Un nouveau préconditionneur pour les problèmes elliptiques à coefficients variables

Mejdi Azaïez, Alain Bergeon, Franck Plouraboué

▶ **To cite this version:**

Mejdi Azaïez, Alain Bergeon, Franck Plouraboué. Un nouveau préconditionneur pour les problèmes elliptiques à coefficients variables. Comptes Rendus Mécanique, 2003, vol. 331, pp. 509-514. 10.1016/S1631-0721(03)00091-3 . hal-00959366

HAL Id: hal-00959366
https://hal.science/hal-00959366

Submitted on 14 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Un nouveau préconditionneur pour les problèmes elliptiques à coefficients variables

Mejdi Azaïez [a,*], Alain Bergeon [b], Franck Plouraboué [c]

[a] *MASTER, École nationale supérieure de chimie et de physique de Bordeaux, 16, av Pey Berland, 33607 Pessac cedex, France*
[b] *IMFT (UMR 5460), Université Paul Sabatier, Toulouse, France*
[c] *IMFT (UMR 5460), CNRS, Toulouse, France*

Résumé

On présente dans cette Note un nouveau préconditionneur pour l'inversion du système algébrique issu de la discrétisation par méthode spectrale d'un problème elliptique du second ordre à coefficients variables et non séparables. Ce préconditionneur est construit en discrétisant un problème similaire à l'original et obtenu par moyenne des coefficients. L'inversion du préconditionneur utilise une méthode directe connue sous le nom de diagonalisation successive.

Abstract

A new preconditonner for general elleptic problems. In this Note we describe a preconditionner for iteratively solving the linear system arising from the discretization of a general nonseparable elliptic problem by spectral element method. This preconditionner is constructed from approximating the original problem with the closest (in some sense) separable elliptic problem. A direct method is then used to invert the preconditionner.

Mots-clés : Mécanique des fluides numérique ; Préconditionneur ; Spectrale ; Gradient conjugué

Keywords: Computational fluid mechanics; Preconditionner; Spectral; Conjugate gradient

* Auteur correspondant.
Adresses e-mail : azaiez@enscpb.fr (M. Azaïez), bergeon@imft.fr (A. Bergeon), plourab@imft.fr (F. Plouraboué).

Abridged English version

In this Note, we investigate a technique which uses a direct method as a preconditionner for solving iteratively the linear system arising from the discretization of the general non-separable elliptic equation (1), (2). Such an equation may be obtained in a wide range of problems like multiphase flows, flows in porous media or inverse problems. It also coincides with the correction step problem involved in projection methods for solving the Navier–Stokes equations for an incompressible viscous flow. For any of these cases, an iterative solver has to be used, the efficiency of which strongly depends on finding a good and effective preconditionner.

A natural preconditionner is the operator $\mathcal{L}_1$ associated to the discretization of the Laplacian. This choice has proved to be effective when the variations of the coefficients $a(\mathbf{x})$ are smooth, that is to say when $\max a(\mathbf{x})/\min a(\mathbf{x})$ is not too large. But the convergence time increases drastically with the ratio $\max a(\mathbf{x})/\min a(\mathbf{x})$ and therefore, a more robust preconditionner must be preferred.

The preconditionner that we propose in this paper is based on the resolution of the problem (3), (4). There are two main reasons for considering this *averaged* version of the problem (1), (2). First, problem (3), (4) transforms the original non-separable problem in a separable one. This property permits to solve very efficiently the auxiliary problem (3), (4) using the so-called tenderized diagonalization direct method [6]. At each preconditionned conjugate gradient iteration, the preconditionner cost is thus negligible compared to the original operator. The second interest is that the averaged problem (3), (4) is an approximation of the original problem (1), (2). Such an approximation is widely used in the context of the volume averaged method in order to solve heterogeneous Darcy problems [3]. The purpose of the volume averaging method is not to solve exactly some partial differential equation problem, but rather to find a correct approximation of its averaged macroscopic behavior. The proposed method is thus a transposition of the volume averaging ideas to the elaboration of an efficient preconditionner.

