**Proceedings of the 42nd IEEE
Conference on Decision and Control
Maui, Hawaii USA, December 2003**

A Newton algorithm for invariant subspace computation with large basins of attraction

P.-A. Absil*† R. Sepulchre* P. Van Dooren‡ R. Mahony§

Abstract

We study the global behaviour of a Newton algorithm on the Grassmann manifold for invariant subspace computation. It is shown that the basins of attraction of the invariant subspaces may collapse in case of small eigenvalue gaps. A Levenberg-Marquardt-like modification of the algorithm with low numerical cost is proposed. A simple strategy for choosing the parameter is shown to dramatically enlarge the basins of attraction of the invariant subspaces while preserving the fast local convergence.

Keywords: invariant subspace, Grassmann manifold, cubic convergence, symmetric eigenproblem, Newton method, global convergence.

AMS Mathematics Subject Classification 2000: 65F15.

1 Introduction

The problem of computing a p-dimensional eigenspace (i.e. invariant subspace) of an n-by-n matrix A is ubiquitous in applied mathematics, with applications in control theory, pattern recognition, data compression and coding, antenna array processing, and a multitude of other domains.

Several methods for subspace estimation were proposed in the late seventies and early eighties. Demmel [Dem87] provides a joint analysis of three of the early methods that refine initial estimates of arbitrary p-dimensional eigenspaces of a (possibly nonsymmetric) $n \times n$ data matrix A. The early methods depend on the various numerical solutions of a common Riccati equation. These methods converge at best quadratically (Chatelin's Newton-based method [Cha84]) even when A is symmetric and involve the solution of a Sylvester equation at each iteration step. Moreover, the iterations defined depend on a choice of normalization used to generate the Riccati equation as well as the present iterative estimate of the eigenspace. More recently, Edelman, Arias and Smith [EAS98] have derived a Newton method on the Grassmann manifold (i.e. the set of p-planes in $\mathbb{R}^n$) that relies on the Riemannian geometry of the manifold. When applied to find critical points of a generalized Rayleigh quotient, this Newton method converges locally cubically to the eigenspaces of $A = A^T$. A similar iteration was derived independently in [LST98]. We will refer to this method as the *Newton-Grassmann* method. A practical implementation was investigated by Lundström and Eldén [LE02]. It is shown in [AMS02] that local (at least) quadratic convergence also holds when $A \neq A^T$, and that convergence is cubic to —and only to— the subspaces that are both eigenspaces of A and A^T. Other iterative methods that operate on the Grassmann manifold and refine estimates of eigenspaces include the multishifted Rayleigh quotient iteration (related to the generalized Rayleigh-shifted QR algorithm [PP73, WE91]) and the block-shifted Rayleigh quotient iteration analysed in [AMSV02].

The first aim of the present paper is to study the global behaviour of the Newton-Grassmann method. In the case where $p = 1$ and only a single eigenvector is computed, the Newton-Grassmann method is known to be equivalent to the Rayleigh quotient iteration (see e.g. [Smi94, ADM$^+$02]) for which the global behaviour is well understood [Par80, BS89]. For the case $p > 1$, existing analytical results focus on the local convergence rate but almost no global analysis has been undertaken. We show that the basins of attraction tend to collapse around the eigenspaces with small *external* eigenvalue gap (the external gap measures the separation between the eigenvalues of A corresponding to the eigenspace and the other eigenvalues of A). However, the quality of the basin of attraction of an eigenspace is not affected by its *internal* gap (the separation between the eigenvalues of A reduced to the eigenspace); this is in contrast with the block-shifted Rayleigh quotient iteration [AMSV02, ASVM03].

Collapsing basins of attraction are undesirable for a method that aims at refining estimates of eigenspaces, since even a good initial estimate of an eigenspace could

*Department of Electrical Engineering and Computer Science, Université de Liège, Institut Montefiore (B28), Grande Traverse 10, B-4000 Liège, Belgium.

‡Department of Mathematical Engineering, Université Catholique de Louvain, Bâtiment Euler (A.119), Avenue Georges Lemaître 4, B-1348 Louvain-la-Neuve, Belgium.

§Department of Engineering, Australian National University, ACT, 0200, Australia.

