We find after some calculation the convergence factor in Fourier to be

$$\rho(k, \omega, h, p, q) = 1 - \frac{2(-\omega^2 + k^2)(2 - \omega^2 h^2 + ph + h(h + q)k^2)}{(p - \omega^2 h + (q + h)k^2)^2}, \tag{30}$$

where we have estimated again the discrete Fourier parameter in the x direction by 0. In the following subsections, we follow the same choices made for the continuous factorization to find suitable parameters p and q to obtain a good preconditioner.

3.1. *Low Frequency Approximation*

We choose, as in the continuous analytic factorization, p such that ρ vanishes at $k = 0$, and q such that ρ vanishes for some $k_2 > \omega$. We find after some computations that the corresponding parameter values for p and q are

$$p = i\omega\sqrt{4 - \omega^2 h^2}, \quad q = \frac{\sqrt{-4\omega^2 + \omega^4 h^2 + 4k_2^2 + k_2^4 h^2 - 2k_2^2 h^2 \omega^2} - i\omega\sqrt{4 - \omega^2 h^2}}{k_2^2}. \tag{31}$$

With these values, the square of modulus of the convergence factor given in (30) for $h < 2/\omega$ becomes

$$R(k, \omega, h, k_2) = \frac{4k^4(k - k_2)^2(k + k_2)^2}{\left(k^2(\omega^2 - k^2)\left((\omega^2 - k_2^2)h^2 - \sqrt{(\omega^2 - k_2^2)(\omega^2 h^2 - 4 - h^2 k_2^2)}h\right) + 2k^4 + 2\omega^2 k_2^2 - 4\omega^2 k^2\right)^2}, \tag{32}$$

which agrees as h goes to zero with the convergence factor found for the continuous analytic factorization given in (7).

Theorem 3.1. *For $h < 1/\omega$, the convergence factor $R(k, \omega, h, k_2)$ given in (32) satisfies $R(k, \omega, h, k_2) \leq 1$ for all k, if and only if*

$$k_2 = \tilde{k}_2 := \omega\sqrt{\frac{2 - \omega h}{1 - \omega h}}. \tag{33}$$

Proof. As in the continuous case, at $k = \omega$ we have $R(\omega, \omega, h, k_2) \equiv 1$, independently of what we choose for the parameter k_2. A lengthy, but not difficult calculation shows that the partial derivative of $R(k, \omega, h, k_2)$ with respect to k evaluated at $k = \omega$ is non zero, except if $k_2 = \tilde{k}_2$ given in (33). Hence, if $k_2 \neq \tilde{k}_2$, then the derivative at $k = \omega$ is non-zero, and since at $k = \omega$ we have $R(\omega, \omega, h, k_2) = 1$, we must have $R(k, \omega, h, k_2) > 1$ for some values of k in the neighborhood of ω by continuity, which shows the only if part. For the if part, we assume that $k_2 = \tilde{k}_2$ given in (33). Again using the partial derivative of $R(k, \omega, h, \tilde{k}_2)$ with respect to k, we find two minima at $k = 0$ and $k = \tilde{k}_2$, and one maximum at $k = \omega$

where $R(\omega, \omega, h, \tilde{k}_2) = 1$. To show that $R(k, \omega, h, \tilde{k}_2) \leq 1$ for all k, it thus remains to analyze $R(k, \omega, h, \tilde{k}_2)$ as $k \longrightarrow \infty$. Taking the limit, we find

$$\lim_{k \to \infty} R(k, \omega, h, \tilde{k}_2) = (1 - \omega h)^2 < 1, \qquad \text{if } h < 1/\omega,$$

which concludes the proof. □

Corollary 3.1. *If $k_2 = \tilde{k}_2$ as given in Theorem 3.1, and $k_{\max} = \pi/h$, then the convergence factor $R(k, \omega, h, \tilde{k}_2)$ given in (32) is for h small and $k \in K$ defined in (9) bounded by*

$$R(k, \omega, h, \tilde{k}_2) \leq R(k_{max}, \omega, h, \tilde{k}_2) = 1 - 2\omega h + O(h^2). \tag{34}$$

Proof. For $k_2 = \tilde{k}_2$ from Theorem 3.1, the convergence factor $R(k, \omega, h, \tilde{k}_2)$ given in (32) attains its maximum on K at $k_{\max}$ for $k_{\max}$ large, and since $k_{\max} = \pi/h$, we have for h small that $R(k, \omega, h, \tilde{k}_2) \leq R(k_{\max}, \omega, h, \tilde{k}_2)$ for all $k \in K$. The result then follows by expanding $R(\pi/h, \omega, h, \tilde{k}_2)$ for h small. □

Note that this result is a big improvement over the result for the continuous factorization shown in Corollary 2.1, where the convergence factor behaved like $1 - O(h^4)$ for h small. As in the continuous case however, we can try to optimize the convergence rate by solving the min-max problem

$$\min_{\omega + \delta\omega < k_2 < k_{\max}} \left(\max_{k \in K} R(k, \omega, h, k_2) \right). \tag{35}$$

If we choose the same strategy as in the continuous case for h small, we get the following theorem.