When approximated by spectral element discretization the algebraic system resulting from solving (3), (4) is discussed in Section 3. Using down-case letters for numerical fields and capital bold faces letters for their matrix counterpart, Eqs. (10) and (11) describe the algebraic structure of the proposed preconditionner. It is interesting to note that one recovers the Laplacian $\mathcal{L}_1$ operator from taking $\mathbf{A} = \mathbf{I}$, i.e., $A_i = 1$ in the proposed algebraic formulation.

The numerical results are illustrated for the class of problems proposed by Shen et al. [8]. The classical Laplacian preconditionner is compared to the new averaged one, for different polynomial orders N and for different values of the power p applied to the test field a defined in Eq. (12). The comparison is illustrated in Fig. 1 (white symbols represent classical Laplacian preconditionner, while dark symbols are for the one proposed here) shows a drastic improvement of the cost through a small number iterations compared to the use of the Laplacian. The numerical experiments are conducted in three dimensions but similar results have been obtained in two dimensions. Moreover, different choices for a have been tested always confirming that the proposed preconditionner is better. Ought to the fact that their computational cost is identical, one concludes it is worth considering the new averaged preconditionner.

A new class of interesting preconditionner could easily be obtained from this work, while considering non uniform average using suitable weighting functions. Another natural extension of this work would be to consider the adaptation of the proposed preconditionner in the case of overlapping spectral elements.

1. Introduction

Le but de cette Note est de présenter un nouveau type de préconditionneur pour inverser le système algébrique issu de la discrétisation du problème suivant : étant donnée deux fonctions a et f définies sur l'ouvert $\Omega =]0, 1[^d$, $d = 2, 3$, trouver p tel que

$$\mathcal{L}p = -\operatorname{div}\big(a(\mathbf{x})\nabla p\big) = f \quad \text{dans } \Omega \tag{1}$$

$$\frac{\partial p}{\partial \mathbf{n}} = 0 \quad \text{sur } \Omega \tag{2}$$

où $\mathbf{x} = \mathbf{x}_i$ pour $i = 1, \ldots, d$.

Ce système intervient dans une large classe de problèmes tels que l'étude des écoulements dans des milieux poreux hétérogènes et l'identification des paramètres dans les problèmes inverses. Il constitue aussi l'étape de projection dans la résolution des équations de Navier–Stokes incompressibles en milieux multiphasiques.

Il est bien connu que la matrice issue de la discrétisation du problème (1), (2) peut être très mal conditionnée notamment lorsque la fonction $a(\mathbf{x})$ varie fortement et/ou rapidement sur le domaine Ω. La littérature est très riche en idées de préconditionneurs et l'on renvoie à [1] et [2] pour une présentation exhaustive dans le cas d'une approximation par méthodes spectrales. L'originalité du préconditionneur que nous proposons réside dans l'optimalité de son efficacité lorsque la fonction $a(\mathbf{x})$ est tensorisée (cas des milieux poreux stratifiés). On montrera néanmoins dans cette note que même dans le cas contraire, la méthode reste intéressante.

Cette Note se compose de trois parties : dans la première, on présente le principe de construction du préconditionneur. Dans la seconde on décrit sa mise en œuvre dans le cadre d'une discrétisation par méthodes spectrales et enfin dans la dernière, on présente quelques résultats numériques illustrant aussi bien l'efficacité que les limites du nouveau préconditionneur en comparant ses performances à un autre préconditionneur fréquemment utilisé dans la littérature. L'exposé se concentrera sur la dimention trois. Ce choix n'est pas restricif, et la transposition algorithmique à deux dimensions est directe.