†Aspirant du F.N.R.S. (Research Fellow with the Belgian National Fund for Scientific Research).

produce convergence to another eigenspace. The second aim of this paper is therefore to propose modifications of the Newton method that enlarge the basins of attraction around the eigenspaces. The modified iterations should moreover preserve the fast local convergence of the pure Newton method and the numerical cost of computing a new iterate should be kept low. The modification we propose, which can be understood as a Levenberg-Marquardt technique, consists in introducing a deformation parameter τ that achieves a continuous transition between the original iteration and a gradient flow with large basins of attraction. This deformation technique is related to line search methods and trust region methods in optimization. We propose a simple choice for τ that dramatically enlarges the basins of attraction around the attractors while preserving the local rate of convergence. The numerical cost of is $O(np^2)$ flops per iteration when A is suitably condensed.

This paper is organized as follows. After a short review of subspaces, eigenspaces and their representations (Section 2), we state the Newton-Grassmann iteration and study its practical implementation and its global behaviour (Section 3). Then in Section 4 we propose modifications of the Newton-Grassmann iteration with enlarged basins of attraction.

2 Subspaces and eigenspaces

This section introduces concepts and notation pertaining to subspaces and eigenspaces. We use $\mathrm{Grass}(p,n)$ to denote the *Grassmann manifold* of the p-dimensional subspaces of $\mathbb{R}^n$ and $\mathrm{ST}(p,n)$ to denote the *noncompact Stiefel manifold*, i.e. the set of n-by-p matrices with full rank. The column space of $Y \in \mathrm{ST}(p,n)$ is denoted by $\mathrm{span}(Y)$. The "span" mapping is an application on $\mathrm{ST}(p,n)$ onto $\mathrm{Grass}(p,n)$ that is nowhere invertible. Given a matrix Y in $\mathrm{ST}(p,n)$, the set of matrix representations of the subspace $\mathrm{span}(Y)$ is

$$\mathrm{span}^{-1}(\mathrm{span}(Y)) = Y\,\mathrm{GL}_p := \{YM : M \in \mathrm{GL}_p\}$$

where GL_p denotes the set of p-by-p invertible matrices. This identifies $\mathrm{Grass}(p,n)$ with $\mathrm{ST}(p,n)/\mathrm{GL}_p := \{Y\,\mathrm{GL}_p : Y \in \mathrm{ST}(p,n)\}$. More details on the Grassmann manifold and matrix representations can be found in [AMS02, Abs03].

Let A be an $n \times n$ matrix. Let $\mathcal{X}$ be a p-dimensional subspace of $\mathbb{R}^n$ and let $Q = [X|X_\perp]$ be an orthogonal $n \times n$ matrix such that X spans $\mathcal{X}$. Then Q^TAQ may be partitioned in the form $Q^TAQ = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right)$ where $A_{11} \in \mathbb{R}^{p\times p}$. The subspace $\mathcal{X}$ is an *eigenspace* (i.e. invariant subspace) of A if and only if $A_{21} = 0$. The *external gap* of the eigenspace $\mathcal{X}$ of A is the shortest distance between the eigenvalues of A_{11} and the eigenvalues of A_{22}. The *internal gap* of $\mathcal{X}$ is the shortest distance between two eigenvalues of A_{11}. The eigenspace $\mathcal{X}$ is called *spectral* if A_{11} and A_{22} have no eigenvalue in common (i.e. nonvanishing external gap). The span of a full-rank $n \times p$ matrix Y is an eigenspace of A if and only if there exists a matrix L such that $AY = YL$, in which case Y is called an *eigenbasis* and L the corresponding *eigenblock* [JS01].

3 The Newton-Grassmann iteration

3.1 Definition

It comes as a direct consequence from the definitions in Section 2 that the p-dimensional eigenbases of A are the full-rank $n \times p$ solutions of the matrix equation

$$F(Y) := \Pi_{Y_\perp} AY = 0 \tag{1}$$

where $\Pi_{Y_\perp} := I - Y(Y^TY)^{-1}Y^T$ is the orthogonal projector onto the orthogonal complement of $\mathrm{span}(Y)$. Let

$$H_Y := \{\Delta \in \mathbb{R}^{n\times p} : Y^T\Delta = 0\} \tag{2}$$

denote the so-called *horizontal space*, orthogonal to the fiber $Y\,\mathrm{GL}_p$. Define a map $J_Y : H_Y \to H_Y$ by projecting the Fréchet derivative of F in a direction $\Delta \in H_Y$ back onto H_Y,

$$\begin{aligned} J_Y : H_Y \to H_Y : \Delta \mapsto \Pi_{Y_\perp} \mathrm{D}F(Y)[\Delta] \\ = \Pi A \Pi\Delta - \Delta(Y^TY)^{-1}Y^TAY \end{aligned} \tag{3}$$

where Π is a shorthand notation for $\Pi_{Y_\perp}$. Using this notation, the Newton-Grassmann equation reads

$$J_Y[\Delta] = -F(Y), \tag{4}$$

which expands into equation (5) below and yields the Newton-Grassmann (NG) algorithm, formally stated as follows:

Algorithm 3.1 (NG) *Iterate the mapping* $\mathrm{Grass}(p,n) \ni \mathcal{Y} \mapsto \mathcal{Y}_+ \in \mathrm{Grass}(p,n)$ *defined by*
1. Pick a basis $Y \in \mathbb{R}^{n\times p}$ *that spans* $\mathcal{Y}$ *and solve the equation*

$$\Pi A \Pi\Delta - \Delta(Y^TY)^{-1}Y^TAY = -\Pi AY \tag{5}$$

under the constraint $Y^T\Delta = 0$, *where*

$$\Pi := I - Y(Y^TY)^{-1}Y^T.$$

2. Perform the update

$$\mathcal{Y}_+ = \mathrm{span}(Y + \Delta). \tag{6}$$

One checks that $\mathcal{Y}_+$ does not depend on the Y chosen in step 1. Indeed, if Y yields the solution Δ of (5), then YM produces the solution ΔM for any $M \in \mathrm{GL}_p$, and $\mathrm{span}(Y+\Delta) = \mathrm{span}((Y+\Delta)M)$.

Algorithm NG admits the following geometric interpretation, valid for arbitrary A. The Grassmann manifold, endowed with the essentially unique Riemannian metric invariant by the action of the orthogonal group, is a Riemannian manifold. In [Smi94], Smith proposes a Newton iteration on abstract Riemannian manifolds. This iteration, applied on the Grassmann manifold in order to solve (1), yields the search direction Δ given by (5), where Δ represents an element of the tangent space $T_Y\mathrm{Grass}(p,n)$ [AMS02]. The update (6) is a simplification of the Riemannian updating procedure

$$\mathcal{Y}_+ = \mathrm{Exp}_{\mathcal{Y}}\Delta \tag{7}$$

consisting in following geodesics on the Grassmann manifold. Assuming $A = A^T$, Algorithm NG (with geodesic update) is also obtained by applying the Riemannian Newton method on $\mathrm{Grass}(p,n)$ for finding a stationary point of a generalized Rayleigh quotient [EAS98].

When $A = A^T$, NG with either geodesic update (7) or projected update (6) converges locally cubically to the p-dimensional spectral eigenspaces of A [AMS02].

3.2 Practical implementation

The major computational load in the NG algorithm is to solve the constrained Sylvester equation (5). Choosing an orthonormal matrix X to represent the current iterate reduces the NG equation (5) to

$$\Pi A \Pi \Delta - \Delta A_{11} = -\Pi A X, \;\; X^T\Delta = 0, \tag{8}$$

where $A_{11} = X^TAX$, and the next iterate is given by $X_+ = (X+\Delta)M$, where M is chosen to orthonormalize X_+ (M can e.g. be obtained by a QR factorization). The first thing to do is to reduce A_{11} to an upper-triangular form. This is cheap since A_{11} is a $p \times p$ matrix and p is in practical applications typically much smaller than n. For simplicity of the argument, we will assume that A_{11} can be reduced to diagonal form, which e.g. happens in the frequently encountered case $A = A^T$. This diagonalization decouples equation (8) into p independent systems of linear equations of the form

$$\Pi(A - \rho_i I)\Pi\delta = -\Pi A x, \;\; X^T\delta = 0. \tag{9}$$

In the general case where A_{11} is upper-triangular, the first equations are independent from the last ones and a substitution technique can be used. An alternative to the approach proposed in [LE02] for solving (9) is to use an idea from [PW79] and write equation (9) as an $(n+p)\times(n+p)$ problem :

$$\begin{bmatrix} A-\rho_i I & X \\ X^T & 0 \end{bmatrix}\begin{bmatrix}\delta \\ m\end{bmatrix} = \begin{bmatrix}-Ax \\ 0\end{bmatrix} \tag{10}$$

For some particular structures of A, efficient methods can be used to solve (10). For example, if A is in tridiagonal form, then (10) can be solved in $O(np^2)$ flops (see [ASVM03] for details).

