http://dx.doi.org/10.1090/conm/218/03026

Contemporary Mathematics
Volume **218**, 1998
B 0-8218-0988-1-03026-0

On the Reuse of Ritz Vectors for the Solution to Nonlinear Elasticity Problems by Domain Decomposition Methods

Franck Risler and Christian Rey

1. Introduction

This paper deals with a Rayleigh-Ritz Preconditioner (RRP) that accelerates convergence for the iterative solution to a series of symmetric positive definite linear systems associated with nonlinear substructured elasticity problems. RRP depends upon CG's superconvergent properties and consists of a suitable reuse of Ritz vectors. Moreover, the Rayleigh-Ritz paradigm can be wisely associated with another acceleration technique, the Generalized Krylov Correction, so as to form the SPARKS (Spectral Approach for the Reuse of Krylov Subspaces) algorithm. Numerical assessment of both RRP and SPARKS is provided on a large-scale poorly-conditioned engineering practice.

2. Solution to Nonlinear Elasticity Problems

We consider computation of the equilibrium of bodies made up of compressible hyperelastic material and that undergo large deformation. A Lagrangian formulation is chosen and all variables are defined in the reference configuration. Moreover, Ω in $\mathbb{R}^3$ and Γ exhibit the domain occupied by the body and its boundary respectively. The equilibrium equations may then be written in a weak form as follows

$$(1)\qquad \begin{cases} \text{Find } u \in \{H + u_0\} \text{ such that} \\ \displaystyle\int_\Omega \frac{\partial \Phi}{\partial F}(u) : \nabla v d\Omega = \int_\Omega f.v d\Omega + \int_{\Gamma_g} g.v d\Gamma \qquad \forall v \in H \\ H = \{v \in H^1(\Omega)^3, v = 0 \text{ on } \Gamma_u = \Gamma - \Gamma_g\} \end{cases}$$

where H denotes the space of kinematically-admissible displacement fields, (:) stands for the double contractor operator between two tensors A and B ($A : B = \mathrm{Tr}(A^T B)$), x are the coordinates of any particle of the domain measured in the reference configuration (Ω) in a fixed orthonormal basis of $\mathbb{R}^3$, u_0 is the imposed

---

1991 *Mathematics Subject Classification.* Primary 65B99; Secondary 65Y05.

Key words and phrases. Domain Decomposition, Nonlinear Elasticity, Superlinear Convergence, Krylov Subspaces, Ritz Values.

Supercomputing facilities were provided by the IDRIS Institute (Institut du Développement et de Ressources en Informatique Scientifique). The two authors are grateful to the IDRIS Institute for this support.

displacement field on part Γ_u of the domain boundary, $v(x)$ is any admissible displacement field in the reference configuration, $u(x)$ is the unknown displacement field, $F(x) = Id + \nabla u(x)$ is the deformation gradient, $g(x)$ is the surface tractions on part Γ_g of the domain boundary, complementary to Γ_u in Γ, $f(x)$ is the density of body forces (we assume that external loadings f and g do not depend on the displacement field u - dead loading assumption), and Φ is the specific internal elastic energy.

The problem given by Eq.(1) is discretized through a finite element method [**13**] and leads to the solution to a nonlinear problem of the form $\mathcal{F}(u) = 0$. Such a discrete problem is solved by means of Newton-type methods that amount to the resolution to a succession of symmetric positive definite linear problems, the right hand sides and the matrices of which are to be reactualized.

The reader may refer to [**1**] for a complete presentation of nonlinear elasticity problems. Moreover, further explanations on Newton-type algorithms can be found in [**5**] or in [**6**].

3. Iterative Solution to a Series of Linear Problems

3.1. The Substructuring Paradigm. By condensing each linear problem on the subdomains interface, non overlapping Domain Decomposition (DD) methods (primal [**7**] or dual [**4**] approach) enable to solve iteratively with a Conjugate Gradient (CG) algorithm the following succession of linear problems,

$$(P^k): \; A^k x^k = b^k \quad , \quad k = 1, \dots, m \tag{2}$$

where A^k denotes the matrix of Schur complement either in primal or dual form depending on the approach chosen, and b^k is the associated condensated right hand sides. Note that the A^k matrix herein considered is symmetric, positive, definite [**9**]. From now on, we will be focusing on the dual domain-decomposition paradigm. The proposed Ritz preconditioner may nevertheless suit to the primal approach, though some characteristic properties of the dual interface operator magnify the positive effects upon the convergence of this preconditioner.