Theorem 3.2. *Let $k_{\max} = \pi/h$, K be the set defined in (9), and let $k_2 = \tilde{k}_2^*$ defined by the equation*

$$R(\omega + \delta\omega, \omega, h, \tilde{k}_2^*) - R(k_{\max}, \omega, h, \tilde{k}_2^*) = 0, \tag{36}$$

where $R(k, \omega, h, k_2)$ is given in (32). Then for h small, we have

$$\tilde{k}_2^* \approx \sqrt{2}(\omega + \delta\omega) - \frac{\sqrt{\omega^2 + 4\omega\delta\omega + 2\delta\omega^2}\,\omega^2(\omega + \delta\omega)}{2\sqrt{2}\delta\omega(2\omega + \delta\omega)}h, \tag{37}$$

and the convergence factor $R(k, \omega, h, \tilde{k}_2^*)$ is bounded for all $k \in K$ by

$$R(k, \omega, h, \tilde{k}_2^*) \leq R(k_{\max}, \omega, h, \tilde{k}_2^*) = 1 - 2\sqrt{\omega^2 + 4\omega\delta\omega + 2\delta\omega^2}h + O(h^2). \tag{38}$$

Proof. There is no closed form solution for $\tilde{k}_2^*$ satisfying (36), but we know from (13) in Theorem 2.2 that as $h \longrightarrow 0$, the optimal parameter is $\bar{k}_2^* = \sqrt{2}(\omega + \delta\omega)$. We therefore insert the ansatz $\tilde{k}_2^* = \sqrt{2}(\omega + \delta\omega) + Ch^\alpha$, $\alpha > 0$ into (36) and expand for small h. We find after a lengthy calculation that asymptotically (36) becomes

$$\frac{2\omega^2\sqrt{\omega^2 + 4\omega\delta\omega + 2\delta\omega^2}}{(\omega + \delta\omega)^2}h + 4\frac{\sqrt{2}(2\omega + \delta\omega)\delta\omega C}{(\omega + \delta\omega)^3}h^\alpha + O(h^2) + O(h^{1+\alpha}) + O(h^{2\alpha}) = 0.$$
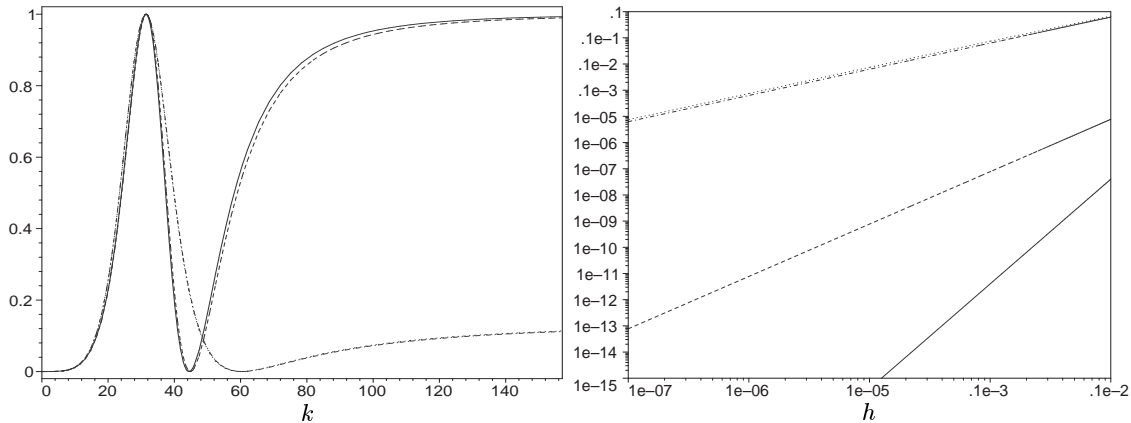
12   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

Fig. 3. On the left comparison of the convergence rates as a function of k, and on the right a comparison of $1 - R$ as a function of h where R attains its maximum.

Balancing the first terms leads to $\alpha = 1$ and

$$C = -\frac{\sqrt{\omega^2 + 4\omega\delta\omega + 2\delta\omega^2}\,\omega^2(\omega + \delta\omega)}{2\sqrt{2}\delta\omega(2\omega + \delta\omega)}.$$

Inserting this asymptotic result for $\tilde{k}_2^*$ into $R(k, \omega, h, \tilde{k}_2^*)$ and evaluating at $k = k_{\max}$ gives the bound (38). $\qquad\square$

Figure 3 shows a comparison of the theoretical convergence factors, for $\omega = 10\pi$ and $\delta\omega = \pi$. On the left, R is shown as a function of k for $h = 1/50$ for the continuous analytic factorization with fixed $k_2 = \bar{k}_2$ (solid line) and with optimized $k_2 = \bar{k}_2^*$ (dashed line), and for the semi-discrete analytic factorization (the lower curves) with fixed $k_2 = \tilde{k}_2$ (dash-dotted line) and optimized $k_2 = \tilde{k}_2^*$ (dotted line). One can clearly see the superior performance of the semi-discrete factorization given by the lower curves and also a small difference between the performance of the two continuous factorizations (solid and dashed). On the right with the same pattern coding we show for the same convergence rates $1 - R$ at $k = k_{\max}$ as a function of h. The asymptotic behavior is as predicted by the analysis: the semi-discrete factorization has the weakest dependence on h, $R = 1 - O(h)$, and using the fixed $k_2 = \tilde{k}_2$ or the optimized $k_2 = \tilde{k}_2^*$ does not make a significant difference (dash-dotted and dotted), whereas for the continuous factorization the difference is significant, $1 - O(h^4)$ in solid and $1 - O(h^2)$ dashed.