3.3 Basins of attraction

Previous convergence results for the Newton-Grassmann algorithm were *local* results [EAS98, LST98, AMS02]. That is, around each p-dimensional spectral eigenspace $\mathcal{V}$ there *exists* a ball B in the Grassmann manifold $\mathrm{Grass}(p,n)$ such that the iteration converges to $\mathcal{V}$ for all initial point in B. An important issue however is the *size* of this ball, since a large ball means that the iteration will converge to the targeted eigenspace even if the initial estimate is not very precise. From now on, unless otherwise stated, we assume that $A = A^T$, which guarantees that the spectral eigenspaces are orthogonal to each other. In the general case, the distance between eigenspaces may be arbitrarily small and therefore large basins of attraction around the eigenspaces can not be expected.

In the case $p = 1$, the Newton-Grassmann algorithm is equivalent to the Rayleigh quotient iteration (RQI) [Par80] as shown e.g. in [Smi94, ADM+02]. The global behaviour of the RQI has been studied in depth by Batterson and Smillie [BS89]. They have shown that the iteration converges to an eigendirection of A from almost all initial point. Moreover, their analysis of the low-dimensional $n = 3$ case shows that the basin of attraction of the dominant eigenvector collapses when the dominant eigenvalue is ill-separated from the second-dominant eigenvalue (see illustration on Figure 1).

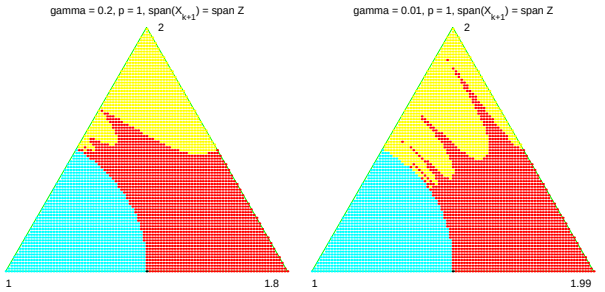

Figure 1: Basins of attraction for NG ($n = 3$, $p = 1$) for $A = \mathrm{diag}(1, 1.8, 2)$ (left) and $A = \mathrm{diag}(1, 1.99, 2)$ (right). The three vertices correspond to the three eigenspaces. A similar figure appears in [BS89]. This figure also applies to NG with $n = 3$, $p = 2$.

When $p > 1$, the NG algorithm, the multi-shifted RQI and block-shifted RQI mentioned in the introduction all differ [ASVM03]. The global behaviour of the NG algorithm in the case $p = 2$, $n = 3$ can be deduced from the following *duality property*: If $\mathcal{X}^k$ is a sequence of iterates generated by NG, then their orthog-

onal complements $\mathcal{X}_\perp^k$ also forms a sequence of iterates of NG [ASVM03]. Therefore, representing the 2-planes by their orthogonal complements yields back the basins of attraction of the case $p = 1$, $n = 3$ illustrated on Figure 1.

We now consider a higher-dimensional case with $p = 3$ and $n = 7$. In order to illustrate the influence of eigenvalue gaps (external and internal) on the basins of attraction, we use the matrix

$$A = \mathrm{diag}(1, 2, 2+\gamma, 2+2\gamma, 3, 4, 5) \qquad (11)$$

where γ is a small number (we choose $\gamma = 0.01$). We select three different eigenspaces in order to illustrate the influence of internal and external gaps. In each case, we pick 10^4 initial points randomly at three given distances of the targeted eigenspace and we count how often the sequence of iterates fails to converge to the target. Here are the results of our experiments:

Init. err.	$\mathcal{V}_{leli}$	$\mathcal{V}_{lesi}$	$\mathcal{V}_{seli}$
$\frac{1}{50}\frac{\pi}{2}$	0%	0%	0%
$\frac{1}{10}\frac{\pi}{2}$	0%	0%	3.35%
$\frac{1}{3}\frac{\pi}{2}$	0.10%	0%	11.80%

The three colums correspond to three different targeted eigenspaces: $\mathcal{V}_{leli}$ with large external and internal gaps (eigenvalues 1, 3, 4); $\mathcal{V}_{lesi}$ with large external gap and small internal gap (eigenvalues 2, $2+\gamma$, $2+2\gamma$); and $\mathcal{V}_{seli}$ with small external gap and large internal gap (eigenvalues 2, 3, 4). The initial error is defined as the largest principal angle between the initial subspace and the targeted eigenspace. These results illustrate that basins of attraction collapse with small external gap but are not affected by a small internal gap. An analytical interpretation is given in [ASVM03].