2. Présentation du préconditionneur

Un choix naturel de préconditionneur consiste à remplacer l'Éq. (1) par la résolution d'un Laplacien muni des mêmes conditions aux limites (2). Or il est bien connu que ce choix n'est intéressant que quand la variation des coefficients de $a(\mathbf{x})$ reste faible. Le préconditionneur que nous proposons repose sur l'écriture d'un problème voisin du problème original possédant de plus la propriété d'être tensoriel. Il s'agit de trouver p solution de

$$-\operatorname{div}\bigl(\mathbf{A}(\mathbf{x})\nabla p\bigr) = f \quad \text{dans } \Omega \tag{3}$$

$$\frac{\partial p}{\partial \mathbf{n}} = 0 \quad \text{sur } \Omega \tag{4}$$

En dimension 3, $\mathbf{A}$ est la matrice diagonale définie par ses coefficients diagonaux $A_i \equiv A_{ii}$:

$$A_1(x_1) = \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} a(\mathbf{x})\,\mathrm{d}x_2\,\mathrm{d}x_3, \quad A_2(x_2) = \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} a(\mathbf{x})\,\mathrm{d}x_1\,\mathrm{d}x_3, \quad A_3(x_3) = \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} a(\mathbf{x})\,\mathrm{d}x_1\,\mathrm{d}x_2 \tag{5}$$

Ce nouveau problème est à coefficients variables mais contrairement au problème initial dont il est issu, il est tensoriel. Nous allons par la suite exploiter cette propriété afin de minimiser le coût de l'inversion du système algébrique issu de sa discrétisation. Outre l'intérêt algorithmique de cette propriété, le choix de moyenner l'opérateur dans chaque direction pour obtenir une approximation du problème initial est inspiré par la méthode de prise de moyenne volumique. Cette méthode est en particulier utilisée pour obtenir des propriétées effectives macroscopiques à partir d'une description moyenne d'équations aux dérivées partielles. Elle a par exemple été mise en œuvre dans le cas du problème de Darcy hétérogène étudié ici (où $a(\mathbf{x})$ représente alors la perméabilité du milieu poreux) pour lequel la solution obtenue sur la pression macroscopique constitue une approximation intéressante de la solution complète [3]. En ce sens, la moyenne de l'opérateur que nous proposons comme préconditionneur constitue une transposition de l'esprit de cette méthode afin d'obtenir une approximation du problème initial. D'autres types de moyennes peuvent aussi être utilisées en fonction du problème à traiter. Différents types de moyenne ont été testés : moyennes arithmétique, harmonique, et géométrique. Aucune différence sensible n'a été observée sur les cas tests étudiés.

3. Description du problème approché

Le problème discret consiste à approcher la formulation variationnelle du problème (1), (2) par un argument de Galerkin avec intégration numérique [4]. Il s'agit donc de trouver $p_N \in \mathbf{P}_N(\Omega)$ polynôme de degré $\leqslant N$ à moyenne nulle vérifiant l'équation

$$\big(a(\mathbf{x})\nabla p_N, \nabla q_N\big)_N = (f, q_N)_N \tag{6}$$

dans laquelle en trois dimensions ($d = 3$), le produit scalaire s'écrit :

$$(p, q)_N = \sum_{i,j,k=0}^{N} p\big((x_1)_i, (x_2)_j, (x_3)_k\big) q\big((x_1)_i, (x_2)_j, (x_3)_k\big) \rho_i \rho_j \rho_k \tag{7}$$

où $x_i = \{(x_i)_j\}_{0 \leqslant j \leqslant N}$ avec $i = 1, 2$ ou 3. Il s'agit de la formule d'intégration de Gauss–Lobatto–Legendre définie par la proposition suivante :