3.2. Definition and Fundamental properties.

3.2.1. *CG characterizing Properties.* The Conjugate Gradient algorithm applied to the solution to the linear problem (P^k) arising from Eq.(2), depends on the construction of a set of w_i^k descent directions that are orthogonal for the dot product associated with the A^k matrix. The Krylov subspace thus generated may be written as

$$K_{r_k}(A^k) = \{ \, w_0^k, w_1^k, \dots, w_{r_k-1}^k \, \} \quad ; \quad K_{r_k}(A^k) \subset \mathbb{R}^n \tag{3}$$

where the subscript r_k denotes the dimension of this latter subspace and n exhibits the number of unknows of the substructured problem to be solved.

A characteristic property of CG is given by

$$\left\| x^k - x_{r_k}^k \right\|_{A^k} = \min_{y \in x_0^k + K_{r_k}(A^k)} \left\| x^k - y^k \right\|_{A^k} \tag{4}$$

where x_0^k is a given initial field and with $\|v\|_{A^k} = (A^k v, v)$.

Consequently, by introducing the A^k-orthogonal projector $P_{K_{r_k}}^{A^k}$ onto the $K_{r_k}(A^k)$ Krylov subspace, the r_k-rank approximation of the solution can be written

$$x_{r_k}^k = x_0^k + P_{K_{r_k}}^{A^k}(b^k - A^k x_0^k) \quad \text{with} \quad P_{K_{r_k}}^{A^k}(x) = \sum_{i=0}^{r_k-1} \frac{(x, w_i^k)}{(A^k w_i^k, w_i^k)} w_i^k \tag{5}$$

3.2.2. *Ritz Vectors and Values.* The Ritz vectors ${y_j}^{(r_k)} \in K_{r_k}(A^k)$ and $\theta_j^{(r_k)} \in \mathbb{R}$ values are defined such that [**8**]

$$A^k y_j^{(r_k)} - \theta_j^{(r_k)} {y_j}^{(r_k)} \perp K_{r_k}(A^k) \tag{6}$$

The convergence of the Ritz values towards a set of r_k eigenvalues of the A^k matrix exhibits the dominating phenomenon, on which the CG's rate of convergence and the so-called superlinear convergence behavior [**8, 12**] depends.

4. The Rayleigh-Ritz Preconditioner

4.1. A Krylov Based Spectral Approach. The purpose of this new preconditioner is to utilize spectral information related to the dominating eigenvalues arising from Krylov subspaces so as to accelerate the resolution of a succession of linear systems of the form given by Eq.(2). The relevance of this approach has been analysed in ([**11**], criterion 2.3) and its validity domain has been defined.

More precisely, we intend herein to very significantly accelerate the convergence of a set of p dominating Ritz values to trigger a superconvergent behavior of CG. The key to the Ritz approach lies in the so-called *effective* condition number that quantitatively weights the rate of the CG's convergence in the course of the resolution process. The *effective* condition number is defined at the j iteration of the CG as the ratio of the largest uncaptured eigenvalue of the A^k matrix to its smallest eigenvalue. Note that a given λ eigenvalue of the A^k matrix is considered captured whenever a Ritz value provides a sufficiently accurate approximation of λ so that the corresponding eigenvector no longer participates in the solution process [**12**].

Therefore, with the Ritz approach, we seek to drastically reduce the *effective* condition number within the first CG's iterations. Besides, the spectrum of the dual interface operator is distinguished by few dominating eigenvalues which are not clustered and are well separated from the smaller ones [**4**]. Consequently, the new algorithm is expected to be even more efficient that some of the intrinsic spectral properties of the linear problems we deal with, magnify its positive effects upon convergence of the dominating Ritz values.