3.2. *A More Balanced Approximation*

Following the ideas from the continuous factorization, we now choose the parameters p and q such that the convergence factor vanishes at two intermediate frequencies, at $k_1 < \omega$ and

at $k_2 > \omega$. This leads after some calculation to the parameters

$$
\begin{aligned}
p &= -\frac{k_1^2\sqrt{(k_2^2-\omega^2)(4+h^2(k_2^2-\omega^2))}+ik_2^2\sqrt{(\omega^2-k_1^2)(4+h^2(k_1^2-\omega^2))}}{k_2^2-k_1^2}, \\
q &= \frac{\sqrt{(k_2^2-\omega^2)(4+h^2(k_2^2-\omega^2))}+i\sqrt{(\omega^2-k_1^2)(4+h^2(k_1^2-\omega^2))}}{k_2^2-k_1^2},
\end{aligned}
\tag{39}
$$

and one can see that as h goes to zero, these parameters converge to the parameters in the continuous case given in (21). With this choice, the new approximate factorization is exact for the two frequencies k_1 and k_2.

Theorem 3.3. *Let $R(k,\omega,h,k_1,k_2)$ denote the square of the modulus of the convergence factor given in (30) for the choice p and q given in (39). Then $R(k,\omega,h,k_1,k_2)$ is bounded by one for all $k \in \mathbb{R}$ if and only if*

$$
k_2 = \tilde{\tilde{k}}_2(k_1) := \sqrt{\frac{2\omega^2 - k_1^2 - h\omega^2\sqrt{\omega^2 - k_1^2}}{1 - h\sqrt{\omega^2 - k_1^2}}}.
\tag{40}
$$

Proof. The proof is following the same steps as the proof of Theorem 2.3. $\square$

As in the case of the continuous factorization, the choice of Theorem 3.3 gives an entire family of semi-discrete preconditioners whose convergence factor is bounded by one for all k. One could also allow the convergence factors to become slightly larger than one outside of the numerical set of frequencies K defined in (9), and then try to find the best parameters k_1 and k_2, which solve the min-max problem

$$
\min_{0 \le k_1 < \omega < k_2}\left(\max_{k \in K} R(k,\omega,h,k_1,k_2)\right),
\tag{41}
$$

but we are not considering this issue further in this paper. The numerical experiments in the next section will show that there would be only little one could gain studying (41) further.

4. Numerical Experiments

We consider for our numerical experiments a two dimensional open cavity problem modeled by

$$
-\omega^2 u - \Delta u = f, \quad \text{in } \Omega = [0,1]\times[0,1],
\tag{42}
$$

with homogeneous Dirichlet boundary conditions on the top, bottom and right boundaries. On the left boundary, which represents the opening of the cavity, we impose the Robin condition $(\partial_x + i\omega)(u) = 0$. We choose for the right hand side the function $f(x,y) = \delta(x - \frac{1}{2})\delta(y - \frac{1}{2})$, which corresponds to a point source in the center of the cavity. We discretize the problem using a standard five point finite difference discretization with the same mesh parameter h both in the x and y directions. This leads to a discrete matrix problem of the form

$$
A\boldsymbol{u} = \boldsymbol{f}
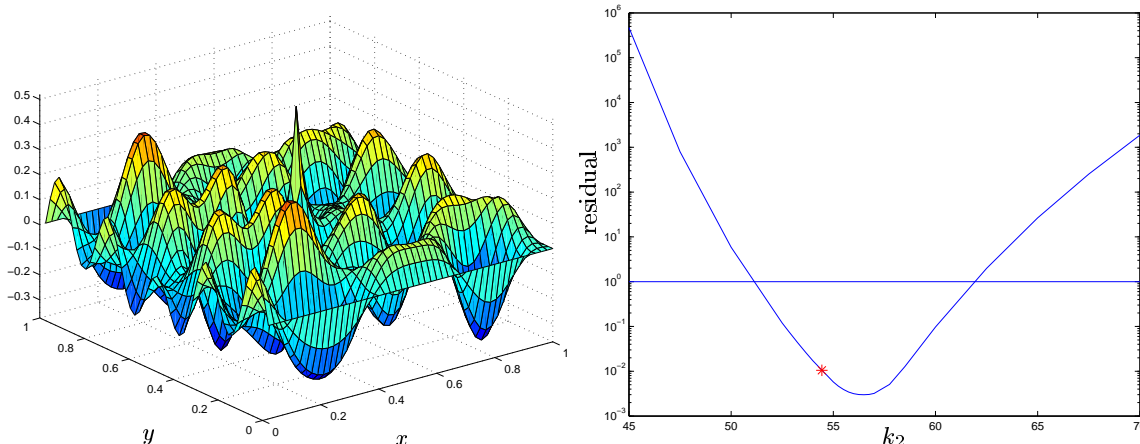\tag{43}
$$

14   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

Fig. 4. Solution of the open cavity problem for $\omega = 9.36\pi$ on the left, and convergence and divergence behavior of AILU used as an iterative solver on the right, depending on the parameter k_2.

where A is a large, sparse, indefinite matrix. We follow in the implementation of the AILU preconditioners presented here the approach using the relation between AILU and the block-LU decomposition [4].