Soit $(x_i)_0 = -1$ *et* $(x_i)_N = 1$ *où* $x_i = \{(x_i)_j\}_{0 \leqslant j \leqslant N}$ *avec* $i = 1, \ldots, d$. *Pour chaque* $i = 1, \ldots, d$, *il existe un unique ensemble de* $N - 1$ *points* $\{(x_i)_j\}_{1 \leqslant j \leqslant N-1}$ *contenus dans* $]-1, 1[$ *et* $N + 1$ *réels positifs* $\{\rho_j\}_{0 \leqslant j \leqslant N}$ *tels que* :

$$\forall \Phi \in \mathbf{P}_{2N-1}\big(]-1, 1[\big), \quad \int_{-1}^{1} \Phi(x_i)\, \mathrm{d}x_i = \sum_{j=0}^{N} \Phi\big((x_i)_j\big) \rho_j \tag{8}$$

Les $(x_i)_j$ $(1 \leqslant j \leqslant N - 1)$ sont les racines de L'_N où $\{L_m\}_{0 \leqslant m \leqslant N}$ sont les $N + 1$ polynômes de Legendre de degré N deux à deux orthogonaux dans $\mathcal{L}^2(]-1, 1[)$ et vérifiant $L_m(1) = 1$.

La mise sous forme matricielle de ce système se fait en prenant comme fonctions tests les polynômes $q_N(x_1, x_2, x_3) = h_i(x_1) h_j(x_2) h_k(x_3)$ avec $h_i(x)$ le polynôme de lagrange de degré N caractéritique du point x_i. L'inconnue du problème est aussi décomposée selon :

$$p_N(x_1, x_2, x_3) = \sum_{r,s,t=0}^{N} p_N(\mathbf{x}_{rst}) h_r(x_1) h_s(x_2) h_t(x_3)$$

Ce choix nous conduit à écrire pour tout (i, j, k) le système matriciel :

$$\begin{aligned} &\sum_{r=0}^{N} p_N(\mathbf{x}_{rjk}) \big(a\big(x_1, (x_2)_j, (x_3)_k\big) h'_r, h'_i\big)_N \rho_j \rho_k + \sum_{s=0}^{N} p_N(\mathbf{x}_{isk}) \big(a\big((x_1)_i, x_2, (x_3)_k\big) h'_s, h'_j\big)_N \rho_i \rho_k \\ &\quad + \sum_{t=0}^{N} p_N(\mathbf{x}_{ijt}) \big(a\big((x_1)_i, (x_2)_j, x_3\big) h'_t, h'_k\big)_N \rho_i \rho_j = f(\mathbf{x}_{ijk}) \rho_i \rho_j \rho_k \end{aligned} \tag{9}$$

Tout comme la forme bilinéaire (6) la matrice intervenant dans l'Éq. (9) admet des valeurs propres toutes positives dont une valeur propre nulle correspondant à la solution $p =$ constante. Une description détaillée de ses propriétés est donnée dans [5] où est notamment discutée l'utilité d'une surintégration améliorant celle que nous avons présentée et utilisée. Le système (9) est inversé par l'algorithme du gradient conjugué préconditionné (PGC).

La même méthode de discrétisation est utilisée pour le préconditionneur (3), (4) et conduit après division de chacune des lignes par le poids $\rho_i \rho_j \rho_k$, au système algébrique :

$$\begin{aligned} &\frac{1}{\rho_i} \sum_{r=0}^{N} p_N(\mathbf{x}_{ijk}) \big(A_1(x_1) h'_r, h'_i\big)_N + \frac{1}{\rho_j} \sum_{s=0}^{N} p_N(\mathbf{x}_{ijk}) \big(A_2(x_2) h'_s, h'_j\big)_N + \frac{1}{\rho_k} \sum_{t=0}^{N} p_N(\mathbf{x}_{ijk}) \big(A_3(x_3) h'_t, h'_k\big)_N \\ &\quad = f(\mathbf{x}_{ijk}) \end{aligned}$$

qui se réécrit sous la forme du système algébrique tensoriel suivant :

$$(\mathbf{A_1} \otimes \mathbf{I} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A_2} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{I} \otimes \mathbf{A_3})\mathbf{P} = \mathbf{F} \tag{10}$$