Let us define a $Q \in \mathbb{R}^{n \times p}$ matrix, the columns of which store an approximation of p eigenvectors of the current A^k operator. Inasmuch as we are aiming to reduce the *effective* condition number, we will prescribe at the i iteration of the Conjugate Gradient algorithm an optional orthogonality constraint that is presented as

$$Q^T g_i = 0 \quad \forall i \tag{7}$$

In terms of Krylov subspaces, that yields

$$K_{r_k}(A^k) \subset Ker Q^T = (Im Q)^{\perp} \tag{8}$$

Note that this orthogonality constraint is similar to the one associated with a new framework that has been recently introduced to speed up convergence of dual substructuring methods [3]. But, while in [3] considerations upon domain decomposition method found the algorithm, the Ritz approach depends on a spectral analysis of the condensed interface matrix. Consequently, apart from this sole formulation similarity, these two methods are based on completely different concepts.

Furthermore, we emphasize the fact that the constraint given by Eq.(7) is optional and, for obvious reasons, does not modify the admissible space to which the solution belongs. Besides, providing that the constraint is enforced at each CG's iteration, it must consequently be verified by the solution to the linear problem. Since the residual vector associated with the final solution is theoretically equal to zero, the orthogonality condition prescribed by Eq.(7) is thus satisfied.

Let's now focus on the construction of the Q matrix in the framework of the resolving to a series of linear problems. We advocate that the approximation of eigenvectors arises from the Ritz vectors associated with the p dominating Ritz values originating from the first system (P^1). Note that the efficiency of the conditioning problem depicted in Eq.(7) is submitted to two main assumptions ([11], Hypothesis 3.1) in (a) the convergence of Ritz vectors, and (b) the perturbation of eigendirections among the family $\{A^k\}_{k=1}^{k=m}$ of matrices.

Moreover, if the number of linear problems to be solved is high and the columns of the Q matrix do not provide a sufficiently accurate approximation of eigenvectors related to dominating eigenvalues of a given A^q matrix $(1 < q \leq m)$, the Q matrix has to be reactualized. Hence, it requires suspending the prescription of the constraint given in Eq.(7) while solving (P^q) and computing the Ritz vector associated with the p dominating Ritz values in order to update the Q matrix. The Ritz conditioning problem is then restored until another reactualization procedure is required.

4.2. Construction of the Rayleigh-Ritz Preconditioner. For the sake of clarity, and since no confusion is possible, the k superscript is herein omitted and the A^k matrix is simply noted A.
In order to prescribe the optional constraint given by Eq.(7), we shall superpose, at each iteration i of the Conjugate Gradient algorithm, the field x_i and an additional field of Lagrange multipliers ξ_i such that

$$(9)\qquad \begin{aligned} x_i \longrightarrow \widetilde{x}_i &= x_i + \xi_i = x_i + Q\alpha_i \\ Q^T\widetilde{g}_i &= 0 \quad \text{with} \quad \widetilde{g}_i = A\widetilde{x}_i - b \end{aligned}$$

Substituting the second equation of Eq.(9) into the first one yields

$$(10)\qquad Q^TA^kQ\alpha_i + Q^TA^kx_i - Q^Tb^k = 0$$

Consequently, α is given by

$$(11)\qquad \alpha_i = -(Q^TA^kQ)^{-1}Q^TA^kx_i + (Q^TA^kQ)^{-1}Q^Tb^k$$

Then, substituting Eq. (11) into Eq.(9) yields

$$(12)\qquad \begin{aligned} \widetilde{x}_i &= (I - Q(Q^TA^kQ)^{-1}Q^TA^k)x_i + Q(Q^TA^kQ)^{-1}Q^Tb^k \\ \widetilde{x}_i &= (I - Q(Q^TA^kQ)^{-1}Q^TA^k)x_i + x_0 \\ \text{with} \quad x_0 &= Q(Q^TA^kQ)^{-1}Q^Tb^k \end{aligned}$$

Let the projector $\overline{P}$ be defined by

$$(13)\qquad \overline{P}: \; x_i \longrightarrow \widetilde{x}_i - x_0 \quad \text{with} \quad \overline{P} = (I - Q(Q^TA^kQ)^{-1}Q^TA^k)$$