4 Modified NG algorithm with enlarged basins of attraction

Large basins of attraction around the eigenspaces are desirable as they ensure that the iteration will converge to the targeted eigenspace even if the initial subspace is a relatively poor estimate. The analysis in Section 3.3 has shown that a small external eigenvalue gap produces a degradation of the basins of attraction of the NG algorithm. For this reason, we now propose a way of improving enlarging the basins of attraction around the eigenspaces.

Let us define a *cost function*

$$f(Y) := \frac{1}{2}\mathrm{trace}((Y^TY)^{-1}F(Y)^TF(Y)) \qquad (12)$$

where F is defined by (1). It is easily checked that $f(Y)$ depends only on the span of Y, and not on the basis Y itself [AMS02]. So, the cost function f defines a scalar field on the Grassmann manifold. This scalar field is zero at the eigenspaces of A and strictly positive everywhere else. An illustration of the level curves of f is shown on Figure 2.

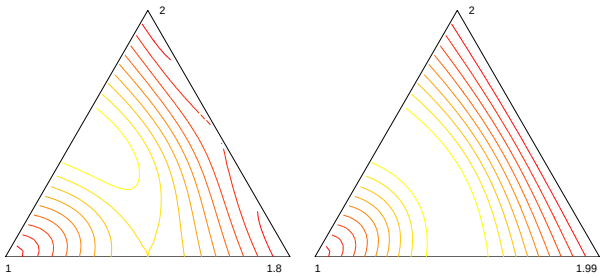

Figure 2: Level curves of the cost function f for $A = \mathrm{diag}(1, 1.8, 2)$ (left) and $A = \mathrm{diag}(1, 1.99, 2)$ (right). The cost function vanishes at the three eigenspaces and only there. The gradient descent flow for f consists in following the steepest descent path of these level curves.

While the basins of attraction of the NG iteration deteriorate in the presence of a small external gap, Figure 2 suggests that the basins of attraction of the steepest descent flow of the cost function f remain broad even when the eigenvalue gap shrinks. A numerical simulation of the steepest descent flow of f for A given in (11) shows indeed that the distance between each eigenspace and the boundary of its basin of attraction is large (greater than $\frac{1}{3}\frac{\pi}{2}$) in all cases. This prompts us to follow the steepest descent of f when the solution is far away from a solution and use the Newton method in the neighbourhood of a solution. It is however difficult to decide when the transition between the two behaviours should occur. If the Newton iteration takes over too soon, the basins of attraction may be collapsed. If the transition occurs late in the iterative process, then more steps will be necessary before obtaining a good approximation of the eigenspace. A remedy is to implement a smooth transition between the two behaviours by means of a deformation parameter, an idea which connects with trust region methods (see e.g. [DS83] or chap. 11 in [NW99]). Let Y be a basis for the current iterate. A formula in [AMS02] yields

$$\mathrm{grad}\, f(Y) = J_Y^T[F(Y)], \qquad (13)$$

where J_Y^T denotes the adjoint of the operator J_Y (3) defined with respect to the inner product $\langle \Omega_1, \Omega_2 \rangle_X = \mathrm{trace}((X^TX)^{-1}\Omega_1^T\Omega_2)$. Note that the gradient of f, $\mathrm{grad}\, f(Y)$, is defined in the sense of the Grassmann manifold (it is the horizontal lift at Y of the gradient of the scalar field on the Grassmann manifold defined by f). On the other hand, the NG equation (5) reads

$J_Y[\Delta] = -F(Y)$, or equivalently

$$J_Y^T \circ J_Y[\Delta] = -J_Y^T[F(Y)].$$

A continuous deformation between the gradient descent flow of f and the Newton method NG is given by the Levenberg-Marquardt formula

$$(J_Y^T \circ J_Y + \tau\,\mathrm{Id})[\Delta] = -J_Y^T[F(Y)]. \qquad (14)$$

If $\tau = 0$ then Δ is the NG-vector given by the NG equation (4). If τ is large then the direction of Δ is close to the negative gradient of f and the iteration is similar to an Euler integration of the gradient descent flow of f.