4.1. *AILU as an Iterative Solver*

As we have seen in the analysis, AILU as an iterative solver for discretized Helmholtz problems is not convergent in general, and we only recommend it to be used as a preconditioner for a Krylov method. Nevertheless, for certain problem configurations, the method can be made convergent, which illustrates our analysis in an interesting way. For example for our domain Ω with homogeneous Dirichlet conditions on top and bottom, the solution can be expanded in a Fourier series with the harmonics $\sin(j\pi y)$, $j \in \mathbb{N}$. Hence the relevant frequencies are $k = j\pi$, $j = 1, 2, \ldots$. They are equally distributed with the spacing π, the lowest frequency is $k_{\min} = \pi$, and choosing $\delta\omega = \pi/2$ would leave precisely one frequency in $(\omega - \delta\omega, \omega + \delta\omega)$ for the Krylov method and treat all the others by the optimization. If ω falls in between the relevant frequencies, say $j\pi < \omega < (j+1)\pi$, then we can get even the iterative method to converge by choosing $\delta\omega = \min(\omega - j\pi, (j+1)\pi - \omega)$. To illustrate this in this subsection, we need to be more precise and consider the spectrum of the discretized operator. Choosing $h = 1/50$, the spectrum of the discrete Laplacian is shifted from 9π to 8.88π and from 10π to 9.84π. If we therefore choose ω to be exactly between those two frequencies of the discrete problem, $\omega = 9.36\pi$ and optimize using $\delta\omega = 0.48\pi$, the iterative AILU should be convergent for an appropriate choice of the parameter k_2 ($k_1 = 0$). We show in Figure 4 on the left the final result of the simulation, and on the right the behavior of AILU used as an iterative solver for various values of the parameter k_2. Note how a certain range of choices for k_2 does indeed lead to a convergent stationary iterative method. The star in Figure 4 on the right indicates the choice recommended by the semi-discrete analysis in Theorem 3.1, which leads to $\tilde{k}_2 = 54.44$ and indeed is close to the best performing pa-

rameter in the fully discrete simulation. With the theoretical parameter $\tilde{k}_2$, AILU used as an iterative method on this problem reduced the initial residual by a factor of 10^{-6} in 596 iterations, starting from a random initial guess, and in 892 iterations starting from a zero initial guess. The results of the fully continuous analysis lead to a parameter $\bar{k}_2 = 41.58$ (Theorem 2.1) and $\bar{k}_2^* = 43.89$ (Theorem 2.2), both of which do not lead to a convergent stationary iterative AILU, as one can see from Figure 4 on the right. At any rate, to really use AILU, one needs a Krylov method, which is preconditioned with AILU, which we will do in the next subsection.

4.2. *AILU as a Preconditioner for a Krylov Method*

We use the same configuration as in the previous subsection, but now use the AILU as a preconditioner for several Krylov methods, and compare its performance to other incomplete LU preconditioners. In Table 1, we show the iteration counts obtained to reduce the initial residual by a factor 10^{-6}, both using a zero initial guess, and a random initial guess. One

Table 1. Comparison of iteration counts when AILU is used as a preconditioner for various Krylov methods, together with results for standard incomplete LU preconditioners.

Iteration counts using a random initial guess and various preconditioners						
Method	none	ILU('0')	ILU(1e-2)	AILU($\bar{k}_2$)	AILU($\bar{k}_2^*$)	AILU($\tilde{k}_2$)
GMRES	647	204	48	69	70	80
QMR	> 2000	> 2000	> 2000	91	91	83
BiCGStab	> 2000	545	77	43	45	53

Iteration counts using a zero initial guess and various preconditioners						
Method	none	ILU('0')	ILU(1e-2)	AILU($\bar{k}_2$)	AILU($\bar{k}_2^*$)	AILU($\tilde{k}_2$)
GMRES	443	210	51	47	47	53
QMR	1718	> 2000	> 2000	59	59	54
BiCGStab	1913	636	70	38	38	44

can see that this small scale model problem is already challenging for an unpreconditioned Krylov method. Only GMRES converged to the solution without preconditioner for the random initial guess in less than 2000 iterations, and with a zero initial guess, the situation is similar, with QMR and BiCGStab barely converging in less than 2000 iterations. Since we use GMRES without restarts, the computational cost for it is high, as one can see in Table 2. The situation is not much better for the methods preconditioned with ILU('0'), with no fill-in. QMR is still not converging in less than 2000 iterations, BiCGStab now converges, with a relatively high iteration count of about 600 iterations, and GMRES converges in about 200 iterations. Comparing the flop count however, BiCGStab is doing much better than GMRES already in this case. The ILU(1e-2) preconditioner finally brings down the iteration count significantly for both GMRES and BiCGStab, and also the flop count is significantly reduced for this small problem, except for QMR which is still not convergent. Note that

16   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

Table 2. Comparison of flop counts in mega flops when AILU is used as a preconditioner for various Krylov methods, together with results for standard incomplete LU preconditioners.

Method	Flop counts in mega flops using a random initial guess
	none	ILU('0')	ILU(1e-2)	AILU($\bar{k}_2$)	AILU($\bar{k}_2^*$)	AILU($\tilde{k}_2$)
GMRES	13014.00	1324.60	107.17	180.98	185.64	235.46
QMR	-	-	-	108.06	108.06	97.52
BiCGStab	-	538.86	117.71	52.41	54.25	63.99

Method	Flop counts in mega flops using a zero initial guess
	none	ILU('0')	ILU(1e-2)	AILU($\bar{k}_2$)	AILU($\bar{k}_2^*$)	AILU($\tilde{k}_2$)
GMRES	6027.80	1401.20	118.14	93.24	93.24	114.34
QMR	-	-	-	69.21	69.21	63.36
BiCGStab	1307.70	628.79	114.26	45.62	45.62	52.92

there is already a significant cost difference for computing the ILU(1e-2) preconditioner, compared to the other ones, as one can see in Table 3, and this discrepancy in cost increases dramatically with problem size to a point, where ILU(1e-2) can not be computed any more, while AILU is still very cheap [6]. And with AILU, all preconditioned methods for all proposed

Table 3. Comparison of flop counts in mega flops for computing the various preconditioners.