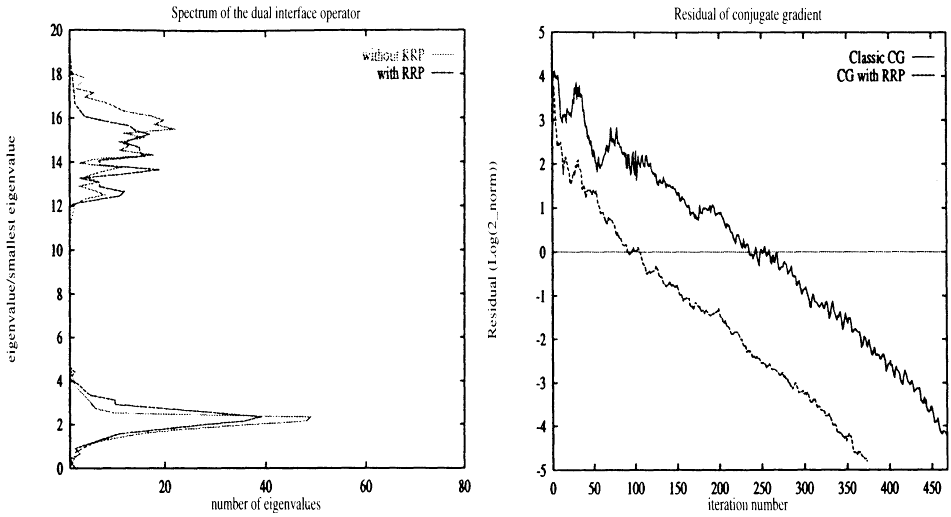

FIGURE 1. Spectral distribution and residual history with or without the RRP (R100) preconditioner

Since the considered projector $\overline{P}$ is not symmetric, the prescription of the Rayleigh-Ritz Preconditioner (RRP) at each CG iteration has then to be performed in two projection steps, by $\overline{P}^T$ and $\overline{P}$ respectively.

5. Application

Numerical efficiency is assessed on a large-scale poorly-conditioned non linear problem: a three-dimensional steel-elastomer laminated structure that distinguishes with great heterogeneousness and high nonlinearity. The considered structure has a parallelepiped geometry and is discretized by hexaedral finite elements (Q1 elements). Besides, an axial compression loading with an imposed displacement is applied. Material behavior is modelled by the Ciarlet-Geymonat specific internal energy Φ [**2**] and the associated equivalent Young modulus and Poisson coefficient are $(E, \nu) = (1.3 \text{ MPa}; 0.49)$ and $(E, \nu) = (2 \times 10^5 \text{ MPa}; 0.3)$ for the elastomer and the steel respectively.

On account of the quadratic convergence of the Newton methods, the stopping criterion of Newton iterations (nonlinear iterations) is set to 10^{-6} while the accuracy requirement for solving each linear problem is 10^{-3}. In all cases, the linear problems are also preconditioned by the classical *Lumped* preconditioner [**4**] and the spectral results have been estimated from the Ritz values for the Krylov space generated by the Conjugate Gradient algorithm, when a number of iterations n_r (close to the dimension n of the problem) is performed. Finally, N_s and N denote from now on the number of processors and the number of unknows of the nonlinear problem considered respectively.

5.1. Numerical Performance. In Table 1, is reported the number of iteration achieved by the Conjugate Gradient algorithm within the Newton iterations and the Figure 1 exhibits the spectral distribution and the residual history of the (P^2) linear problem. In all cases, *Classic CG* means that RRP is not applied and *CG with RRP (R_p)* indicates that a RRP whose size is p is prescribed. On the

TABLE 1. Numerical performances with various preconditioners

N_s	N	solver	Newton Iterations		
			N_1	N_2	N_3
15	85680	Classic CG	294	307	367
15	85680	CG with RRP (R50)	294	254	312
15	85680	CG with RRP (R100)	294	212	272
15	85680	CG GKC	294	120	71
15	85680	CG with SPARKS (R50)	294	73	51

spectral distribution, we observe that, not only the condition number κ of the RRP preconditioned matrix is reduced, but also the dominating values are fewer and more spread out. It thus paves the way for a fast-convergence of Ritz values towards the dominating eigenvalues – and hence a drastic reduction of the *effective* condition number – during the first CG iterations, what is supported by the chart of the residual history. On the other hand, this latter curve shows that, in a second phase of the resolution process, when Ritz values have converged towards the eigenvalues associated with the eigenvectors, an approximation of which is provided by the columns of the Q matrix, the rate of convergence is decelerated and becomes close to the one of the Classical CG.