Assuming $A = A^T$ and taking Y orthonormal, equation (14) expands to

$$\Pi A\Pi A\Pi\Delta + \Delta Y^T AYY^T AY - 2\Pi A\Pi\Delta Y^T AY + \tau\Delta = -(\Pi A\Pi AY - \Pi AYY^T AY), \quad (15)$$

which has to be solved under the horizontality constraint $Y^T\Delta = 0$. Like in the case of the orginial NG algorithm, one first diagonalizes the small $p \times p$ matrix $A_{11} := Y^T AY$. This decouples (15) into p individual systems of linear equations of the form

$$((\Pi A\Pi - \rho_i I)^2 - \tau I)\delta = -g,\ Y^T\delta = 0. \qquad (16)$$

Solving this equation is expensive even if A is structured, because the projected matrix $\Pi A\Pi$ does in general not preserve the structure. A remedy is to consider instead

$$\Pi((A - \rho_i I)^2 - \tau I)\Pi\delta = -g,\ Y^T\delta = 0, \qquad (17)$$

which can be rewritten in the structured form

$$\begin{bmatrix} (A-\rho_i I)^2 - \tau I & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \delta \\ m \end{bmatrix} = \begin{bmatrix} -g \\ 0 \end{bmatrix}. \qquad (18)$$

In the case where A is tridiagonal, the QR decomposition of $(A - \rho_i I)$ requires $O(n)$ flops. System (18) is then replaced by

$$\begin{bmatrix} R^T R - \tau I & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \delta \\ m \end{bmatrix} = \begin{bmatrix} -g \\ 0 \end{bmatrix}, \qquad (19)$$

where R has three diagonals. Solving system (19) (possibly with iterative refinement to ensure stability) requires $O(np^2)$ flops.

The left-hand sides of (16) and (17) only differ by the term $\Pi AYY^T A\Pi\delta$ which yields a discrepancy between the two solutions δ of (16) and (17) that depends cubically on the distance to the targeted eigenspace. Therefore, a property of local cubic convergence will be valid for both methods. As shown in [ASVM03], the method defined by (17) can be interpreted as a matrix generalization of the homogeneous Newton method studied by Dedieu and Shub [DS00].

Classical strategies for choosing τ involve several parameters that the user can choose at his convenience [DS83, NW99]. In the present case, the very simple choice $\tau := f$ preserves the local cubic convergence and considerably enlarges the basins of attraction around the eigenspaces.

Indeed, local cubic convergence with $\tau = f$ follows immediately: τ is quadratic in the distance between the current iterate $\mathcal{Y}$ and the target eigenspace $\mathcal{V}$, while the right-hand side of (15) is linear in the distance. Consequently, the perturbation on the solution Δ of (15) induced by $\tau = f$ is cubic.

The global behaviour of (15) with $\tau = f$ is illustrated in Figure 3 by the low-dimensional example utilized in Section 3.3. We obtained a similar figure for the modification (17). Comparison with Figure 1 shows that the basins of attraction have been considerably enlarged around the three eigenspaces. The improvement is even more spectacular in the larger dimensional example (11): the modified iterations invariably converged to the targeted eigenspace for the specified initial distances. We had to choose the initial error greater than $\frac{1}{2.2}\frac{\pi}{2}$ in order to observe convergence to a wrong eigenspace.

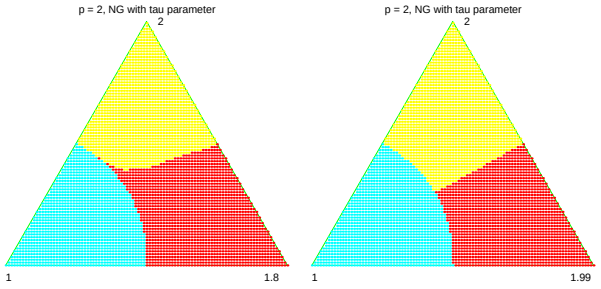

Figure 3: Attraction basins for NG with $\tau := f$ (15) in the case $p = 2$, $n = 3$. Compare with original NG on Figure 1. Local cubic convergence is preserved.

5 Conclusion

The Grassmann-Newton method [EAS98, LST98, LE02, AMS02] was known to coverge locally cubically (under mild assumptions) to the eigenspaces of a symmetric matrix A. In the present paper, the global behaviour of this iteration on the Grassmann manifold has been studied. Using a Levenberg-Marquardt technique, a modified Grassmann-Newton method has been proposed that displays enlarged basins of attrac-

tion while preserving the local cubic convergence and maintaining a low complexity. The paper [ASVM03] presents these results in more details and also considers the case of the multi-shifted and block-shifted Rayleigh quotient iterations.