Les vecteurs **P** et **F** contiennent respectivement les valeurs de p et de f aux points du maillage. **I** désigne la matrice identité et les matrices $\mathbf{A_i}$, $\mathbf{i} \in \{1, 2, 3\}$ sont données par :

$$(\mathbf{A_i})_{jk} = \frac{1}{\rho_j}(A_i h'_k, h'_j)_N \tag{11}$$

Ce dernier système est inversé en utilisant la méthode dite de diagonalisations successives [6]. Il s'agit d'une méthode directe dont le coût est équivalent à un gradient conjugué convergeant en 6 itérations et ce indépendamment de la valeur de N [7].

4. Résultats numériques

Le but de cette section est d'illustrer les performances de notre préconditionneur en les comparant avec celles obtenues lorsque le préconditionneur est le Laplacien. La donnée f choisie correspond à la solution proposée par Shen et al. [8] :

$$\begin{aligned} p(x_1, x_2, x_3) &= \cos(x_1)\cos(x_2)\cos(x_3) \\ a(x_1, x_2, x_3) &= 1 + 100x_1^2 + x_2^2 + 10^n x_3^2 \end{aligned} \tag{12}$$

On notera que pour cet exemple la variation $\max|a(\mathbf{x})|/\min|a(\mathbf{x})|$ est de l'ordre de $\mathcal{O}(\sup(100, 10^n))$.

Dans le Tableau 1 nous avons présenté le nombre d'itérations utilisé par le (PGC) pour converger vers la solution attendue, pour différents degrés polynomiaux N des fonctions de base. Le critère de convergence a été fixé égal à 10^{-10}. On observe clairement le gain apporté par le nouveau préconditionneur en comparaison avec le cas du Laplacien pour lequel $\mathbf{A} = \mathbf{I}$.

Nous n'avons pas représenté ici de comparaison avec des préconditionneurs élémentaires de type *diagonal* car le préconditionneur Laplacien, en terme de performances, leur est déjà très supérieur [2]. Le nouveau précondionnement que nous proposons réduit significativement le nombre d'itérations de gradient conjugué au regard du préconditionnement par le Laplacien et ce comportement se confirme lorsque n augmente. La Fig. 1 illustre précisément ce fait à savoir que le gain en nombre d'itérations est d'autant plus important que la puissance n est grande. Des calculs en deux dimensions ont confirmé la prééminence du préconditionneur proposé sur le même cas test. Par ailleurs, différents cas tests ont été essayés pour comparer les deux méthodes en présence de fort contrastes du champ $a(\mathbf{x})$. Dans tous les cas, le préconditionneur proposé s'est avéré plus performant.

Tableau 1
Comparaison du nombre d'itérations du PGC (en ordonnée) pour atteindre la convergence en fonction du degré polynomial N (en abcisse) pour un préconditionneur égal au Laplacien ($\mathbf{A} = \mathbf{I}$) et le préconditionneur proposé ici (Éq. (10)) pour $n = 4$

Table 1
Comparaison between our preconditionner and the classical one ($\mathbf{A} = \mathbf{I}$)

N	8	16	24	32	48	64
$\mathbf{A} = \mathbf{I}$	20	51	75	98	137	156
Nouveau préconditionneur	8	14	20	26	36	44