Computational results in terms of CPU time are not addressed in this paper since they highly depend on implementation issues and would require further explanations. Nonetheless, numerical assessments show that the provided acceleration of convergence is not offset against the computational overheads involved by the RRP prescription.

5.2. The SPARKS algorithm. Whereas the Rayleigh-Ritz preconditioner is based on a condensation of information related to the upper part of the spectrum, the Generalized Krylov Correction[**10**] reuses in a broader and a non-selective way information originating from previously generated Krylov subspaces. We associate those two latter algorithm within an hybrid (RRP-GKC) preconditioner, which we will be calling from now on SPARKS (Spectral Approach for the Reuse of Krylov Subpaces).

Numerical experiments show that the RRP has a dominating contribution during the first iterations, when the Conjugate Gradient (CG) explores spectral subspaces related to the dominating eigenvalues. Afterwards, the rate of convergence is mainly ruled by the GKC preconditioner which enables to speed up the capture of phenomena associated with lower frequencies. A numerical assessment is provided in Table 1 and further validations have shown that the very significant acceleration of convergence provided by RRP within the SPARKS algorithm goes far beyond the computational overcost generated. Moreover, SPARKS distinguishes with an increasing efficiency when the size n of the problem grows.

6. Conclusion

We have presented in this paper a Raleigh-Ritz preconditioner, that is characterized by the reuse of spectral information arising from previous resolution processes. Principles and construction of this preconditioner have been addressed and numerical performance of the Ritz approach has been demonstrated on a large-scale poorly-conditioned engineering problem. Moreover, a new hybrid Krylov-type preconditioner, known as SPARKS, and deriving from both the Rayleigh-Ritz and the Generalized Krylov Correction has been introduced and has proved outstanding numerical performances.

References

1. P.G. Ciarlet, *Mathematical elasticity*, North-Holland, Amsterdam, 1988.
2. P.G. Ciarlet and G. Geymonat, *Sur les lois de comportement en elasticité non linéaire compressible*, C.R. Acad. Sci. Paris **T. 295, Série II** (1982), 423–426.
3. C. Farhat, P-S. Chen, F. Risler, and F-X. Roux, *A simple and unified framework for accelerating the convergence of iterative substructuring methods with Lagrange multipliers*, International Journal of Numerical Methods in Engineering (1997), in press.
4. C. Farhat and F-X. Roux, *Implicit parallel processing in structural mechanics*, vol. 2, Computational Mechanics Advances, no. 1, North-Holland, june 1994.
5. H.B. Keller, *The bordering algorithm and path following near singular points of higher nullity*, SIAM J. Sci. Stat. Comput. **4** (1983), 573–582.
6. P. Le Tallec, *Numerical analysis of equilibrium problems in finite elasticity*, Tech. Report CEREMADE 9021, Cahier de la décision, Université Paris-Dauphine, 1990.
7. ______, *Implicit parallel processing in structural mechanics*, vol. 1, Computational Mechanics Advances, no. 1, North-Holland, june 1994.
8. B. N. Parlett, *The symmetric eigenvalue problem*, Prenctice-Hall,Englewood Cliffs, New Jersey, 1980.
9. C. Rey, *Développement d'algorithmes parallèles de rrésolution en calcul non linéaire de structures hétèrogènes: Application au cas d'une butée acier-elastomère*, Ph.D. thesis, Ecole Normale Supérieure de Cachan, 1994.
10. C. Rey and F. Léné, *Reuse of Krylov spaces in the solution of large-scale nonlinear elasticity problems*, Proceeding of the Ninth International Conference on Domain Decomposition Methods, Wiley and Sons, in press.
11. C. Rey and F. Risler, *The Rayleigh-Ritz preconditioner for the iterative solution of large-scale nonlinear problems*, Numerical Algorithm (1998), in press.
12. A. Van der Luis and H. A. Van der Vorst, *The rate of convergence of conjugate gradients*, Numerische Mathematik **48** (1986), 543–560.
13. O. Zienkiewicz, *The finite element method*, McGraw-Hill, New-York, Toronto, London, 1977.

Laboratoire de Modélisation et Mécanique des Structures - U.R.A 1776 du C.N.R.S, U.P.M.C - TOUR 66 - 5ième Etage - 4, Place Jussieu 75252 Paris Cedex 05

E-mail address: risler@rip.ens-cachan.fr, rey@ccr.jussieu.fr