6 Acknowledgements

We thank M. Petre for his contribution to the Matlab script that generates pictures of basins of attraction. This paper presents research partially supported by the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. This work was partially fulfilled while the first author was a guest at the Mathematisches Institut der Universität Würzburg under a grant from the European Nonlinear Control Network. The work was completed while the first and second authors were visiting the department of Mechanical and Aerospace Engineering at Princeton University. The hospitality of the members of both departments is gratefully acknowledged. The second author thanks N. Leonard and E. Sontag for partial financial support under US Air Force Grants F49620-01-1-0063 and F49620-01-1-0382.

References

[Abs03] P.-A. Absil, *Invariant subspace computation: A geometric approach*, Ph.D. thesis, Faculté des Sciences Appliquées, Université de Liège, Secrétariat de la FSA, Chemin des Chevreuils 1 (Bât. B52), 4000 Liège, Belgium, 2003.

[ADM+02] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub, *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA J. Numer. Anal. **22** (2002), no. 3, 359–390.

[AMS02] P.-A. Absil, R. Mahony, and R. Sepulchre, *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, submitted to Acta Appl. Math., November 2002.

[AMSV02] P.-A. Absil, R. Mahony, R. Sepulchre, and P. Van Dooren, *A Grassmann-Rayleigh quotient iteration for computing invariant subspaces*, SIAM Review **44** (2002), no. 1, 57–73.

[ASVM03] P.-A. Absil, R. Sepulchre, P. Van Dooren, and R. Mahony, *Cubically convergent iterations for invariant subspace computation*, submitted to SIAM J. Matrix Anal. Appl., January 2003.

[BS89] S. Batterson and J. Smillie, *The dynamics of Rayleigh quotient iteration*, SIAM J. Numer. Anal. **26** (1989), no. 3, 624–636.

[Cha84] F. Chatelin, *Simultaneous Newton's iteration for the eigenproblem*, Computing, Suppl. **5** (1984), 67–74.

[Dem87] J. W. Demmel, *Three methods for refining estimates of invariant subspaces*, Computing **38** (1987), 43–57.

[DS83] J. E. Dennis and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice Hall Series in Computational Mathematics, Prentice Hall, Englewood Cliffs, NJ, 1983.

[DS00] J.-P. Dedieu and M. Shub, *Multihomogeneous Newton method*, Math. Comp. **69** (2000), no. 231, 1071–1098.

[EAS98] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl. **20** (1998), no. 2, 303–353.

[JS01] Z. Jia and G. W. Stewart, *An analysis of the Rayleigh-Ritz method for approximating eigenspaces*, Math. Comp. **70** (2001), no. 234, 637–647.

[LE02] E. Lundström and L. Eldén, *Adaptive eigenvalue computations using Newton's method on the Grassmann manifold*, SIAM J. Matrix Anal. Appl. **23** (2002), no. 3, 819–839.

[LST98] R. Lösche, H. Schwetlick, and G. Timmerman, *A modified block Newton iteration for approximating an invariant subspace of a symmetric matrix*, Linear Algebra Appl. **275-276** (1998), 381–400.

[NW99] J. Nocedal and S. Wright, *Numerical optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.

[Par80] B. N. Parlett, *The symmetric eigenvalue problem*, Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632, 1980, republished by SIAM, Philadelphia, 1998.

[PP73] B. N. Parlett and W. G. Poole, *A geometric theory for the QR, LU and Power Iteration*, SIAM J. Numerical Analysis **10** (1973), no. 2, 389–412.

[PW79] G. Peters and J. H. Wilkinson, *Inverse iteration, ill-conditioned equations and Newton's method*, SIAM Review **21** (1979), no. 3, 339–360.

[Smi94] S. T. Smith, *Optimization techniques on Riemannian manifolds*, Hamiltonian and gradient flows, algorithms and control (Anthony Bloch, ed.), Fields Institute Communications, vol. 3, American Mathematical Society, 1994, pp. 113–136.

[WE91] D. S. Watkins and L. Elsner, *Convergence of algorithms of decomposition type for the eigenvalue problem*, Linear Algebra Appl. **143** (1991), 19–47.