	ILU('0')	ILU(1e-2)	AILU($\bar{k}_2$)	AILU($\bar{k}_2^*$)	AILU($\tilde{k}_2$)
mega flops	0.21	8.78	0.22	0.22	0.22

choices of parameters are convergent with a low iteration count, and the flop counts are significantly lower than for the other preconditioners, except for GMRES when starting with a random initial guess. This situation however changes as soon as the problem size increases [6]. BiCGStab shows the lowest iteration count, with AILU, but each iteration costs two matrix vector multiplies, and if this is taken into account, it is comparable to the performance of QMR. It is also interesting to note that for QMR, the semi-discrete parameter of AILU(3) leads to the best performance, whereas for GMRES and BiCGStab, the simple parameter derived from the continuous analysis is slightly better.

To investigate this issue further, we tested on the same small model problem the AILU preconditioner which is exact for two frequencies k_1 and k_2. We varied these two frequencies, and each time ran the iterative version of AILU and also AILU as a preconditioner for GMRES, QMR and BiCGStab, in each case with a random initial guess, for 26, 52 and 78 iterations, to see for which choice of the parameters k_1 and k_2 the methods performed best in the course of the iterations. The results are shown in Figure 5 for AILU as an iterative solver on the left and for AILU as a preconditioner for GMRES on the right, and in Figure 6 for AILU used as a preconditioner for QMR and BiCGStab. In each case, we started with a random initial guess and show a level set plot of the log base 10 of the
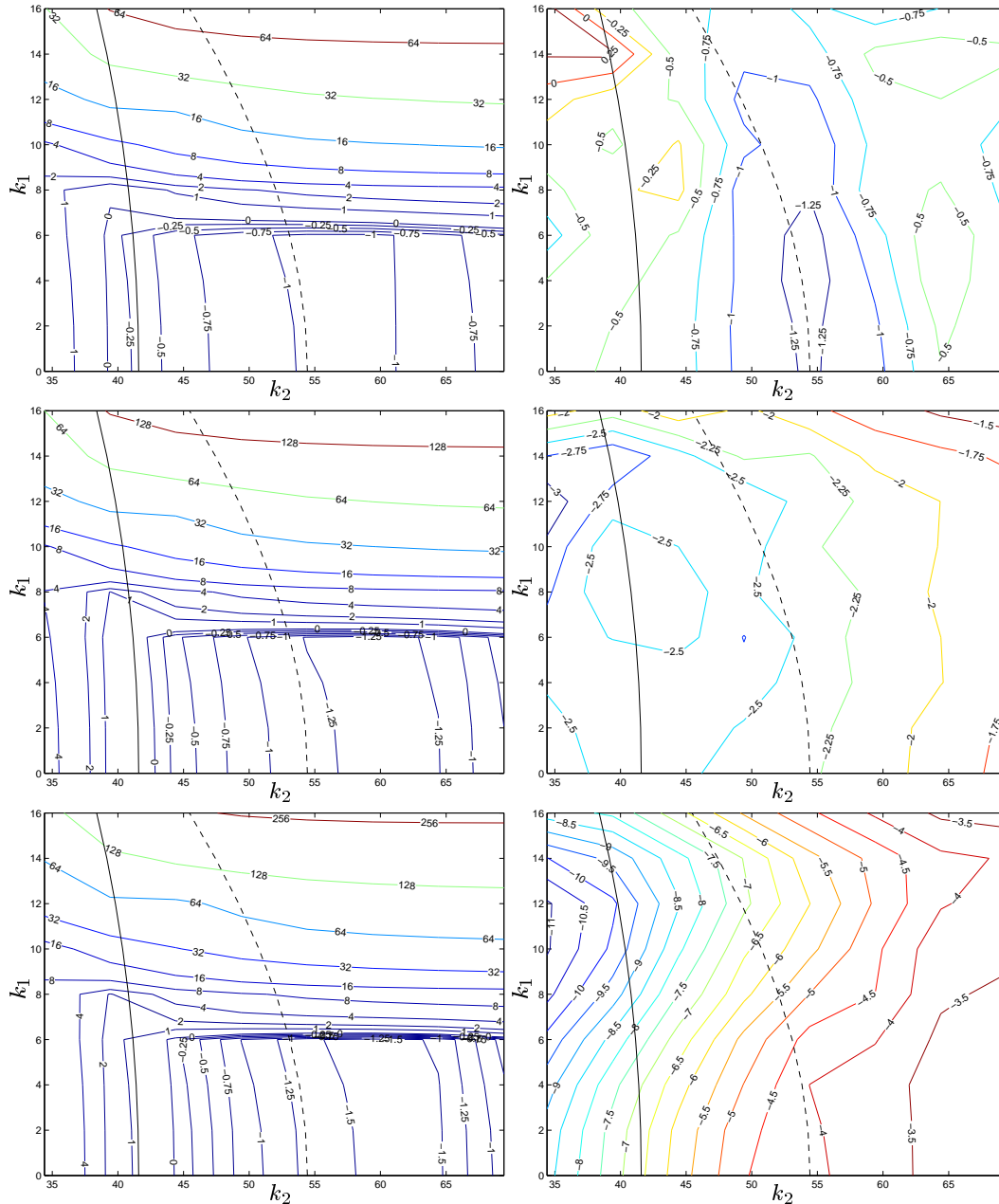
Fig. 5. Residual obtained after 26, 52 and 78 iterations of AILU as an iterative solver on the left, and as a preconditioner for GMRES on the right, for various values of the parameters k_1 and k_2. The dashed line indicates the choice $k_2 = \tilde{k}_2(k_1)$ from Theorem 3.3 which guarantees a semi-discrete convergence factor bounded by one, and the solid line indicates the choice $k_2 = \bar{\bar{k}}_2(k_1)$ from Theorem 2.3 which guarantees a continuous convergence factor bounded by one.
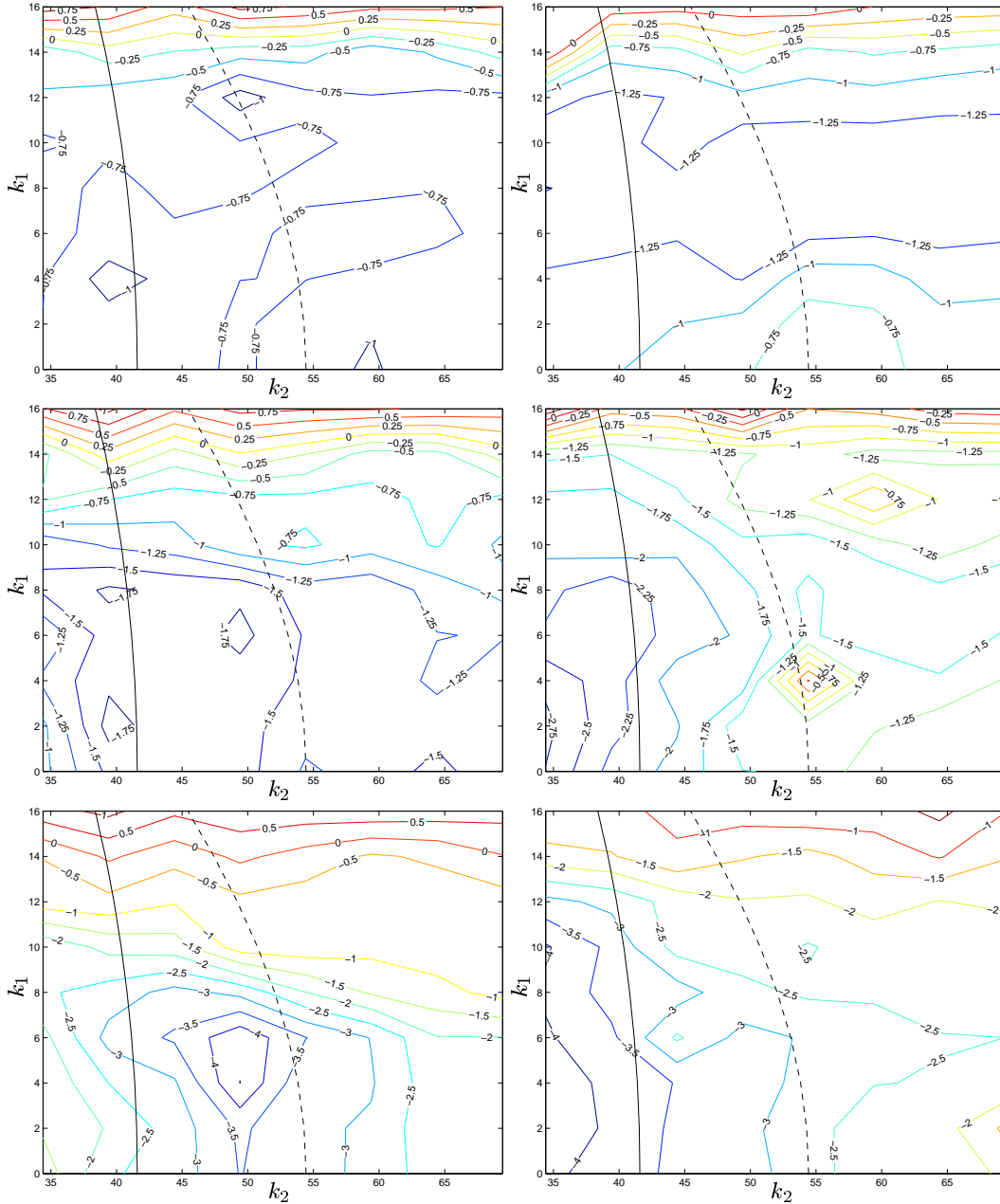
18   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