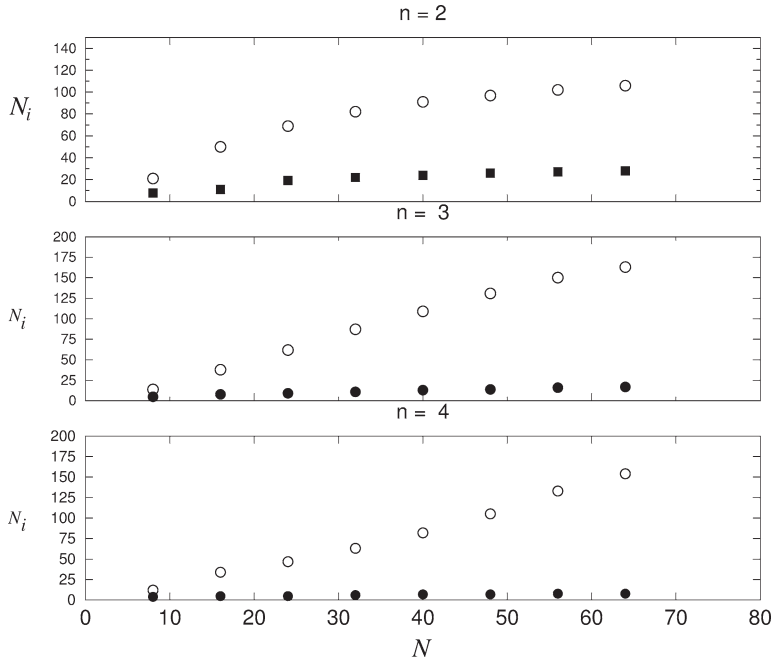

Fig. 1. Evolution du nombre d'itérations N_i avec le degré polynomial N pour différentes valeurs de n (Éq. 12). Les symboles blancs correspondent aux résultats obtenus avec le préconditionneur Laplacien et les symboles noirs avec le préconditionneur proposé ici (Éq. 10).

Fig. 1. PCG's iterations number versus polynomial orders for different values of the power n (white symbols represent classical Laplacian preconditionner, while dark symbols are for the one proposed here).

5. Conclusion

Dans cette Note, nous avons présenté un nouveau préconditionneur pour l'inversion du système algébrique issu de la discrétisation par méthode spectrale d'un problème elliptique du second ordre à coefficients variables et non séparables. Les expériences numériques en deux et trois dimensions montrent l'apport en terme du nombre d'itérations lorqu'un algorithme itératif est utilisé. Ce travail a permis de montrer qu'en utilisant une moyenne uniforme sur l'opérateur initial, le gain par rapport à la méthode de référence (préconditionnement par le Laplacien) peut être considérable. D'autres types de pondérations moyennes pourraient être utilisées en fonction du problème à préconditionner. La suite de ce travail sera consacrée à son extension dans le cadre des éléments spectraux.

Références

[1] C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang, Spectral Methods in Fluid Dynamics, Springer-Verlag, 1987.
[2] M. Deville, E. Mund, E. Ronquist, High-Order Methods for Incompressible Fluid Flow, Cambridge Univ. Press, New York, 2002.
[3] M. Quintard, S. Whitaker, Écoulement monophasique en milieu poreux : effet des hétérogénéités locales, J. Méc. Theor. Appl. 6 (1987) 691–726.
[4] C. Bernardi, M. Maday, Approximations Spectrales de Problèmes aux Limites Elliptiques, Springer-Verlag, 1992.
[5] Y. Maday, E. Ronquist, Optimal error of spectral methods with emphasis on non-constant coefficients and deformed geometries, in: Proceedings of the ICOSAHOM '89 Conference, 1994, pp. 91–115.
[6] R.E. Lynch, J.R. Rice, D.H. Thomas, Direct solution of partial difference equations by tensor product methods, Numer. Math. 6 (1964) 185–199.
[7] M. Azaïez, M. Dauge, M. Maday, Méthodes spectrales et des éléments spectraux, in : G. Cohen (Ed.), Méthodes numériques d'ordre élevé pour les ondes en régime transitoire, in : Collection Didactique, I.N.R.I.A., 1994, Chapter IV.
[8] J. Shen, F. Wang, J. Xu, A finite element multigrid preconditionner for Chebyshev-collocation methods, Appl. Numer. Math. 33 (2000) 471–477.