Fig. 6. Residual obtained after 26, 52, 78 iterations of AILU as a preconditioner for QMR on the left, and for BiCGStab on the right (counting half iterations as iterations, to have the same number of matrix-vector multiplies), for various values of the parameters k_1 and k_2. The dashed line indicates the choice $k_2 = \tilde{\tilde{k}}_2(k_1)$ from Theorem 3.3 which guarantees a semi-discrete convergence factor bounded by one, and the solid line indicates the choice $k_2 = \bar{\bar{k}}_2(k_1)$ from Theorem 2.3 which guarantees a continuous convergence factor bounded by one.
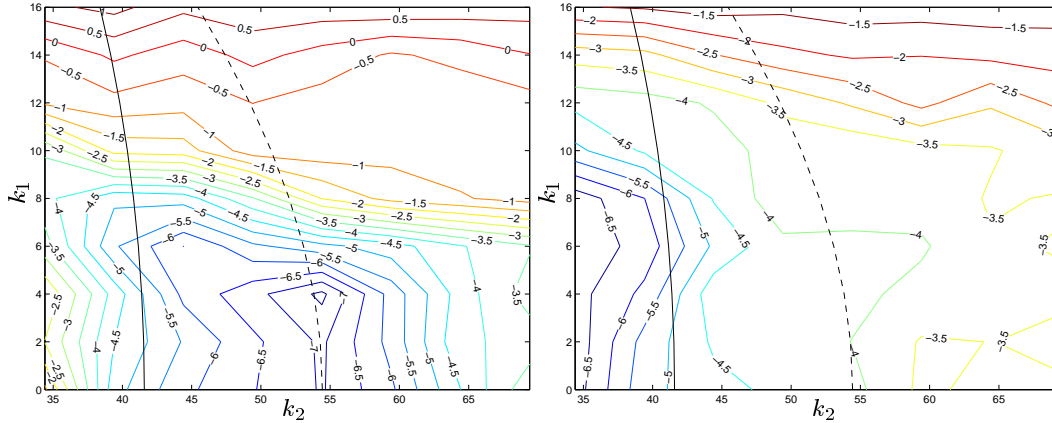
Fig. 7. Residual obtained after 104 iterations of AILU as a preconditioner for QMR on the left, and for BiCGStab on the right (counting half iterations as iterations, to have the same number of matrix-vector multiplies), for various values of the parameters k_1 and k_2. The dashed and solid lines are as before.

residual obtained, scaled with h^2, so that the residual corresponds to the physically relevant one. The black solid line shows the link between the parameters $k_2 = \bar{\bar{k}}_2(k_1)$ given by Theorem 2.3 from the continuous analysis, while the dashed line shows the link between the parameters $k_2 = \tilde{k}_2(k_1)$ given by Theorem 3.3 from the semi-discrete analysis. There are several interesting observations: first, AILU as an iterative method converges in a rectangular region clearly visible in Figure 5 on the left, and the picture does not change much as the iteration progresses from top to bottom. As soon as one leaves this region, divergent modes seem to appear and the method fails. The relation $k_2 = \tilde{k}_2(k_1)$ from Theorem 3.3 predicts the valley of convergence quite well, but not the sharp interface in k_1, where the method starts to diverge, which does not seem to be captured by the semi-discrete analysis. When AILU is used as a preconditioner for a Krylov method, it converges for a much wider range of parameters. It is very interesting to note, that early in the iteration, the relation $k_2 = \tilde{k}_2(k_1)$ from Theorem 3.3 predicts the valley of best convergence extremely well, while later, a different convergence mechanism related to GMRES seems to set in, and a better choice of parameters would be more corresponding to the relation $k_2 = \bar{\bar{k}}_2(k_1)$ given by Theorem 2.3 from the continuous analysis. Understanding the complex interplay between GMRES and the AILU preconditioner would be invaluable to further improve the performance of the preconditioned method.

Looking at Figure 6, we see that QMR has for a wide range of parameters about the same performance early in the iteration, but later the choice indicated by the semi-discrete analysis seems to better and better indicate the best choice. This becomes even more apparent at iteration 104 shown in Figure 7, where one can clearly see that in contrast to GMRES, QMR has the best asymptotic performance for the parameters derived in the semi-discrete analysis leading to $k_2 = \tilde{k}_2(k_1)$ in Theorem 3.3. This is different for BiCGStab, whose region of best performance early in the iteration is a valley orthogonal to the one predicted

by the analysis, and which later moves more toward the region indicated by the continuous analysis, $k_2 = \bar{\bar{k}}_2(k_1)$ given by Theorem 2.3, very similar to GMRES preconditioned with AILU. It is tempting to conjecture that QMR is asymptotically closer to the stationary iterative method and its semi-discrete analysis, as far as the choice of optimal parameters is concerned, whereas GMRES and BiCGStab somehow seem to sense the underlying problem in this case and the optimal parameters derived from the continuous analysis. We repeated the same set of experiments also on a refined mesh for $h = 1/100$, and also for a problem where ω was doubled, and the situation remained similar.

We now turn our attention to how the AILU preconditioner scales with the mesh size and ω. We first do a scaling experiment where we keep $\omega = 10\pi$ fixed, but refine the mesh. In Table 4, we show the iteration counts needed to reduce the initial residual by a factor 10^{-6} when starting with a zero initial guess. We can see that for ω constant, the iteration

Table 4. Increase of the iteration counts when the mesh is refined for a fixed ω.

	GMRES			QMR			BiCGStab		
h	$\bar{k}_2$	$\bar{k}_2^*$	$\tilde{k}_2$	$\bar{k}_2$	$\bar{k}_2^*$	$\tilde{k}_2$	$\bar{k}_2$	$\bar{k}_2^*$	$\tilde{k}_2$
1/50	47	49	54	60	55	57	36	39	42
1/100	52	54	55	58	54	55	37	39	40
1/200	60	62	61	58	63	60	39	43	42
1/400	71	73	72	78	78	79	44	47	44
1/800	-	-	-	121	112	113	57	57	56

counts remain low as the mesh is refined. Only GMRES becomes too memory intensive on the finest mesh with 640'000 complex unknowns, where 1G of memory did not suffice any more to compute a solution. QMR and BiCGStab still solve the highly resolved acoustics problem using about 120 AILU preconditioned matrix vector products. None of the standard ILU preconditioners can be used to solve the refined problems [6].

We next show a scaling experiment, where ω increases as the mesh is refined, and hence a more and more difficult Helmholtz problem is solved. We start with a coarse mesh $h = 1/50$ and $\omega = 5\pi$ and then increase ω by keeping the ratio ωh constant. In Table 5, we show the iteration counts needed to obtain a reduction of the residual by a factor 10^{-6} when starting with a zero initial guess. We can see that the iteration counts now reflect the

Table 5. Iteration count dependence on ω for a fixed ratio ωh, where h is the mesh size.

		GMRES			QMR			BiCGStab		
ω	h	$\bar{k}_2$	$\bar{k}_2^*$	$\tilde{k}_2$	$\bar{k}_2$	$\bar{k}_2^*$	$\tilde{k}_2$	$\bar{k}_2$	$\bar{k}_2^*$	$\tilde{k}_2$
5π	1/50	27	26	26	28	27	26	18	19	19
10π	1/100	52	54	55	58	54	55	37	39	40
20π	1/200	146	150	155	216	172	161	101	108	112
40π	1/400	-	-	-	stag	stag	531	313	322	353

increasing difficulty of the problem. For GMRES, the iteration counts of the largest problem are too high to allow the method to proceed with 1G of memory. QMR converges for the parameter choice in AILU stemming from the semi-discrete analysis, but stagnates for the other parameter values on the largest problem. BiCGStab still converges with all three parameter choices in AILU, even for $\omega = 40\pi$, and turns out to be the most robust method in this experiment.

5. Conclusions

We presented and analyzed two variants of the AILU preconditioner for the Helmholtz equation: a continuous one and a semi-discrete one. Both are incomplete LU preconditioners based on the analytic factorization of the continuous or semi-discrete Helmholtz operator and local approximations of the non-local analytic factors. We derived parameters in both cases of AILU which led to stationary iterative methods that have convergence factors bounded by one, and we have also derived optimized parameters when the method is used in a numerical setting. In this case, the convergence factors are strictly less than one, and we analyzed the asymptotic performance of the AILU variants when the mesh is refined. We then presented extensive numerical experiments to investigate the performance of the AILU variants as preconditioners for Krylov methods. We found that AILU as a preconditioner for QMR leads to a method with a parameter dependence similar to the stationary AILU and corresponding to the semi-discrete analysis, whereas GMRES and BiCGStab preconditioned with AILU show a different parameter dependence, more related to the continuous AILU analysis. It would be invaluable to obtain more insight into the interplay between AILU and the Krylov methods used, in order to be able to better predict the best parameters for the performance of Krylov method preconditioned with AILU in the context of acoustics problems: this paper is only a first step into that direction.

References

1. Y. Achdou and F. Nataf. Dimension-wise iterated frequency filtering decomposition. *Num. Lin. Alg. and Appl.*, 10:511–539, 2003.
2. A. Buzdin. Tangential decomposition. *Computing*, 61:257–276, 1998.
3. J. F. Claerbout. *Fundamentals of geophysical data processing with applications to petroleum prospecting.* McGraw-Hill, 1976.
4. M. J. Gander and F. Nataf. AILU: A preconditioner based on the analytic factorization of the elliptic operator. *Num. Lin. Alg. and Appl.*, 7(7–8):543–567, 2000.
5. M. J. Gander and F. Nataf. AILU for Helmholtz problems: A new preconditioner based on an analytic factorization. *C. R. Acad. Sci. Paris*, I(331):261–266, 2000.
6. M. J. Gander and F. Nataf. AILU for Helmholtz Problems: a New Preconditioner Based on the Analytic Parabolic Factorization. *Journal of Computational Acoustics*, 9(4):1499–1506, 2001.
7. M. J. Gander. Preconditioners for Acoustic Problems: AILU, *Special issue of the journal Revista de Acústica*, Vol. XXXIII, 2002, [CDROM].
8. E. Giladi and J. B. Keller. Iterative solution of elliptic problems by approximate factorization. *Journal of Computational and Applied Mathematics*, 85:287–313, 1997.

22   *MARTIN J. GANDER and FRÉDÉRIC NATAF*

9. F. Nataf. Résolution de l'équation de convection-diffusion stationaire par une factorisation parabolique. *C. R. Acad. Sci., Paris*, I 310(13):869–872, 1990.

10. F. Nataf, J. P. Loheac, and M. Schatzman. Parabolic approximation of the convection-diffusion equation. *Math. Comp.*, 60:515–530, 1993.

11. J. Stoer and R. Bulirsch. *Numerical Analysis*. Springer, Berlin, 1993.

12. C. Wagner. Tangential frequency filtering decompositions for symmetric matrices. *Numer. Math.*, 78(1):119–142, 1997.

13. C. Wagner. Tangential frequency filtering decompositions for unsymmetric matrices. *Numer. Math.*, 78(1):143–163, 1997.

14. G. Wittum. An ILU-based smoothing correction scheme. In *Parallel algorithms for partial differential equations, Proc. 6th GAMM-Semin., Kiel/Ger. , Notes Numer. Fluid Mech.*, 31:228–240, 1991.

15. G. Wittum. *Filternde Zerlegungen. Schnelle Löser für grosse Gleichungssysteme*. Teubner Skripten zur Numerik, Stuttgart, 1992.